Fisheries and Oceans Canada / Pêches et Océans Canada
Ecosystems and Oceans Science / Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Proceedings Series 2015/051

Pacific Region

Proceedings of the Pacific regional peer review on Coastwide Assessment of Arrowtooth Flounder (*Atheresthes stomias*) in 2014 and recommendations for 2015

May 12-13, 2015
Nanaimo, British Columbia

Chairperson and Editor: Rowan Haigh

Fisheries and Oceans Canada
Science Branch
3190 Hammond Bay Road
Nanaimo, BC V9T 6N7

December 2015

Foreword

The purpose of these Proceedings is to document the activities and key discussions of the meeting. The Proceedings may include research recommendations, uncertainties, and the rationale for decisions made during the meeting. Proceedings may 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.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

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

ISSN 1701-1280

Correct citation for this publication:

DFO. 2015. Proceedings of the Pacific regional peer review on Coastwide Assessment of Arrowtooth Flounder (*Atheresthes stomias*) in 2014 and recommendations for 2015; May 12-13, 2015. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2015/051.

TABLE OF CONTENTS

SUMMARY

These Proceedings summarize the relevant discussions and key conclusions that resulted from a Fisheries and Oceans Canada (DFO) Canadian Science Advisory Secretariat (CSAS) Regional Peer Review (RPR) meeting on May 12-13, 2015 at the Pacific Biological Station in Nanaimo, British Columbia (BC).

Arrowtooth Flounder (*Atheresthes stomias*) is a significant species within the groundfish trawl fishery. Since 2006 the fishery has been subject to a coastwide trawl Total Allowable Catch of 15 000 mt. Arrowtooth Flounder is also a minor component of the groundfish hook and line fisheries, and subject to vessel trip limits. The combined hook and line fisheries catch in 2013 was estimated to be 73 mt. Arrowtooth Flounder is the second most frequently released species in the Pacific Halibut (*Hippoglossus stenolepis*) hook and line fishery (est. 12 mt) and consequently was identified as a component of the non-target catch in the Marine Stewardship Council certification process for the Pacific Halibut fishery. The last official assessment of Arrowtooth Flounder was published in 2001 (Fargo and Starr 2001) and provided analyses of biological, survey, commercial catch and catch-per-unit effort data. These proceedings summarize the discussions from the RPR process that reviewed the first model-based assessment ever performed for the BC coastwide stock of Arrowtooth Flounder.

Meeting participants included representatives of DFO Science, DFO Fisheries Management, the Alaska Fisheries Science Center (NOAA, USA), the Canadian Groundfish Research and Conservation Society, the Commercial Industry Caucus (trawl), the BC Trawl Fishery, the Sport Fishery Advisory Board, the David Suzuki Foundation, and the Centre for Science Advice – Pacific (CSAP).

The conclusions and advice resulting from this review will be provided in the form of a Research Document and a Science Advisory Report providing advice from Science to managers and other clients. These documents will be publicly available on the DFO's CSAS website.

Compte rendu du processus d'examen par les pairs de la Région du Pacifique sur l'évaluation du stock de plie à grande bouche (*Atheresthes stomias*) sur toute la côte en 2014 et recommandations pour 2015.

SOMMAIRE

Le présent compte rendu résume les discussions pertinentes et les principales conclusions qui découlent de la réunion d'examen par les pairs de la Région du Pacifique, du Secrétariat canadien de consultation scientifique de Pêches et Océans Canada (MPO), qui a eu lieu les 12 et 13 mai 2015 à la Station biologique du Pacifique de Nanaimo, en Colombie-Britannique.

La plie à grande bouche (*Atheresthes stomias*) est une espèce importante pour la pêche au poisson de fond au chalut. Depuis 2006, un total autorisé des captures de 15 000 tm est appliqué à la pêche au chalut sur l'ensemble de la côte. La plie à grande bouche est également un élément mineur de la pêche au poisson de fond avec ligne et hameçon, pour laquelle les bateaux sont assujettis à des limites de sortie. En 2013, le poids combiné des prises de pêches avec ligne et hameçon a été estimé à 73 tm. La plie à grande bouche est la deuxième espèce le plus souvent rejetée dans le cadre de la pêche au flétan du Pacifique (*Hippoglossus stenolepis*) avec ligne et hameçon (environ 12 tm) et, par conséquent, elle a été désignée comme faisant partie des prises non visées par le processus de certification du flétan du Pacifique du Marine Stewardship Council. La dernière évaluation de la plie à grande bouche a été publiée en 2001 (Frago et Starr 2001) et contenait des analyses de données biologiques, de données provenant des relevés et de données sur les prises commerciales et sur les captures par unité d'effort. Le présent compte rendu résume les discussions tenues au cours du processus régional du premier examen par les pairs de l'évaluation par modèles jamais effectuée au sujet du stock de plie à grande bouche (*Atheresthes stomias*) sur toute la côte de C.-B.

Les participants à la réunion provenaient du Secteur des sciences du MPO, de la Gestion des pêches du MPO, du Alaska Fisheries Science Center (NOAA)(États-Unis), de la Canadian Groundfish Research and Conservation Society, du regroupement commercial du secteur industriel (pêche au chalut), des pêches au chalut de fond de la Colombie-Britannique, du Conseil consultatif sur la pêche sportive, de la Fondation David Suzuki et du Centre des avis scientifiques – Région du Pacifique (CASP).

Les conclusions et les avis qui découlent de cet examen seront présentés sous forme d'un document de recherche et d'un avis scientifique offrant des conseils du Secteur des sciences aux gestionnaires et à d'autres clients. Ces documents seront accessibles au public sur le site Web du SCCS du MPO.

INTRODUCTION

This proceedings document summarises the main discussion points arising from the regional peer review (RPR) meeting on May 12-13, 2015 held at the Pacific Biological Station in Nanaimo to review an assessment for Arrowtooth Flounder (*Atheresthes stomias*). The Terms of Reference appear in Appendix A. Working paper written reviews appear in Appendix B. The agenda and participants for the RPR meeting appear in Appendices C and D, respectively. The RPR meeting participants (Appendix D) are collectively called "participants" herein. Unless otherwise noted, figure and table references are to those in the working paper.

The conclusions and advice resulting from this RPR will be provided in the form of a Canadian Science Advisory Secretariat (CSAS) Science Advisory Report (SAR) to inform fishery management decisions. The SAR and supporting Research Document will be made publicly available on the CSAS website.

REGIONAL PEER REVIEW MEETING

Working Paper:	Arrowtooth Flounder (*Atheresthes stomias*) stock assessment for the west coast of British Columbia, CSAP WP 2013GRF01.
Authors:	Chris Grandin and Robyn Forrest
Reviewers	Joanne Morgan, Fisheries and Oceans Canada, St. John's NL, Carey McGilliard, Alaska Fisheries Science Center, Seattle WA,
Chairperson:	Rowan Haigh (Groundfish, MEAD, PBS, DFO)
Rapporteur:	Melissa Nottingham (Groundfish, MEAD, PBS, DFO)
Presenter:	Chris Grandin (Groundfish, MEAD, PBS, DFO)
Meeting:	May 12-13, 2015, Seminar Room, Pacific Biological Station, Nanaimo BC

PRESENTATION OF THE WORKING PAPER

Throughout this document (excluding Appendices), sentences in regular non-italic font either contribute to editorial flow or reflect the comments and questions from reviewers and participants. Italicised sentences reflect the comments and responses by the authors. Sentences in bold font and preceded by a ★ symbol flag recommendations that arose from the RPR meeting.

During the main presentation, participants occasionally asked for or provided clarification.

- One participant noted that the working paper stated that pre-1996 commercial data came from "voluntary" logbooks, but logbook recording had been mandatory since 1987. Individual vessel quotas (IVQs) were introduced in 1997, and total allowable catch (TAC) for this species was first applied in 2006. The primary reason for not reporting prior to 1996 was not the lack of TAC, but the lack of observer coverage. The authors were unaware that full reporting had been mandatory since 1987. Limited data also stem from a logbook that only offered a small (0.5 inch square) space to record releases of all species.
- A participant asked why the final-year recruitment was so much higher than that for previous years.

- ◊ *Authors responded that it was due to the high recent survey estimates combined with increased recent levels of catch and that real-world stock recruitment would not be that high.*

- Participants requested clarification on TAC and status quo.
  - ◊ *Authors clarified that status quo refers to the TAC (15,000 t/y) set by groundfish management for the last few years; in the working paper's decision table this only includes females. The female amount needs to be multiplied by a factor to calculate TAC (male and female). The authors requested that participants could help on a method for doing this.*
  - ♦ A participant added that there is carry-over of TAC (~30%) that varies annually.

The participants were asked for questions on the presentation, including points for clarification.

- One participant asked: why use a catch-at-age model when there were too few ages sampled?
  - ◊ *The authors agreed about the ages, but originally didn't expect a problem because 6000 otoliths had been processed. However, these age data spread out over all the surveys and the commercial fleet provided insufficient coverage. The problem was discovered after the model was applied; the authors were unaware of the uncertainty until too far into the assessment work.*
  - ♦ A participant added that age data contain information that biomass data cannot (e.g., selectivity relative to maturity). The age data suggests that this species is selected when larger and older, which might be missed by a biomass-only model.
- One participant asked why the model does not include ageing error.
  - ◊ *According to the Sclerochronology Laboratory (SCL), percent age agreement between agers for Arrowtooth Flounder are 80% plus or minus 2 years of age. It must be noted that percent agreement should not be used as a measure of age bias or ageing error. The SCL uses percent agreement as a measure for structure difficulty only. For age estimates below 80% +/-2 the entire sample is re-aged by another technician. Neither the otolith nor the break and burn technique have been validated for Arrowtooth Flounder. Pacific Halibut otoliths via the break and bake technique have been previously bomb-carbon validated and are used as a proxy for Arrowtooth Flounder age estimates.*
  - ♦ The participant stated that while not formally included in models, other researchers do look at ageing error. The authors asked if this should be added to the recommendations, but the participant indicated that it was not a big problem. Another participant added that ageing error tends to smooth out ages and is more about precision not bias. This smoothing tends to squash initial recruitment. If error isn't very large, incorporating ageing error does not add anything useful to the model.
- Participants asked for clarification on why pre-1996 catch data were not included.
  - ◊ *The decision to exclude these data was made by the technical working group because catch data were incomplete. Arrowtooth Flounder was released at sea frequently and in high proportions without being recorded, and so a reconstruction would need to estimate high levels of releases based on relatively small amounts of reported landings. A reconstruction takes time and resources, which were not available for this assessment.*
  - ♦ A participant disagreed that an industry advisor would recommend against using age data. The participant recalled that the technical working group thought there was

inconsistency, but did not believe there was a recommendation to dismiss the data. The same participant remained confident that there are a number of industry people who could guide a pre-1996 catch reconstruction, but it would be hard work.

- ♦ Another participant noted that the pre-1996 reconstruction is more difficult for this species than for other species. Arrowtooth Flounder is abundant, but it deteriorates quickly when caught and becomes useless very soon after landing. For this reason, industry would make a tow at the end of a trip and take the fish directly to the processing plant. Arrowtooth taken earlier in the trip were not kept due to low economic value and the opportunity cost of space in the hold that was better used for more valuable species. These factors meant a variable at-sea release rate (0-90 %), which leads to huge uncertainty in estimating pre-1996 catch.

- ◊ *It was deemed better to start with reliable data for an initial population model. Catch reconstruction would be unsatisfactory at this point in time.*

- ♦ Another participant added that these issues provide many research possibilities.

★ **The authors agreed to consult industry and implement a catch reconstruction for Arrowtooth Flounder in future stock assessments.**

- A participant noted that there was a very disconnected group of biomass indices to fit rather than using a classic continuous series. The model might have trouble resolving these biomass indices. There is also the issue of starting from a non-equilibrium position, which is difficult for a surplus model when there is potentially lots of removal prior to 1996. Ages would need to be acquired to set up the initial population structure. The participant thought the decisions made in this assessment were reasonable, and would make the same ones in hindsight.

- A participant asked if there is information that shows variation of males and females when comparing survey and fisheries. The authors stated that the document contains this information (Table C.1).

WRITTEN REVIEWS

REVIEWER 1 – JOANNE MORGAN (DFO)

The written review by Joanne Morgan appears in Appendix B. Here, responses by authors to the reviewer are presented. Joanne attended the meeting via webinar.

- The reviewer stated that the data may not be sufficient for an-age structured model.

- ◊ *The authors echoed the reviewer's comment and suggested that one of the purposes of the RPR meeting was to decide if an age-structured model and its advice are appropriate.*

- ◊ *The Groundfish section experiences limitations on resources and personnel as well ageing request allocations. The authors would like a discussion on how to move forward in a resource-limited environment. For example, is pursuing complex age models sustainable? The authors agreed that the limited age composition data do not appear to track cohorts. However, the Hecate Strait survey sample size was larger but the data still did not show cohort patterns. Is this due to sampling or does it truly reflect population structure? We don't know at this stage.*

- The reviewer wanted to see more presentation in the working paper of priors and posteriors for the various sensitivity runs. This would enhance interpretation.

- ◊ *The authors mentioned that the trawl sensitivity run saw u_{MSY} decrease a little. The authors agreed that the priors on steepness strongly contributed to reference points. The biggest contribution comes from selectivity, and the authors demonstrated this by forcing selectivity to the left (to match maturity). Another problem is the lack of total catch data prior to 1996. If there were estimated historical catches, the authors would have a much better handle on productivity.*

There was a call for questions from the participants for the reviewer, but silence prevailed.

REVIEWER 2 – CAREY MCGILLIARD (NOAA)

The written review by Carey McGilliard appears in Appendix B. Here, responses by authors to the reviewer are presented. Carey attended the meeting via webinar.

- The reviewer commented that there are lots of male Arrowtooth Flounder in the western Aleutian Islands (Alaska); therefore, maybe males do not experience higher mortality. The reviewer asked about the validity of the assumed sex ratio.
  - ◊ *The authors responded that proportions ♀:♂ varied between surveys and think that DFO should explore proportions of females by area as a research recommendation.*
  - ◊ *The topic is understudied and there is no hypothesis for sex distribution. The authors agreed that plotting the information onto maps was a good idea. There are drawbacks to a female-only model. The authors followed the lead of a few other rockfish assessments (e.g, Rock Sole, Lepidopsetta bilineatus, DFO 2014). Due to sexual dimorphism the authors avoided some problems but needed to discard age data for males. The authors agreed that uncertainty in the proportion of females should be explored using sensitivity analysis.*
  - ◊ *Authors requested comment from industry on sex distribution.*
  - ♦ One industry participant was unsure about general distribution, but stated that in winter the fish aggregate for spawning. The participant also mentioned that males do not occur in large aggregations during the fall and spring-summer fisheries.
  - ♦ Another participant asked if the authors checked seasonality of distribution and the authors agreed it was a good idea. A participant added that Arrowtooth Flounder occur at different depths seasonally but aggregations often comprise a single sex.
  - ♦ There was some discussion about reasons to expect that proportion of females might vary in samples:
    - (i) surveys are generally conducted in summer vs. a fishery that has switched to winter in the past 4-5 years;
    - (ii) surveys don't necessarily cover the primary fishing locations;
    - (iii) codend sizes likely vary, and
    - (iv) these fish migrate.
  - ♦ A participant pointed out that the proportion of females does not appear to be a reflection of season because the fishery and surveys see the same ratio (Table C.1). A suggestion was made to conduct collaborative research on how/why this species forms spatial aggregations that become predominantly one sex or the other.
  - ♦ A question was posed to the reviewer – does aggregation in Alaska occur for spawning? The reviewer was unsure. Another participant noted that it is common with flatfish to see lots of males or lots of females.

- The authors were asked if historical ratios are good indicators of the future should there be a change in habitat area or fish size.
  - ◊ *The authors responded that length data are not being used but could inform selectivity, i.e., conditional age-at-length. The Hecate Strait Multi-Species Assemblage Survey is the longest running survey but is only used as an index. Length data from this survey could help to inform selectivity. If the model starts before1996, this survey could be valuable; however, these data were not used because the model started in 1996. Length could be used to construct length-age keys; a conditional age-at-length model would be more time consuming.*
  - ♦ A participant responded to the reviewer's comments on length-at-age stating that the the way length data are used in age-structured models often provides a poor approximation to age, which introduces bias. To use length, one needs a length-structured model that ignores age because fish grow at different rates. However, this kind of model only approximates the way fish grow. In the end, a model that incorporates both length and age is needed. Could the authors use a delay difference model such as that used by the IPHC for Pacific Halibut?
- The reviewer noted that freezer trawl ages were poorly sampled. Could pooling age data make selectivity less precise? How does selectivity differ among fleets?
  - ◊ *The authors responded that there were no freezer-trawl ages in the database before the assessment. But if there were, splitting between fleets and then by sex would reduce the number of ages available to derive model input. Also, the number of samples was already low. However, available length data for freezer-trawl and shore-side fleets were compared and the results appeared to be similar between the fleets.*
- Growth rate was estimated using data back to the 1950s. If excluded, would the growth curve be similar? Has growth changed over time?
  - ◊ *The authors replied that for this assessment they only looked at data from 1996 on due to decisions made early on, and agreed that it is a good idea to check data from prior years and look at changes between time periods.*
- There were questions about the variability in the WCVI survey and a low 2012 index point.
  - ◊ *Authors showed maps of the survey locations and asked for comments from survey participants on why the 2012 index might be low. Participants noted that in 2012 there were other species that also had low index points. There may be a need to check simultaneous movements in survey indices among species and perhaps use this as an indicator of environmental covariance. Despite the occasional anomalies, the index plots showed quite clearly that Arrowtooth was well-captured.*
  - ◊ *Authors asked the reviewer if survey catchability and temperature are incorporated into Alaska assessments, and suggested that in the future DFO might look at species distribution and its correlation with temperature.*

★ **There appeared to be general agreement to conduct spatial analyses of Arrowtooth Flounder distribution to explore reasons for sex-specific aggregations, including environmental covariates.**

FEEDBACK FROM THE SCLEROCHRONOLOGY LAB

Authors introduced the supervisor of the Sclerochronology Laboratory (SCL), Stephen Wischniowski, to demonstrate how age data are used and to summarize the resource pressure in terms of groundfish requests. The authors requested these comments to contextualize the reliance on age data and to ask what is reasonable for us to expect from the ageing lab.

- The sclerochronology participant described challenges to his unit – the number of clients, preparation time, increased processing time for long-lived species, and prioritization. The demands on the SCL are great; however, utilizing an age request forecasting model to determine what species require age estimates can greatly increase the numbers of structures aged. Using Arrowtooth Flounder as an example, with a 5-year advanced notice the SCL could have easily produced 15,000 ages. It's unrealistic to expect a high number of age estimates with only a one or two year notice.
  - ◊ *The authors noted one reviewer's point on the lack of age information in the commercial fishery (freezer trawl) and questioned whether catch-at-age structure is appropriate, given the lack of data before 1996.*

GENERAL DISCUSSION

One participant summarized the problems of the model adopted for this assessment – short index series, ageing data without cohort signals, absence of pre-1996 catch data, and the absence of structure that the model doesn't contemplate (migration, seasons, sex distribution, etc.).

MODEL SUITABILITY

Should a catch-at-age model be used at this time? What are the advantages of using a catch-at-age model?

- ○ *The authors referred to earlier comments (during the meeting) on the advantages of using a catch-at-age model over a simpler model: it can estimate fishing selectivity on a population and can integrate diverse sets of data into one framework. There are significant uncertainties: missing catch data prior to 1996, an unknown scale of productivity, and missing cohort structure. The previous assessment (Fargo and Starr 2001) provided an analysis of different data streams and concluded that harvest rates at that time were likely sustainable given the stable catches and non-truncated nature of the age composition. In this assessment, the authors conducted many sensitivity analyses but could not force the population below upper stock reference points. All survey indices of abundance appeared to be stable (without trend). Under the assumption of a single fleet, catch since 1996 comes from a stable population because selectivity was fixed. A significant caveat is that selectivity might have changed with the introduction of freezer trawl boats. Another significant caveat is that productivity scale was not known because catch data prior to 1996 were not available.*
  - ♦ Note: Many of these problems exist for any model structure, not just an age-structured model.
- ○ *The authors also highlighted the ability to make statements on stock status with respect to various reference points. They could have made qualitative statements like those provided 15 years ago, but were requested to make definitive stock status statements. If the model is not accepted then such statements cannot be made. The authors were confident that the Arrowtooth Flounder stock is above all traditional reference points. The decision tables provide probabilities on this confidence.*
- A participant added that the pre-1996 data problems do not go away with other models. The difference between maturity and selectivity is highly influential. The age model allows us to understand this.

- ◊ *The authors looked at biological data, and the selectivity curves used in 2001 are similar to the current ones. There was not a big shift in productivity.*
- ♦ One reviewer added that she agreed with most comments. If one can use an age-structure model, there is more opportunity to learn about the population. However, if the data aren't good enough, then the primary researchers (authors) need to make the decision on which model to use. The reviewer also cautioned that while moving forward, we should be mindful that the Arrowtooth Flounder fishery has been changing (to freezer trawlers) and catch has increased substantially. The full impact of these changes has likely not been felt yet. She recommended using caution and taking uncertainties into account.

AGE AND COHORTS

There was discussion about age proportions and cohort patterns.

- Participants considered other regional areas and asked authors if they had looked at assessments from Alaska or the Pacific USA. If so, did the assessments show strong cohorts followed by weak ones? In this assessment, recruits seem to be consistent, but year classes don't seem to be strong (Fig. 3).
  - ◊ *The authors had not studied the US assessments.*
- The US reviewer mentioned that the Gulf of Alaska assessment (Spies and Turnock 2013, Fig 14, p.584) shows a clear recruitment pattern, cycling since 1956. The reviewer did not know if the model incorporates autocorrelation.
  - ◊ *The authors indicated that there is a slight cohort pattern in this assessment, less obvious than the US pattern. The authors did not model autocorrelation.*
  - ♦ The participant suggested that the observed cohort patterns might be confounded by factors other than biology.
  - ♦ A participant noted the appearance of 1-year old fish in the WCVI Synoptic survey (Fig. 4).
- The chair posed the question: if ageing more otoliths still results in no cohort patterns, is it worth using the age data?
  - ♦ One participant reminded the group about resource allocation: if it isn't useful, the ageing lab can work on other species.
  - ♦ Another participant recommended simulating more age data to see if cohort patterns can be clarified, then decide whether or not to age more Arrowtooth Flounder.

CATCH AND BIOMASS

There was a discussion about catch and biomass.

- Looking at catch in Fig 2, a participant pointed out that he did not see order-of-magnitude changes compared to the past. Increasing catches are not the biggest problem.
  - ◊ *The authors agreed and said it was an example of how the model can be useful. They also emphasized the need to exercise caution when using the decision tables.*
- One reviewer stated that these fish are not selected until age 9 and the increase in catch began less than 10 years ago, so we may not see the results yet.

- ◊ *The authors agreed. One benefit may be to see what will happen over time as more years of catch data are added. Estimates of vulnerable and spawning biomass have both dipped a bit in the last few years.*

- One participant pointed out that the BC Arrowtooth Flounder fishery had declined notably (after 2005) based on marketability, and that industry was having a difficult time recapturing market share. The crash, however, was based on poor product, not abundance.

  - ◊ *The authors agreed that if catch increases, it will be due to a better product that is more marketable.*

- **Note from the editor**: Figure 20 presents MCMC trajectories of spawning and vulnerable biomass. The 95% envelopes appear to be largely separate, which seemed to foster the mistaken belief in the RPR meeting that the spawning stock would be entirely invulnerable. While a large proportion of the spawning stock is invulnerable due to a selectivity curve far to the right of the maturity curve (see Fig. 40), there will be a proportion of the spawning stock that is vulnerable. Figure 20 simply portrays two populations of different magnitude.

  - ♦ One participant asked the group if anyone could identify other flatfish stocks that were "completely" invulnerable (re: Fig 20), and what, if anything, had resulted from those assessments?
  - ♦ One reviewer pointed out similar results in the Gulf of Alaska (GoA) Rex Sole (*Errex zachirus*) stock. In that assessment, the model was used to set quotas based on the biomass estimate but not to set fishing mortality. (Some participants were not sure about this distinction.) The reviewer continued that the US assessment team was not making policy recommendations, just reporting on the stock status. The review panels advised trusting the biomass estimates but not the fishing mortality.
  - ◊ *The authors asked if the GoA Rex Sole panel trusted selectivity.*
  - ♦ The reviewer believed that they did, and like Arrowtooth Flounder the maturity curve lay well to the left of the selectivity curve.
  - ◊ *The authors discussed precautionary fishing practices: sustainability dictates that selectivity should fall to the right of maturity, but now that we have this situation exaggerated, the advice is deemed not very precautionary. Suddenly, we are afraid to accept it because the risk of being wrong is high.*
  - ♦ A participant responded that it is easier if the estimate of selectivity is solid, but we don't seem to believe this.

- One participant pointed out that the Alaska assessment on Arrowtooth Flounder had similarities to this one, with approximately 70% of survey and commercial catch being female. He also noted that the information from modern recorded releases at sea may be valuable; releases are similar within all depth strata. There is information on catch composition from trips of 3-5 days duration. Often there is no Arrowtooth market, but if a market appears, retention of this species will increase, but only on the last day of each fishing trip.

- Participants discussed at-sea releases as indices, which can show a trend over time. Arrowtooth Flounder is ubiquitous and caught in all strata. The fish could be counted before releasing them.

  - ◊ *The authors will look at this in the future.*

- ♦ A participant asked if release data included hook and line releases.
- ◊ *The authors responded that H&L releases of Arrowtooth are less than 0.5%.*
- ♦ Another participant mentioned using IPHC surveys in future.
- ◊ *The authors looked at hook and line data, but Arrowtooth catch was not significant. It is one of the biggest releases in the Pacific Halibut (Hippoglossus stenolepis) fishery but constitutes a very small percentage of total catch.*
- ♦ One participant asked if the amount released was simply an amount necessary to operate the fishery. The participant responded that he would need to look at the target by depth strata. When there is a market and a targeted fishery, there is no releasing. The percentage of males caught may be higher when fishing other species because the vessel master is not fishing schools. Males in releases may be higher as a consequence.

ANALYSES AND REFERENCE POINTS

The authors went through various sensitivity analyses to illustrate the impact on reference points. Some off the key analyses included moving selectivity left towards maturity so that more of the population was vulnerable. The model, however, simply scaled estimated biomass and reference points down (Fig. 41).

- A participant asked for clarification in Fig. 19. Are the reference points calculated from the reference curve or the reference model? The participant also suggested using reference points in the figures that are consistent with different sensitivity runs.
  - ◊ *The authors responded that the reference points ($0.4B_{MSY}$, $0.8B_{MSY}$, $0.2B_0$, $0.4B_0$) are medians of the reference case. Each of the 1000 MCMC sample vectors B_t has its own B_0 and B_{MSY}, and so the probability of $B_{2015}<B_0$, say, is the proportion of times that 1000 sampled B_{2015} fall below their respective B_0.*
  - ♦ For this reason, a participant expressed the need to plot biomass B_t/B_0 (or B_t/B_{MSY}) so that the reference points can indeed be represented by a single line.
- ○ *For Scenario 12 (selectivity ogive = maturity ogive), the authors plotted spawning and vulnerable biomass trajectories as in Fig. 20 (not presented in the working paper) and found that the vulnerable and spawning biomass envelopes overlap more than for the reference case. (Editor: as noted earlier, this observation says nothing about the proportion of the spawning biomass that was vulnerable.)*
- ○ *The authors posed a question to the managers: is a one-year projection sufficient considering that the Terms of Reference (App. A) requested a long-term projection?*
  - ♦ Participants wondered why only one year was presented when there will not likely be another assessment for at least 5 years.
  - ◊ *The authors noted that there is often a mismatch of needs and projection ability. They felt uncomfortable projecting further than one year. The assessment exercise highlighted that much is unknown about the stock, including why there is no cohort signal.*
- A participant noted that in Fig. 15, the posterior of $\ln(R_{\mathrm{init}})$ was pushing up against the lower bound of -2.
  - ◊ *The authors acknowledged this and agreed that it might have a significant impact on model results. Their best guess was a scale down of the population and lower reference points. The authors ran a quick MPD which showed no obvious change; however, an*

MCMC run overnight indicated that the posterior indeed shifted lower than the initial lower constraint.

- ♦ The participants agreed that this change should be incorporated into a new reference case for the model.

★ **The authors will re-run the model MCMC with wider bounds on the R_{init} prior and use the results as a new reference case for Arrowtooth Flounder.**

ACCEPTING THE MODEL

The discussion eventually posed the question: is this model acceptable for advice?

- One participant stated that the available data are not robust enough. That is, despite the best efforts of the assessment team, can these data be used to provide scientific advice? The Terms of Reference outlined an alternative path.
  - ♦ Another participant responded by asking how does one decide whether the data are good enough? This participant stated that the job of science is to offer models and outline the limitations to the managers.
  - ♦ Managers can make decisions based on the understanding that they have the results and the caveats.
  - ♦ The participant continued asking how to evaluate effectiveness, what is the benefit.

○ *The authors asked the US reviewer if Alaska assessments have criteria for accepting or rejecting advice?*

  - ♦ The reviewer indicated that yes, but the criteria depended on which tier a stock falls into, and provided a few examples of species that fall into each tier. The Alaska tier system is predefined, so decisions are easier. There is no need to accept or reject models, only determine their reliability, which leaves fewer things to decide.

- A participant noted that the 2005 catch was the highest on record (Fig. 2). The decision table (Table 14) offers constant catch policies up to 30,000 t, which exceed the highest catch in 2005.
  - ◊ *Authors pointed to vulnerable biomass (Fig. 20), which sits below 100,000 t. A projection further into the future will not yield much useful information. The authors expressed no confidence in projecting further out than one year. The longer time introduces more uncertainty; however, recruitment here seems average so perhaps a longer projection might be done.*
  - ♦ A participant noted that data-rich fisheries like Pacific Hake (*Merluccius productus*) and Pacific Halibut do not project far forward (Hake 1 year, Halibut 3 years), but assessments for these species are performed annually.
- A participant asked if there were potentially limiting interactions with other species.
  - ♦ The responses yielded no other issues, but one participant expected an increase in the number of other species as bycatch.
  - ♦ Another participant disagreed because the fishery is chiefly a combination fishery. Vessels pursuing fresh-caught fish target multiple species because the higher value of species other than Arrowtooth Flounder increases the daily value. Freezer boats, on the other hand, would choose lower bycatch areas and process more target species.

- One participant asked if there had been thought given to reconstructing commercial CPUE.

  ◊ *The authors responded that they did construct a CPUE series but the technical working group advised against using it because the fishery is conducted in a way that does not reflect directed fishing effort. Many other factors, especially fisher behaviour, confound the CPUE signal.*

★ **The participant requested that the final paper include a discussion on why CPUE was not used in this assessment.**

There were significant concerns by participants about the uncertainties in the assessment, primarily

(i) missing catch data prior to 1996,

(ii) potentially poor estimation of selectivity, and

(iii) two different commercial fleets represented as a single fleet.

- Some discussion revolved around the possible rejection of the assessment and its consequences. Significant resources in the form of monitoring and science had been devoted to this (and any) assessment.

  ♦ Industry participants were highly concerned that another groundfish stock would not make it through the assessment process to provide guidance on sustainable quotas (e.g., Redbanded Rockfish, *Sebastes babcocki*, DFO 2015a). They need third-party validation for MSC (Marine Stewardship Council) certification, which labels fisheries sustainable. The participants worry that DFO always seems to be reaching for higher standards than the data can support.

- A participant raised the issue of conservation. If the decision tables provide no constraints on the fishery (because the estimated B is always above the reference points), then should we be considering alternative triggers?

  ♦ Participants and authors were reluctant to adopt *ad hoc* rules (e.g., Pacific Cod, *Gadus macrocephalus*, 2014).

  ♦ One participant suggested that the model and decision tables be accepted but that the SAR (Science Advisory Report) include auxiliary information such as highlighting the lack of declining trends in surveys.

  ♦ The CSAP representative asked: what should occur between now and the next assessment (in 2020)?

  ◊ *The authors responded that in the short-term (one year), the fishery is not in a crisis situation. If the model is believable, then a large portion of the spawning biomass is invulnerable to the fishery. However, it is important to reconstruct pre-1996 catch, validate the estimate of selectivity, and explore whether more age data would reduce uncertainty.*

  ♦ In the mean time, CSAP still needs checks and balances that stand until the next assessment, some way of determining whether a full stock assessment should occur sooner than 2020.

  ♦ Science participants noted that given all the species on the assessment list, the trigger would need to be very powerful. This also affects researchers' work plans.

- Participants expressed the need to stop reinventing the process for data-limited species and establish rules to deal with assessments. Since scheduling began (5 years ago?), more species have come up that have never been assessed. This is likely to occur repeatedly. Each assessment adds to a large list of recommendations, but science still doesn't have better data, and resources are still strapped.
- A participant asked if the authors were looking to the RPR meeting to decide how to take the female-based model results and apply them to decision tables for total catch.
  - ◊ *The authors responded yes, based on what has been done in other assessments using female-based models.*
  - ♦ After a discussion on commercial catch where the last 4 years changed to the dominance of freezer trawlers, participants requested two more columns in decision tables adjusting for sex ratios calculated from the last 4 years and another adjusted for sex ratios for the entire time series (see Table C.1).

★ **The authors agreed to add two additional columns to the decision tables showing what total catch would be given the proportion of females in commercial samples – one using the average proportion females for the last 4 years and the other using the average proportion females for the entire time series.**

- One participant pointed out that the 2005 fishery was not a test fishery, but rather a response to market and scaling up to target Arrowtooth. The market picked up significantly until quality decreased and the market collapsed.

★ **The authors agreed to remove references to Port Simpson and "test fishery" from the final paper.**

Despite all the concerns expressed, the participants agreed that the working paper provides useful information, reference points, and appropriate decision tables for management. Additionally, the research recommendations provide guidance for the future.

★ **There was consensus to accept the working paper with caveats.**

CONCLUSIONS

CONSENSUS ON PAPER ACCEPTABILITY

There was consensus to accept the working paper with caveats.

Caveats: Major uncertainties include

(i) missing at-sea release data prior to 1996,

(ii) potentially poor estimation of selectivity,

(iii) two different commercial fleets represented as a single fleet, and

(iv) lack of cohort signal in the age data.

INSTRUCTIONS TO AUTHORS

- Re-run the model MCMC with wider bounds on the prior for $\ln(\bar{R}_{\text{init}})$ [-5,6] and use the results as a new reference case for Arrowtooth Flounder.
- Add two additional columns to the decision tables showing what total catch would be given the proportion of females in commercial samples – one using the average proportion

females for the last 4 years and the other using the average proportion females for the entire time series.

- Include a discussion on why CPUE was not used in this assessment.
- Remove references to Port Simpson and "test fishery" from the final paper.
- Add a 50 kt catch policy for projections. Authors can determine catch-policy increments (if any) between 30 kt and 50 kt.
- Run MCMCs a bit longer to deal with convergence statistics.
- Add B_{2015}/B_0, V_{2015}, V_0, V_{2015}/V_0 in Table 6, where V = vulnerable biomass. Make sure all model symbols are defined in Table E.1 (e.g., B_0, $B_{t,}$ etc.). Ensure that current biomass B_{2015} and fishing mortality F_{2014} have correctly subscripted years throughout the paper.
- Include a reference to the IPHC which has a long time series of non-directed Arrowtooth catch.
- Check whether catch entries are duplicated for survey trips in in the databases FOS and GFBio.
- Address reviewers' comments and edits fully, wherever appropriate.

RECOMMENDATIONS

RESEARCH

- Age Arrowtooth Flounder otoliths that were sampled from commercial freezer trawlers.
- Consult industry and build a catch reconstruction for Arrowtooth Flounder.
    - The reconstruction should be developed by a technical working group, subject to peer review, and include all flatfish species (if possible).
    - The project would need to go through CSAP for proper resource allocation.
- Research and develop more robust selectivity ogives for the commercial fleets (regular and freezer trawls) as well as for surveys.
- Monitor survey trends for changes in Arrowtooth Flounder abundance.
- Conduct spatial analyses of Arrowtooth Flounder distribution to explore reasons for sex-specific aggregations, including environmental covariates.

SCIENCE ADVISORY REPORT (SAR)

- Request a photo of a BC Arrowtooth Flounder from the Surveys group.
- Include a map of survey locations.
- Include all columns and catch intervals (up to 30 kt) in decision tables.
- Specify whether years are fishing years or calendar years.
- Include depletion B_{2015}/B_0 in Table 1.
- In the preamble, discuss why the model is female-only and how this can be implemented in decision tables. Include the words "female only" in each caption.
- Adopt terminology used by management:

- change the term “discards” to “at-sea releases”;
- add label “at-sea release” to tables and figures where appropriate;
- change the term “bycatch” to “non-directed catch” (as opposed to directed).

- Make changes to Fig. 3 –
  - female spawning biomass should be termed “relative”,
  - change left axis label to “Catch (x 10,000 t), and
  - specify “female” for exploitation rate.
- Change Fig 4 (horizontal bar plot of current status) to include all B_{MSY} and B_0 reference points, similar to that presented in the SAR for Lingcod (*Ophiodon elongates*, DFO 2015*b*).
- Confirm age at 50% maturity (~7 y) in the last paragraph of the Uncertainties section. Maybe include some results from the sensitivity analysis that uses the selectivity = maturity ogive.
- Use total catch biomass instead of female biomass for increments in the decision table.
- Change instances of “0.000” to “0”.
- Clarify the catch data assumptions (in the Introduction):
  - mortality only results from commercial bottom trawl, not mid-water trawl;
  - released Arrowtooth Flounder experiences 100% mortality;
  - outline other sources of mortality;
  - state that catch includes landings and releases.

ACKNOWLEDGEMENTS

Melissa Nottingham was the rapporteur for this RPR. Her organized notes helped tremendously in the formation of the proceedings document. The authors provided feedback and clarification where needed. The reviewers, Joanne Morgan and Carey McGilliard, contributed greatly to the RPR process. Participants, thank-you all.

REFERENCES

DFO. 2014. Stock Assessment and Harvest Advice for Rock Sole (*Lepidopsetta* spp.) in British Columbia. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2014/039.

DFO. 2015*a*. Proceedings of the Pacific Regional Peer Review on Redbanded Rockfish (*Sebastes babcocki*) Stock Assessment for the Pacific Coast of Canada in 2014; December 9, 2014. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2015/032.

DFO. 2015*b*. Stock assessment for Lingcod (*Ophiodon elongatus*) for the Strait of Georgia, British Columbia in 2014. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/014.

Fargo, J and Starr, P.J. 2001. Turbot stock assessment for 2001 and recommendations for management in 2002. Can. Stock Assess. Sec. Res. Doc. 2001/150. 70 p.

Spies, I. and Turnock, B.J. 2013. Section 7. Assessment of the arrowtooth flounder stock in the Gulf of Alaska, p.541-610. *In* NPFMC Gulf of Alaska Stock Assessment and Fishery Evaluation (SAFE) Reports, North Pacific Fishery Management Council, P. O. Box 103136, Anchorage, AK 99510.

APPENDIX A: TERMS OF REFERENCE

Coastwide Assessment of Arrowtooth Flounder (*Atheresthes stomias*) in 2014 and recommendations for 2015

Regional Peer Review – Pacific Region

May 12-13, 2015 Nanaimo, British Columbia
Chairperson: Rowan Haigh

Context

Arrowtooth Flounder (*Atheresthes stomias*) is a significant species within the groundfish trawl fishery. Since 2006 the fishery has been subject to a coastwide trawl Total Allowable Catch of 15 000 mt. Arrowtooth Flounder are also a minor component of the groundfish hook and line fisheries and where vessels are subject to trip limits. The combined hook and line fisheries catch in 2013 was estimated to be 73 mt. Arrowtooth Flounder is the second most frequently discarded species in the Pacific Halibut (*Hippoglossus stenolepis*) hook and line fishery (est. 12 mt) and consequently was identified as a component of the non-target catch in the Marine Stewardship Council certification process for the Pacific Halibut fishery. The last assessment of Arrowtooth Flounder was published in 2001 (Fargo & Starr 2001) and provided analyses of biological, survey, commercial catch and catch-per-unit effort data. A formal, model-based assessment has never been done for this stock.

Updated stock status and harvest advice would assist in determining whether current harvest levels are sustainable and compliant with the Sustainable Fisheries Framework (SFF). One policy under the SFF, *A Fishery Decision-making Framework Incorporating the Precautionary Approach* (PA Framework) (DFO 2009), outlines the methods for applying the Precautionary Approach (PA) in the management of Canadian fisheries. Application of the PA usually implies estimation of fishery reference points and evaluation of current stock status relative to those reference points. Estimation of these quantities requires development of quantitative models, conditioned on available fishery, survey and biological data.

Landings and discard data for the Arrowtooth Flounder trawl fishery are unreliable prior to 1996. Since 1996, 100% observer coverage in the commercial groundfish trawl fishery has provided reliable reporting of catch and discards. Prior to 1996, however, landings of Arrowtooth Flounder were characterized by high proportions of discards due to low market demand, low value and the known rapid deterioration of product due to the presence of the same parasites prevalent in Pacific Hake. In many cases, the entire Arrowtooth Flounder catch in a trawl would be discarded, and in many cases Arrowtooth Flounder discards were not properly recorded in fishing logbooks. As a result of this inconsistent record keeping, trends in catch statistics prior to 1996 are biased and it will be impossible to reconstruct historical catches prior to 1996. A model-based approach to the assessment of stock status, conditioned on estimates of historical catches before 1996, is therefore unlikely to be successful and has been dropped as a possible approach to the stock assessment of this species. An alternative approach, which initializes the model from a non-equilibrium starting point and which is based on the reliable catch and discard time series beginning in 1996, may be possible. However, such models are often difficult to resolve, due to the need to estimate additional parameters that may not be supported by the available data. Consequentially, attempts at a model-based stock assessment may fail.

The Fisheries Management Branch of Fisheries and Oceans Canada requested advice on (a) the status of Arrowtooth Flounder coastwide and (b) harvests consistent with DFO policy.

Objectives

Guided by the DFO Sustainable Fisheries Framework, particularly the *Fishery Decision-making Framework Incorporating the Precautionary Approach* (DFO 2009), meeting participants will review the following working paper to provide the basis for discussion and advice on the specific objectives outlined below.

Grandin, C., Forrest, R. and Starr, R. Assessment of Arrowtooth Flounder (*Atheresthes stomias*) *in 2014 and recommendations for management in 2015. CSAP Working Paper 2013GRF001*

- Investigate the feasibility of conducting a model-based stock assessment.
- If model-based assessment **is** possible:
  - Estimate reference points consistent with the DFO Precautionary Approach. Include the biological considerations and rationale used to make such a determination.
  - Evaluate the current status of Arrowtooth Flounder relative to the estimated reference points.
  - Evaluate the consequences of varying constant catches on future population status, providing decision tables and figures of long-term projected biomass.
- If model-based assessment **is not** possible:
  - Provide reasons why it is not possible.
  - Present analyses of biological and fishery data.
  - Investigate and provide alternative advice to managers, based on available data.
  - Recommend research options to move toward a model-based assessment.
- Recommend an interval for re-assessment of Arrowtooth Flounder.

Expected Publications

- CSAS Science Advisory Report
- CSAS Research Document
- CSAS Proceedings

Participation

- DFO (Science, Fisheries Management, Oceans, Habitat)
- Aboriginal communities
- Province of British Columbia
- External reviewers
- Industry
- Non-governmental organizations and other scientists and stakeholders.

References

DFO. 2009. A fishery decision-making framework incorporating the Precautionary Approach,

Fargo, J and Starr, P.J. 2001. Turbot stock assessment for 2001 and recommendations for management in 2002. Can. Stock Assess. Sec. Res. Doc. 2001/150. 70 p.

APPENDIX B: WORKING PAPER REVIEWS

REVIEW – JOANNE MORGAN (DFO)

Date:	April 29 2015
Reviewer:	Joanne Morgan, Fisheries and Oceans
CSAS Working Paper:	2013GRF001
Working Paper Title:	Arrowtooth Flounder (*Atheresthes stomias*) stock assessment for the west coast of British Columbia

The authors have done a good job in conducting this assessment and in its presentation. The work is extensive and given the data has likely not been easy. They should be commended.

The purpose of the working paper is clearly stated. The methods and results are generally sufficiently explained.

In this review I will discuss first the main topics and then will give some of the more minor details. I have italicized by major recommendations/questions, but there are others in the text below as well.

The Data

This is a highly uncertain assessment and the main issue is the data. There is little or no information in the age composition data. There is no evidence of any ability to track cohorts. This could be because there has been no variation in cohort strength but it could also be related to a number of other things including sample size. If the numbers on the figures actually represent the number of individuals aged as the caption says then this is not sufficient. In 1996 there are 35 otoliths spread across 10 ages, in 2009 29 otoliths across 12 ages! In 2013 there are 75 otoliths from almost 14000 t of catch. This is not adequate and would certainly impact the ability to follow cohorts.

Contrast Figure 3 from the current assessment with the clear cohorts visible in the assessment of Martell et al which developed the method on which the current assessment is based.

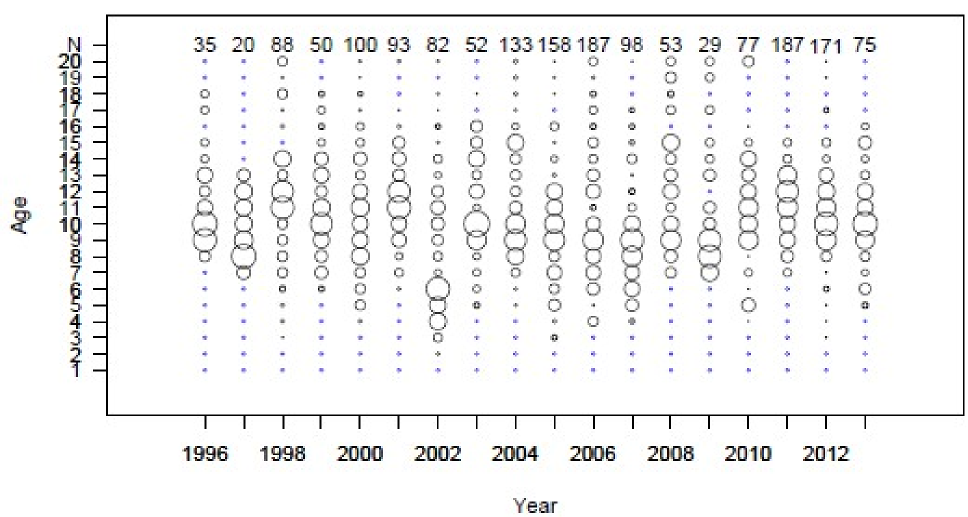


Figure 3. Age composition data for the Reference Case for the commercial trawl fishery. Blue points represent zero samples. N values represent the number of specimens aged in each year.

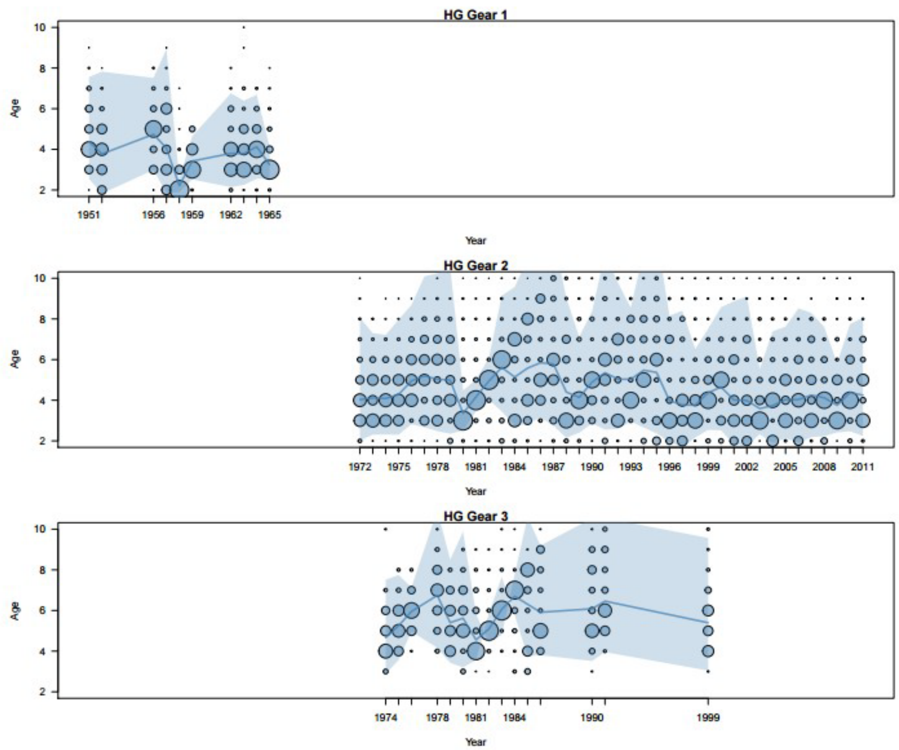
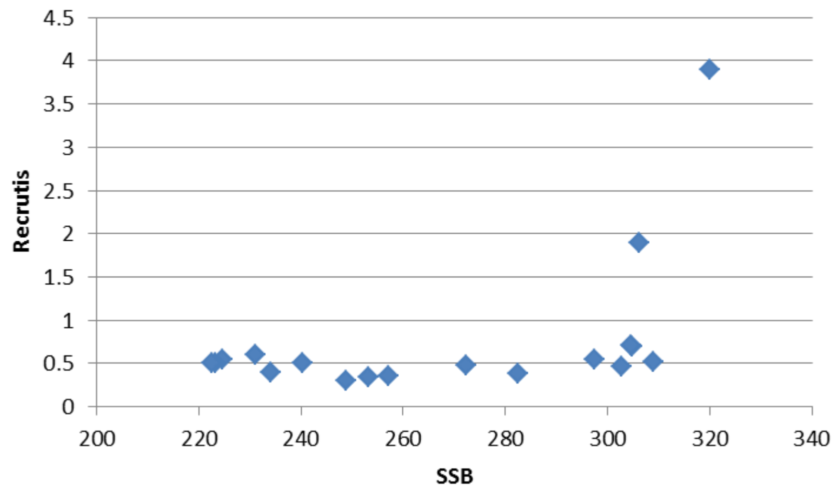


Figure 15. Proportions-at-age versus time for the winter purse seine fishery (top), seine roe fishery (middle) and the gillnet fishery (bottom) in Haida Gwaii. The area of the circle reflects the proportion-at-age, each column sums to 1, zeros are not shown, and age 10 is a plus group. Also shown is the mean age of the catch (line) and the approximate 95% distribution of ages (shaded polygon) for each year.

The same lack of contrast in cohort strength is evident in the survey indices. Have the authors analyzed the data for significant year effects? The figures of age composition would indicate that the data may not provide a sufficient basis for an age structured model.

This lack of information in the age composition permeates the results. The authors acknowledge this issue in a number of places but never follow it through to its logical conclusion. The lack of any stock recruit relation, the dependence of h on its prior, probably also the apparent constant population size over the time frame of the assessment and the scaling of the population, will all be impacted by this.



Another possibly major problem with the data is the lack of information prior to 1996. A 19 year period is relatively short when trying to devise reference points based on stock recruit. But perhaps more importantly important stock dynamics could be missing with this short time series which does not include potentially important fishing mortality prior to 1996.

There needs to be serious consideration given as to whether or not these data should be used in an age structured model. The data should be examined to see if significant cohort effects are present.

So the main issues with the data are the lack of contrast in the age composition data, the low sample size for the commercial otoliths and the lack of data prior to 1996. Other issues are the lack of representative sample from the freezer trawlers (as the authors point out) and possibly the changing selectivity of the fishery. How much does the model depend on commercial selectivity being stable?

Reference points

As the authors point out reference points based on MSY will be affected by selectivity and so they propose other biological reference points. (The rationale for this is however not well described [page10] and should be better described). However, the do not discuss other F reference points. This same uncertainty should affect F_{msy} as well (in fact they mention uncertainty in F_{msy} in this same paragraph). *Some consideration should be given to suggesting other reference points which could perhaps be better estimated (such as YPR or SPR reference points) or to suggesting perhaps some percent of F_{msy} (say ½ or 2/3) as a limit in order to hedge against the uncertainty.*

Model diagnostics and Sensitivity Analyses

The authors have done a reasonable job of presenting and interpreting model diagnostics but I have a few comments.

In the trace plots in figure 8 it is very difficult to see what is happening. Even still there seems to possibly be some pattern in some of the parameters, particularly the recruitment parameters. *Are there other convergence statistics that could be examined?* I am not familiar with the implementation of Bayesian inference in ADMB but perhaps the sampler running mean and/or Gelman and Rubin shrink factors could be examined for the major parameters.

Table 4 lists 11 parameters with priors. But only 9 priors and posteriors are shown in figures (for 6 parameters). *The selectivity priors and posteriors could be informative and the authors should consider showing them.*

The interpretation of sensitivity analyses would benefit greatly from some presentation of the priors and posteriors from these runs, particularly for the parameter of interest in the given run but also for all the main parameters being estimated. In this way one could see if one of the runs improved, for example, the ability of the model to extract information about steepness.

The sensitivity analyses for q include a run with a higher mean for the prior. However, the posterior distributions on q from the reference case show that the q ‘wants’ to be lower than the prior. Why not have a sensitivity run with a lower mean on q? Again presentation of the prior and posterior for the sensitivity analyses would help in the interpretation here.

Advice

The decision table includes female only TAC options. Of course the catch will be both male and female (seems to vary from 10-30% male). Would the decision table be converted to total catch and if so how? Would it be wiser to not convert it given that there does not seem to be a way to predict how many males would be caught and this would also result in a lower TAC given the uncertain assessment?

All TAC options show little or no probability of being below any of the biological limit reference points examined. The authors correctly state 'We therefore urge caution with the advice presented in Tables 14 and 15. The magnitude of catch and discards prior to 1996 is a major source of uncertainty in this assessment. The short time series of catch and surveys that were available in this assessment were relatively uninformative about the scale of the population, evidence for which can be seen in the high parameter correlations (Figure 16) and the influence of the prior probability distributions for the survey catchability parameters (Section 2.4.4). If reliable pre-1996 catch data had been available, the stock assessment and harvest advice may have been very different to those presented here.' This assessment is highly uncertain. Does this warning suffice?

More minor comments

The fits to the biological data (growth, maturity) seem good. However, female length at age does not seem to have quite levelled off by age 20. Are there otoliths beyond age 20 that could be used to confirm the Linf for females?

Sample size for ageing is given as 6000 but I count 4500. Are the sample sizes on the age composition figures correct?

I do not understand what is meant by 'unsorted/discard samples only'.

A number of things are referenced as being based on Holt et al 2014 but this paper is not yet available on the CSAS website (at least I couldn't find it!). The authors should consider if it is necessary to describe in more detail those things that are dependent on this reference.

Figure 5 presents the length quantiles of samples taken from 2 different fleets. It is difficult to interpret this without the 2 fleets being on one figure. In fact, the length frequency distributions themselves might be easier to interpret here. An additional question is whether the distribution of age across length differs between the two. It would also be helpful to present the length frequency distribution for males and females in the commercial data at least for a few representative years to see how much overlap there is in size and hence vulnerability to the fishery.

Biological parameters are assumed to have been constant over time. Has this assumption been examined?

'This prior was based on a literature review on steepness parameters for Pacific flatfish species done by Holt et al. (2014). A review of steepness estimates for flatfish species by Maunder (2012) suggested that flatfish steepness using a Beverton-Holt stock-recruit relationship may be around 0.94 (where h approaching 1.0 implies recruitment is independent of spawning biomass).' I found this confusing. Was the prior based on Holt or Maunder?

Q is sometimes subscripted 1-4 sometimes 2-5. This discrepancy is present in text, tables and figures.

The plots for the sensitivity analyses are difficult to interpret in that I was not able to tell which credible interval belonged to which run…and I am not colour blind! Is there some other way to present these?

There is some issue with the x-axis labels in figure 42.

There is no plot of the estimate stock recruit function and how it relates to the 'observed' recruitment. What does the Beverton-Holt actually look like?

Future Research

Pre 1996 catch data

Greater sample size for otoliths

Production model if age composition continues to show no contrast.

REVIEW – CAREY MCGILLIARD (NOAA)

Date:	May 7, 2015
Reviewer:	Carey McGilliard, National Oceanic and Atmospheric Administration, Alaska Fisheries Science Center
CSAS Working Paper:	2013GRF001
Working Paper Title:	Arrowtooth Flounder (*Atheresthes stomias*) stock assessment for the west coast of British Columbia

Summary:

This is the first statistical catch-at-age model for B.C. Arrowtooth Flounder. The model uses a Bayesian errors-in-variables approach. I was impressed by all of the work that was invested to develop the model. I found it useful that the authors estimated (rather than fixed) steepness, natural mortality, and catchability parameters in a Bayesian context. This was helpful for identifying areas of uncertainty, confounding among parameters, assessing what and how much information is contained in the data sources that are available, and identifying the limitations of the data. The topics for discussion below are all ideas for improving current or future assessments.

Topics for discussion:

1. What are the leading hypotheses as to why Arrowtooth catch is 80-90% female? Are males and females found in different locations? Are differences in growth in combination with selection of only larger fish by the surveys and fishing fleet influencing the sex ratios? Do males have a higher natural mortality rate than females?
2. What are the advantages and drawbacks of the female-only model? Given that this is a female-only model where it is estimated that females made up 80-90% of the catch, it seems that the assumption that catches are known perfectly might be a stretch. This could influence estimates of the scale of the stock (absolute biomass).
3. The authors mention that freezer trawl ages are under-represented in the fishery age composition data. My understanding is that there are 2-4 Arrowtooth fishing sectors contributing to the fishery age composition data: non-freezer trawls (where Arrowtooth is a discard fishery), freezer trawls that entered the fishery in 2006, the short-lived test fishery at Port Simpson, and a hook-and-line fishery (bycatch in the Pacific halibut fishery); is this correct? Why model one aggregated fishing fleet rather than 2-4 separate fleets? The concern that the authors bring up that the freezer trawl ages are under-represented in the age composition data suggests that the authors think that there could be differences in age composition between fishing sectors. If this is the case, it seems that it may be advantageous to model the freezer trawls as a separate fleet. This would alleviate the problem of having an under-representation of freezer trawl ages, in part, as the model would fit a separate selectivity curve to the freezer trawl age data and this would provide information on whether there are differences in selectivity or availability between fishing sectors. Also, is the proportion of female catch the same in each sector? This relates to Topic 1: why is Arrowtooth catch 80-90% female?

4. The authors mention that there are no age data for the Hecate Straight Multi-species Assemblage Survey; are length data available? Could length data be incorporated to help inform selectivity for this survey?
5. Inclusion vs. exclusion of older data:
   a. The Hecate Straight Multispecies Assemblage Survey started in 1984, but the present assessment only includes observations from 1996 onward; wouldn't inclusion of older data (particularly length composition data if it exists) help to inform estimates of B0 and relative biomass? In addition, could the length data inform an estimate of a historical fishing mortality rate to account for the uncertain discards of Arrowtooth prior to 1996? What were the reasons for excluding the older years of the survey?
   b. Growth was estimated using data dating back to the 1950s, while the rest of the data in the assessment dates back to 1996. Has any analysis been done to assess whether changes in growth rates have occurred since the 1950s and whether growth curves would be the same or similar if data from 1996 onward were used?
6. Did you try out using length-based selectivity, rather than age-based selectivity? If differences in sex ratios are due to size selectivity, then perhaps size-based selectivity would be more representative of the fishing process.
7. Page 43, panel d: Do you have ideas about why the WCVISS survey observations vary so much in recent years? Is catchability changing over time for this survey? What processes might be contributing to these patterns? It is thought that catchability of several Bering Sea-Aleutian Island flatfish species varies with temperature (e.g. Wilderbuer et al. 2002). Could something similar be going on here?
8. A map identifying the locations of the survey areas would be helpful.
9. A plot of age composition fits and residuals aggregated over years for each of the surveys and for the fishing fleet would be helpful to gauge how well the model is fitting to the age composition data overall for each of these data sources.

Further ideas for future research and assessments

1. For future assessments and research: I agree with the authors that a stock structure analysis would be very helpful. The surveys are specific to particular areas, therefore this assessment has aspects of an areas-as-fleets model. The survey selectivity curves that are estimated correspond to sub-areas within the range of the stock that is modeled. However, the assessment model is not spatial, so it assumes that these selectivity curves represent selectivity for survey fleets that are covering the entire range of the stock. Simulation studies have shown that the assumption of areas-as-fleets can lead biased biomass estimates in a number of circumstances, such as when directed movement occurs among areas and fishing mortality rates vary among areas (related references: Punt et al. 2015, Punt et al. 2014, Waterhouse et al. 2014, Punt and McGilliard 2008). Are commercial catches available by area? Does fishing mortality vary substantially among areas? Perhaps a stock structure analysis and potentially a spatial assessment approach would be useful, at least as an exploratory model.
2. Consider including estimates of ageing error and ageing bias in future assessments. I don't think it's necessary to change this year's assessment to include this. However, assuming that ages are perfectly known can be a source of bias in determination of stock status by way of influencing fits to age composition data.

3. Catchability experiments for the current surveys seem like they would be worthwhile to try to pin down priors on catchability a bit better (which may then change estimates of natural mortality and selectivity).
4. Have you considered a conditional age-at-length model where growth is estimated within the model? The advantages to doing this are that uncertainty about growth estimates can be propagated through the model. In addition, both length- and age-at-length data can be directly used, and this length data may help inform estimates of parameters other than growth parameters (e.g. length-based selectivity).

References:

Punt, A.E., Haddon, M., Tuck, G.N. 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth, and recruitment? A case study based on pink ling in Australia. Fish. Res. 168: 85-99.

Punt, A.E., Hurtado-Ferro, F., Whitten, A.R., 2014. Model selection for selectivity in fisheries stock assessment. Fish. Res. 158, 124–134.

Punt, A.E., McGilliard, C., and R. Hilborn. 2008. Evaluation of the impact of Marine Protected Areas on stock assessment-based fisheries management. Final Report to Washington Sea Grant (109pp).

Waterhouse, L., Sampson, D.B., Maunder, M., and Semmens, B.X. 2014. Using areas-as-fleets selectivity to model spatial fishing: asymptotic curves are unlikely under equilibrium conditions. Fisheries Research. 158: 15-25 (special issue).

Wilderbuer, T.K., A.B. Hollowed, W.J. Ingraham, Jr., P.D. Spencer, M.E. Conners, N.A. Bond and G.E. Walters. 2002. Flatfish recruitment response to decadal climatic variability and ocean conditions in the eastern Bering Sea. Progress in Oceanography. 55:235-247.

APPENDIX C: AGENDA

Canadian Science Advisory Secretariat

Centre for Science Advice Pacific

Regional Peer Review (RPR) Meeting

Coastwide Assessment of Arrowtooth Flounder (*Atheresthes stomias*)
in 2014 and recommendations for 2015

May 12-13, 2015

Seminar Room, Pacific Biological Station, Nanaimo BC
Chair: Rowan Haigh
Rapporteur: Melissa Nottingham

DAY 1 - Tuesday, May 12

Time	Subject	Presenter
0900	Introductions Review Agenda & Housekeeping CSAS Overview and Procedures	Chair
0915	Review Terms of Reference	Chair
0930	Presentation of Working Paper	Authors
1030	***Break***	
1050	Overview Written Reviews	Chair + Reviewers & Authors
12:00	***Lunch Break***	
1300	Identification of Key Issues for Group Discussion	RPR Participants
1330	Discussion & Resolution of Technical Issues	RPR Participants
1445	***Break***	
1500	Discussion & Resolution of Results & Conclusions	RPR Participants
1630	Develop Consensus on Paper Acceptability & Agreed-upon Revisions	RPR Participants
1700	Adjourn for the Day	

DAY 2 - Wednesday, May 13

Time	Subject	Presenter
0830	Introductions Review Agenda & Housekeeping Review Status of Day 1	Chair
0845	*(As Necessary)* Carry forward outstanding issues from Day 1	RPR Participants
1000	Develop Consensus on Paper Acceptability & Agreed-upon Revisions	RPR Participants
1030	***Break***	
1050	Science Advisory Report (SAR) Develop consensus on the following for inclusion: • Sources of Uncertainty • Results & Conclusions • Additional advice to Management (as warranted)	RPR Participants
1200	***Lunch Break***	
1300	Science Advisory Report (SAR) • Continued	RPR Participants
1445	***Break***	
1500	Next Steps – Chair to review • SAR review/approval process and timelines • Research Document & Proceedings timelines • Other follow-up or commitments (*as necessary)*	Chair
1545	Other Business arising from the review	Chair & Participants
1600	***Adjourn meeting***	

Note:
All participants are “reviewers” and are asked to be familiar with the material provided (Terms of Reference, working paper(s), supporting material) and participate fully in review discussions.

APPENDIX D: PARTICIPANTS

Last Name	First Name	Affiliation
Ackerman	Barry	DFO Fisheries Management – Groundfish
Chalmers	Dennis	Province of BC
Edwards	Andrew	DFO Science, Groundfish Section
Forrest	Robyn	DFO Science, Groundfish Section
Grandin	Chris	DFO Science, Groundfish Section
Haigh	Rowan	DFO Science, Groundfish Section
Hargreaves	Marilyn	DFO Science, Centre for Science Advice Pacific
Holt	Kendra	DFO Science, Groundfish Section
Keizer	Adam	DFO Fisheries Management – Groundfish
King	Jackie	DFO Science, Groundfish Section
Krishka	Brian	DFO Science, Groundfish Section
Kronlund	Rob	DFO Science, Groundfish Section
Mann	Shannon	Trawl Fishery Representative
McGilliard☼	Carey	Alaska Fisheries Science Center, NOAA, Washington
Morgan☼	Joanne	DFO Science, NAFC, Newfoundland
Mose	Brian	Commercial Industry Caucus (CIC) – Trawl
Nottingham	Melissa	DFO Science, Groundfish Section
Olsen	Norm	DFO Science, Groundfish Section
Starr☼	Paul	Canadian Groundfish Research & Conservation Society
Turris	Bruce	Canadian Groundfish Research & Conservation Society
Wallace☼	Scott	David Suzuki Foundation
Wischniowski	Stephen	DFO Sclerochronology Laboratory
Workman	Greg	DFO Science, Groundfish Section
Yamanaka	Lynne	DFO Science, Groundfish Section

☼Offsite participants connected via webinar or teleconference.