Fisheries and Oceans Canada / Pêches et Océans Canada
Science / Sciences

C S A S / **S C C S**

Canadian Science Advisory Secretariat / **Secrétariat canadien de consultation scientifique**

Proceedings Series 2011/062 / **Compte rendu 2011/062**

Pacific Region / **Région du Pacifique**

A review of the Pacific herring assessment framework and stock assessment and management advice for Pacific herring 2011 status and 2012 forecasts

Examen du cadre d'évaluation du hareng du Pacifique, avis sur l'évaluation et la gestion des stocks de hareng du Pacifique : état en 2011 et prévisions pour 2012

September 7-9, 2011
Nanaimo, BC

Septembre 7-9, 2011
Nanaimo, C.B.

Linnea Flostrand, Chairperson

Linnea Flostrand, présidente de réunion

Fisheries and Oceans Canada / Pêches et Océans Canada
Pacific Biological Station / Station biologique du Pacifique
3190 Hammond Bay Road
Nanaimo, BC / C.-B. V9T 6N7

February 2012

Février 2012

Canada

Foreword

The purpose of these Proceedings is to document the activities and key discussions of the meeting. The Proceedings include research recommendations, uncertainties, and the rationale for decisions made by the meeting. Proceedings also document when data, analyses or interpretations were reviewed and rejected on scientific grounds, including the reason(s) for rejection. As such, interpretations and opinions presented in this report individually may be factually incorrect or misleading, but are included to record as faithfully as possible what was considered at the meeting. No statements are to be taken as reflecting the conclusions of the meeting unless they are clearly identified as such. Moreover, further review may result in a change of conclusions where additional information was identified as relevant to the topics being considered, but not available in the timeframe of the meeting. In the rare case when there are formal dissenting views, these are also archived as Annexes to the Proceedings.

Avant-propos

Le présent compte rendu a pour but de documenter les principales activités et discussions qui ont eu lieu au cours de la réunion. Il contient des recommandations sur les recherches à effectuer, traite des incertitudes et expose les motifs ayant mené à la prise de décisions pendant la réunion. En outre, il fait état de données, d’analyses ou d’interprétations passées en revue et rejetées pour des raisons scientifiques, en donnant la raison du rejet. Bien que les interprétations et les opinions contenues dans le présent rapport puissent être inexactes ou propres à induire en erreur, elles sont quand même reproduites aussi fidèlement que possible afin de refléter les échanges tenus au cours de la réunion. Ainsi, aucune partie de ce rapport ne doit être considérée en tant que reflet des conclusions de la réunion, à moins d’indication précise en ce sens. De plus, un examen ultérieur de la question pourrait entraîner des changements aux conclusions, notamment si l’information supplémentaire pertinente, non disponible au moment de la réunion, est fournie par la suite. Finalement, dans les rares cas où des opinions divergentes sont exprimées officiellement, celles-ci sont également consignées dans les annexes du compte rendu.

Proceedings Series 2011/062

Pacific Region

Compte rendu 2011/062

Région du Pacifique

A review of the Pacific herring assessment framework and stock assessment and management advice for Pacific herring 2011 status and 2012 forecasts

Examen du cadre d'évaluation du hareng du Pacifique, avis sur l'évaluation et la gestion des stocks de hareng du Pacifique : état en 2011 et prévisions pour 2012

September 7-9, 2011
Nanaimo, BC

Septembre 7-9, 2011
Nanaimo, C.B.

Linnea Flostrand, Chairperson

Linnea Flostrand, présidente de réunion

Fisheries and Oceans Canada / Pêches et Océans Canada
Pacific Biological Station / Station biologique du Pacifique
3190 Hammond Bay Road
Nanaimo, BC / C.-B. V9T 6N7

February 2012

Février 2012

ISSN 1701-1272 (Printed / Imprimé)
ISSN 1701-1280 (Online / En ligne)

Published and available free from:
Une publication gratuite de :

Fisheries and Oceans Canada / Pêches et Océans Canada
Canadian Science Advisory Secretariat / Secrétariat canadien de consultation scientifique
200, rue Kent Street
Ottawa, Ontario
K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/

CSAS-SCCS@DFO-MPO.GC.CA

Correct citation for this publication:

DFO. 2012. A review of the Pacific herring assessment framework and stock assessment and management advice for Pacific herring 2011 status and 2012 forecasts, September 7-9, 2011. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2011/062.

TABLE OF CONTENTS

SUMMARY

Participants from Fisheries and Oceans Canada (DFO) Science Branch and Fisheries and Aquaculture Management Branch and external participants from the Province of British Columbia Ministry of Environment, the Herring Conservation Research Society, the Haida First Nation, the Heiltsuk First Nation, the University of Simon Fraser, the University of British Columbia and Parks Canada and invited biological consultants, attended a CSAS review on September 7, 8 and 9th to assess and develop advice on the following Research Document working paper:

- Moving Towards the Sustainable Fisheries Framework for Pacific herring: Data Models, and Alternative Assumptions; and Stock Assessment and Management Advice for the British Columbia Herring Stocks: 2011 Assessment and 2012 Forecasts

Discussions and comments on Part 1 and Part 2 of the working paper and on the development of the Science Advisory Report are presented in these Proceedings. The document was accepted subject to revisions. Products of the meeting will be one CSAS Research Document and a CSAS Science Advisory Report.

SOMMAIRE

Les délégués de la direction des sciences et de la direction de la gestion des pêches et de l'aquaculture du ministère des Pêches et des Océans (MPO) et les participants externes du ministère de l'Environnement de la Colombie-Britannique, de la Herring Conservation Research Society, des Haïdas, des Heiltsuks, de l'université Simon Fraser, de l'université de la Colombie-Britannique et de Parcs Canada, ainsi que les conseillers biologistes invités ont assisté à un examen du SCCS les 7, 8 et 9 septembre afin d'évaluer le document de recherche suivant et de formuler des conseils à ce sujet :

- Vers un cadre pour la pêche durable au hareng du Pacifique : données, modèles et hypothèses de rechange; Évaluation des stocks de hareng de la Colombie-Britannique et avis pour la gestion : évaluation de 2011 et prévisions pour 2012.

Les discussions et les remarques sur les parties 1 et 2 du document de travail et sur l'élaboration de l'avis scientifique sont présentées dans ce compte rendu. Le document est adopté sous réserve de rectifications. Un document de recherche et un avis scientifique du SCCS seront émis à la suite de la réunion.

INTRODUCTION

The Centre for Science Advice – Pacific (CSAP) review was held September 7-9, 2011 at the Pacific Biological Station in Nanaimo, British Columbia. External participants representing First Nations, industry and the British Columbia government were present. Each day the Subcommittee Chair (Linnea Flostrand) opened by welcoming participants and reviewing the agenda and terms of reference.

The working paper was reviewed in two parts:

- Part 1: Moving Towards the Sustainable Fisheries Framework for Pacific herring: Data Models, and Alternative Assumptions
- Part 2: Stock Assessment and Management Advice for the British Columbia Herring Stocks: 2011 Assessment and 2012 Forecasts

REVIEWS

Working Paper: "Moving towards the sustainable fisheries framework for Pacific herring: data models, and alternative assumptions; and stock assessment and management advice for the British Columbia herring stocks: 2011 assessment and 2012 forecasts" *by Steven Martell, Jake Schweigert, Jaclyn Cleary and Vivian Haist*

** Paper was accepted subject to revisions*

PART 1: Moving towards the sustainable fisheries framework for Pacific herring: data models, and alternative assumptions

The information reported in Part 1 of the research document attempted to address specific technical components of the newly developed Pacific herring Statistical Catch Age Model (iSCAM) related to:

- Model structure and parameterization
- Natural mortality (M) and spawn survey catchability (q)
- Gear selectivity, catchability, recruitment and productivity estimation
- Model priors, likelihoods and decision rules
- Estimation of unfished biomass and stock depletion

Steven Martell presented information on the integrated Statistical Catch Age Model (iSCAM) platform that was used to model the time series of the five major and two minor British Columbia herring populations for the 2011 assessment. He described aspects of the input data, modeling flexibility and assumption options. He described how error and variability were depicted in catch, spawn survey, recruitment and age-composition distributions, as well as structural assumptions associated with modeling seine and gillnet selectivity, age-2 recruits, fishing, natural mortality, fecundity and the main components of the objective function (likelihoods for the data, likelihoods for structural assumptions, phase penalties and prior densities for model parameters). He presented methods and results associated with simulation testing of the Strait of Georgia stock to evaluate precision and bias when estimating key parameters. He also described how the iSCAM platform was able to provide spawning biomass time series trends similar to HCAM for most major stock areas despite structural differences associated with weighting of age-composition data using a multivariate logistic function, pooling age-composition bins when an age group is less than 2% of the overall proportion , use of an informative prior on spawn survey catchability, *q*, and updating the prior for Beverton-Holt steepness to a Beta distribution.

Several points of clarification were raised. It was pointed out that a number of data sources from the spawn on kelp fishery are not represented in the assessment and that information on spawn on kelp catch, mortality, and selectivity may provide value to annual assessments and determination of reference points. It was noted that allocation of catch among gear types is essential to the calculation of MSY-based reference points.

There were questions related to fishing selectivity. It was asked how the model represents changes in market preferences of fish size (and resulting changes in fishing behaviour, for example if there is an increase in catch in the winter food and bait fishery relative to other fisheries. Another question was asked about how a reduction in gillnet mesh size, which has occurred, would affect the model output. The response included a description of dome versus asymptotic selectivity and that in cases when fishing behaviour changes over time, ideally, time invariant selectivity shouldn't be assumed. Although the iSCAM model does not account for the gear changes over time, the gillnet selectivity can change with changes in fish weight via an additional estimated parameter.

There was a question about the implications of estimating q and how this may affect model fits and biases. The response given was that in most cases, fixing q was perceived to improve error patterns from past HCAMs because past HCAMs showed some poor fits to spawn data (e.g. SOG and PRD positive residuals).

Part #1 Formal review by Nathan Taylor

The reviewer submitted a written review (focused on Part 1 of the working paper) and was present to participate in meeting discussions. He commended the authors for making the herring iSCAM platform and data accessible to test. He noted that there was ambiguity associated with incorporating sustainable fisheries framework policy / theory into the paper because its purpose for being included was unclear, especially if there is no reason to discount B_0 theory applicable to herring. Several points of clarification were made with regards to text describing modeling methods and the authors agreed to clarify these points. For example, the reviewer questioned how selectivity can be greater than 1. The answer is that the selectivity values are ratios relative to the average selectivity, not absolute selectivity. The reviewer posed questions around the purpose of the review and the changes to assessment methods. He emphasized that alternate hypotheses (differing modeling assumptions) would have different uncertainty, risk metrics and decision tables which should be compared. He thought that the greatest differences were based on differing assumptions of spawn survey catchability (q=1 versus q not equal to 1) and gear selectivity (especially gillnet) and he asked whether meeting participants agree that assumptions are suitable for application in Part 2 "the assessment".

Part #1 Formal review by Sean Cox

The reviewer submitted a written review (focused on Part 1 of the working paper) but was not present to participate in meeting discussions. In the reviewer's absence, the written review was read out so the authors could address individual points. Several of the reviewer's concerns related to how the modeling applied to MSY theory, and the reviewer concluded that the scope of the paper was too broad and should have focused more on a complete evaluation of the iSCAM assessment under a range of conditions, including the assumptions used in the final assessment. The reviewer provided several suggestions to improve the focus of the paper, especially related to background and why modeling refinements were undertaken. The authors indicated that they will address these points.

The reviewer also questioned why an equilibrium unfished biomass (B_0) is used if herring stock abundance fluctuates so greatly between relatively short periods. He commented that a Management Strategy Evaluation should be developed. The reviewer also suggested that Part 1 should have included a thorough set of simulation tests under a range of assumptions at least as wide as those used in the real assessment and that DIC (Deviance Information Criterion) and performance measures for judging bias (e.g. mean relative and absolute errors, root mean square errors, etc.) be included to evaluate outcomes. The lead author agreed with these points but noted that time constraints prevented this from being incorporated in the paper.

Other suggestions that the reviewer made which the authors and meeting participants thought should be further explained in a revised text are related to: 1) how the harvest strategies were developed, 2) how is "conservative" defined in terms of comparing 0.4BMSY with $0.25B_0$, 3) the spawn survey data, especially with respect to *q*, 4) non-informative versus vague priors, and 5) modeling time-varying M. The reviewer questioned the criteria used to pool age classes when an age class represents 2% or less of the age composition, and further questioned how the effective sample size varies as a result. The lead author responded that 2% was an arbitrary decision, not based on exploring the effects of this method (or alternative percentages) and should be explored further. The reviewer also mentioned that the alternative of having an age class with 0% may affect the modeling in ways that may be undesirable (hake was used as an example of where the 2% pooling method is used).

On the subject of the Bayesian prior for the dive survey *q*, the reviewer showed concern about the use of the prior and how it relates to "moving towards a sustainable fisheries framework". The reviewer expressed specific concern over the analyses described in Appendix C.3 and the rigour of the studies from which data were gathered in order to calculate *q* prior. He also indicated there is no discussion on factors that might have changed over time and space, including spawning habitat. He suggested that future work examine the effects of allowing *q* to vary over time and fixing M and selectivity. He showed concern for the lack of a plan to evaluate the impact on the resource of using the proposed *q* prior and noted that *q*=1 may be safeguard to assessment and management errors. He thought the application of the proposed *q* prior should be rejected. Some meeting participants felt that the language used in the written review on this subject was too strong.

Part 1: General Discussion

Following the two formal reviews, Vivian Haist gave a presentation on the Bayesian prior for the dive survey index proportionality constant *q*", related to information presented in Appendix C.3 of the working paper. She presented information on the process of identifying factors affecting *q* and their distributions (egg loss prior to the survey, bias in mean egg density and drift in dive survey observations) and listed publications associated with egg loss rates. She also showed a summary table of estimates for the annual mean days between mid spawn and mid survey dates (ranging from 6.4 to 9.2 days for the five major and two minor herring areas) and explained statistical components of the proposed *q* prior. She stated that using an informative prior on *q* and estimating *q* by area (where *q* does not equal 1) is an unbiased approach that is risk neutral.

A general discussion followed the presentation by Vivian Haist. There was discussion on the topic of using an assessment model that is perceived to be risk averse versus an approach that is perceived to be risk neutral. Because applying *q*=1 likely generates minimum spawner biomass estimates, it was asked “why should science be providing minimum estimates to minimize risk if risk is considered in management decisions?”

A concern was raised around the issue of true *q* being non-stationary among years and the difficulty with estimating time varying *q*. It was suggested that some type of meta-analysis might be applicable to modeling time varying *q* and although there is significant variability in spawn deposition there is also some regularity. There was some discussion on the role of modeling relative or absolute fecundity and how temporal changes in female size and fecundity affect estimates of *q*. There was also some concern over whether the literature on egg loss might be dated since it is from over 20 years ago and not necessarily from comparable conditions of spawning habitat, stock sizes and stock structure. Some points were raised about varying egg density and predation levels at varying depths.

Another comment was made pertaining to how past assessment reports included appendix tables that enabled annual and area comparisons between spawner biomass and spawn survey indices and that these tables were thought to have some value. For example, spawner biomass estimates were generally always greater than spawner index, which was thought to provide some insight into how the model was fitting the data.

Some participants believed that the model should be accepted unless there is a clear scientific reason to reject it and that management risks should not be a determining factor but other participants were not comfortable with this approach. In trying to resolve this issue it was suggested that work presented in Part 1 can be accepted without having to endorse the stock assessment results associated with the new prior. If it is agreed that the new model is considered an incremental improvement over the past model, that doesn’t necessarily require changes to be implemented this year. This suggestion caused some confusion since it was unclear what form of science advice would result.

It was also stated that there are First Nations concerns around the Central Coast that are not being heard or implemented when managing the fishery and that the science is perceived to be insufficient in capturing conservation issues of poor stock productivity.

Another concern was raised over how the harvest rates are affected from changes in scaling of biomass estimates (resulting from changing *q*). Concern was expressed for the Haida Gwaii, Central Coast, and West Coast of Vancouver Island stocks since their stock productivity has continued to be low. There was consensus that there is considerable uncertainty as to the causes for low productivity of these three stocks (HG, CC and WCVI) and sufficient reason to be precautionary when planning commercial harvests.

Part 1: Conclusions

There was agreement that the information for the new *q* prior has value because it incorporates relevant information on egg loss rates. However, sensitivity tests and further investigation into characterization of *q* over each time series are required.

There appeared to be consensus to accept the parameterization of gillnet selectivity, although some people were unclear on the technicalities and how it impacted results.

Several text revisions are required that mainly relate to suggestions from formal reviewers to clarify methods and provide more background on some topics.

Recommendations for future work were identified, relating to:

- Investigating effects of allowing *q* to vary over the time series (fixing M and other parameters)
- Investigating effects of changes to gear selectivity.
- Sensitivity tests associated with varying proportions associated with pooling of age classes?

PART 2: stock assessment and management advice for the British Columbia herring stocks: 2011 assessment and 2012 forecasts

The information reported in Part 2 of the research document was in response to questions relating to:

- What is the estimated 2011 spawning biomass for Pacific herring by each major and minor stock assessment area?
- How are herring stocks in these areas changing over time?
- What is the forecasted pre-spawning mature stock biomass for 2012?
- Are there specific concerns that Fisheries Management should be aware of, and if so, what are those concerns?

Steven Martell presented information on the 2011 stock assessment and forecasts for 2012, including an overview of the data and the analytical methods used for the five major and two minor areas. Data sets included catch by gear, spawning survey indices, age-compositions and weight-at age. Analytical methods and assumptions described included estimating spawn survey index proportional constants (*q*); fecundity modeling; gear selectivity; pooling of age samples when <2% of sample age composition, and assumed error and variability distributions for various estimates. Results by area were associated with modeled estimates of spawning biomass and residuals; natural and fishing mortality estimates; estimates of age-2 recruits, and retrospective analysis (depicting data removal over 10 years). Additional results were presented on marginal posterior and prior distributions of estimated parameters; pairwise relationships of posterior samples from leading parameters and derived variables; newly derived unfished equilibrium biomass estimates (B_0); 2011 biomass estimates; 95% confidence intervals and marginal distributions for B_0 and B_t, and forecasted maximum allowable catch options based on two sets of B_0 estimates. Highlighted were three outstanding issues related to 1) sensitivity of reference points (i.e. B_0) to methods/assumptions used for calculating non-stationary parameters (e.g. M, selectivity and growth), 2) strong retrospective bias for PRD highlights poor modeling abilities and high uncertainty for this area, and 3) the need for formal evaluations of alternative hypotheses.

Steven Martell encouraged points of clarification and comments throughout his presentation. There was a question about combining versus separating the test and commercial fishery sample data for the assessment, and the response was that future work should be done on this issue. A comment was made that some mortality from fishing activities may be underestimated, such as collateral damage from gillnet and SOK fisheries as well as from by-catch in other fisheries. A meeting participant questioned the reason for the sharp decline in biomass around 2005 with increasing estimates of M, especially for the CC stock.

Ron Tanasichuk presented results from the recent August WCVI trawl survey and recruitment forecasts for the WCVI and SOG stock assessment regions, based on age composition data (see Appendix). The methodology for forecasting recruitment to these areas was previously approved (Tanasichuk 2000, 2002) and includes a regression model based on the relationship between proportions of age 2+ herring observed in the trawl survey and proportions of age 2+ herring estimated by the assessment model (e.g. iSCAM) for the subsequent pre-fishery or pre-spawning season. He concluded that the recruitment forecast for 2012 season is “poor” for the WCVI and “good” for the SOG.

Jake Schweigert presented preliminary herring catch information from a recent July-August nocturnal WCVI sardine surface trawl survey. Future work consists of formal analyses to compare herring catch per unit effort and length distributions between the summer herring and sardine trawl surveys but the time series of herring biological information from the sardine survey is currently limited. He also commented on some preliminary results from the SOG juvenile herring survey and stated that the 2009 year class strength is expected to be average.

Part 2: Formal review by Sean Cox

The reviewer submitted a written review (focused on Part 2 of the working paper) but was not present to participate in meeting discussions. In the reviewer’s absence, the written review was read out so the authors could address the points individually.

The reviewer had several concerns with the use of the harvest control rules, decision rules, and application of currently defined LRP (limit reference point) and USR (upper stock reference), doubting their effectiveness in prescribing to the Sustainable Fisheries Framework (SFF) and precautionary principle. Some of the reviewer’s points were beyond the scope of the paper’s objective because the paper was not meant to evaluate harvest control and decision rules, LRP and USR in a MSE (management strategy evaluation) context. The authors will reword “Decision Rules” relative to $0.25B_0$ and their use as a commercial fishing cut-off and not explicitly a conservation limit.

The reviewer had some concerns related to the use of input data and posed questions related to spawn survey methods, treatment of test fishery sample data, and use of mean weight-at-age data. Some of the concerns reflected a lack of understanding of how the assessment methods are applied, which highlighted a need for some revisions in the paper. With regards to the use of mean weight-at-age data and possible bias, the authors responded that the input data used for representing weight-at-age is only from seine samples (commercial and test) which are assumed to be unselective. Authors noted that representation of ages may vary between these samples and preliminary work has identified other confounding effects of grouping test and commercial seine data. Future work is being planned to separate and evaluate this.

With regards to providing catch advice, the reviewer questioned the recruitment forecasting protocol and lack of direction for managers in selecting a recruitment option which has direct bearing on catch options. To address this concern, the authors presented results (not included in the working paper) related to three risk probability metrics which the authors proposed fishery managers can use when setting allowable catches . The proposed risk metrics are: 1) the probability that a harvest option will result in an effective harvest rate that exceeds 20%; 2) the probability that a harvest option will induce a stock to fall below 0.25 B_0; and 3) the probability that the spawning biomass of a stock will decline in the subsequent year. Probabilities associated with each risk metric will be included as tables in the final working paper (see Appendix).

The reviewer also questioned the inclusion of residual and bubble plots as likely being unnecessary but the authors disagreed and emphasized that it is important to be able to reference trends in these plots.

The reviewer questioned the credibility in the range of F_{MSY} estimates (0.54-2.36/ yr, Table 2.3 in working paper) and the high value of 2.36 for HG. He also questioned to what degree the priors are driving estimation, since the posterior distributions for some estimates (h- recruitment steepness) are similar to the priors. The authors noted that F_{MSY} values are maximum likelihood estimates (MLEs) with long tails in their distribution thus showing considerable uncertainty but noted that F_{MSY} estimates are not used to give catch advice. The authors agreed that greater evaluation of the priors is necessary (i.e. evaluate policy results related to the influence of priors) and when there is a lot of uncertainty in the estimates, the priors have more influence. It was also noted that information from priors still may be valid.

The reviewer questioned why some time series estimates of SOG B_t were greater than B_0,and why the SOG stock is not modeled to drop to $0.4B_{MSY}$ (as a LRP) when the stock collapsed in the 1960s (unlike other modeled stocks). He noted that there should be some discussion about this in the paper. Some meeting participants pointed out that the stock can be very productive and herring populations can fluctuate greatly.

With regards to the pairs plot (Fig 2.33 and 2.34) the reviewer suggested that the paper show what is correlated with spawn survey catchability (q) since the results are heavily weighted by the prior for q and is likely the "leading" parameter. He also suggested clarification in the text and graphs to inform whether there is lack of correlation in the posteriors or lack of correlation in the priors. Authors agreed to provide updated pairs plots in the revised text and noted that annual recruitment estimates (based on R_0) are dependent on q.

The authors clarified that age composition (from seine samples) has the greatest effect on estimating mortality but that weight-at-age also has some influence. The reviewer suggested that some positive bias would affect biomass estimates because of gear selectivity. The authors believe positive bias would be relatively small (and likely for younger fish) since seine selectivity is thought to be negligible.

The reviewer agreed with the "outstanding issues" described in the paper and suggested that the discussion be further developed to include plans addressing these and how managers might be able to cope with so many uncertainties. The authors agreed with these suggestions and emphasized the role of the risk metrics and decision tables that were presented (see Appendix).

Part 2: Formal review by Nathan Taylor
The reviewer stated he found it difficult to separate the technical review of the methods (Part 1) from the review for harvest advice (Part 2), especially without clear statements on management objectives and suitability of the control rules and reference points. For Part 2, he stated there was a lack of presentation of uncertainty related to management options as well as assessment parameterization and assumptions. He suggested the use of a decision table showing probabilities of key stock responses given management outcomes, which the authors supplied and presented earlier that day. He especially emphasized that changes to the model represent competing hypotheses about how to represent population dynamics and stated results from competing hypotheses should be evaluated. He also commented on the uncertainties of relating the assessment to a precautionary approach (PA) with risk tolerance and suggested that the assessment apply a decision table showing probabilities for different reference points. The lead author responded that this cannot be done without a formal gear allocation agreement but future modeling could take this into account.

Part 2: General Discussion

It was clarified that estimating reference points and implementation of PA and sustainable fisheries framework (SFF) theory is under development and was not intended to be an integral part of review objectives nor meeting conclusions.

It was again emphasized by authors and participants that incidental mortality such as from SOK and by-catch are not accounted for which may bias mortality estimates and compromise the effectiveness of the assessment.

Concern was expressed that managers need biomass and catch options to plan fisheries and it is apparent that more work is required to evaluate changing *q* even though participants may accept the way the *q* prior was developed. It was further stated that it was unclear how management is to use the risk probability metrics versus the harvest control rule outputs presented in the paper.

There was debate as to whether the group accepts the model and its results. There is agreement that future work is required to evaluate the hypotheses and evaluate implications on a management framework. There is agreement that the developed risk probability / decisions tables be included in the paper and referred to by management for developing a management plan. It was pointed out that the decision tables are based on assessment methods that have *q* estimated and it is difficult for participants to evaluate this model over other options. Again the question was raised whether authors and science advice need to make modeling choices that are risk adverse or whether managers should assess risk and make choices. It was pointed out that some people do not have a full understanding of the model and rely on the authors decisions and wonder how the current results compare with those from previous HCAM methods. The authors indicated that HCAM had biases related to gillnet selectivity that were addressed in the new iSCAM. It was pointed out that one of the formal reviewers does not agree that the model should be accepted. It was also noted that this reviewer was not present to get clarification on some of the technicalities and to participate in discussions.

Concern was expressed for the three stocks with recent low productivity (HG, CC and WCVI) and harvest control rule catch options. A question was posed: "Do we need to worry about new cutoffs since at poor recruitment the WCVI would be open?". A response to this was that participants should advise against commercial fishing despite the forecasted biomass. A current decision rule stipulates the maximum harvest is the difference between a forecast and the cutoff. It was noted that risk probability metrics provide a way to evaluate that harvest option. Someone suggested that maybe in the cases of HG, CC and WCVI the assessment method should be to apply q=1. But it was pointed out, any results made available with q=1 would not be reviewed and would be equivalent to doing another assessment.

Someone emphasized that the focus of the concerns should not be on the modeling methods and newly derived estimates of B_0 and B_t but rather on advising whether to proceed with the current harvest strategy in relation to the stock levels. There is uncertainty associated with biomass estimates and with forecasting associated with HG, CC and WCVI and the continued low productivity for these stocks, despite the absence of commercial fishing. The authors stated that the science advice to management should be not to fish in the HG, CC and WCVI regions until there is a better understanding for the recent low productivity. It was also suggested that a management decision should be able to be based on information gaps not captured in the model that are thought to affect stock productivity (i.e. ecological interactions).

It was stated that information such as the forecast rule, summer trawl survey forecasts and risk probabilities need to be referred to when giving advice and should be included in the Science Advisory Report.

Part 2: Conclusions

The paper was accepted subject to revisions, the main revisions being related to: 1) including risk probability tables presented at the meeting with sufficient background and discussion on their role and application; 2) inclusion of 95% credibility intervals for estimates of B_{2011}, B_0 and forecasts of B_{2012}, and 3) clarifying methods and adding background on topics as outlined in submitted reviews.

Forecasts of age-3 recruits for the 2012 season using the previously accepted forecasting methods were agreed upon, whereby forecasts for the SOG and WCVI are based on the summer trawl survey observations, forecasts for the three other major areas are based on assignment rules, and forecasts for the minor areas are always "average". Stock recruitment forecasts are "poor" for the Haida Gwaii (2E), Central Coast and west coast of Vancouver Island, "good" for the Strait of Georgia, and "average" for the Prince Rupert District and the two minor areas (Area 2W and Area 27).

There is a high level of concern and uncertainty associated with the continued low productivity in the Haida Gwaii (2E), Central Coast and West Coast of Vancouver Island stock areas and a lack of understanding as to the causes.

There is an especially high level of uncertainty associated with the biomass estimates for the Prince Rupert District where retrospective analysis suggest positive bias (overestimation) from model results; therefore caution is advised in planning fisheries in this area.

Several recommendations for future work were identified, relating to:

- Examining effects of assigning test fishing as a separate gear in the model
- Quantifying spawn on kelp mortality associated with adult fish in closed pens
- Quantifying egg removal associated with spawn on kelp fisheries
- Examining effects of non-stationary selectivity
- Examining effects of ageing drift errors over time (study underway)
- Examining causes of the positive retrospective bias in estimates of PRD biomass

ACKNOWLEDGEMENTS

Thank you to Janeane MacGillivray for assisting in the coordination and organization of the meeting and a special thank you to Marilyn Joyce for managing the webinar and assisting with final meeting preparations. Thank you to Jake Schweigert and Kristen Daniel for being the rapporteurs during the review discussions.

APPENDIX A: AGENDA

A review of the Pacific Herring Assessment Framework

Stock Assessment and Management Advice for the Pacific Herring: 2011 Status and 2012 Forecast
Pacific Regional Science Advisory Process

September 7-9, 2011
Pacific Biological Station, Nanaimo, BC

Chairperson: Linnea Flostrand

Day 1 – Wednesday September 7th

9:00	Welcome & Introductions	Linnea Flostrand
9:15	Review Agenda & Housekeeping	Linnea Flostrand
9:30	CSAS Overview & Meeting Procedures	Linnea Flostrand
9:45	Review of Terms of Reference	Linnea Flostrand
10:00	Presentation of Part 1 of Working Paper- *Pacific Herring Framework: data, models and alternative assumptions*	Steve Martell
10:40	***Break***	
11:00	Part 1: Reviewer Presentation & Author Response	Sean Cox
11:40	Part 1: Reviewer Presentation & Author Response	Nathan Taylor
12:20	***Lunch Break***	
1:30	Part 1: Group Discussion to identify issues and topics needing further discussion - *Pacific Herring Framework*	RAP Participants
2:45	***Break***	
3:00	Part 1: Discussion and resolution - Issues & Topics	RAP Participants
4:30	***Adjournment***	

<u>Day 2 – Thursday September 8th</u>

8:30	Introductions & Housekeeping	Linnea Flostrand
8:45	Review Day 1 & Confirm Agenda for the day	Linnea Flostrand
9:00	Presentation of <u>Part 2 of Working Paper</u>- *Stock assessment and management advice for the B.C. herring stocks: 2011 assessment & 2012 forecasts*	Steve Martell
9:40	Brief summary of recruitment forecasts from La Perouse summer trawl survey	Ron Tanasichuk
9:50	<u>Part 2:</u> Reviewer Presentation & Author Response	Sean Cox
10:20	***Break***	
10:40	<u>Part 2:</u> Reviewer Presentation & Author Response	Nathan Taylor
11:10	<u>Part 2:</u> Group Discussion to identify issues and topics needing further discussion - *Stock assessment and management advice for the B.C. herring stocks*	RAP Participants
12:05	***Lunch Break***	
1:15	<u>Part 2</u>: Discussion and resolution - Issues & Topics	RAP Participants
2:45	***Break***	
3:00	<u>Science Advisory Report (SAR)</u>: Develop Consensus on: • Key findings & conclusions • Uncertainties • Ecosystem Considerations • Advice for Management / application of protocols • Recommendations for future work • Other	RAP Participants
4:30	***Adjournment***	

* Day 3 –Friday September 9th

8:30	Introductions & Housekeeping	Linnea Flostrand
8:45	Review Day 1 and 2 & Confirm Agenda for the day	Linnea Flostrand
9:00	Continue Science Advisory Report: Develop Consensus on: • Key findings & conclusions • Uncertainties • Ecosystem Considerations • Advice for Management / application of protocols • Recommendations for future work • Other	RAP Participants
10:30	***Break***	
11:00	Finalize Science Advisory Report	RAP Participants
12:00	***Lunch Break and Adjournment***	

APPENDIX B: TERMS OF REFERENCE

Pacific Herring Assessment Framework; Stock Assessment and Management Advice (2011 Status and 2012 Forecast)

Pacific Regional Science Advisory Process

September 7-9, 2011
Pacific Biological Station, Nanaimo, BC
Chairperson: Linnea Flostrand

Context

Annually, an assessment of Pacific herring abundance and forecasts for the coming year is generated for each of the five major and two minor stocks in British Columbia, using a statistical catch–age-model. The assessment framework integrates data from sampling the population with analytical methods to model population dynamics and harvest control rule components. The annual assessment is reviewed through a Canadian Science Advisory Secretariat Regional Advisory Process (RAP) and harvest advice is provided to Fisheries and Aquaculture Management each fall to inform the development of the Integrated Fisheries Management Plan (IFMP).

Refinements to the herring statistical catch-age model have occurred on an ongoing basis since its earliest version (Haist and Stocker 1984), with the last major review in 2008 (Christensen et al 2009) and with some updates in 2009 and 2010 (Cleary et al 2009, DFO 2009a, DFO 2010). However, recent RAP reviews have identified specific aspects of the current model's structural assumptions and decision rules that warrant a more in depth review (DFO 2009a). In addition to the formal RAP reviews, the Herring Conservation and Research Society (HCRS) sponsored a herring stock assessment model review in June of 2010, which also produced recommendations to explore a number of data issues and alternative structural assumptions in the model.

In order to be more compliant with DFO's Sustainable Fisheries Framework (SFF) policy "*A fishery decision-making framework incorporating the Precautionary Approach*" (DFO 2009b), The Herring Assessment Team, led by DFO Science, has developed an alternative modeling framework to address some of the recommendations that have come forth at RAP reviews concerning model assumptions and decision rules. This RAP will review changes to the statistical catch-age model, which includes updates to reference point used in the herring harvest control rule, and comparisons to results from the preceding version of the statistical catch-age model will be made available. This RAP will also review 2011 biomass estimates and forecasts for the 2011/2012 fishing season.

One research document is being presented and reviewed, which is structured in two parts and the meeting will also be structured as two Parts: (1) formal review of revised stock assessment model and framework, and (2) provision of science advice for the 2011/12 fishing season.

Objectives

Working paper to be reviewed:

Moving towards the sustainable fisheries framework for Pacific herring: data, models, and alternative assumptions and Stock Assessment and Management Advice for the British Columbia Herring Stocks: 2011 Assessment and 2012 Forecasts - Steven Martell, Jake Schweigert, Jaclyn Cleary and Vivian Haist,

Part 1
Moving towards the sustainable fisheries framework for Pacific herring: data, models, and alternative assumptions

Specific technical components of the Pacific herring statistical catch-age population assessment model assumptions and structure to be evaluated include:

- Model structure and parameterization
- Confounding effects of natural mortality (M) and spawn survey catchability (q),
- Gear selectivity, catchability, recruitment, and productivity estimation.
- Model priors, likelihoods, and decision rules
- Estimation of unfished biomass and stock depletion levels (time series of annual biomass over unfished biomass).

Part 2
Stock Assessment and Management Advice for the British Columbia Herring Stocks: 2011 Assessment and 2012 Forecasts

- What is the estimated 2011 spawning biomass for Pacific Herring by major and minor stock assessment area?
- How are herring stocks in these areas changing over time?
- What is the forecasted spawning biomass for 2012?
- Are there any specific concerns that Fisheries Management should be aware of, and if so, what are those concerns?

Expected publications

- CSAS Proceedings (1)
- CSAS Science Advisory Report (1)
- CSAS Research Document (1)

Participation

DFO Science Branch
DFO Fisheries and Aquatic Management Branch
BC Ministry of Agriculture- Marine Fisheries Branch
Commercial and recreational fishing interests
First Nations organizations
Non-government organizations
Academia
For further information on participation in the peer review process: http://www.dfo-mpo.gc.ca/csas/csas/Process-Processus/ExtPart-PartExt/Ext-Part-RAP_e.htm

References and Additional Information

Christensen, L.B., V. Haist and J. Schweigert. 2009. Modeling herring population dynamics. Herring catch-at-age model version 2. Can. Sci. Adv. Secr. 2009/073: 65p.

Cleary, J.S., Schweigert, J.F., Haist, V. 2009. Stock assessment and management advice for the British Columbia herring fishery: 2009 assessment and 2010 forecasts. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/079. vii + 81 p.

DFO. 2009a. Proceedings of the Pacific Scientific Advice Review Committee (PSARC) Pelagic Subcommittee Meeting: Stock assessment and management advice for BC herring fishery, 2009 assessment and 2010 forecasts and herring multi-stock analysis; September 2, 2009. DFO Can. Sci. Advis. Sec. Proceed. Ser.2009/037.

DFO. 2009b. A fishery decision-making framework incorporating the Precautionary Approach. http://www.dfo-mpo.gc.ca/fm-gp/peches-fisheries/fish-ren-peche/sff-cpd/precaution-eng.htm.

DFO 2010. Science Advisory Report. Stock Assessment Report on Pacific Herring in British Columbia in 2010. 2010/064. http://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2010/2010_064-eng.html

Haist, V., and M. Stocker. 1984. Stock assessment for British Columbia herring in 1983 and forecasts of the potential catch in 1984. Can. MS Rep. Fish. Aquat. Sci. 1751: 50p.

APPENDIX C: ATTENDEES

Last name	First name	Affiliation	Sept 7	Sept 8	Sept 9
Boldt	Jennifer	DFO, Science	Web	Web	Web
Cleary	Jaclyn	DFO, Science	Yes	Yes	
Daniel	Kristen	DFO, Science	Yes	Yes	
Flostrand	Linnea	DFO, Science	Yes	Yes	Yes
Forrest	Robyn	DFO, Science	Yes		
Fort	Charles	DFO, Science	Yes	Yes	Yes
Hall	Peter	DFO, FM (North Coast)	Yes	Yes	Yes
Joyce	Marilyn	DFO, Science	Yes	Yes	Yes
Kanno	Roger	DFO, FM	Yes	Yes	
Mah	Jordan	DFO, FM		Yes	
Midgley	Peter	DFO, Science	Yes	Yes	Yes
Mijacika	Lisa	DFO, FM	Yes	Yes	Yes
Ryall	Paul	DFO, FM	Yes		
Schweigert	Jake	DFO, Science	Yes	Yes	Yes
Spence	Brenda	DFO, FM	Yes	Yes	Yes
Tanasichuk	Ron	DFO, Science	Yes	Yes	Yes
Taylor	Nathan	DFO, Science	Yes	Yes	Yes
Carpenter	Steve	Heiltsuk Gladstone Reconciliation	Yes	Yes	Yes
Chalmers	Dennis	BC Ministry of Fisheries	Yes	Yes	Yes
Haist	Vivian	Haist Consulting / HCRS	Yes	Yes	
Hamer	Lorena	Herring Conservation & Research Society	Yes	Yes	Yes
Hay	Doug	Emeritus / HCRS	Yes	Yes	
Haycroft	Carly	BC Ministry of Fisheries	Yes	Yes	
Irving	Nicholas	Parks Canada	Yes		
Jones	Russ	Council of Haida Nation	Yes		
Martell	Steve	University of British Columbia	Yes	Yes	
McFarlane	Gordon (Sandy)	Emeritus	Yes	Yes	
Morley	Rob	Canadian Fishing Company	Yes	Yes	
Newman	Earl	Heiltsuk Gladstone Reconciliation	Yes	Yes	
Safarik	Ed	Herring Conservation & Research Society	Yes	Yes	

* Web = participated by Webinar

APPENDIX D: WCVI AND SOG HERRING RECRUITMENT FORCASTING FROM TRAWL SURVEY RESULTS

Provided by Ron Tanasichuk and Jennifer Boldt, Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, B. C.

Sampling

Data for the 2012 forecasts are from an offshore mid-water trawl survey conducted during August 2-8, 2011 using the R/V W. E. Ricker off the southwest coast of Vancouver Island (Appendix D Figure 1). Twenty trawl tows were made and sufficient sample sizes of herring were collected in 13 of the tows (Appendix D Table 1). We excluded data from tows with catches less than 100 kg because the samples were assumed to be unrepresentative. A random sub-sample of 150 fish was taken from the 13 tows with sufficient herring catches. Standard length (mm) and mass (g) were measured for each fish in a sub-sample and scales from the first 100 fish were aged at the Pacific Biological Station.

Herring length frequency distributions differed from those in 2010 but small herring continued to be exceptionally abundant. There was a very large aggregation in deep water (>120 m) below the southeast edge of Swiftsure Bank. Aggregations also occurred on Finger and 40 Mile banks and at the Southwest Corner. There were numerous small, diffuse schools off the Banks, between Finger and Swiftsure banks.

Analysis

The recruitment forecast is made using the methodology described in Tanasichuk (2002). The method has two parts. The first part forecasts the proportion of age 2+ (recruit fish) in the incoming pre-spawning biomass. It is based on the linear relationship between two measures of recruit abundance. The first measure is the proportion of age 2+ herring trawled during August off the south-west coast of Vancouver Island. The second measure is the estimated proportion of age 2+ fish in the biomass for the subsequent spring pre-fishery season (iSCAM, Integrated Statistical Catch Age Model; Martell et al. 2011). Proportion data are transformed using the logit transformation (Sokal and Rohlf 1995). Predictive regressions are re-expressed as geometric mean regressions (GMR) because both variables were measured with error (Ricker 1973). The time series (Appendix D Table 2) is updated annually with the proportion age 2+ in the August trawl samples, and model estimates of proportion of age 2+ fish in the pre-fishery biomass for the most recent fishing season.

The second part of the methodology consists of calculating the number of incoming age 2+ recruits. It uses the forecasted proportion age 2+ from the regression, and the number of returning adults (age 3+ and older, age 3++) forecasted by the iSCAM. Number of age 2+ fish (R_t) in the pre-fishery biomass is estimated as:

$$(1)\quad R_t = (N_t \cdot (1 - p_{t-1})^{-1}) - N_t\ ,$$

where t is fishing year, N is the number of age 3++ forecasted to be in the pre-fishery biomass, and p is the proportion of age 2+ fish forecasted from the offshore survey. The logic of the calculation is as follows. The offshore survey samples the entire pre-fishery biomass (age 2+ and age 3++ (age 1+ recruitment is assumed to be negligible)) but the stock assessment model forecasts age 3++ only. Based on the offshore survey, the number of age 3++ forecasted by the stock assessment model is 1-p_{t-1} of the entire number of forecasted fish in the pre-fishery biomass. By multiplication then, the total number of fish in the pre-fishery biomass is

$N_t \cdot (1 - p_{t-1})^{-1}$. The number of recruits (R_t) is the forecasted total number of fish in the pre-fishery biomass minus the forecasted number of age 3++.

Identification of age 2+ (recruit) herring

Fish were assigned to age 2+ or not age 2+ using the methods described in Tanasichuk (2002). Age-length data were pooled over tows and stratified by 1 mm length intervals. For the pooled data, the proportion of age 2+ fish in each length interval (l) was estimated as:

$$(2)\ P_{2+,l} = N_{2+,l} \bullet N_l^{-1}.$$

Number of age 2+ fish ($N_{2+,y}$) in tow *y* was then estimated as:

$$(3)\ N_{2+,y} = \sum P_{2+,l} \bullet N_{l,y}$$

Proportion of age 2+ fish in a sample in a given tow was estimated by dividing N_{2+} by the number of fish measured. We excluded data from Tows 4, 6, 11, 14, 19 and 20 from the forecasting data because length-frequency distributions for these tows suggested that there were no adult fish in the catches, so we assumed the catches were from schools of immature fish.

Performance of recruitment forecasts

Numbers of recruits are re-expressed as Poor, Average or Good recruitments depending on where they occur in the distribution of the model's estimated recruitment time series. Boundaries between Poor and Average, and Average and Good recruitment, are calculated as the 33rd and 67th percentiles respectively of the cumulative probability distributions of the age 2+ abundance time series.

Appendix D Table 3 shows the performance of the recruitment forecasts for WCVI and SOG herring. The 2011 forecasted recruitment category for the WCVI was Poor and the observed recruitment, based on the stock assessment model estimate, was Poor. The recruitment forecast has been correct in 7 of 10 years. The 2011 forecasted recruitment category for the SOG was Good and the observed recruitment, based on the stock assessment model estimate, was Good. Results show that the forecasts were accurate in 6 of 9 years.

2012 recruitment forecast

The mean proportion of age 2+ herring in the 2011 survey, weighted by CPUE, was 0.26. Consequently, using the regression based on data to the 2011 fishing season inclusive, 0.49 of the fish in the WCVI 2012 pre-fishery biomass are forecasted to be age 2+. The stock assessment model forecast of the number of age 3++ fish in the 2012 pre-fishery biomass is 86 x 10^6. Therefore, the forecasted number of age 2+ herring is 83 x 10^6. The current breakpoints between Poor/Average and Average/Good recruitments are 137 x 10^6 and 254 x 10^6 fish respectively. Consequently, recruitment for WCVI herring in 2011 is forecast to be Poor.

The recruitment forecast for Strait of Georgia herring is Good. The forecasted proportion of age 2+ fish in the 2011 Strait of Georgia pre-fishery biomass is 0.52. The forecasted number of age 3++ fish is 910 x 10^6 and the forecasted number of age 2+ fish is 986 x 10^6. The current breakpoints between Poor/Average and Average/Good recruitments are 415 x 10^6 and 735 x 10^6 fish respectively.

Table 1. August 2011 trawl tows, herring catch, catch per unit effort (CPUE) and proportion of age 2+ herring from trawl tows off the south-west coast of Vancouver Island (immature fish).*

Tow	Subarea Name	Catch (kg)	CPUE ($kg \cdot m^{-3}$)	Prop. Age 2+
1	40 Mile Bank	121	0.00052	0.52
2	40 Mile Bank	0	0	.
3	40 Mile Bank	0	0	.
4	40 Mile Bank	724	0.00278	0.26*
5	12 Mile Bank	251	0.00070	0.38
6	Swiftsure Bank	828	0.00335	0.22*
7	Swiftsure Bank	357	0.00144	0.56
8	Swiftsure Bank	1	0.00000	.
9	Swiftsure Bank	106	0.00025	0.42
10	Swiftsure Bank	5	0.00000	.
11	Swiftsure Bank	964	0.00427	0.21*
12	Swiftsure Bank	506	0.00107	0.28
13	Finger Bank	69	0.00032	.
14	SW Corner	908	0.00334	0.29*
15	Potholes	21	0.00000	.
16	40 Mile Bank	1593	0.00707	0.41
17	Eddy	0	0	.
18	Swiftsure Bank	6810	0.02290	0.18
19	Potholes	101	0.00031	0.39*
20	SW Corner	662	0.00234	0.31*

Table 2. Time series of data for estimating recruitment forecasting regressions. Prefishery estimates are from the iSCAM assessment model.

Summer trawled	Fishing season	Trawled	WCVI prefishery	SOG prefishery
1987	1988	0.46	0.71	0.71
1988	1989	0.25	0.29	0.38
1989	1990	0.25	0.35	0.74
1990	1991	0.10	0.43	0.44
1991	1992	0.62	0.58	0.67
1992	1993	0.16	0.46	0.47
1993	1994	0.24	0.39	0.66
1994	1995	0.27	0.27	0.31
1995	1996	0.25	0.43	0.65
1996	1997	0.34	0.77	0.68
1997	1998	0.19	0.31	0.56
1998	1999	0.14	0.30	0.37
1999	2000	0.42	0.43	0.58
2000	2001	0.38	0.57	0.64
2001	2002	0.61	0.64	0.59
2002	2003	0.17	0.53	0.53
2003	2004	0.09	0.43	0.39
2004	2005	0.41	0.59	0.51
2005	2006	0.19	0.52	0.50
2006	2007	0.73	0.70	0.69
2007	2008	0.06	0.50	0.38
2008	2009	0.73	0.69	0.56
2009	2010	0.09	0.46	0.14
2010	2011	0.66	0.76	0.76

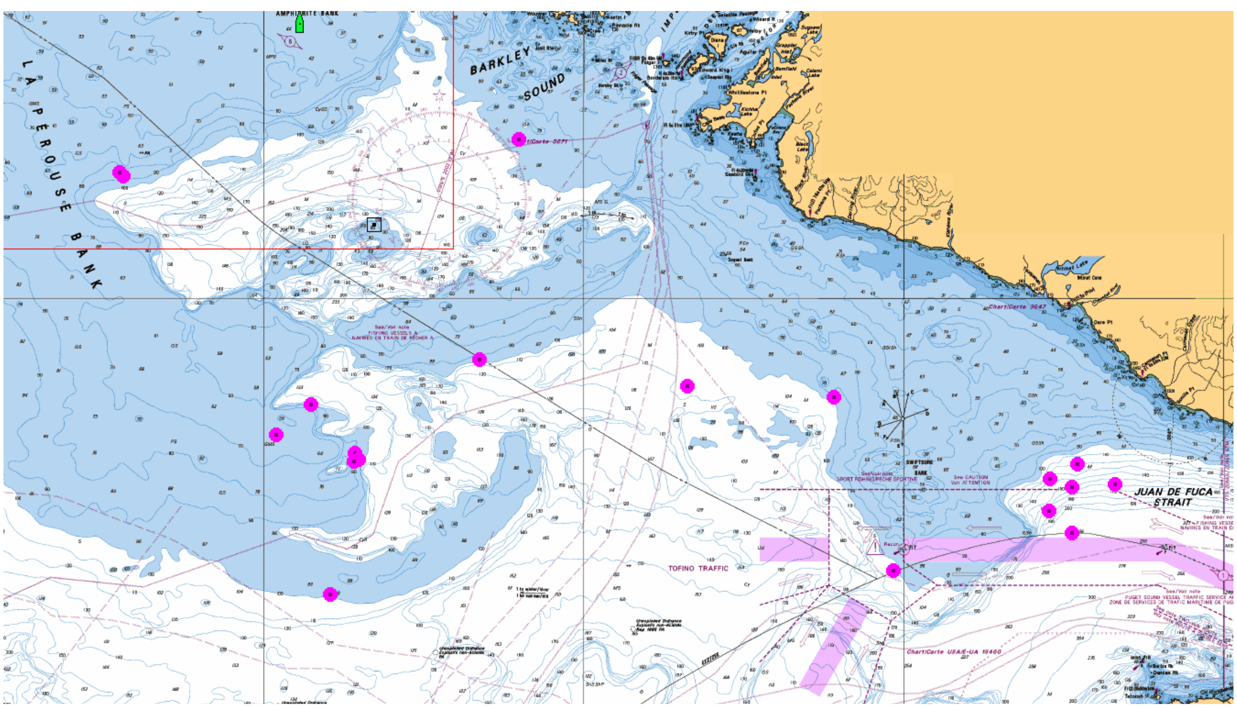


Figure. 1. La Perouse study area and midwater tow sites (pink circles).

Table 3, A. Recruitment forecasts for WCVI herring, 2002-2011. B. Recruitment forecasts for SOG herring, 2002-2011. Numbers of fish x 10^6. All observed estimates are iSCAM output. Numbers of fish estimated by the model are as of July 1 of the year before the fishery and are multiplied by e^{-M} x lambda (age-independent and year-dependent survival x age-dependent and year-independent availability) to generate pre-fishery estimates. Recruitment distribution breakpoints for Poor/Average (p=0.33) and Average/Good (p=0.67) are from age 2+ time series for each of the 2002 through 2011 iSCAM forecasts. Forecasting began with the 2002 fishing season, when the regressions became statistically significant

A	Proportion		Forecast		Observed	
Season	Forecast	Observed	Number	Category	Number	Category
2002	0.79	0.64	332	GOOD	176	AVERAGE
2003	0.31	0.53	67	POOR	214	AVERAGE
2004	0.20	0.43	61	POOR	133	POOR
2005	0.61	0.59	239	AVERAGE	152	POOR
2006	0.38	0.52	61	POOR	24	POOR
2007	0.83	0.70	41	POOR	22	POOR
2008	0.19	0.50	1	POOR	19	POOR
2009	0.79	0.69	36	POOR	81	POOR
2010	0.29	0.46	23	POOR	26	POOR
2011	0.73	0.76	58	POOR	128	POOR

B	Proportion		Forecast		Observed	
Season	Forecast	Observed	Number	Category	Number	Category
2002	0.84	0.59	5893	GOOD	1320	GOOD
2003	-	-	-	-	214	AVERAGE
2004	0.30	0.39	798	GOOD	946	GOOD
2005	0.67	0.51	3023	GOOD	1078	GOOD
2006	0.47	0.50	968	GOOD	506	AVERAGE
2007	0.84	0.69	1743	GOOD	353	POOR
2008	0.28	0.38	39	POOR	267	POOR
2009	0.80	0.56	908	GOOD	335	POOR
2010	0.36	0.14	120	POOR	115	POOR
2011	0.77	0.76	1293	GOOD	1233	GOOD

REFERENCES

Martell, S., Schweigert, J., Cleary, J. and V. Haist. 2011. Moving towards the sustainable fisheries framework for Pacific herring: data models, and alternative assumptions; and stock assessment and management advice for the British Columbia herring stocks: 2011 assessment and 2012 forecasts. CSAP Working Paper. *DRAFT*.

Sokal, R. R. and F. J. Rohlf. 1995. Biometry. 3rd edition. W. H. Freeman and Co. New York, NY. 887p.

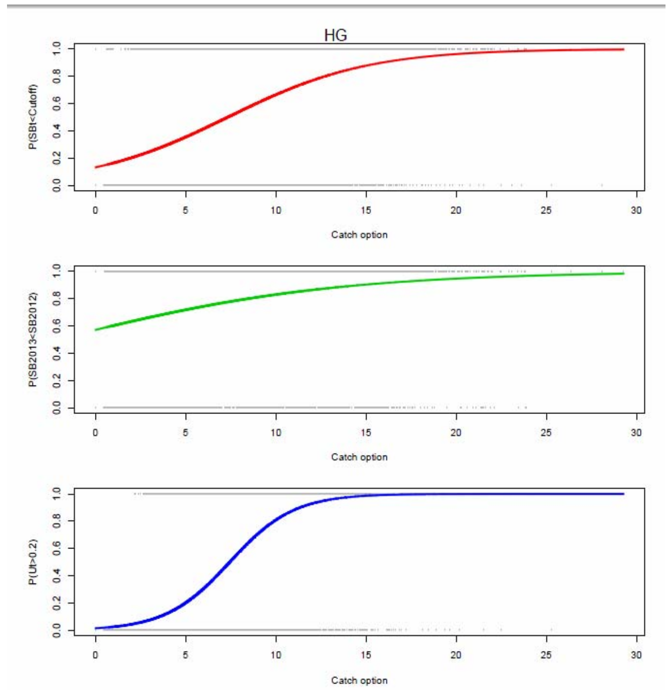
Tanasichuk, R. W. 2002. An evaluation of a recruitment forecasting procedure for Strait of Georgia herring. PSARC Working Paper P2002-01. 24p.

APPENDIX E: RISK PROBABILTY METRICS FOR MAJOR STOCK AREAS

SB_{2013}< Cutoff = Spawning stock biomass in 2013 falling below the cutoff level (0.25B0).
SB_{2013}<SB_{2012}= Spawning stock in 2013 being less than the spawning stock biomass in 2012.
U_{2012}>0.2 = 2012 harvest rate (catch option /3+ biomass) being greater than the target harvest rate of 0.2.

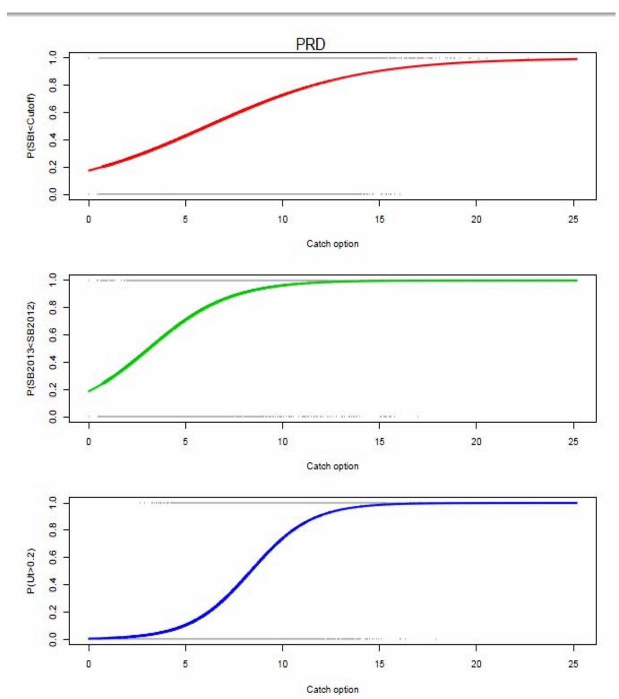
Haida Gwaii (2E)

	Catch option (tonnes)		
Risk Probability	**SB_{2013} <Cutoff**	**SB_{2013} <SB_{2012}**	**U_{2012} >0.2**
0.05	0	0	2,228
0.10	0	0	3,552
0.15	583	0	4,372
0.20	1,940	0	4,989
0.25	3,060	0	5,498
0.30	4,039	0	5,944
0.35	4,928	0	6,348
0.40	5,760	0	6,727
0.45	6,558	0	7,090
0.50	7,340	0	7,445
0.55	8,121	0	7,801
0.60	8,919	907	8,164
0.65	9,751	2,547	8,542
0.70	10,640	4,299	8,947
0.75	11,619	6,229	9,392
0.80	12,740	8,439	9,902
0.85	14,096	11,113	10,519
0.90	15,898	14,666	11,338
0.95	18,809	20,404	12,662



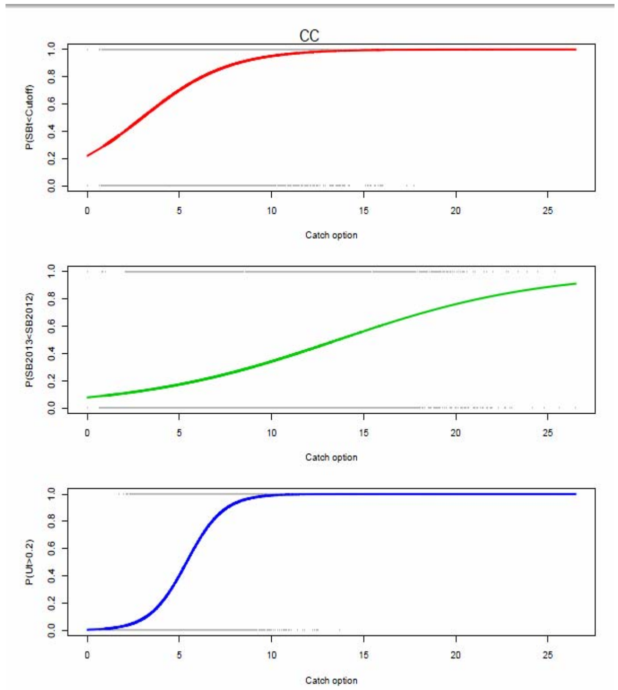
Prince Rupert District

	Catch option (tonnes)		
Risk Probability	**SB_{2013} <Cutoff**	**SB_{2013} <SB_{2012}**	**U_{2012} >0.2**
0.05	0	0	3,743
0.10	0	0	4,914
0.15	0	0	5,639
0.20	655	182	6,185
0.25	1,788	787	6,636
0.30	2,777	1,315	7,030
0.35	3,675	1,795	7,388
0.40	4,516	2,244	7,722
0.45	5,322	2,675	8,043
0.50	6,112	3,097	8,358
0.55	6,902	3,518	8,673
0.60	7,708	3,949	8,994
0.65	8,549	4,398	9,328
0.70	9,447	4,878	9,686
0.75	10,437	5,406	10,080
0.80	11,569	6,011	10,531
0.85	12,940	6,743	11,077
0.90	14,761	7,716	11,802
0.95	17,702	9,287	12,973



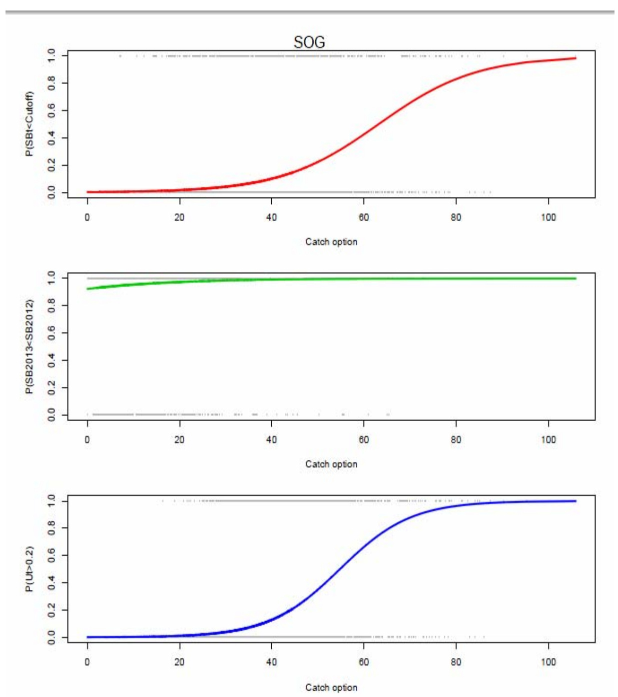
Central Coast

	Catch option (tonnes)		
Risk Probability	**SB_{2013} $<Cutoff$**	**SB_{2013} $<SB_{2012}$**	**U_{2012} >0.2**
0.05	0	0	2,450
0.10	0	1,521	3,195
0.15	0	4,063	3,657
0.20	0	5,977	4,005
0.25	402	7,557	4,292
0.30	994	8,938	4,543
0.35	1,530	10,192	4,770
0.40	2,033	11,366	4,984
0.45	2,515	12,491	5,188
0.50	2,987	13,593	5,388
0.55	3,459	14,696	5,589
0.60	3,940	15,821	5,793
0.65	4,443	16,995	6,006
0.70	4,980	18,249	6,234
0.75	5,571	19,629	6,485
0.80	6,247	21,210	6,772
0.85	7,067	23,124	7,119
0.90	8,155	25,665	7,581
0.95	9,913	29,771	8,327



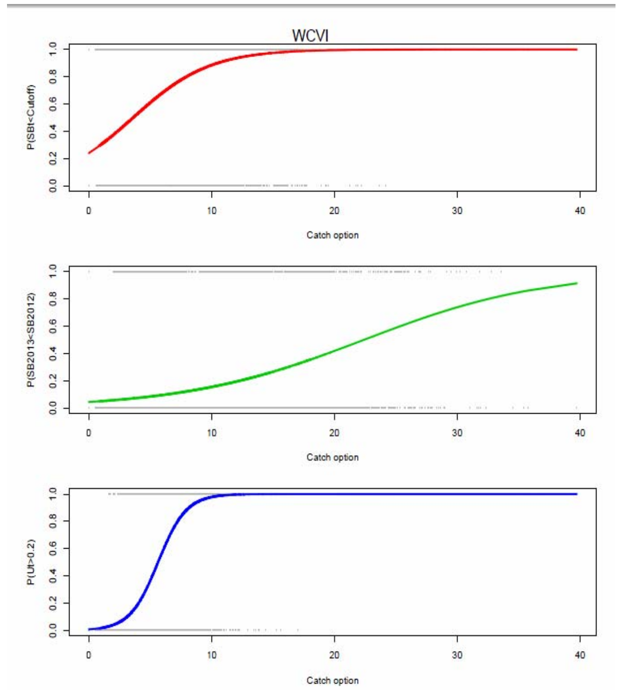
Strait of Georgia

	Catch option (tonnes)		
Risk Probability	**SB_{2013} $<Cutoff$**	**SB_{2013} $<SB_{2012}$**	**U_{2012} >0.2**
0.05	32,082	0	32,080
0.10	39,969	0	37,840
0.15	44,852	0	41,406
0.20	48,528	0	44,091
0.25	51,565	0	46,308
0.30	54,218	0	48,246
0.35	56,627	0	50,005
0.40	58,882	0	51,651
0.45	61,043	0	53,230
0.50	63,161	0	54,777
0.55	65,280	0	56,324
0.60	67,441	0	57,902
0.65	69,696	0	59,549
0.70	72,105	0	61,308
0.75	74,758	0	63,245
0.80	77,794	0	65,463
0.85	81,471	0	68,148
0.90	86,354	0	71,714
0.95	94,241	8,031	77,474



West Coast of Vancouver Island

	Catch option (tonnes)		
Risk Probability	**SB_{2013} <Cutoff**	**SB_{2013} <SB_{2012}**	**U_{2012} >0.2**
0.05	0	907	2,269
0.10	0	6,355	3,125
0.15	0	9,728	3,655
0.20	0	12,267	4,054
0.25	184	14,365	4,384
0.30	969	16,197	4,671
0.35	1,682	17,861	4,933
0.40	2,349	19,418	5,178
0.45	2,989	20,912	5,412
0.50	3,616	22,375	5,642
0.55	4,243	23,838	5,872
0.60	4,882	25,331	6,106
0.65	5,549	26,888	6,351
0.70	6,262	28,552	6,613
0.75	7,047	30,385	6,900
0.80	7,946	32,482	7,230
0.85	9,034	35,022	7,629
0.90	10,478	38,395	8,159
0.95	12,812	43,842	9,015



APPENDIX F: SUBMITTED REVIEWS

Review of Martell et al. 2011: "Part 1: Moving Towards the Sustainable Fisheries Framework for Pacific herring: Data Models, and Alternative Assumptions **"**
by Nathan Taylor (Reviewer 1), Groundfish Section, Marine Ecosystems and Aquaculture Division, Fisheries and Oceans Canada

I successfully downloaded, compiled and after some modifications to the output code ran the software on a 64 bit PC running Windows XP. Using it, I was successfully able to reproduce the assessment using version 1.1. I note that, line 2588, if(last_phase() && PLATFORM =="Linux" && !retro_yrs) effectively prevents the output from being written if users are not using Linux, or, I presume a Linux emulator like Cygwin. This problem was relatively easy to fix by modifying removing the control statement and changing the system command to a windows-compatible "COPY". I appreciate the difficulties of cross-platform development and given the superior efficiency and speed of using AD model builder on Linux, I would prefer to use it. However, Fisheries and Oceans Canada computers must use Windows so this is a problem that would be convenient to tackle. As a minor aside, own efforts to use PLATFORM as a flag to write either Windows or Linux type output failed.

I did not have time to independently reproduce the assessment using other software(s). This would have been my preference; using the source code provided, which I did, does not necessarily show that the approach is valid; it demonstrates that readers can follow the instructions for downloading and running the software. However, in addition to using ISCAM on the herring data, I also applied it five rockfish species and found the results consistent with analyses that I had already conducted. While not within the scope of the document, I'd have liked to have seen a more comprehensive set of simulations done, with less informative data. The value of being able to run the assessment model myself, and in particular, use the GUI to view model input and output cannot be overstated.

It is a big achievement that reviewers and others can easily run the model used for the assessment and more generally, it is excellent to have a set of stock assessment tools developed by a broader community of assessment scientists. I think that there are several advantages: producing a common set of outputs for DFO's may increase the comprehensibility and production-efficiency of fisheries assessments and, among other things, reduce coding errors. I commend the authors for their initiative in this respect.

I have a few general and specific comments about the paper. Several sections (the title, 1.2.2, 2.5, 2.8.10, etc) refer to using the DFO's precautionary approach to fisheries management for reference points but it is not clear what that means. Some of this ambiguity is not the authors' fault. The precautionary approach document is itself unclear with table 1 implying a control rule that depends on both recent stock trajectory and zone, including high and low in the cautious zone, yet the first figure of that same document implies a rule based exclusively on zones. Furthermore the PA approach seems to allow for reference points to be determined case-by-case. Annex 1b states that "Actual reference points for a stock may use other metrics and be set lower or higher than these (default) references but should be demonstrably appropriate for the stock and be consistent with the intent of the PA". If the paper means moving towards *Bmsy-based* reference points it should be stated as such.

Notwithstanding the ambiguity surrounding what it actually means to be PA compliant, I am not yet convinced that that moving "towards DFO's SFF (is actually) a necessary next step" as the paper claims in section 2.8.10. Firstly, the herring fishery apparently seems to have already defined the essential elements of the PA approach notably: LRPs, USRs, and a control rule (section 2.7.5); this is consistent with the PA approach. Secondly I cannot tell if the authors mean moving towards a *Bmsy*-based control rule or *Bmsy*-based reference points. If it's the latter, I cannot immediately see that moving toward a *Bmsy*-based control rule is demonstrably superior to the B_0 -based one given some problems with using *Bmsy* that the assessment incisively points out: the dependence of *Bmsy* on selectivity and fleet partitioning; moreover, there are others that it does not: risk tolerance of being in the critical zone given recruitment variability; how would a harvest control rule that scaled down fishing mortality between 0.8*Bmsy* and 0.4*Bmsy* perform in a stock where the thickness of that zone is very small?; are assessment model precise enough to use *Bmsy*-based control rules that didn't shut off fishing at below 10 % B_0 ? It's interesting to note the B_0 -based control rule has done a good job at keeping the stock out of the *Bmsy*-based cautious and critical zones (Fig. 2.19). Moving toward the PA approach might imply a suite of nuanced issues beyond estimating *Bmsy* i.e.: whether or not *Bmsy*-based reference points should be adopted, what the basis for that choice would be and if applicable, what the new control rule would be.

I have several minor comments that I list below.

- I don't understand the concept of selectivities being greater than one. How selectivity can be greater than, is an issue that needs discussion.

- It's not clear why several prior distributions were chosen. These should be documented. For example, A.1.15, the variance partitioning parameter ρ~Beta (15,60). In general, the rational for the choice of priors needs to be documented, even if the reason is to improve convergence behaviour.

- Fig. 2.23 and others like it are undecipherable unless I increase my screen magnification to 300 %. Also the y-axis has no scaling whatsoever.

- I had the impression iSCAM is parameterized in terms of steepness *h* and not the Goodyear compensation ratio κ (as defined in table A-4) so I am confused about whether or not κ is an estimated parameter in ISCAM. The simulation evaluation portion of the assessment suggests that the estimated parameter vector consists of: *log(Ro)*, *h*, *log(m)*, *log(Rbar)* and *log(Rinit)*. However in the ISCAM code, κ appears as: the seventh element of the estimated parameter vector (along with steepness h as the second element of theta on line 509), and as a global variable type (line 568) that gets transformed from steepness on lines 685-693.

- Furthermore, line 20 of the "SOGHerring2011q1.ctl" control file kappa appears to be bounded between 0.01-5 with the comment " (precision)" beside it; pairs plots of the leading parameters (Fig. 2.33 for example) have both κ and *h*; in the model documentation, κ appears on page 81 to define *so*; κ defines θ, in tables A-1, A-3 implying that it is estimated but *h* does not in tables A-1 or A-3 at all; finally, section 2.8 describes the seventh element of the parameter vector as being the total variance □. It is essential that the model presentation clarify this, and make the notation consistent throughout the document and ideally, the computer code.

- To be consistent with the description of B_0 in terms of spawning stock biomass, it might be useful to describe the reference points 0.4 and 0.8*Bmsy* in terms of *SSBmsy* also. It may already be so, but I note that how *Bmsy* is derived from the *Fe* and *Ce* does not appear in tables A-1 and A-3.

- The caption of Fig 2.19 defines critical and cautious zones in terms of *Bmsy/*B_0 whereas the PA approach defines the Limit Reference Point where *Bt*< *0.4BMSY*. Does the paper mean *Bt/Bmsy*<0.4 instead of *Bmsy/B0*<0.4? If I take the values from table 2.3 and compute *0.8Bmsy/*B_0 , it appears to be consistent with what is on this figure (see below). As a minor aside, does the *Bt/B0* row in Table 2.3 mean *B2011/B0*?

Stock	**HG**	**PRD**	**CC**	**SOG**	**WCVI**
No.	159	206	190	235	174
FMSY	2.36	0.54	1.31	1.4	0.98
MSY	8,761	6,669	9,104	27,442	10,260
B_0	40,684	68,761	59,365	135,523	57,462
$0.25B_0$	10,171	17,190	14,841	33,881	14,366
BMSY	8,708	18,600	11,514	28,211	11,281
0.8BMSY	6,966	14,880	9,211	22,568	9,025
0.4BMSY	3,483	7,440	4,605	11,284	4,512
Bt	16,723	27,288	14,624	129,070	14,909
Bt/B_0	0.41	0.4	0.25	0.95	0.26
$0.8Bmsy/B_0$	0.171222	0.216402	0.155159	0.166525	0.15706
$0.4Bmsy/B_0$	0.085611	0.108201	0.077571	0.083263	0.078521

Review of Martell et al. 2011: "Part 1: Moving Towards the Sustainable Fisheries Framework for Pacific herring: Data Models, and Alternative Assumptions**"**
by Sean Cox, Simon Fraser University (Reviewer 2)

Terms of Reference for review
Specific technical components of the Pacific herring statistical catch-age population assessment model assumptions and structure to be evaluated include:
Model structure and parameterization,
Confounding effects of natural mortality (M) and spawn survey catchability (q),
Gear selectivity, catchability, recruitment, and productivity estimation,
Model priors, likelihoods, and decision rules, and
Estimation of unfished biomass and stock depletion levels (time series of annual biomass over
unfished biomass).

This document attempts to cover considerable ground in "moving towards the sustainable fisheries framework". I think the scope is too broad at the moment, because there are many open questions and some critical uncertainties the need further attention. If producing a "best assessment model" is the goal, which appears to be the case here, then the paper should focus on a more complete evaluation of the iSCAM assessment model under a range of conditions encompassing the assumptions used in the final assessment. Basically, this assessment appears to be a work-in-progress and I make some recommendation for revisions that would may hopefully be of help.

It is probably easiest to go section-by-section of the paper to cover key points that I think need work.

Specific comments
Section 1.1: The paper reads like an instruction manual in several places, rather than a research document. Specific instructions on how to execute the ADMB program should be left to an appendix.

Some background for the herring fishery assessments would have helped provide necessary context for this particular assessment. Throughout the review I'm wondering why a new assessment model is necessary, what the perceptions of stock status and over-fishing were based on previous models, etc.

The statement: "These cutoffs are currently thought to be more conservative than the current default Limit Reference Point of $0.4B_{MSY}$", which appears in a couple of places in the document, is a good example of the need for some background. It is not clear "who" thinks this and why. Is it these authors' opinion? I doubt that the reference provided (i.e., DFO 2006) has any specific opinion on the herring fishery. How is conservative defined? The point of clearly articulated harvest policies is to make these definitions clear so that everyone understands them. What is "conservative" for industry objectives may not be conservative in the view of other stakeholders or from a biological perspective.

Section 1.2.2 (SFF): This section needs to provide a more accurate description of how harvest strategies are developed and also a more complete description of the SFF. DFO's particular approach to developing harvest strategies is not "The general framework" as stated in the first sentence. The DFO approach is a "specific" example of an attempt to develop a precautionary strategy (which also happens to be untested and not well-articulated). Note that "Table 1" of the SFF, which describes acceptable tolerances for stock decline depending on stock status and recent abundance trends, is apparently mandatory to include in the harvest strategy. No reference is made to that here.

I'm not sure why the remaining discussion of equilibrium-based reference points for multi-gear fisheries is in this section. I agree that non-stationary parameters increases the difficulty of estimating reference points, but the paper provides no indication of how to deal with that problem. In fact, the paper ultimately uses equilibrium-based reference points, stationary parameters, and a constant allocation of quota among fleets. So, I wonder why the paper derives reference points this way, when the whole point of developing a "new" stock assessment model is to track non-stationarity parameters. From the assessment viewpoint (as stated on page 6), any policy recommendations based on equilibrium assumptions will be wrong the moment they are applied.

Section 1.2.3 (Simulation tests): First, I would drop the simulation tests based on perfect information. This is only a test of whether the ADMB fitting algorithm works under error-free data conditions, which (hopefully) has been well established by now. If the same model is used for simulation and estimation on error-free data, then it is impossible to tell whether the model is "correct" or not.

This assessment is extremely complex and, from what I think I know about herring, important consequences may follow from assessment errors. Therefore, the purpose of Part 1 should be a very thorough simulation test of the model under a **range of assumptions at least as wide as those used in the real assessment**. The simulation test in the paper does not appear related to what is actually done in the assessment. For example, (a) the assessment model is given the true proportion of observation error, which is not done in the real assessment, (b) the total standard deviation is 0.4, which seems somewhat low given the levels of observation and process error in the real assessment, (c) key parameters like M and selectivity function coefficients are constant, and (d) the data-generating distributions are identical to the likelihoods used in estimation (which is common). I don't know what happens to weight-at-age or how it is used in the simulation test, but it seems like it might be important to the actual assessment. Ultimately, one cannot use these simulation tests to critically evaluate the bias and precision of the assessment model because they are not the same. The results shown in Figure 1.3 should not be used as an indication that this model is "unbiased and precise" because the situation is not particularly realistic. It is certainly encouraging, even under optimistic conditions, but I wouldn't stake the fishery on this single test. In a revision, I would expect to see a simulation test of every model considered in the assessment under (1) correctly specified structural model (e.g., time-varying M, observation, process errors, etc.), (2) high and low levels of observation to process error ratios, etc. Although B_0 x-plots are useful for particular graphical comparisons, the simulation tests need more specific performance measures for judging bias and precision (e.g., mean relative and absolute errors, root mean square errors, etc. may provide more succinct summaries).

Section 1.2.4 (pooling age classes): Pooling age-classes that have low representation in the age-composition makes sense. But, why use a cutoff of 2%? And, what happens to the effective sample size when several age classes are pooled? The model specification in Appendix 1 (pg 91) is not what is actually used in the computer code – these should be a much closer match.

Section 1.2.4 (commercial catch): Why treat the commercial catch as continuous, when the fishery operates in a short pulse? The assumed errors in catch are small relative to everything else that is going on in this model. Besides, it would be relatively easy to test the impacts of catch errors on a discrete catch model (in the revised simulations).

Section 1.2.4 (spawn survey data): First, I am surprised that an in-depth description of the spawn surveys is missing given the importance of this survey to the assessment. That needs to be provided.

Second, the plots always show pre- and post-1988 spawn surveys on the same graphs and this seems a bit misleading to me. These two surveys are quite different, so there should not be an impression of continuity.

Third, presentation of the precision parameters needs to be more consistent. In some places, precision is 1/variance (the correct way) and in others it is 1/std deviation. Std deviations are not additive as implied in the last sentence ("To implement…rho = 0.35/1.15). Hopefully, this is done correctly in the code, I have not checked.

Section 1.2.4 (prior distributions): The paper should be clearer about non-informative vs vague priors. Non-informative priors imply only a known range, whereas vague priors include distributional information. A log-normal distribution for M may be appropriate on numerical grounds, but there is usually plenty of information in the data to conclude that M > 0. So, a normal prior may do just as well.

The description of how time-varying M is modeled raises an important general issue about the likelihood function for this model. In particular, it appears that only part of the process error is included in the ratio of process to observation errors. Time-varying M and time-varying selectivity both contribute to total variance as process errors, but they are not included in the total variance. Instead, the random-walk variances are assumed known, which means the variance estimates for recruitment and observation errors are conditional on variances for M and selectivity even when the total variance is estimated in Part 2. Also please see the bottom of the 3rd-to-last paragraph and check whether these weights for input to HCAM are computed correctly. As noted above, the ratio parameter "rho" must be computed from variances, not std deviations.

Why is the prior for stock-recruitment steepness so informative for both models? Is this assessment more interested in explaining noise in the data rather than estimating a fundamental production parameter? Whether the prior steepness is 0.5, 0.6, 0.7 can have pretty important implications for harvest policies, especially when the data are not contributing anything further information as indicated later.

Finally, the whole section should be edited for typos, wording, redundancies, etc.

Section 1.3 (Results)
Table 1.1: First, this table should be expanded to show the complete set of results following from each assessment model. That is, all components of the harvest control rule, including spawning biomass depletion, Fmsy, Bmsy, recommended exploitation rate, etc., should be provided along with the output quota recommendation. The differences shown in the table may trade-off and ultimately result in the same exploitation rate relative to Fmsy and escapement relative to Bmsy.

The text describing the table is not very informative about the differences between models, e.g., What is it about the different error structure assumption that would lead to such large differences between models?

Please explain the residuals in Figure 1.5 – they don't seem to correspond to the fits in Figure 1.4.

Section 1.3.3 (Alternative assumptions…): As suggested for the table above, Table 1.2 should provide a complete list of harvest control rule parameter estimates and output quantities.

The number of parameters in Table 1.2 really aren't all parameters because there are different priors involved in each one. It would be better to use DIC-based calculations of the effective number of parameters.

Section 1.3.4 (Preliminary assessments…): The previous sections described alternative models without making a clear choice for going forward with these preliminary assessments.

Other Editorial
1. I think $2\theta' = \theta$ should be $\theta' = 2\theta$ and $0.5\theta' = \theta$ should be $0.5\theta' = \theta$. In addition, the label of Figure 1.3 shows log2(true/estimate), which is different from what is presented in the text.

P4 1.2.2 – The line "In the case of a single fishing fleet…" needs to be re-worded. Bmsy does not just apply to "fixed gear" fisheries.

I would rather see the actual gear types presented in the report figures rather than generic "Gear 1", "Gear 2", etc.

Appendix A.1 Technical Description of iSCAM

Section A.1.2: The equilibrium model is not the actual one used in the reference point calculations. The actual one has multiple gears, an allocation of yield to each, stock-recruitment is parameterized differently, the Ricker model is not considered, etc.

In the parameters to Table A-1, what is M_a? If it is age-specific, then where is the model for how M varies with age?

In this section, I would rather see examples of equilibrium yield and reference point calculations for each of the management areas for a baseline model rather than a generic example as shown in Figure A.46.

Page 84: delete "arbitrarily" in the first sentence since I doubt allocation was determined without any reasons. Why use the 20-year average rather than something more recent? Is there a fixed allocation in place in the management plan?

I don't understand why 100% of the total mortality is assumed to take place prior to spawning? What is the assumed date of spawning? What are the start and end dates for this model? Jan 1 – Dec 31? If all mortality is assumed to take place prior to spawning, why use a continuous fishing model that averages abundance over the year to account for continuous mortality?

The caption to Table A-3 is incorrect. This model is not parameterized as stated.

Now I am beginning to worry about the assumed observation error proportion of total error – see T3.2. It is based on standard deviations. This is not consistent with the likelihood in equation A.8. **And check the ADMB code too.**

Equation T3.7 does not specify the initial M at t=1.

Equation T3.11 has a misplaced error η_t : doesn't that apply to the whole sum?

I have a cheap equation editor, so I can't reproduce some of those funky symbols, but please provide the distributional assumptions for all the error terms (deviates) – I counted 5 error terms plus the distribution for M_1.

Similar to the equilibrium calculations, the age-composition likelihood is not the actual one used in the assessment – eqs A.9-A.10 do not account for pooling. Again, the word "arbitrary" in choosing minimum proportions of 2% does not instill confidence…

The residual in Equation A.11 is not consistent with the recruitment model. The function $f(B_{t-k})$ should be $\log(f(B_{t-k}))$.

The stock-recruitment section does not describe the parameterization used in the assessment.

Figure A.49: please show the actual priors used in the assessment. Generic examples are not useful here.

Section A.1.16 (Survey priors): The first part of the first sentence is incorrect. The survey q **IS** being estimated – there just happens to be a closed-form MLE.

I don't want to nit-pick, but does a reference from 1973 (i.e., Hardwick 1973) provide much confidence about the time-invariance of herring eggs per gram? Isn't this easy to measure today?

Appendix C.3 Bayesian prior for the dive survey…*q*

I have too many concerns about the prior developed in this section to endorse its use for assessment. My first, obvious question, is WHY? How does "Moving towards a sustainable fisheries framework…" imply "develop an informative prior on catchability from scant literature and some large assumptions"?

Second, it is not a rigorous scientific approach to the problem. There is no discussion of the well-known problems and risks in using "expert knowledge", or a very limited available literature, which is what "expert" seems to mean in this case. There is no discussion of how the "factors" might have changed over time or space or spawning habitat, or vegetation changes, etc.

If there is so much concern about catchability, why not try allowing *q* to vary over time and let *M* and selectivity remain fixed? If herring spawn in different locations each year, then why should *q* remain constant?

Finally, there is not plan to assess the potential impacts of using this prior on the future of this resource. Assuming q=1 at least produces a "minimum" biomass estimate so that any other assessment and management errors are buffered. Using this prior would remove any such safeguards against severe over-fishing and would, in fact, increase the risk of stock decline by several-fold. The fact that none if this is even mentioned in the text warrants rejecting it outright.

Review of Martell et al. 2011: "Part 2: Stock Assessment and Management Advice for the British Columbia Pacific Herring Stocks: 2011 Assessment and 2012 Forecasts**"** by Nathan Taylor (Reviewer 1), Groundfish Section, Marine Ecosystems and Aquaculture Division, Fisheries and Oceans Canada

It is not possible to completely separate the technical review of the paper (part I) from the review of the harvest advice. Naturally, problems related to: the validity of the modeling approach; the formulation of the control rule; statistical likelihoods; data treatment; etc. propagate through to the stock assessment and management advice for BC Herring stocks. I won't discuss these any further in this section but will assume that any unresolved technical and/or presentation problems from part I of this assessment remain germane to the second. Some clear statement of the management objectives; the suitability of the control rule; upper and lower limit reference points remain key items.

My main comment about part II is that harvest advice requires a more thorough presentation of uncertainty. One the management side, there appears to be two sets of reference points in play. On the stock assessment side, there is considerable uncertainty in both parameter and structural dimensions. I am not sure what the history of presenting management advice has been in the herring fishery but the decision table 2.6 (or its sister for minor areas) is impoverished. Typically such a table would present in each column, the probability of a range of key management outcomes in this case for example: the probability that the stock is below or above cut-off levels/zones; the probability that the stock will be in a particular zone given a range of harvest levels. The presentation of the harvest advice in probabilistic terms is particularly important in this case. Modeling with ISCAM has attempted to capture: estimating q; the time varying effects of selectivity and natural mortality so the parameter uncertainty is surely considerable. I recognize that sampling from the recruitment percentiles in table 2.6 attempts to capture some of this uncertainty but high, medium and low recruitment classification seem arbitrary. Moreover, there appear to be several competing hypotheses about how to represent the population dynamics: fixing q at one; using the prior on q; and the minimum proportion in the age composition likelihood. It would be helpful if the decision table presented the probabilities of being in any of the old, or new stock status zones for each hypothesis about the status of the stock. Finally, the PA approach also attempts to address risk tolerance (see table 1) and it's not possible to determine if a particular quota will result in low, medium or high probabilities of being in, for example, the critical zone. To the extent that is practically possible, the assessment needs such a decision table showing the probabilities of being in: old, new cut-off zones, PA critical, cautious and healthy zones, and probabilities that F>Fmsy for each hypothesis about the status of the stock.

Review of Martell et al. 2011: "Part 2: Stock Assessment and Management Advice for the British Columbia Pacific Herring Stocks: 2011 Assessment and 2012 Forecasts" by Sean Cox, Simon Fraser University (Reviewer 2)

Terms of Reference for review

What is the estimated 2011 spawning biomass for Pacific Herring by major and minor stock assessment area?

- *How are herring stocks in these areas changing over time?*
- *What is the forecasted spawning biomass for 2012?*
- *Are there any specific concerns that Fisheries Management should be aware of, and if so, what are those concerns?*

Before getting into Part 2, I want to make it clear that the modelling approach developed in Part 1 is a unique and state-of-the-art platform for stock assessment. The range of options, as well as simulation capability, allows for a very thorough and in-depth study of the reliability of stock assessment models using both maximum likelihood and Bayesian paradigms. With that being said, I found it hard to believe that four authors contributed to these papers. Most of my comments and critiques were issues that should have been identified, discussed, and worked out well in advance of the final version of the paper. Of course, I may be completely wrong, but regardless, I encourage all the paper authors to develop a plan for collaborating on the project in a way that allows the importance and value of the new modelling platform to be exploited to the fullest. What is presented in the papers really only scratches the surface of iSCAM's capability, so I suggest further analyses that would improve it's application to the herring stock assessment.

In this review, I will focus on only a few major issues. All editorial and lesser importance comments are on the hardcopy, which I will forward to the authors when I return to Vancouver later this week.

Harvest Control Rule and Decision Rules

Section 2.7.5 (Catch advice): I don't think one can apply both the harvest control rule described at the top of page 36 **and** the decision rule described in the three bullet points just below it. These appear to be mutually exclusive. The decision rule in the bullet points is what has actually been used to manage this fishery in the past. The harvest control rule using a LRP and USR is DFO's default policy.

The "decision rule", as it is described in the three bullets, violates the principles of the SFF and the Precautionary Approach (PA). When the stock is between the LRP and USR, the "total allowable catch is based on a reduced harvest rate **that would deplete the stock to the LRP level**" [my emphasis]. The point of the SFF and PA is to avoid the LRP; making the LRP a target when the stock is in the cautious zone is the exact opposite of what these two policies prescribe. If that is actually the way harvest decisions are made, then the stock assessment model is not the biggest issue to deal with for this fishery.

I cannot really blame the authors for not providing a clear and specific interpretation of the SFF since I don't know of another fishery in Canada besides sablefish that has actually tried to implement it. And we found that the SFF is is not self-consistent.

There are two difficulties in this paper in applying the SFF. First, the LRP and USR should not be used to define a harvest control rule: those reference points are used to establish conservation objectives only. The harvest control rule, which can take any form, should be defined so as to avoid the LRP with some high probability. The USR serves little purpose except to define the upper bound of the "Cautious Zone". The Target stock status, which is somewhere in the Healthy Zone, is defined by the Reference Removal rate, e.g., F_{msy}.

The second difficulty here is attempting to force the existing herring decision rule into the SFF harvest control rule form. The reference points used in the two approaches are very different. According to two of this assessment's authors (Cleary et al. 2010), the herring decision rule is a hybrid strategy involving a "minimum escapement, fixed exploitation rate" defined by

$$U_t = \begin{cases} \min\left(\dfrac{fSSB - 0.25B_0}{fSSB}, 0.2 \right) & fSSB > 0.25B_0 \\ 0 & fSSB \le 0.25B_0 \end{cases}$$

where U_t is the target exploitation rate and *fSSB* is the forecast spawning biomass. This rule says that, when the stock is above the "cutoff", the fixed exploitation rate of 20% applies, and when the stock is below the cutoff the minimum escapement applies (note: this formula should be confirmed and put into the paper to add clarity). Note that the min() function avoids "depleting the stock to the cutoff" under all conditions as implied in the paper's description of the decision rule (see emphasis above). The actual rule limits the exploitation rate to less than 20% under all circumstances.

In contrast, the harvest control rule described in the SFF is of the form,

$$U_t = \begin{cases} 0.2 & fSSB \ge 0.8B_{\mathrm{MSY}} \\ 0.2\left(\dfrac{fSSB - 0.4B_{\mathrm{MSY}}}{0.4B_{\mathrm{MSY}}} \right) & 0.4B_{\mathrm{MSY}} \le fSSB < 0.8B_{\mathrm{MSY}} \\ 0 & fSSB < 0.4B_{\mathrm{MSY}} \end{cases}$$

This rule incorrectly mixes conservation objectives up with operational rules, so it really shouldn't be applied this way. The most recent sablefish assessment describes how we defined harvest control rules in quasi-SFF terminology and how we used SFF references points and risk tolerances to define fishery objectives.

There is nothing wrong, in principle, with the herring decision rule. It is has a target stock size associated with a 20% average harvest rate, and should promote stock recovery by reducing the harvest rate when the stock is below this target level, provided that (i) the target associated with 20% harvest rate is at least B_{msy} and (ii) B_{msy} is greater than the cutoff of $0.25B_0$. This just needs to be demonstrated in the paper, perhaps, as I suggested in Part 1, by providing specific equilibrium analyses to identify all these specific reference points.

The only way to demonstrate that the herring decision rule is consistent with the SFF is to show that the rule avoids the LRP, e.g., $0.4B_{msy}$, at least 95% of the time over some reasonable timeframe. This is beyond the scope of this paper because it requires closed-loop simulation. However, it might be possible to do some basic simulation work to examine whether the herring decision rule will deplete the stock below $0.4B_{msy}$ in the short term given expected error levels and across several scenarios/models. Or, one could draw samples of all harvest control rule parameters from their joint posterior distribution to compute the probability of post-fishery biomass being below $0.4B_{msy}$.

Section 2.7.1 (Input data): As I mentioned in Part 1, I am not very comfortable with my level of knowledge about the data used in this assessment. This assessment should have a complete and detailed description of the sampling procedures and any transformations, pooling, etc. that happens prior to assembling the final assessment data set.

Concern 1: risks associated with sampling being directed to reported spawning activity. Has there been any evaluation of whether this protocol is prone to cause changes in mean survey catchability or precision?

Concern 2: use of the age composition derived from test fishery charters. There is no Gear = TestFishery, so I'm not clear on how these samples are used. Are they combined with purse-seine or gillnet samples? Is there any reason to suspect that test fishery selectivity is different from either commercial type? Are changes in selectivity estimated by the model associated with more or less influence of test fishery age composition?

Concern 3: use of mean weight-at-age data as input to assessment. I'm assuming that mean weight-at-age data come from test fishery samples, but that is not obvious. Regardless, what is the rationale for using this data directly in biomass calculations and selectivity functions? Has anyone showed, by simulation, the types of biases this introduces? The data are already subject to some form of size-/age-/maturity-dependent selectivity, so how does this interact with size-dependent selectivity models used later?

An alternative approach is to allow for, e.g., a cohort-specific L_∞ parameter and fitting the model to observed weight-at-age data from a specific gear (and its selectivity function). There is plenty of data, so these parameters will probably be reasonably well-determined. I haven't though too much about this, so I would not expect a solution for this paper. However, there needs to be some more depth to the discussion of the existing approach and some alternatives.

Section 2.7.5 (Catch advice): I might as well bring this up here, but it also applies to the Result and Discussion. What is the purpose of providing catch advice under an arbitrary selection of Poor, Good, and Average recruitment? Is this paper leaving it up the managers to decide what recruitment category to use? What information will they use to make that choice? I would like to see an appraisal of the decision matrix these managers face give these scenarios along with the alternative parameterizations.

Section 2.8 (Results): First, pages and pages of residual and bubble plots are distracting and not particularly informative about the key issues involved in this assessment. Most of the text associated with these plots provides only qualitative descriptions anyway – some actual statistics like runs tests, autocorrelation, etc. would be more informative.

Table 2.3: This table is the most important piece of information in the entire assessment and should therefore be discussed and analyzed in greater depth. For instance, the range of Fmsy estimates (0.54/yr to 2.36/yr) is astounding. What I find most interesting, though, is that PRD has the lowest (Fmsy = 0.54/yr) and it is really the only stock for which there is any hint of information about the stock-recruitment steepness parameter. Figure 2.38 shows that posterior distributions for steepness (h) are almost identical to the priors with PRD being the most different. Therefore, for most stocks, the fishing mortality reference points are more functions of the priors than functions of the data. Estimates for HG, in particular, are too high to be credible. I also wonder how much information there really is about the key scale parameter Ro and whether that is just the result of an informative prior on catchability.

A minor point - Reference points in this table should be shown relative to B_{msy} or B_0 so they are comparable across stocks the way Fmsy is.

I find it difficult to believe that the pre-fishery SOG herring stock is **at or above the unfished level**. I'm equally concerned that there is no mention or discussion of this estimate in the paper. Are there any fishery or ecosystem indicators suggesting that herring are at pre-fishing levels? Why would this model generate estimates like this? Is it "precautionary"?

According to Figure 2.19, the SOG stock never entered the Critical Zone when, by all accounts, it collapsed in the 1960s. Is this just a poor SFF default choice of $0.4B_{msy}$ as a LRP for this stock, or is the fitted SOG model not capable of generating stock sizes that depleted? Clearly, the other stock biomass estimated come much closer to the LRP during collapse, so perhaps it is just an SOG issue.

Figure 2.19 also shows that the WCVI and CC stocks have been below the cutoff. This suggest that, in hidsight, the herring decision rule may not have worked as intended in the mid-2000s. It would be informative to show how the estimated cutoffs have changed over time. This could be done easily from the existing retrospective analysis by just outputting estimated cutoff values each year.

For the pairs plots (Figs 2.33, 2.34), I would be most interested in seeing what is correlated with catchability. The caption indicates that only leading parameters are shown, but I disagree with this. Clearly, **the assessment has put everything on the prior for catchability so it must be considered "leading" in this particular case**. For others, it is difficult to tell whether those plots are showing lack of correlation in the posteriors or lack of correlation in the priors. For instance, as mentioned above, posteriors for steepness are mostly identical to priors, so I would not expect much correlation with any other model parameters.

Section 2.8.3 (Estimates of mortality): As mentioned before, it is difficult to define the number of parameters for these models when parameters have different prior precision. It is also difficult to know how many effective parameters are involved in 12 or 60 nodes for the spline functions. You still have a time-varying parameter, so the interpolated values must count to some degree as parameters.

Are the trends in estimated natural mortality associated with the changes in mean weight-at-age?

Figure 2.18: multiplying numbers-at-age by empirical weight-at-age doesn't produce total biomass as stated. The weight-at-age is dependent on a selectivity function, so the estimated total biomass is actually positively biased with respect to the true population biomass.

Section 2.10 (Outstanding issues): I appreciated most of the issues raised in this section, since many issues caught my attention earlier. I think the discussion should be expanded to go into greater depth on these issues, as well as to propose a plan for moving forward in dealing with them. It is also important to discuss how managers might be able to cope with so many types of uncertainty.