
**Report on the 2006 Assessment and
Harvest Policy of the International
Pacific Halibut Commission**

**NIWA Client Report: WLG2007-55
July 2007**

NIWA Project: MIA07301

Report on the 2006 Assessment and Harvest Policy of the International Pacific Halibut Commission

Author
R I C C Francis

Prepared for

University of Miami

NIWA Client Report: WLG2007-55
July 2007

NIWA Project: MIA07301

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point, Wellington
Private Bag 14901, Kilbirnie, Wellington, New Zealand
Phone +64-4-386 0300, Fax +64-4-386 0574
www.niwa.co.nz

Contents

Executive Summary	iv
1. Background	1
2. Review Activities	1
3. Findings	1
3.1 Data	2
3.2 Model structure	6
3.3 Log likelihood	8
3.4 Alternative fits	12
3.5 Area apportionment	12
3.6 Harvest policy	14
3.7 Uncertainty	18
4. Conclusions	21
4.1 The seven questions	21
4.2 Suggestions for any future workshops	22
APPENDIX 1: Statement of Work	23
APPENDIX 2: Materials Provided	25

Reviewed and Approved for release by:

Dr Rosemary Hurst

Executive Summary

A public workshop was held in Seattle, Washington, June 27-28, 2007 to consider the 2006 stock assessment and harvest policy of the International Pacific Halibut Commission (IPHC). Presentations were made on both the assessment and harvest policy, additional analyses were requested and carried out, and the workshop discussed the results. Following the workshop, the independent reviewers met informally with IPHC staff to discuss technical issues arising from the workshop.

The workshop was well run and IPHC staff were clear in their presentations and helpful in their response to queries.

The following recommendations are offered. High priority should be given to completion of the analysis of the PIT tag data, with the aims of explaining the low recapture rates and improving estimates of migration rates. Future assessment modelling should be spatially structured so that 1) the migration estimates can be used in the assessment, 2) all data can be used in the area apportionment, 3) various current assumptions can be tested, and 4) the interaction between migration and harvest policy can be investigated. There is little evidence to support the survey-based area apportionment procedure, but there is no obvious alternative to it without a spatially structured assessment model. The Commission should consider modifying its harvest policy in response to new information about adult migration. Some suggestions are made which may improve some aspects of the pre-model data analysis, log likelihood, estimation of uncertainty, and any future workshops.

1. Background

This report reviews, at the request of the University of Miami (see Appendix 1), the 2006 assessment and harvest policy of the International Pacific Halibut Commission (IPHC). The author was provided beforehand with links to various documents (Appendix 2) and participated in the workshop which considered the assessment and harvest policy.

2. Review Activities

A public workshop was held at the Nexus Hotel in Seattle, Washington, on June 27-28, 2007 to consider the 2006 stock assessment and harvest policy of the IPHC. The workshop was chaired by Dr Steve Martell (University of British Columbia) and those attending included IPHC staff, several IPHC Commissioners, fishing industry representatives, people from various U.S. and Canadian fisheries agencies and universities, and two external reviewers (including the author).

IPHC staff made a series of presentations to the workshop concerning both the stock assessment and harvest policy. One presentation on the second day reported on additional analyses carried out in response to matters arising on the first day. After the workshop, the two reviewers met informally with IPHC staff to discuss some technical issues arising from the workshop.

3. Findings

Overall, I was impressed by the documents I read, the structure of the workshop and presentations by IPHC staff, and the courtesy and helpfulness of these staff in responding to my queries and those of other workshop participants. They showed high professional standards, as well as strong initiative in detecting and responding to the thorny problem of adult migration.

All my comments and suggestions below should, of course, be considered subject to qualification or rebuttal by IPHC staff. There is a tremendous volume of detail associated with the data and analyses that I was asked to consider, and it was clearly not possible for me to comprehend it all in the time available. I hope that I have correctly identified, and understood, all important details, but I may well have missed some.

The seven parts of this section correspond to the seven questions that I was asked to address (see Appendix 1), which are repeated at the beginning of each subsection. In

some places I refer to the documents I was provided by their file names, as given in Appendix 2 (e.g., sa06.pdf, 2k6rara04.pdf).

3.1 Data

Are the stock assessment data adequate? If not, what more is needed?

Five sets of data were used in the assessment model: 1) catch numbers at age and sex; 2) catch per unit effort (CPUE) from the commercial setline fishery; 3) CPUE from the setline survey; 4) proportions at age and sex from the setline survey; and 5) numbers at length in the bycatch.

My impression, after reading the documents and seeing the workshop presentations, is that that these data, and the procedures followed in collecting and processing them, are of the high standard one would expect from an internationally recognised agency like the IPHC. In common with many other agencies, the IPHC found that their earlier age data (from otolith surface readings) were biased and imprecise, but they have dealt with this well using a misclassification matrix.

The most important of the assessment data sets are those from the setline survey. Because they are fishery-independent they are extremely valuable as assessment tools, and I commend the IPHC for its continuing commitment of substantial resources to the survey. I was impressed with the effort IPHC staff devoted, both in design and data analysis, to ensure the quality of the survey data by standardising effort and rejecting stations whose catch rates were affected by predation. During the workshop it was suggested that the grid design of the survey may bias results, and that randomly located stations would be better. I acknowledge this weakness but am, on balance, in favour of the grid design for this survey. Given the large number of stations, any bias is likely to be small, and the grid design reduces variance and provides substantial logistical advantages, and also some analytical advantages which I will discuss in the next section.

3.1.1 Possible data refinements

I have some suggestions concerning refinements in the analysis of the survey data and commercial CPUE. I suspect that these will have little effect on the stock assessment, although a large effect is possible. Even if their effect is small, what is to be gained is a greater confidence that any important patterns in the data are not being overlooked.

With regard to the survey data I have two suggestions. First, the survey area should be stratified for the analysis of both the CPUE indices and proportions at age and sex. I can see nothing to be lost from stratification, and there are two possible gains: removal of any bias that may arise when the number of stations is not exactly proportional to stratum area (slide 11 of Bill Clark's area apportionment talk showed that this is important in Area 2A), and reduction in variance (if there are significant between-stratum differences in catch rates). Depth contours are obvious candidates for stratum boundaries, but I would suggest also using lines perpendicular to the shoreline (e.g., some or all of the lines shown in slide 6 of Claude Dykstra's presentation).

My second suggestion is to carry out an analysis of all the survey data (as suggested by Paul Medley) with the aim of identifying any factors (e.g., time of day, tidal phase, vessel, etc) that might affect catchability. The aims of this analysis are twofold: to increase understanding of setline catchability, and (possibly) improve the survey data used in the stock assessment. For this analysis, the grid design is an advantage because it gives us more power to detect such factors. What I have in mind is something like the following. Define

$$Y_{sy} = \frac{X_{sy} / M_{sy}}{\text{mean}_{y'} (X_{sy'} / M_{sy'})}$$

where X_{sy} is the catch rate from station s in year y and M_{sy} is the mean catch rate from all stations in year y that are in the same stratum as station s . Dividing by M_{sy} removes any effects of large-scale spatial and temporal changes in halibut abundance; the other division removes any effect due to the individual station location (the latter is possible only because of the grid design of the surveys). Now use a GLM (or GAM) to look for factors that explain variation in the Y_{sy} , and thus in halibut catchability. The use of standard gear and fishing methods, and the large number of year-station combinations available for this analysis, make this a powerful data set for such an analysis.

For the commercial CPUE data I suggest standardisation using the usual methods, which are described, for example, in several papers in the 2004 review issue of *Fisheries Research* (volume 70, issues 2-3). I am aware that the need for such standardisation in the halibut assessment is limited because the associated catchability is allowed to vary with time. However, it seems to me a bad practice not to remove the effect of any factors that can be shown to affect commercial catchability. It is a difficult task for our assessment models to extract the abundance signal from the considerable noise that is typical in fisheries data, and we should do all we can to help them. Further, such standardisation may improve our knowledge of factors affecting catchability.

Finally, I would like to express my reservations about the use of area weighting to combine the area-specific commercial CPUE indices for use in the coast-wide model. I do not mean to suggest that there is a clearly better way of combining these data. Nor do I want to imply that this method of combination has not been often used in the fisheries literature. I simply wish to point out that area weighting requires two strong assumptions: that the CPUE in the a th area, X_a , is proportional to the density of fish in that area (i.e., $X_a = q_a B_a / A_a$, where B_a is the biomass in that area and A_a is its area – in n.m.², say) (we usually assume X_a is proportional to the *biomass*, rather than the *density*); and that the constant of proportionality, q_a , is the same in all areas. One of my reasons for advocating a different model structure (see Section 3.2.1) is that it both avoids the necessity for these assumptions and allows us to test them.

3.1.2 PIT tag data

The data from the current PIT tag experiment were not used directly in the stock assessment. However, they provided preliminary information on migration which was crucial in the decision to recommend a change in the form of the assessment, from closed-area to coast-wide, with the key result being evidence that a non-trivial proportion of halibut continue to migrate long after recruiting to the fishery. I will argue below that this result also has important implications for the harvest policy (see Section 3.6.3).

I was impressed by the design and scope of this experiment. I think it is comparatively rare that such experiments achieve the ideal of tagging in each area in proportion to abundance as well as the present one. This is particularly important when we see that a high percentage of recaptures from the survey occurred near the tagging location.

There is a pressing need to understand why recapture rates were so low, so that estimated fishing mortalities from the tagging data were only a fraction of those from the assessment. Of particular concern is the possibility that whatever phenomenon is responsible for this under-estimation may bias estimates of migration. However, I do not know how likely this is.

I have a couple of comments on the analysis that attempted to explain the low recapture rates by “small-scale mismatches between the distribution of released tags and the distribution of commercial fishing” (p. 140, 2k6rara04.pdf). First, I do not see how this hypothesis explains low recapture rates. My thinking is that if the tagged fish are restricted to, say, a quarter of the area, then they are available to only a quarter of commercial sets, but within the area they occupy their average density is four times what it would have been had they been distributed across the entire area. Don’t the

factors of one quarter and four balance out, so that the expected number of tagged fish caught is the same as it would have been had they not been restricted in area? What the mismatch hypothesis *does* imply is that the distribution of numbers of tags per landing should be contiguous, and it is only this which is being tested in the analysis of Table 1 (p. 141, 2k6rara04.pdf). With regard to this analysis, either I have misunderstood it (quite possible!) or the expected numbers in this table are wrong, and the difference between observed and expected numbers is greater than it appears. Here, for the record, are my calculations of the expected numbers for area 2B. I calculated the mean for the Poisson distribution in this area as the total number of tags recovered ($242 = 108 + 31 \times 2 + 13 \times 3 + 2 \times 4 + 5 \times 5$) divided by the total number of landings ($727 = 568 + 108 + 31 + 13 + 2 + 5$). Using the Splus function `dpois`, I calculated the expected numbers in the table as `round(727 * dpois(0:5,242/727))`, which gives values of 521 173 29 3 0 0. In each area my estimates of the expected number of landings with no tags should be over-estimates, because I treated the observed landings in the column '5+' as containing exactly 5 tags. However, my estimates of these numbers were always less than those in the table.

With regard to the analysis of migration rates, I am concerned that the present analysis, which is on the scale of the regulatory areas, may be biased because it ignores the substantial spatial heterogeneity in scanning rates (see Figure 1, p. 142, 2k6rara04.pdf). I wonder if the analysis could be done on a smaller spatial scale. One approach would be to estimate the following parameters: the probability that an individual will migrate within the next year, and, for those fish that do migrate, the mean and standard deviation of the eastward distance migrated. These parameters could be treated as smooth functions of age and/or length and/or position around the coast. They give a more biologically meaningful description of the migration than does a matrix of between-regulatory area migration rates. That matrix can, of course, be calculated from the above parameters, together with an initial population distribution from the tagging surveys.

There's an important distinction to be made between migration rates, and their effects on the underlying population. For example, preliminary estimates of annual percentage rates of migration between regulatory areas (Table 4, p. 136, 2k6rara04.pdf) suggest approximately equal rates of migration from 3A to 3B (7.7%) and vice versa (9.9%). However, we cannot know the net effect of these migrations without knowing what the pre-migration distribution of the population was. According to both of two alternative estimates of this distribution, the net effect is that Area 3A is a source of fish, whereas 3B is a sink (Table 1). From a management point of view, I think the net effects shown in this table are more important than the migration rates. Table 1 also illustrates the point that the migration may be more complicated than a simple eastward movement (in which Area 3B would not be a

sink). I do not mean to suggest here that the migration *is* complicated in this way, because I know that the migration rate estimates on which this table are based are very preliminary. However, it does not do any harm to be reminded that nature is often (usually?) more complicated than we might like to think it is, and halibut migration may not be exclusively eastward.

Table 1: Annual percentage change in population size assuming the migration rates estimated from PIT tag data (see Table 4, p. 136, 2k6rara04.pdf) and two alternative estimates of the pre-migration distribution of the population (from Table 1, sa06.pdf).

Source of pre-migration population distribution	Area				
	2B	2C	3A	3B	4A
Closed-area assessment in 2006	13	-5	-5	18	-9
Survey area apportionment in 2006	14	-2	-3	6	-11

3.2 Model structure

Is the structure of the assessment model appropriate? If not, what changes should be made?

Two alternative models were used in the 2006 assessment: closed-area and coast-wide. These models differed only in their input data; their structures were essentially the same. I have no criticisms of this common structure. It dealt well with the available data and allowed the modeller a great deal of flexibility in investigating the shape of selectivity curves and how these curves, and the associated catchabilities, might vary over time.

IPHC staff presented some good arguments against the use of the closed-area model. For example, there is the 2C/3B paradox: “Area 3B is twice the size of area 2C and has a higher survey CPUE, but the [closed-area] assessment says there is more biomass in Area 2C” (p. 83, 2k6rara04.pdf). I am happy to accept these arguments. The problem for me (and for the IPHC, I believe) is that the coast-wide model requires some way of apportioning the estimated current biomass amongst the regulatory areas. It is important to distinguish between accepting the coast-wide model over the closed-area, and accepting the area-apportionment scheme. I know that the IPHC staff are well aware of this distinction but I am not confident that this was true of all workshop participants.

3.2.1 A spatially-structured model

I would like to describe, and advocate, a model that is intermediate between the closed-area and coast-wide models. Its advantages over these models are: 1) it uses information from the PIT tag data; 2) it does the area apportionment in a way that is consistent with both current and historical biomass distributions as well as migration rates; 3) it provides a method of testing various assumptions, including the key one underlying the current area-apportionment scheme (area-independent survey catchability); and 4) it provides a tool to deal with the interaction between migration and harvest policy. Its main disadvantage is that it is markedly more complex, and it might take a year or two before it is stable.

The spatially-structured model divides the population amongst the regulatory areas. The five data sets currently used in the existing models (see Section 3.1) would enter the model separately by area. Thus, for example, there would be a survey and commercial CPUE series for each area. The population dynamics for each regulatory area (i.e., the equations describing how the number of fish in a cohort decline from year to year because of fishing and natural mortality) would be exactly as they are in the existing models with one addition: equations describing movement between areas.

I would expect that migration parameters would be estimated outside the model and then fixed in the model. In this situation, the parameters estimated would be exactly the same as were estimated in the closed-area models (see pp 14-15, sr83.pdf). In fact, we could mimic a simultaneous run of all the closed-area models by setting all migration rates to zero (a useful test of the new model). But this is only the starting point; we can do much better than this with the spatially-structured model. The current closed-area models necessarily estimate different survey and commercial catchabilities and selectivities in each area. With the spatially-structured model we can test a series of hypotheses about between-area differences in these quantities. The obvious hypothesis to test (because it underlies the current area apportionment) is that survey catchability is the same in all areas. If we found (as I would expect) that this hypothesis was rejected when there was no migration, we could ask the question what rates of migration do we need to assume so that this hypothesis is not rejected, and are these migration rates consistent with the PIT tag data? (Bill Clark briefly described – but did not present – informal analyses he did with the current models aimed at addressing precisely this question, treating migration as a change in natural mortality; I believe the spatially-structured model would provide a much more robust, and theoretically sound, approach to this problem).

The spatially-structured model avoids a great weakness of the coast-wide model: its inability to use the considerable quantity of area-specific information. Because

management has been by area we might expect that population trends would be different in the different areas. If the current area apportionment is correct these differences would be large, because recent exploitation rates have been very different (see Fig. 6, p. 157, 2k6rara04.pdf). Are these exploitation rates consistent with the area-specific data on catches and catch rates? The coast-wide model provides no way of answering this question, but the spatially-structured model does. Perhaps more importantly, with the spatially-structured model the area apportionment is an integral part of the assessment, and uses all available data. With the coast-wide model, the area apportionment is external to the model, uses only part of the data (the survey CPUE), and, in its current form, requires an untested assumption (area-independent catchability).

The spatially-structured model is also an obvious tool to investigate more fully the effect of migration on the harvest policy, which I discuss in Section 3.6.3.

3.3 Log likelihood

Is the log likelihood used to fit the model appropriate? If not, what should be used?

The objective function lies at the heart of all modern stock assessment models. This function is of great importance because it controls the estimation of parameters (i.e., the model fitting): by definition, the final parameter estimates are those which minimise the value of the objective function. Two types of terms were added together to make the objective function used in the halibut assessment: log likelihoods and penalties (in Bayesian models, a third type – prior distributions – is also used).

The penalty terms used in the halibut assessment seemed to me reasonable and relatively uncontroversial, and so I will say no more about them. Nor will I say anything about the robustification function used to reduce the effect of outliers (Fig. 11, p. 54, sr83.pdf), except that I think it is a good idea.

The role of the log likelihood terms in the objective function is twofold. First, they provide, for each individual observation, a measure of the difference between the observed value, X_{obs} , and the value predicted by the model, X_{model} , given a set of parameter values. The form of this measure is determined by the statistical distribution assumed for each observation. In the halibut model all observations were assumed to be normally distributed, so the measure of difference was always $(X_{\text{model}} - X_{\text{obs}})^2$. The second role of the likelihood terms is to assign a weight to each individual observation, which is a way of telling the model how close X_{model} should be to X_{obs} . A high weight encourages the model to find parameter values which make X_{model} very

close to X_{obs} . Most of what I have to say about the halibut log likelihood concerns these weights.

In the halibut assessment these weights were assigned in a three-step process. First, a sampling standard deviation, s , was either calculated for, or assigned to, each observation using a variety of methods (see pp 8-9, sr83.pdf). This assigned a weight of $1/s^2$ to the observation. The model was then fitted using these weights and root mean square errors, τ , were calculated for each type of observation (see Table 1, p. 41, sr83.pdf). The weight assigned to each observation was then changed to $1/(s\tau)^2$, and the model was run again with the new weights. If the resulting fit was considered inadequate for some type of observation, an arbitrary additional factor W was assigned to it to encourage a better fit to it (e.g., a factor of 10 was used for survey and commercial CPUE observations in some assessments – see p. 18, sr83.pdf). Exactly how this factor was applied in the likelihood function was not described, but I believe the effect was to change the weight on these observations to $W/(s\tau)^2$.

Data weighting is often a difficult and controversial issue in stock assessments. It is controversial because different weightings can produce quite different assessment results. It is difficult because it is not possible to provide an objective set of rules that will guarantee the best weighting in all assessments, and so it is hard to avoid a subjective component to the weighting decisions. Subjective weightings lead to the undesirable possibility that the results of assessments carried out by different modellers (with different subjective judgements) could be quite different.

I have several suggestions that I think could improve the halibut assessment by making the data weighting more theoretically (statistically) sound and less subjective. These suggestions may or may not make a significant change to the assessment results, but I think they could make the assessment statistically more sound, and thus more defensible.

3.3.1 Suggested changes to the log likelihood

The approach I suggest is based on viewing the total error associated with each observation, $e_{\text{total}} = (X_{\text{model}} - X_{\text{obs}})$, as being the sum of two parts: $(X_{\text{true}} - X_{\text{obs}}) + (X_{\text{model}} - X_{\text{true}})$, where X_{true} is the true value (in the real world, not in the model) of the observed quantity. The first part is what is usually called the observation error, and the second I call process error (some people use the term model error), so we can write $e_{\text{total}} = e_{\text{obs}} + e_{\text{proc}}$. The important point to notice is that although many types of observation contain information about e_{obs} they almost never, by themselves, contain any information about e_{proc} . For example, for survey CPUE we can estimate e_{obs} from the between-station variation in catch rates, but these catch rates contain no information (by

themselves) about the random year-to-year variations in survey catchability which I believe form the main contribution to e_{proc} . (I am assuming here that the model assumes no year-to-year variation in this catchability). For the survey proportions at age (and sex), the observations contain abundant information about e_{obs} , but not about e_{proc} , which is often very substantial. One obvious source of process error for this type of observation is the common model assumption that natural mortality does not vary with either year or age. Depending on the type of observation and the assumed sampling distribution, we can describe the size of the two types of errors using either standard deviations (s.d.s, s_{obs} and s_{proc}) or coefficients of variation (c.v.s, c_{obs} and c_{proc}). Whichever approach we use, these quantities add as squares: $s_{\text{total}} = (s_{\text{obs}}^2 + s_{\text{proc}}^2)^{0.5}$ and $c_{\text{total}} = (c_{\text{obs}}^2 + c_{\text{proc}}^2)^{0.5}$. If we apply this view of error structure to the halibut log likelihood it is easy to see that, at least when no additional W factors are used, the halibut s corresponds roughly to my s_{obs} , and τ is associated with s_{proc} .

My first suggestion to improve the log likelihood is to consider changing from the current multiplicative model, $s_{\text{total}} = s_{\text{obs}}\tau$ (or $s_{\text{total}} = s_{\text{obs}}\tau/W^{0.5}$), to a more theoretically sound (I think) additive model – either $s_{\text{total}} = (s_{\text{obs}}^2 + s_{\text{proc}}^2)^{0.5}$ or $c_{\text{total}} = (c_{\text{obs}}^2 + c_{\text{proc}}^2)^{0.5}$. Note that for a given set of observations (e.g., survey proportions at age) there will usually be a different observation error for each individual observation, but a single process error that will be added to all observations within that set. This means that the effect of changing to additive errors will be greatest for sets of observations within which the variation in observation error is large (e.g., proportions, or numbers, at age or length). An example will help to show how different the additive and multiplicative approaches can be. Suppose we base our observation errors for a set of proportions at age on a multinomial distribution with $N = 10\,000$, and the proportions vary from 0.001 to 0.1, so c_{obs} will vary from 0.316 to 0.030, respectively. If c_{proc} is 0.3 (a typical value for my own assessments) then c_{total} will vary from 0.436 to 0.301. Thus, the effect of adding this process error is very substantial for the high proportions (the c.v. increases from 0.030 to 0.301) and much less so for the low proportions (c.v. increase from 0.316 to 0.436). With multiplicative errors and $\tau = 2.5$, the corresponding changes are much greater for the low proportions (from 0.316 to 0.790) and less for the high proportions (from 0.03 to 0.075).

My second suggestion is to check which error model is best. Regardless of any theoretical grounds, the best justification for any approach to modelling error distributions is to show that it is consistent with the data and model. This is fairly easily done with residual plots. For example, with the proportions at age observations, I would standardise all the residuals so that their expected s.d. (according to the assumed error model) is 1, plot the absolute residuals against log proportion, and fit a smooth line through the plotted points. This line should be approximately horizontal if the error model is appropriate. Such an approach can be used to decide between

multiplicative and additive errors, and, if additive errors are used, whether to base these on s.d.s or c.v.s.

My third suggestion is to fix the errors for those sets of observations that you wish to be sure of fitting well (primarily the survey CPUE, possibly also the commercial CPUE) and estimate process error within the model for all other sets of observations (i.e., include a c_{proc} or s_{proc} as an estimable parameter for each set – or a τ , if you can show that multiplicative errors are better). I thought that the use of smoother to estimate a total c.v. for halibut commercial CPUE (p.9, sr83.pdf) was an excellent idea, and suggest doing the same for the survey CPUE. I acknowledge there is a subjective element in this, because you have a choice of how smooth to make your smooth line. However, you can at least be upfront about this by including in your assessment document a plot showing the smooth line you fitted to each CPUE data set. That way, it is easy for a reviewer to make a judgement as to whether they believe your line is too smooth or not smooth enough (this is essentially a judgement about how smooth we should expect biomass trajectories to be). Once the model has been fitted you should compare the variance of the CPUE residuals with that predicted by your assumed errors. If this variance is too large it is an indication of a conflict between data sets, or between them and model assumptions.

When faced with this sort of evidence of conflict there are three main responses. The first choice (though often difficult) is to change the model assumptions until the conflict disappears. The second choice is to try to determine which data sets are in conflict and try model runs in which some data sets are omitted. Failing all else, up-weight those observations that you are failing to fit (by reducing s_{total} or c_{total}) and run the model again. The aim is to add sufficient weight to these observations so that the variance of the residuals is about what should be expected according to your original assumed errors (not the up-weighted errors). This up-weighting is not desirable, but at least there is a reasonably objective means of deciding how much additional weight to apply.

My fourth suggestion is to avoid double entering of observations (including total proportions at age, as well as proportions at age for males and females; including CPUE in number as well as in weight; and including CPUE at age, as well as total CPUE). I think I understand why this was done (as an attempt to make sure of good fits to all aspects of the data) but it seems to me quite indefensible from a statistical point of view and it interferes, in a way that I find hard to predict, with the estimation of uncertainty (see Section 3.7).

My final suggestions concern the composition observations (proportions or numbers at age or length). These observations are not derived from simple random samples, so

the use of the multinomial distribution to calculate s_{obs} or c_{obs} is inappropriate. For example, the calculation of proportions at age from the setline survey involves two-stage sampling – stations within a stratum, and otoliths within a station – not a simple random sample. A more reliable way of estimating observation c.v.s (or s.d.s) is to use resampling methods on the raw data. Also, I wonder if a normal error distribution is appropriate for the composition data. A normal distribution with a c.v. exceeding about 0.4 will include a reasonable proportion of negative numbers, which do not make sense for composition data. In the composition data I'm used to dealing with c_{total} often exceeds 0.4, so I use a lognormal distribution, rather than a normal. I note that plots like Fig. 10 in sr83.pdf have been used to justify the use of the normal distribution. My guess, but I could be wrong, is that these plots are a mixture of fairly symmetric distributions, for the larger proportions, and positively skewed distributions for the smaller proportions.

3.4 Alternative fits

Is the suite of alternative fits adequate?

I think that the suite of alternative fits presented in the assessment (Table 2, p. 12, sa06.pdf) was perfectly adequate to explore the range of plausible fits to the coast-wide model.

I liked the use of retrospective runs as a diagnostic tool but note that it can be difficult to decide whether a retrospective trend is of concern (indicating a serious fault in the model structure) or simply what could be expected because of a chance pattern of residuals in the observations. Sets of retrospective runs based on simulated data can provide some help in making this decision.

3.5 Area apportionment

Is the area apportionment procedure correct?

For me, this was the weakest part of the assessment.

I do not mean this statement to be a criticism of the IPHC staff, because the approach they took had a logic to it that is hard to escape. They started with the model that had been used previously (the closed-area model) and found that this was inconsistent with the recent information about migration (from the PIT tags). There were also some internal inconsistencies (e.g., the “2C/3B paradox” – see above). Their decision then to switch to a coast-wide model seems utterly reasonable (I doubt that it would have

been possible to develop a more complex model, like that in Section 3.2.1, within the time available). Having done that, they needed a method to apportion the total exploitable biomass, as estimated from the coast-wide model, amongst the regulatory areas. The choice they made was the obvious one – to use the fishery-independent survey data – and this required some assumption about the survey catchability in each area. The usual scientific approach is to apply Occam’s razor. That is, to make the simplest assumption that is consistent with the data. In this case that assumption is that the survey catchability is the same in all areas. This is not because there was strong (or even much) evidence of equality. Rather it is because there was no plausible way to estimate separate catchabilities by area.

What is the best decision from a scientific point of view is not necessarily the best to use in managing a fishery. I think the Commissioners were prudent to reject the new assessment because it would have led to a substantial change in the allocation of quota amongst regulatory areas. While there were reasonably good grounds to believe that the previous method of allocation (using the closed-area assessment) was flawed, the evidence to support the new method was weak. In such a situation a pragmatic approach is to stay with something like the status quo until the scientific picture becomes clearer.

I have four reasons to doubt the assumption of area-independent survey catchability, though none is strong, and none leads to a clear alternative assumption. The first derives from the area-specific recapture rates for PIT tags in the 2006 survey (bottom part of Table 3, sa06.pdf). A simple chi-square test applied to this table shows a highly significant departure from the assumption of equal catchability ($P = 0.005$) [I could not understand the assertion that this was only “marginally significant” (bottom of p. 9, sa06.pdf)]. The second reason is the comparison between setline and trawl CPUE in Areas 3A, 3B, and 4A (Fig. 3, prospect.pdf). This was used by IPHC staff to reject the closed-area model, which implied that setline survey catchability was much lower in Area 3A than in 3B and 4A (Fig. 1, prospect.pdf). However, the same data also rejects the assumption of area-independent catchability because, for example, it indicates that setline catchability in Area 3B was 50-60% of that in 4A (the fact that the plotted 3B points in Fig. 3 are lower than those for 4A in all eight length classes makes this difference statistically significant regardless of the associated confidence intervals). I do not consider this strong evidence, because it requires the assumption that trawl catchability is the same in 3A, 3B, and 4A, and I am less willing to make that assumption than the IPHC staff seem to be. The third reason is that the recent Area 2 exploitation rates estimated using the assumption of area-independent survey catchability are very high (see Fig. 6, p. 157, 2k6rara04.pdf), particularly in Area 2B. It seems hard to believe that the effects of such high exploitation rates would not have been noticed before now, particularly by fishers. Finally, I note that catchability is a

function of both fisher behaviour and fish behaviour. Much is done to ensure that fisher behaviour is the same throughout the survey area, but it is not possible to standardise fish behaviour. As the environment, biotic and abiotic, changes from Oregon via the Pacific coasts of British Columbia and Alaska to the Bering Sea, I would expect halibut behaviour to change in response, and this change may well affect its catchability.

3.6 Harvest policy

Is the harvest policy appropriate; i.e., does it strike a proper balance between utilization and precaution? If not, how should it be modified?

The IPHC management of the halibut fishery in recent years has been based on what is called the Constant Harvest Rate (CHR) policy. The aim of this policy has been to apply the same harvest policy in all areas. In fact, the actual harvest rate has not been constant, either in time or across areas. There are two main reasons for this. First, the target harvest rate has changed over time (varying between 0.20 and 0.35 since 1985) as understanding of the stock dynamics has evolved. This is to be expected, and is not unreasonable. Second, in setting quotas for each regulatory area in each year, various adjustments have been made to the harvest rates. I will comment first on the calculation of target harvest rates, and then on the adjustments to these rates. Finally, I will discuss the effect of migration on this harvest policy.

3.6.1 Calculation of the target harvest rate

I was impressed by the simulation approach used to calculate the target harvest policy (pp 30-36, sr83.pdf). This explicitly allows for periodic regime shifts affecting productivity, density-dependent growth, length-specific selectivity, and time-invariant maturity at age. As a sensible conservative measure, parallel simulations are done assuming that the current low growth rates will persist, even if biomass is reduced. Four performance measures are calculated for each a range of possible target harvest rates: average annual yield, average spawning biomass, average actual harvest rate, and the proportion of years the spawning biomass falls below a threshold (Table 5 and Fig. 29, sr83.pdf). The only potential weakness I could see in these simulations is that they did not allow for various adjustments in harvest rates (i.e., they assumed that the actual harvest rate was the same as the target harvest rate, except for random errors in estimating current biomass).

As is common in simulations of this type, the procedure for choosing the 'best' target harvest rate from the results of these simulations involved a trade-off between utilization (as measured by the realised yield) and precaution (as measured by the

probability of falling below a biomass threshold). The risk associated with falling below the biomass threshold is an increase in the probability of recruitment failure. The larger the target harvest rate the greater this risk is, but the higher the average yield (assuming no recruitment failure).

I interpret the question at the head of this section as asking whether the IPHC staff found the optimal trade-off between utilization and precaution in choosing the target harvest rate from the simulation results. I can offer no opinion on this because what is optimal depends on how risk-averse the Commission wishes to be, and I have no information about that.

3.6.2 Adjustments to the harvest rate

In the documents I saw I found evidence of three types of adjustment that have, or could have, been made to target harvest rates. The first relates to the reduction that is supposed to occur when the spawning biomass falls below the threshold and/or limit reference points (as illustrated in Fig. 27, p. 73, sr83.pdf). This seems to me a very sensible, and prudent, type of adjustment, though it is a moot point as to whether it is actually part of the IPHC harvest policy (as opposed to the simulation experiment of Section 3.6.1) because, as far as I am aware, it has never been applied.

The second type of adjustment is when IPHC staff either reduce a harvest rate because the assessment is uncertain or there is reason for concern in a particular regulatory region (e.g., the reduction from 0.20 to 0.15 in Areas 4B and 4CDE in 2006), or increase it to reduce the effect of an otherwise large reduction in quota (e.g., the increase from 0.20 to 0.25 in Area 2 in the 2006 coast-wide assessment). While I can understand the reason for making such adjustments I am concerned by their ad hoc nature, particularly when a harvest rate is increased. I could see no reason why the increase in Area 2 should have been to 0.25, and not 0.3, or 0.225.

The third type of adjustment is as a result of the “slow up-fast down” (SUFDD) policy (p. 149, 2k6rara04.pdf). I found no precise definition of this policy in the documents I read, but I understood it to mean that the agreed yield for year y , $Y_{\text{agreed},y}$, would be calculated from the recommended yield, $Y_{\text{recommended},y}$, using an algorithm something like the following

$$Y_{\text{agreed},y} = \begin{cases} 0.33Y_{\text{recommended},y} + 0.67Y_{\text{agreed},y-1} & \text{if } Y_{\text{recommended},y} > Y_{\text{agreed},y-1} \\ 0.5Y_{\text{recommended},y} + 0.5Y_{\text{agreed},y-1} & \text{if } Y_{\text{recommended},y} < Y_{\text{agreed},y-1} \end{cases}$$

I was unsure about the status of this policy. I could not understand why it appeared to play no part in the calculation of the yields recommended after the 2006 assessment (Table 1, sa06.pdf). Also, it was unclear to me why, given the existence of this policy, it was deemed necessary to increase the recommended harvest rate in Area 2 from 0.20 to 0.25.

It concerned me that only the first of these three types of adjustment was allowed for in the simulations to calculate target harvest rates. If SUFD is formally part of the IPHC harvest strategy I can see no reason for not including it in these simulations. Perhaps it was not included because it is only intermittently applied? As long as the second type of adjustment is ad hoc, it will not be possible to include it in the simulations. I am not bothered by downward adjustments, because the effect of these is precautionary. However, upward adjustments mean that there is more risk to the stock than is indicated by the simulations which underlie the choice of target harvest rate.

3.6.3 Effect of migration on the harvest policy

If halibut migration rates turn out to be of the order indicated by preliminary estimates from PIT tags the Commission may want to consider modifying their harvest policy. The modification could reduce harvest rates in eastern areas, and increase them in western areas. To see why this change might be advisable we need to consider the rationale behind the current policy.

As I understand it, the harvest policy is intended to protect the spawning stock. Spawning occurs in most, if not all, of the regulatory areas, but it is not known which areas are most important in producing recruits to the coast-wide population. Thus it is prudent to offer the same protection to the spawning population in all areas, and that is achieved by applying the same harvest rate in all areas. One index of the level of protection given is the current spawning biomass expressed as a percentage of what it would have been were there no fishing. To simplify the following discussion I will call this index *PI* (Protection Index). The smaller *PI* is, the less protection is being given to the spawning stock (but note that the index is not linear, so we cannot say that a *PI* of 20 indicates half the protection of a *PI* of 40). The intent of the harvest policy appears to be that, on average, *PI* should be (approximately) the same in all areas.

The effect of migration is that the policy of applying the same harvest rate in all areas no longer achieves the goal of having the same *PI* in all areas. With an eastward migration and equal harvest rates, *PI* will, on average, be lower in the east and higher in the west. This is illustrated in Table 2, which I constructed using results from the migration modelling carried out by IPHC staff (pp 148-149 and Fig. 4, 2k6rara04.pdf)

assuming an annual migration rate eastward of 6% (i.e., $T = 0.06$). For example, fishing with $F = 0.2$ produces a lower PI in Area 2B (24) than in Area 4A (33). The contrast is stronger with a higher rate of fishing: with $F = 0.3$, PI is 16 in Area 2B and 25 in Area 4A. I made two assumptions in calculating these figures: that the total unfished biomass in areas 2B to 4A would be 500 M lb, and that if there were no migration the PI for all areas would be 29 at $F = 0.2$, and 20 at $F = 0.3$. Neither of these assumptions affect the main point of this table. That is, the ratio of the PI values in the two extreme areas (2B and 4A) – 24:33 with $F = 0.2$, and 16:25 with $F = 0.3$ – would not change under different assumptions.

Table 2: Equilibrium spawning biomass and protection index (PI) by area under three levels of fishing (unfished, $F = 0.2$, $F = 0.3$) assuming an annual migration rate eastward of 6% ($T = 0.06$) and an arbitrary total unfished biomass of 500 M lb. PI is the equilibrium spawning biomass expressed as a percentage of what it would be were there no fishing.

Area	Unfished ($F = 0.0$)		$F = 0.2$			$F = 0.3$		
	% of all ¹	M lb	% of all ¹	M lb	PI	% of all ¹	M lb	PI
2B	23	115	19	28	24	18	18	16
2C	21	105	20	29	28	19	19	18
3A	32	160	35	51	32	36	36	23
3B	16	80	17	25	31	18	18	23
4A	8	40	9	13	33	10	10	25
All	100	500	100	145	29 ²	100	100	20 ²

¹ From Fig. 4, p. 155, 2k6rara04.pdf; ² Average values from Table 5, p. 44, sr83.pdf

The effect of migration on harvest policy may be greater than is shown in Table 2, because migration rates may be higher than the 6% I assumed there. In the preliminary estimates of migration rates (Table 4, p. 136, 2k6rara04.pdf) the annual percentage of fish leaving each of areas 4A to 2C ranged from 9% to 15%.

Things are a bit more complicated than shown in Table 2 because the current harvest policy does not actually provide exactly the same protection for the spawning stocks in each area. For example, with a harvest rate of 0.2, PI is expected to be 24 in Area 2B, 27 in 2C, and 36 in 3A (Table 5, p. 44, sr83.pdf) (the figure of 29 used in Table 2 is the average of these values). However, the migration modelling on which the calculations of Table 2 was based ignored these differences. I conclude that if we assume no migration, the current policy offers somewhat less protection to the spawning stock in Area 2B than in Area 3B. The effect of migration would be to increase this difference between areas. The question the Commission must address is whether this increase is sufficiently large to be of concern. If so, there will be a need to consider changing the harvest policy.

3.7 Uncertainty

Does the assessment adequately measure and report the uncertainty of the yield recommendations? If not, what more should be done?

The problem of reporting uncertainty in stock assessments and yield recommendations is a difficult one to which there is no simple answer. In discussing this problem I will describe two types of uncertainty (within- and between-model) and then consider whether the correct reporting of uncertainty really matters. To illustrate some points I will use results from some recent assessments of New Zealand hoki.

3.7.1 Within-model uncertainty

The main, and perhaps only, quantitative description of uncertainty presented in the 2006 halibut assessment appears to be a c.v. of 7% for the estimate of current coast-wide exploitable biomass (p. 6, sa06.pdf). It was also reported that this was half the corresponding values for the closed-area assessments. These c.v.s were calculated from the Hessian matrix evaluated at the point estimate. Roughly speaking, the c.v. of 7% implies that we can be reasonably confident that the true exploitable biomass was within +/-14% of the estimated value. Uncertainty in the associated yield recommendations does not seem to have been explicitly quantified. For the closed-area assessments it would be reasonable to assume that the c.v.s for the recommended yields would be similar to those for the exploited biomass. However, for the coast-wide assessment the yield c.v.s would have been much greater than 7% (because of additional uncertainty associated with the area apportionment) but difficult to calculate (because there is no obvious way of quantifying the uncertainty associated with the assumption of area-independent survey catchability).

This use of the Hessian matrix is a standard way of expressing uncertainty in quantities estimated by maximum likelihood. The only criticism I have of the c.v.s presented in the 2006 assessment relate to the changes in the log-likelihood which I suggested in Section 3.3.1. That is to say, I believe we would have more confidence in these c.v.s if these changes were made. I do not support the ad hoc multiplication of 2 to deal with double counting of observations (see bottom of p. 20, sr83.pdf).

C.v.s calculated using the Hessian matrix represent what I call within-model uncertainty. This is hard to define exactly, but roughly speaking it is that uncertainty which exists if the model assumptions are “broadly” true and the specified errors (i.e., S_{total} or c_{total} – see Section 3.3.1) are correct. The looseness of this definition derives

from the word “broadly”, which is needed because the process errors (represented by S_{process} or C_{process}) make some allowance for departures from the model assumptions.

Unfortunately, different approaches to estimation can produce very different estimates of within-model uncertainty. For example, some stock assessments use Bayesian estimation, rather than maximum likelihood. A joint posterior distribution is estimated for all model parameters and the uncertainty in any estimated quantity (such as exploitable biomass) is calculated from this posterior. In one Bayesian assessment I did for New Zealand hoki I compared Hessian-derived variances for all model parameters with those from the Bayesian posterior and found that these sometimes differed by more than an order of magnitude in either direction, though the Bayesian variances were usually larger.

3.7.2 Between-model uncertainty

In some stock assessments it is clear that not all the uncertainty can be expressed by estimates of within-model uncertainty from a single model. A good example of this is provided by the 2006 assessment of two stocks of New Zealand hoki, in which three alternative models produced quite different views of the range of uncertainty in certain biomass estimates (Figure 1). The main differences between the three models concerned the method of explaining the relative lack of old fish (either domed selectivity or age-dependent natural mortality) and whether there is natal stock fidelity (i.e., whether a fish always belongs to the same stock as its parents). In this assessment none of the three models was clearly superior to the others in either plausibility or fit to the observations.

Some measure of between-model uncertainty was presented informally in the 2006 coast-wide halibut assessment in the table of estimates from seven alternative models (p. 5, sa06.pdf), although some of these models were clearly markedly inferior to others, and so should be excluded for this purpose.

It is not easy to be confident in any particular assessment that the range of models considered captures all significant between-model uncertainty. Given sufficient time, it is very likely that the IPHC staff could have constructed other, more complex models which extended this uncertainty. One obvious example is the spatially-structured model I advocate in Section 3.2.1 (though I repeat that I do not think that this could have been easily done within the time frame of the 2006 assessment).

There are some drawbacks to presenting between-model uncertainty to fisheries decision makers. These decision makers often require a single estimate of each quantity of interest (e.g., current exploitable biomass), and a single c.v. representing

the total uncertainty in that quantity. Unfortunately, it is usually not possible to produce such single estimates from multiple models. To do this requires assigning a weight to each alternative model, which is far from straightforward.

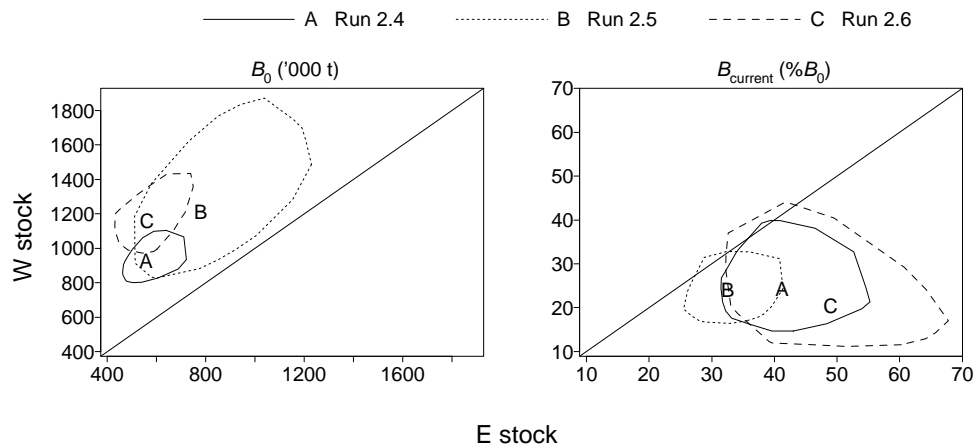


Figure 1: Illustration of between-model uncertainty in the 2006 assessment of two stocks of New Zealand hoki. Each panel shows point estimates ('A', 'B', 'C') of either unfished biomass (B_0 , left panel) or current biomass (B_{current} as % B_0 , right panel) for the two stocks (E and W) and the associated approximate 95% confidence regions (polygons) from three alternative models: 2.4, 2.5, and 2.6.

3.7.3 Does uncertainty matter?

In most fisheries agencies charged with producing stock assessments it goes without saying that it is desirable, where possible, to provide some measure of uncertainty for key estimates, such as current biomass and yields. I think it also widely acknowledged that our ability to estimate uncertainty is usually poor. This view was well expressed by IPHC staff: "For lack of anything better, such estimates [i.e., Hessian-based estimates of variance] are often reported anyway" (p. 20, sr83.pdf). I have tried in the preceding discussion to provide some explanation for this situation.

However, my view is that our inability to estimate uncertainty well probably does not matter very much. Consider, for example, what might have happened if the PIT tagging had not been done, so that there was no reason to doubt the closed-area assessment in 2006. The Commissioners would have been provided with yield estimates for each regulatory area, and these would have had c.v.s of about 15%. The question is, is it likely that the quota decisions made by the Commissioners would have been different had these c.v.s been different (say 5%, or 25%)? I suspect not.

4. Conclusions

4.1 The seven questions

Are the stock assessment data adequate? If not, what more is needed?

The data currently used in the assessment model seem adequate, although I have some suggestions for refinements in how the survey data and commercial CPUE are analysed outside the model (Section 3.1.1). There is a pressing need to complete the analysis of the PIT tag data, with the aims of explaining the low recaptures rates and improving the estimation of migration rates (Section 3.1.2).

Is the structure of the assessment model appropriate? If not, what changes should be made?

The structure of the assessment model (essentially the same in the closed-area and coast-wide assessments) was appropriate given the time frame of the 2006 assessment.

For future assessments I suggest the development of a spatially-structured model (described in Section 3.2.1) which has the following advantages: 1) it uses migration information from the PIT tag data; 2) it uses all data (rather than just survey CPUE) to do the area apportionment; 3) it provides a method of testing various assumptions, including the key one underlying the current area-apportionment scheme (area-independent survey catchability); and 4) it provides a tool to deal with the interaction between migration and harvest policy.

Is the log likelihood used to fit the model appropriate? If not, what should be used?

I offer several suggestions to improve the log likelihood (Section 3.3.1). These suggestions may or may not make a significant change to the assessment results, but I think they could make the assessment statistically more sound, and thus more defensible.

Is the suite of alternative fits adequate?

Yes.

Is the area apportionment procedure correct?

While there is clear evidence that the previous area apportionment procedure (using the closed-area assessments) is flawed, there is little evidence that the new procedure is correct, and some grounds to believe that it is not. Unfortunately, there was no obvious alternative at the time of the assessment. The spatially-structured model offers an alternative for the future.

Is the harvest policy appropriate; i.e., does it strike a proper balance between utilization and precaution? If not, how should it be modified?

The simulation procedure used to support the selection of a target harvest rate is sound. I cannot say whether the selected target harvest rate strikes “a proper balance between utilization and precaution” because what is “proper” depends on what level of risk is acceptable to the Commission, and I have no information on that. I found the practice of adjusting harvest rates for each year and area (Section 3.6.2) hard to understand, and note that upward adjustments compromise the target harvest rates derived from the above simulation procedure.

Does the assessment adequately measure and report the uncertainty of the yield recommendations? If not, what more should be done?

My suggestions for improving the log likelihood should also improve the current estimate of stock assessment uncertainty, but I note that this type of uncertainty is not well estimated anywhere.

4.2 Suggestions for any future workshops

I have two small suggestions that may be of use in planning any future workshops. First, it makes the reviewers’ task much simpler when material presented in workshop is (at least mostly) restricted to that in documents provided beforehand (the obvious exception being analyses requested during the workshop). Second, I thought that 1.5 days was the absolute minimum duration, given the amount of material to present, digest, and discuss. Longer would have been better.

APPENDIX 1: Statement of Work

This appendix contains the Statement of Work that formed part of the consulting agreement between the University of Miami and the author.

General

The International Pacific Halibut Commission seeks an independent review of its stock assessment and harvest policy. The assessment is an age- and sex-structured model, coded in AD Model Builder, which is similar in most respects to the groundfish assessments done by the NMFS Alaska Fisheries Science Center. The harvest policy is based on stock and fishery simulations that include environment-dependent recruitment and density-dependent growth as reported in previous published analysis. The reviewers should be fully competent in modern stock assessment methods, in particular the use of AD Model Builder software and contemporary statistical catch-at-age analysis.

Specific

The reviewer's work will be as follows:

1. The reviewer will read Scientific Report 83, which describes the stock assessment and harvest policy in detail.

<http://www.iphc.washington.edu/halcom/research/sa/papers/sr83.pdf>

2. The reviewer will read the 2006 stock assessment documents.

<http://www.iphc.washington.edu/halcom/research/sa/papers/sa06.pdf>

<http://www.iphc.washington.edu/halcom/research/sa/papers/prospect.pdf>

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara04.pdf>

3. The reviewer will attend a public assessment workshop in Seattle, Washington, from June 27-28, 2007, where the material in Scientific Report 83 will be presented and discussed in detail with attendees from other agencies, the industry, and the public. The IPHC will arrange an independent chair for this workshop.
4. The reviewer will meet with IPHC staff during the meeting to go over any questions arising during the meeting.
5. No later than July 13, 2007, the reviewer will submit an independent report via electronic mail to Dr. David Die (ddie@rsmas.miami.edu) and Mr. Manoj Shivlani (mshivlani@rsmas.miami.edu). The report must include, but not be restricted to, answers to the following specific questions:

- (i) Are the stock assessment data adequate? If not, what more is needed?
- (ii) Is the structure of the assessment model appropriate? If not, what changes should be made?
- (iii) Is the log likelihood used to fit the model appropriate? If not, what should be used?
- (iv) Is the suite of alternative fits adequate?
- (v) Is the area apportionment procedure correct?
- (vi) Is the harvest policy appropriate; i.e., does it strike a proper balance between utilization and precaution? If not, how should it modified?
- (vii) Does the assessment adequately measure and report the uncertainty of the yield recommendations? If not, what more should be done?

APPENDIX 2: Materials Provided

Before the workshop the reviewer was provided with the following links to relevant documents.

Necessary Material

The workshop is intended be technical, rather than educational, and participants are expected to be familiar with the IPHC assessment process. Relevant documents are IPHC Scientific Report 83, which describes the assessment model and harvest policy:

<http://www.iphc.washington.edu/halcom/research/sa/papers/sr83.pdf>

as well as documents describing the 2006 IPHC stock assessment and yield recommendations:

<http://www.iphc.washington.edu/halcom/research/sa/papers/sa06.pdf>

<http://www.iphc.washington.edu/halcom/research/sa/papers/prospect.pdf>

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara04.pdf>

<http://www.iphc.washington.edu/halcom/research/sa/papers/indepth.pdf> (new addition)

<http://www.iphc.washington.edu/halcom/research/sa/papers/hook.pdf> (new addition)

Background Material

General reference list on the Commission

Please note the Commission's website contains all historical publications in our Scientific and Technical report series, as well as most Reports of Assessment and Research Activities from 1991-2006. Literature on the IPHC website is at:

<http://www.iphc.washington.edu/halcom/literatu.htm>

Pertinent general references on Commission mandate, history, and management.

McCaughran, D.A. and S.H. Hoag. 1992. The 1979 Protocol to the Convention and Related Legislation. 32 p.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0026.pdf>

The Pacific Halibut: Biology, Fishery, and Management. 1998.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0040.pdf>

The 2007 Pacific Halibut Fishery Regulations.

<http://www.iphc.washington.edu/halcom/pubs/regs/2007iphcregs.pdf>

Kong, T.M., H.L. Gilroy, and R.C. Leikly. 2004. Definition of IPHC statistical areas. 72 p.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0049.pdf>

Clark, W.G., B. A. Vienneau, C. L. Blood, and J. E. Forsberg. 2000. A review of IPHC catch sampling for age and size composition from 1935 through 1999, including estimates for the years 1963-1990. 40 p.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0042.pdf>

IPHC Grid Surveys

Soderland, E, C.L. Dykstra, T. Geernaert, A.M. Ranta, and E Anderson. 2007. 2006 Standardized stock assessment survey Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006: 335 - 365.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara09.pdf>

References on Commercial Fishery

Commercial catch

Forsberg, J.E. 2001. Aging manual for Pacific halibut: procedures and methods used at the International Pacific Halibut Commission (IPHC). Int. Pac. Halibut Comm. Tech. Rep. No. 46: 56 p.

Page down to Technical Report 46

<http://www.iphc.washington.edu/halcom/pubs/technical.htm>

Forsberg, J.E. 2005. Age distribution of the commercial halibut catch for 2006. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006: 75–80.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

Gilroy, H.L., J.E. Forsberg and W.G. Clark. 1995. Changes in commercial catch sampling and age determination procedures for Pacific halibut, 1982 to 1993. Int. Pac. Halibut Comm. Tech. Rep. No. 32: 44 p.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0032.pdf>

Gilroy, H.L., L.M. Hutton, and K.A. Gravel. 2007. 2006 commercial fishery and regulations changes. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006: 35 – 46.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

Hutton, L.M. and K.A. Gravel. 2007. Commercial catch sampling. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006: 67- 73.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

Wastage from the commercial halibut fishery

Gilroy, H.L. 2007. Wastage in the 2006 Pacific halibut fishery. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006: 55 -58.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

References on Other Removals

Bycatch

Clark, W. G., and S. R. Hare. 1998. Accounting for bycatch in management of the Pacific halibut fishery. No. Amer. J. Fish. Mgmt. 18:809-821.

http://www.iphc.washington.edu/Staff/hare/html/papers/bycatch/abst_byc.html

Williams, G. H., C. C. Schmitt, S. H. Hoag, and J. D. Berger. 1989. Incidental catch and mortality of Pacific halibut, 1962-1989. Int. Pac. Halibut Comm., Tech. Rep. No. 23. 94 p.

<http://www.iphc.washington.edu/halcom/pubs/techrep/tech0023.pdf>

Williams, G. H. 2007. Incidental catch and mortality of Pacific halibut, 1962-2006. Report of Assessment and Research Activities 2006:163-174.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara05.pdf>

Sport fishery

Blood, C. L. 2007. 2006 sport fishery. Report of Assessment and Research Activities 2006:47-54.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

Meyer, S. C. 2003. Recreational halibut fishery statistics for Southcentral Alaska (Regulatory Area 3A), 1995-1999. A report to the International Pacific Halibut Commission. AK Dept. Fish and Game, Spec. Pub. No. 03-06, Anchorage.

<http://www.sf.adfg.state.ak.us/FedAidPDFs/Sp03-06.pdf>

Personal Use (Subsistence)

Fall, J. A., Koster, D., and Davis, B. 2006. Subsistence harvests of Pacific halibut in Alaska, 2005. AK Dept. Fish and Game, Tech. Paper 320. 182 p.

<http://www.subsistence.adfg.state.ak.us/TechPap/tp320.pdf>

Williams, G. H. 2007. The personal use harvest of Pacific halibut in 2005. Report of Assessment and Research Activities 2006:59-62.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara03.pdf>

References on PIT Tag Recoveries

Clark, W.G. 2006a. Analysis of PIT tag recoveries through 2005. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2005:123-134.

<http://www.iphc.washington.edu/halcom/pubs/rara/2005rara/2k5rara04.pdf>

Clark, W.G. 2006b. Possible causes of low PIT tag recovery rates in 2004. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2005:285-292.

<http://www.iphc.washington.edu/halcom/pubs/rara/2005rara/2k5rara07.pdf>

Clark, W.G. 2007. Further investigations of low PIT tag recovery rates. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:139-144.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara04.pdf>

Forsberg, J.E. 2007. Portside sampling for recovered PIT tags in Pacific halibut. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:277-297.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara08.pdf>

Webster, R. A. and Clark, W.G. 2007. Analysis of PIT tag recoveries through 2006. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2006:129-137.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara04.pdf>

References on Harvest Policy

Clark, W. G., and S. R. Hare. 2007. Motivation and plan for a coastwide stock assessment. Report of Assessment and Research Activities 2006: 145-160.

<http://www.iphc.washington.edu/halcom/pubs/rara/2006rara/2k6rara04.pdf>

Clark, W. G., and S. R. Hare. 2006. Assessment and management of Pacific halibut: data, methods and policy. Scientific Report 83: 45-65

<http://www.iphc.washington.edu/halcom/pubs/scirep/SciReport0083.pdf>

Clark, W. G. 2004. Effects of gear type, hook spacing and hook size on commercial selectivity and catchability. Report of Assessment and Research Activities 2005: 145-150 on Area 4B.

<http://www.iphc.washington.edu/halcom/pubs/rara/2005rara/2k5rara04.pdf>

Clark, W. G., and S. R. Hare. 2004. Updated simulation analysis of the CCC harvest policy with separate accounting of males and females. Report of Assessment and Research Activities 2004: 171-183 on YPR and 185-197 on Area 4CDE.

<http://www.iphc.washington.edu/halcom/pubs/rara/2004rara/2k4RARA05.pdf>

Links to more in-depth material

http://www.iphc.washington.edu/halcom/pubs/outside/Clark_Hare_1998.pdf
(Treatment of bycatch)

http://www.iphc.washington.edu/halcom/pubs/outside/Clark_Hare_2002.pdf (Climate impacts on halibut growth and recruitment)

http://www.iphc.washington.edu/halcom/pubs/outside/Clark_Hare_2004.pdf (The CCC harvest policy)

http://www.iphc.washington.edu/halcom/pubs/outside/mantua_hare_2002.pdf (The PDO)