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 2016/031

Newfoundland and Labrador Region

Proceedings of the Northern Cod Framework Review Meeting

November 30 – December 4, 2015
St. John's, NL

Chairperson: Don Power
Editor: Emilie Novaczek

Science Branch
Fisheries and Oceans Canada
PO Box 5667
St. John's, NL A1C 5X1

September 2016

Canada

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. 2016. Proceedings of the Northern Cod Framework Review Meeting ; November 30 – December 4, 2015. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2016/031.

TABLE OF CONTENTS

SUMMARY

The Northern Cod Framework Meeting was held November 30-December 4, 2015 in St. John's, Newfoundland and Labrador (NL). The goal of the Framework was to determine which data sets and models will be considered at the March 2016 Regional Peer Review (RPR) as a basis for providing management advice. Management advice was beyond the scope of the framework meeting. Participants included representatives from Fisheries and Oceans Canada (DFO) – Science and Fisheries Management Branches, NL Region, Memorial University of Newfoundland (MUN), the Fish Food and Allied Workers Union (FFAW), and four invited reviewers:

1. Dr. Joanna Mills-Flemming, Associate Professor of Statistics, Dalhousie University (Halifax, Canada);
2. Dr. Anders Nielsen, Senior Research Scientist, National Institute of Aquatic Resources (DTU Aqua, Copenhagen, Denmark);
3. Dr. Chris Legault, Supervisory Research Fishery Biologist, National Marine Fisheries Service (NMFS, Woods Hole, United States); and
4. Dr. Doug Swain, Research Scientist, Fisheries and Oceans Canada (Moncton, Canada).

Various data sets were reviewed, with participants and reviewers providing input on which data are appropriate for a stock assessment model. One dataset, from the offshore acoustic biomass surveys conducted by the Centre for Fisheries Ecosystems Research (CFER) of the Marine Institute of MUN, involved questions beyond the scope of this meeting and has been recommended for further peer review. Two models were presented to the meeting: the Northern Cod Assessment Model (NCAM) and Statistical Catch at Age (SCA) model. NCAM was custom built for the Northern cod stock and includes inputs from a wide number of data sources, including tag data. As presented, the meeting participants were confident that NCAM could be applied to the 2016 Northern cod stock assessment. The SCA model, a more simple and standard approach, would also be acceptable for use in this assessment if concerns with the residual patterns were addressed. A bioenergetics model was also presented that provides a possible explanation for trends in natural mortality estimated in the NCAM and SCA models.

Compte rendu de la Réunion d'examen du cadre sur la morue du Nord

SOMMAIRE

La réunion d'examen du cadre sur la morue du Nord a eu lieu du 30 novembre au 4 décembre 2015 au Centre des pêches de l'Atlantique nord-ouest, à St. John's (Terre-Neuve et Labrador). Le cadre visait à déterminer les ensembles de données et les modèles qui seront envisagés au cours de l'examen régional par les pairs de mars 2016 afin de fournir des avis concernant la gestion. La fourniture d'avis concernant la gestion a dépassé la portée de cette réunion sur le cadre. Ont participé à la réunion des représentants des directions des Sciences et Gestion des pêches de la région de Terre-Neuve-et-Labrador de Pêches et Océans Canada (MPO), de l'Université Memorial de Terre-Neuve, de la Fish, Food and Allied Workers Union, et quatre examinateurs invités :

1. Joanna Mills-Flemming, Ph. D., professeure agrégée de statistiques, de l'Université Dalhousie (Halifax, Canada);
2. Anders Nielsen, Ph. D., chercheur principal, de l'institut DTU Aqua (Copenhague, Danemark);
3. Chris Legault, Ph. D., superviseur, biologiste en recherche halieutique, National Marine Fisheries Service (Woods Hole, États-Unis) ; et
4. Doug Swain, Ph. D., chercheur, Pêches et Océans Canada (Moncton, Canada).

Divers ensembles de données ont été examinés, et les participants et les examinateurs ont fourni des commentaires sur les données jugées appropriées pour un modèle d'évaluation des stocks. Un ensemble de données tiré des relevés acoustiques de la biomasse effectués au large par le Centre for Fisheries Ecosystems Research (CFER) du Marine Institute de l'Université Memorial de Terre-Neuve, a suscité des questions allant au-delà de la portée de la réunion et a été recommandé aux fins d'examen plus approfondi par les pairs. Deux modèles ont été présentés à la réunion : le modèle d'évaluation de la morue du Nord et le modèle statistique fondé sur les prises selon l'âge. Le modèle d'évaluation de la morue du Nord a été construit sur mesure pour le stock de morues du Nord et il intègre des données provenant d'un grand nombre de sources, y compris les données fournies par les étiquettes. Sa présentation à la réunion a convaincu les participants qu'il pourrait être appliqué à l'évaluation du stock de morues du Nord en 2016. Le modèle statistique fondé sur les prises selon l'âge, une approche plus simple et normalisée, serait également acceptable aux fins d'utilisation dans le cadre de cette évaluation, à la condition que les préoccupations concernant les profils résiduels soient prises en compte. Un modèle de la bioénergétique a également été présenté qui fournit une explication possible des tendances de la mortalité naturelle estimées dans le modèle d'évaluation de la morue du Nord et le modèle statistique fondé sur les prises selon l'âge.

INTRODUCTION

The Northern Cod stock was most recently assessed in 2013 (DFO 2013). The purpose of the Framework Review meeting was to determine which data sets and models will be considered at the March 2016 Regional Peer Review (RPR) as a basis for providing management advice. The full Terms of Reference are included as Appendix 1. Datasets were reviewed for robustness of collection methods, changes in the survey design over time, data analysis, estimated uncertainty and utility to the stock assessment process. Stock assessment models were presented and critically reviewed by a panel of experts, including the model authors, meeting participants and external reviewers. Participants included representatives from Fisheries and Oceans Canada (DFO) - Science and Fisheries Management Branches, Newfoundland and Labrador (NL) Region, Memorial University of Newfoundland (MUN), the Fish Food and Allied Workers Union (FFAW), and four invited reviewers:

1. Dr. Joanna Mills-Flemming, Associate Professor of Statistics, Dalhousie University (Halifax, Canada);
2. Dr. Anders Nielsen, Senior Research Scientist, DTU Aqua (Copenhagen, Denmark);
3. Dr. Chris Legault, Supervisory Research Fishery Biologist, NMFS (Woods Hole, United States); and
4. Dr. Doug Swain, Research Scientist, Fisheries and Oceans Canada (Moncton, Canada).

REVIEW OF DATA SOURCES

Meeting participants critically examined the history, collection methods, analysis, and application of various available data sources for a Northern cod stock assessment. Experts on respective datasets were contacted in advance of the Framework Review with short questionnaires on the data collection and the potential for this data to inform stock assessment as an estimate or index of stock size (Appendix 4).

In general terms, the following questions were discussed for each dataset:

1. How are the data collected?
2. Has the design changed over time?
3. How are the data analyzed?
4. Are there estimates of uncertainty?
5. Does this dataset provide an index of stock size?
6. How do these data contribute to assessment of stock status?
7. Are the analysis methods, including uncertainty estimates, appropriate?
8. Are there suggestions for improvements or recommendations for further research?

During the meeting, discussion on the input data was generally focused on arriving at consensus text for Questions 5-8 for each data source that could be used to guide future assessments of the stock. The principal investigator/data custodian associated with each dataset was requested to draft text to guide discussion.

FISHERIES AND OCEANS CANADA RESEARCH VESSEL SURVEY

Presented by: Brian Healey

The DFO Research Vessel (RV) survey represents a well-known and widely used data source in stock assessments. Fishery-independent, multi-species trawl surveys have been conducted by DFO since 1949. The original line transect survey design was replaced by depth stratified tows in 1971, and the area surveyed was expanded significantly in 1977. Additional in-shore strata were added to the survey in the mid-1990s, however these are not among the index strata used to monitor the cod stock, and have not been sampled in every year.

Throughout the time-series, there are some gaps in coverage (missed strata or missed sets) due to limited vessel capacity or poor weather. For example, the 2004 and 2008 surveys delivered reduced coverage, largely due to vessel limitations. Multiplicative models have been used to estimate missing strata in the past, however considering coverage has usually been extensive this has not been applied to Northern cod in recent years. It is important to note that this survey includes many strata, and the proportion that is missed remains very low. However, the ratio of completed to allocated tows within strata is not consistent; a stratum in which only three of ten allocated tows are retrieved will still be considered complete.

There is also some variation in the time of the survey: in some years the fall survey (typically October-December) has extended into the following winter, as late as the first days of February in 2006. Due to variation in the survey timing and variation in Northern cod seasonal migrations, participants noted that it is very important to consider what portion of the stock is covered by the survey. Related to this, it was discussed that stock decline was accompanied by changes in distribution of Northern cod.

The survey gear has changed three times since 1949. The most recent change from Engel to Campelen trawls was instituted in an effort to improve the enumeration of fish below 25 cm, which are not efficiently caught by the larger-mesh Engel trawl. Conversion factors were calculated for Northern cod when the Campelen trawl was introduced (Stansbury 1997), however these studies did not include fish larger than 85 cm as no large fish were sampled when the comparative fishing experiments were conducted in the post-collapse period. Several meeting participants raised concern that these calculations introduce some uncertainty to the RV survey as abundance of larger size classes of Northern cod increase. There is no documented evidence that the Campelen trawl does not efficiently catch large cod. Questions regarding Campelen catch efficiency for larger fish arise because the conversion studies in the Newfoundland Region were originally conducted at a time when the abundance of large cod were in significant decline and may not have been available to the survey.

The consensus of meeting participants was that the DFO RV survey remains a very important dataset for any future Northern cod stock assessments. Meeting participants suggest that this dataset could be improved with additional comparative fishing studies that reflect current trends in size and aggregation density of Northern cod.

EU-SPAIN RESEARCH VESSEL SURVEY

Presented by: Don Power (input received from colleagues at Instituto Español de Oceanografía, Vigo, Spain)

The EU-Spain RV Survey takes place in NAFO Div. 3L, beyond the Canadian Exclusive Economic Zone (EEZ). The survey takes place in July and August annually, beginning in 2003, including tows along the Flemish pass since 2006. Fundamental survey design is similar to the DFO RV Survey: stratified random tows using a Campelen trawl. The surveys differ in tow duration (Spanish RV tows 30 minutes at 3 knots, DFO RV tows 15 minutes at 3 knots), allocation of sets per strata size, and configuration of trawl bridles. For more details on this data, see Román et al. 2015a.

It was noted that this survey appears to sample an area of higher Northern cod density than the DFO RV survey. During winters with expansive ice cover, Northern cod have been recorded in dense aggregations and the area may represent an important area for juveniles (Román et al. 2015b). This dataset has not been applied directly in previous Northern cod stock assessments. The EU-Spain RV Survey includes a very small fraction of the area occupied by this stock, and there is no evidence that a significant portion of the stock is sampled. Therefore, this dataset does not provide an index of stock size, however it may be compared to the DFO RV Spring Survey to inform estimates of seasonal movements of Northern cod.

FISHERIES AND OCEANS CANADA TAG DATA

Presented by: John Brattey

Tagging studies have been conducted since 1948 to collect distribution and movement data for Northern cod. Various tags have been used in this time period, however use of external Floy tags has been used for Northern cod ≥ 45 cm consistently since 1995, with a reward offered for the return of any caught tags since 1997. Double tagging studies have been conducted to provide estimated tag loss and high rewards (assuming 100% return of the high reward tags) have been offered to estimate reporting rate. Tags have been deployed during targeted experiments using CCG vessels and also opportunistically using other platforms throughout the stock area, with a total of 68,000 deployed tags in the post-moratorium period. Until 2007, tag deployments were limited to the inshore. More recently, tags have been deployed offshore to be recaptured inshore and at great distances from the tag point.

Since 2006, tag return data includes indication of whether the recapture occurred in either the stewardship or the recreational fisheries. In both fisheries (mainly gillnet and hand-line, respectively), it is unlikely that tags are being missed by harvesters as fish are handled individually. According to telephone surveys, reporting rates appear to vary without trend from year to year (50-75% in 2007, 70-100% in 2009, and 50-75% in 2012). Reporting rates appear to be slightly lower in the recreational fishery, which accounts for about 40% of total catch, according to tag reports. Participants expressed concern that a significant amount of information would be lost if reporting drops below 50%. Tag returns are currently worth $10 to harvesters for a normal tag, and $100 for a high reward tag. These values have not changed since 1995, and it was suggested during the meeting that these values may be increased to match inflation.

FISHERIES AND OCEANS CANADA ACOUSTIC TELEMETRY

Presented by: John Brattey

In addition to conventional tagging, there is a shorter time series of acoustic tag data for Northern cod telemetry. The majority of tagged fish are ≥ 60 cm using a variety of VEMCO ™ transmitters and for larger transmitters, extended battery life allows data collection up to and well beyond 900 days. The tags emit a unique acoustic signal for each fish, which is recorded with time and date every time a tagged fish swims within 800 m of a hydrophone. It requires a significant amount of field work to retrieve data from the receivers.

The hydrophone and tag network generate a huge amount of data, in excess of 10 million data points to date. We are getting better and longer data series from individual fish, compared to a single data point at recapture for conventional tagging. The acoustic telemetry is also related to

Floy tagging; in acoustic tagging deployments, fish are outfitted with both types of tag which provides a record of the point of release, size, and possible capture.

The main limitation to this data is the number and distribution of hydrophone locations; only eight hydrophones are currently installed offshore (in 3K Bonavista Corridor), plus about six in Hawke Channel. However, the offshore hydrophones are already delivering interesting preliminary data. Tagged Northern cod were not as widely dispersed as expected; authors expected to record one or two fish, however, 22 came within detection range of the small offshore hydrophone network during 2012-13. Inshore, where the spatial scale of fish movements is smaller, hydrophone networks have been extremely effective. With 100 tagged fish in Smith Sound, 88% survival could be confirmed throughout 2006 (the first year following release) and 66% survival in 2007, despite a fishery. This provides valuable context for interpretation of the spatial differences in sentinel fishery catch rate data and may even suggest that the predicted 20% natural mortality rate for Northern cod was an over-estimate, at least for large fish in the inshore.

NORTHERN COD SATELLITE TAG DATA

Presented by: Sherrylynn Rowe

Ninety one-year pop-off satellite tags were attached to large Northern cod ($\geq$ 85 cm, most > 95 cm) between 2012 and 2015. Although not all tags last the full year, and some tags fail, the successful tags typically transmit hourly data on temperature and depth as well as estimated daily location. Data are processed by the manufacturer and preliminary results could be available in May 2016. Location data are derived from light, and are therefore more accurate and easier to interpret when the fish are in shallow water. Derivation of location data for Northern cod offshore and in deep water may be limited. Even if temperature and depth data are used in conjunction with oceanographic models, location estimates may not be reliable due to the relatively homogenous conditions across the shelf.

No results from this study are yet available for review. Pending analysis, tag data may support stock assessment by indicating Northern cod distribution and movement, particularly during surveys and fisheries. However, this dataset is currently limited to relatively few data points from very large fish, which make up a small fraction of the stock. It was indicated that while more tags should be deployed and data retrieved before conclusions are drawn from this dataset, it may offer a promising resource to future Northern cod stock assessments.

NORTHERN COD ACOUSTIC BIOMASS SURVEY

Presented by: Sherrylynn Rowe

Offshore Acoustic Survey (Bonavista Corridor)

Acoustic biomass surveys have been conducted for Northern cod in the Bonavista Corridor since 1990, during the late winter and spring of the year. Latitudinal transects extend across the shelf between 250-500 m depth. Directed trawling is conducted to identify species, size distribution and other characteristics of detected fish aggregations. Trawling may be triggered by a change in the acoustic signal, detection of a high density aggregation, or to confirm an empty signal. Additional mid-water trawls were conducted when aggregations extended far off the bottom. This survey aims to estimate spawning biomass of Northern cod in the study area; very few juveniles are retrieved by trawl sets. Juvenile Northern cod are expected to take cover near bottom during spawning aggregations.

As the technology of acoustic data collection and analysis have advanced, the equipment, vessel, and survey designs have changed. The authors consider the backscatter data to be consistent and comparable across these developments, however directed trawling gear types have varied (Campelen, GOV) and area surveyed is not consistent year to year. Since 2012 there have been no changes in equipment, timing, or vessel. Recent changes relate to survey coverage (e.g. the 2015 survey expanded significantly Northwards). A question was raised as to whether or not the acoustic data collected during the various experiments could be considered a time-series. Furthermore, some participants questioned the extent to which it may be reproducible.

All acoustic biomass surveys are limited by an "acoustic dead zone" (ADZ) in the echogram where biomass is extremely difficult to distinguish from the bottom echo. For this dataset, biomass of the ADZ is estimated by a linear expansion of the backscatter in the bottom 5 m of the signal (see Rose and Rowe 2015 for more detail). It was noted by participants, external reviewers and the presenting author that this assumption has not been validated. The dead zone estimate has ranged between < 10-20% of the signal and could readily be excluded if preferred, although this approach would almost certainly underestimate the numbers. Additionally, there was discussion over the frequency of trawling to validate the results on the echo sounder, and the fact that the number of tows conducted was quite limited in some years. The presenter indicated that trawling was generally conducted when the acoustic signal showed significant change.

The Framework Meeting was convened to review input data sets and stock assessment models. Most data sources reviewed here have been debated extensively during previous assessment meetings and are well understood by the meeting participants. However, the Framework Meeting did not have an acoustician or acoustic biomass data expert present to provide an in-depth review of the relatively new acoustic biomass data. After discussion on questions 1-4 for this data source, several questions were raised on the technical details of both the data collection and analysis methods.

Inshore Acoustic Survey (Smith Sound)

Portions of Smith Sound were surveyed annually to estimate biomass of overwintering Northern cod. In and out movements were first observed in 1995 and the subsequent survey design was based on these observed movement patterns to capture the maximum aggregation in late winter-spring. In many years, multiple surveys were conducted to monitor finer scale movements. Participants pointed out that other work also indicates movement of some portion of the population in and out of the sound seasonally. For example, following a natural mortality event in 2013, studies indicated that few fish were present, and many returned three weeks later. As with any survey, it is important to critically consider what portion of the population is available to sample.

Participants agreed that there is much more confidence in this dataset as a time-series (compared to the offshore acoustic biomass survey) due to the small geographic extent of the survey area, and very limited variation in survey methods, gear types and coverage.

Application of Acoustic Biomass in Stock Assessment

The acoustic survey is presented here as a direct estimate of biomass, however it was suggested by several meeting participants that these data may be more accurately considered as an index of biomass. A working paper was presented which provided a summary of how acoustic surveys are included in various stock assessments within the International Council for the Exploration of the Seas (ICES). Thirty stocks were found that include information on acoustic biomass surveys (Table 1). In general, the surveys covered most or all of the stock area, were conducted with a fairly high density of transects, and employed standard designs that did not vary from year to year. Of the 30 stocks identified by the survey, only two capelin stocks treated the acoustic survey as an absolute estimate of biomass. In these cases the acoustic estimates were used with uncertainty derived from bootstrapping procedures. The majority of stocks (26) treated the acoustic survey as tuning indices in population models or as indices of general trends, independent of the type of model used, the species considered, or the ICES working group. Two stocks were found that described acoustic surveys but did not use them in the assessment model. Reviewers and participants agreed that the acoustic biomass survey data should be presented with estimates of survey catchability, particularly if acoustic data are used to calibrate a population model. This addition would help guide decisions of interpretation of the data as an index or absolution value.

Given the lack of expertise in fisheries acoustics amongst the framework review panel, it was beyond the scope of this meeting to draw conclusions on the utility of the acoustic biomass survey for Northern cod in stock assessment. Meeting participants recommend further review of this dataset (further details are provided in Meeting Conclusions and Recommendations).

Table 1. The use of acoustic surveys in stock assessments conducted by ICES. Only those stocks with acoustic survey information are shown. The table presents the stock, how the acoustic survey is used in the assessment and the assessment method. The assessment working group is given as well.

Stock	How is acoustic survey used?	Stock assessment method
NWWG	-	-
Icelandic summer spawning herring	Tuning index	ADAPT
Icelandic East Greenland capelin	Absolute with uncertainty from bootstrap	Model based on acoustic survey and prediction six months ahead to calculate spawning biomass. The model estimates maturity, growth, and mortality. Target escapement strategy used.
Shallow Pelagic S. mantilla Irminger Sea	Index of trends	Trends based
Deep Pelagic S. mantilla Irminger Sea	Index of trends	Trends based
AFWG	-	-
Norwegian Coastal Cod	Index of trends	Trends based
Barents Sea cod	Tuning index	XSA
Northeast Arctic haddock	Tuning index	XSA
Northeast Arctic saithe	Tuning index	SAM
Redfish S. mantilla I and II	Not used	SCAA
Barents Sea capelin	Absolute with uncertainty from bootstrap	Model based on acoustic survey and prediction six months ahead to calculate spawning biomass. The model estimates maturity, growth, and mortality. Target escapement strategy used.
HAWG	-	-
North Sea herring	Tuning index	FLSAM
Herring Div IIIa Subdiv 22-24	Tuning index	FLSAM
Herring Celtic Sea	Tuning index	ASAP
Herring Div VIa VIIbc	Tuning index	FLSAM
Herring North Irish Sea	Tuning index	FLSAM
Sprat North Sea	Tuning index	SMS
Sprat Div IIIa	Index of trends	Trends based

Table 1. Continued.

Stock	How is acoustic survey used?	Stock assessment method
WGHANSA	-	-
Anchovy Bay of Biscay	Tuning index	Bayesian two stage biomass
Anchovy Div IXa	Index of trends	Trends based
Sardine Div VIIa,b,d SA VII	Index of trends	Trends based
Sardine VIIA IXa	Tuning index	SS3
WGBFAS	-	-
Herring SA 25-39 & 32	Tuning index	XSA
Herring 30	Tuning index	SAM
Herring Gulf of Riga	Tuning index	XSA
Sprat 22-32	Tuning index	XSA
WGDEEP	-	-
Greater smelt SA I, II, IV & Div IIIa	Index of trends	Trends based
WGWIDE	-	-
Boarfish VI-VIII	Tuning index	Exploratory Bayesian SPM
Herring SA I, II & V Div Iva & XIVa	Tuning index	TASACS
Mackerel SA I-VII, XIV & Div VIIIa-e & IXa	Not used	SAM
Blue whiting SA I-IX, XII, XIV	Tuning index	SAM

BEACH SEINE SURVEY (NEWMAN SOUND)

Presented by: Bob Gregory

Seine sampling was conducted at high frequency (11 times per year) in Newman Sound to investigate coastal recruitment of Northern cod. Newman Sound provides important eel grass habitat for juvenile cod. The seine data collected in Newman Sound has been compared to similar efforts in Smith Sound and Leading Tickles, with relatively consistent recruitment patterns across the sample sites.

The primary goal of this dataset is not stock assessment; the Beach Seine Survey was designed to address questions of juvenile cod behavior, however there are indications that this dataset corresponds well to trends in the DFO RV survey. A comparison of the cohort information at age 1 from the Newman Sound beach seine survey and the cohort information at age 3 from a sequential population analysis (SPA) for the inshore central area revealed a strong correlation (DFO 2007); however, the inshore SPA was discontinued after 2007 and it is unclear whether this correlation has persisted.

The author of this dataset hypothesizes that the nearshore eelgrass is preferred juvenile habitat that supports high densities of young Northern cod, even when the stock is low. As the stock recovers, the relationship between the Newman Sound Beach Seine Survey and Northern cod abundance may deteriorate.

INSHORE MOBILE GEAR SURVEY

Presented by: Don Power

A stratified random survey of coastal Divs. 2J3KL covering depths less than 200m was carried out between 2006 11. The goal of this survey was to collect data shoreward of the offshore DFO RV Survey. Five small vessels with standardized otter trawls were employed over the time period and were able to sample as shallow as 15 m.

The results over the five-year study period indicate high variability due to dependence of the survey on movement of migratory cod. This dataset cannot be used to track cohorts. For these reasons, the Inshore Mobile Gear data is not considered to be a reliable index of Northern cod stock status.

NORTHERN COD SENTINEL SURVEY

Presented by: Brian Healey

The Sentinel Survey in the NL Region of DFO is an extension of the Northern Cod Science Program of Index Fishers with modifications to allow for science activities under a fishing moratorium. The Sentinel Survey projects were formally announced by the Minister of Fisheries and Oceans Canada in October 1994. The primary goal of the Sentinel Survey is to collect data for stock assessment, including catch rates, temporal-spatial distribution of Northern cod, length frequencies, additional biological data (e.g. stomach contents, by-catch rates and composition) and long term oceanographic data for complementary research. Annual surveys are conducted inshore during traditional fishing times, with very little overlap with DFO RV survey strata. Various fishing gears have been used in the Sentinel Survey, including gillnets (5.5" and 3.25" mesh in tandem), line trawls, cod traps (phased out by 2002), and hand lines. Hand lines were used mostly in conjunction with nets or trawls as a means of acquiring cod for tagging purposes or to confirm presence of cod when nets provided low or no catch.

Size and duration of the Sentinel Survey have varied. In 1994, the survey included 66 sample sites, which was gradually reduced to a minimum of 43 sites by 2008. Since 2012, 48 sites have been surveyed. The reduction of survey sites has occurred in two ways: some sites were discontinued strategically in the early 2000s, and additional sites have been discontinued as harvesters retired and left the Sentinel Survey. Duration of sampling has ranged from 15 weeks (1995) to 8 weeks (1999). In recent years, the survey has been conducted over 10 weeks. Inshore and offshore movements are central to the interpretation of the Sentinel Survey; a peak in observed abundance in 1998 is likely to be due to changes in distribution and availability, rather than an increase in the stock.

The Sentinel Survey sampling strategy (experimental vs fixed sites) was designed to account for the local knowledge and ability of the harvester to follow and target the most productive areas. Statistically significant differences in catch rates between the two site types have not been examined to date, but such analysis could present very valuable information. The relationship between survey results and overall stock size remains unclear, and there was some general discussion about the 'effective reach' of the Sentinel Survey. Some participants questioned what the relative value of the Sentinel Survey might be (with respect to stock status determination) if the stock continues to show significant growth in the offshore. Notwithstanding this, it was suggested as the stock increases, larger mesh sizes should be added to the design.

STEWARDSHIP CATCH

Presented by: Brian Healey

Commercial removals do not provide an index of stock size but are an essential data requirement to understand population dynamics. Dock-side monitoring of Northern cod stewardship fishery catches has been conducted since the mid-1990s. Catch weight is recorded by independent monitors with varying coverage throughout 300 landing points. If a landing is not monitored at sea, the harvester is required to report ("call-in") as they approach shore, at which point they will be given an authorization number or they will be directed to a dock-side monitor. The authorization number permits the harvester to offload without a monitor present, however these catches may still be subject to random checks. A tiered system of dockside monitoring was implemented in 2010. Tier 1 docks (highest landing areas) are subject to 100% monitoring. Between 62-75% of landings at Tier 2 and 3 docks have a monitor present, and approximately 17% of authorized catches are subject to random checks when not monitored at sea or at arrival. Tier 4 docks have no monitoring but are also subjected to random checks by Conservation and Protection Officers on some occasions.

Harvesters and industry have reported several limitations to this system, including:

1. Abuse of authorization numbers;
2. Inaccurate reported weights;
3. Abuse of catch portion designated as "personal use;"
4. Dumping/discard at sea;
5. High grading for selected sizes (this occurred when price was linked to fish size, which is no longer the case);
6. Misreporting of species or stock area (mainly reported at the time of the moratorium) and;
7. Monitor avoidance and/or information trading on monitor location.

The extent and impact of these issues is not clear at this time. It was noted during the meeting that a request for validation of catch data would be filed with the Fish Harvesters Resource Center, the group responsible for the Dockside Monitoring Program and that the results would be shared with DFO. Information on discrepancies between reported and randomly checked catch weights, and violations noted by conservation officers will be examined to estimate self-reporting bias, however these details were not available at the time of this meeting. It was reported that DFO Science Branch has initiated discussion with colleagues in the Conservation and Protection and Resource Management Aboriginal Fisheries Branches to examine these issues. Where possible, plans will be implemented to either validate or refute any perceived problems, resulting in improved estimates of catch.

COMMERCIAL CATCH AT AGE

Presented by: Danny Ings

The commercial fishery and sentinel survey contribute to a combined analysis of catch at age for Northern cod. Additionally, Conservation and Protection Officers collect length measurements, which are included in catch at age for the food fishery in some years. Length frequencies are calculated by cells composed of source (i.e. Sentinel survey or commercial catch), month, gear type, and unit area. The number at age is calculated as the total catch divided by the average weight at age and multiplied by proportions at age (Gavaris and Gavaris 1983). Significant changes were made to the analysis of catch at age in 1993. Length frequency bins were changed from 3 cm to 1 cm intervals and otolith collection was reduced to one per length group.

This dataset does not provide an index of stock size. In previous assessment models (i.e. ADAPT), average weight at age estimates generated by the catch at age analysis were used to derive stock biomass. The 2013 assessment model (SURBA) uses average weight at age data from the autumn survey to predict stock biomass at the beginning of the year (Cadigan 2013).

RECREATIONAL FISHERY REMOVALS

Presented by: Brian Healey

All recreational fishing was unregulated until the moratorium. Since 1992, the recreational, or food fishery, has undergone several changes:

1. 1993-95: No recreational fishery;
2. 1996-2001: Limited weekend fishery subject to bag and boat limits;
3. 2001-02: Limited license-based fishery with retention tags (30 tags in 2001, 15 in 2002);
4. 2002-05: No recreational fishery;
5. 2006-present: Seasonal fishery (32-35 days) with no required license; 5 fish bag limit, 15 fish boat limit.

There are no available measures of the amount of Northern cod removed by the recreational fishery at present. Recent estimates based on tag reporting suggest that the recreational fishery represents as much 40% of total Northern cod removals in NL.

In 2007 two estimates of recreational fishery catch were calculated. Conservation and Protection provided an estimate based on reported catch and catch checked at sea. DFO contracted an independent consulting firm to conduct a telephone survey of participants in the recreational fishery. The results were radically different: 371 t and 2000 t, respectively. The discrepancy was explored and neither estimate could be validated. This knowledge gap remains a significant challenge for stock assessments.

INSHORE HARVESTER QUESTIONNAIRE

Presented by: Erin Carruthers

In the mid-2000s regional harvester meetings were replaced with a telephone survey as a means of gathering local knowledge on fishery performance to supplement research survey data. However, the questionnaire responses consistently diverge from the scientific survey and the two have not been reconciled. In previous stock assessments, the questionnaire results are presented in summary. To date, there have been no additional analyses beyond regional summaries. The telephone survey follows a stratified random design using the 2008 Groundfish License Holders list. The call list for each NAFO division is based on a randomized order of license holders. In 2014,135 license holders were surveyed in total; of these, 36 license holders were surveyed in Div. 2J, 32 in Div. 3K, and 67 in Div. 3L.

It was noted that questionnaire respondents tend to be positive when asked to compare the current catch to previous years and also the 1980s. For example, since inception of this survey, many harvesters indicate that cod abundance improves in almost every year and that current catches are greater than in the 1980s. This is inconsistent with all other sources of data.

Meeting participants discussed possible adjustments to the questionnaire. For example, it may be more informative to ask harvesters about their best or average catch, instead of asking them to compare the current year to the previous year in general terms. This may allow the questionnaire to capture and track changes in effort or fishing practice that are omitted by other data sources. The current harvester questionnaire is only applied to harvesters who have taken 1000 lbs of cod but it may be informative to include smaller harvesters. In NAFO Subdiv. 3Ps, a license holder survey includes queries to refusal; it was suggested that information on why harvesters are choosing *not* to fish may be useful throughout the stock area. Although possible improvements to the study were discussed, it is not yet clear what utility this dataset may have for a quantitative model of Northern cod. At present, this dataset does not provide an index of stock size.

COMMERCIAL LOGBOOK DATA (INSHORE < 35')

Presented by: Brian Healey

A science logbook was introduced in 1998 for inshore vessels < 35 ft. Harvesters are asked to fill in the daily catch record, providing information on estimated weight caught, location fished, type and amount of gear used, and by-catch details. At the end of the fishing season, harvesters are required to return the logbooks to DFO. Over time, additional data fields were added to the logbook (e.g. bycatch of *Species at Risk Act* [SARA]-listed species, fields for recapture detail of tagged cod).

This dataset does not estimate stock size. In the past, this dataset has been considered as an assessment of fishery performance, which is influenced by many factors beyond management scope: seasonal fish migration, presence of by-catch species, operational costs, weather conditions, etc. In other stock areas, commercial logbooks have been very valuable, due to video and Vessel Monitoring System (VMS) auditing. The audit process comes at an increased program cost, but also provides validated logbook data and includes spatial references that are not otherwise collected. This spatial data is of increasing importance, as more emphasis is placed on the study and conservation of marine species at risk. However, the Northern cod logbook dataset includes small vessels (< 35'); it was noted that the cost of VMS installation is roughly equal to the full annual Northern cod quota value per vessel. At present, about 50-60% of logbooks are returned, including very little positional information.

The meeting did not reach a conclusion on how this dataset may inform future Northern cod assessments. It was suggested by one reviewer that it may be time to discontinue logbook collection, however this dataset offers the only measure of inshore Catch Per Unit Effort (CPUE). If offshore CPUE is deemed relevant to a stock assessment, participants felt that the inshore logbooks should be similarly included. Further discussion about the utility of CPUE to stock assessment was beyond the scope of this meeting.

NORTHERN COD ASSESSMENT MODEL (NCAM)

INTRODUCTION

The Northern Cod Assessment Model (NCAM) is an integrated state-space model created by Dr. Noel Cadigan of the Centre for Fisheries Ecosystems Research (CFER) of the Marine Institute of Memorial University of Newfoundland with support from the Research and Development Corporation of Newfoundland Labrador. Models which integrate across multiple data sources are preferred approaches when possible, as they avoid subjective choices about which data sources may be most appropriate for addressing questions of interest, and they can potentially permit answers to problems which cannot be determined from individual data sets. Since the original publication of the NCAM model (Cadigan 2016), there have been several updates which are detailed below.

The most recent Northern cod stock assessment (2013) was based on a cohort model that produced estimates of relative stock size (i.e. trends) estimated using only the DFO RV Survey. This type of survey cohort model has a limited ability to provide catch advice to managers and projections of stock trends into the future. Other assessment problems for this stock involve the natural mortality rates and survey catch efficiency. There are indications that the natural mortality rate (M) increased substantially around the time of the moratorium, but the level and duration of this increase remains uncertain (Lilly 2008). Survey catch efficiency (Q), an index of stock availability to survey gear, is related to stock size (N) and the survey results (I). Q should be constant over time if the survey is used directly to infer stock trends but there are good reasons to think that Q changed since the moratorium.

These limitations were the motivation for NCAM development. As an integrated state-space model, NCAM uses a wide range of relevant information on population dynamics in order to address limitations of individual datasets. NCAM was built using Template Model Builder (TMB) software as a mixed-effects model including fixed parameters (θ) and process errors (ψ). The model is largely self-weighted.

NATURAL AND FISHING MORTALITY

Presented by: Noel Cadigan

The foundation of NCAM is a conventional population dynamics model. Total mortality (Z) is divided into natural mortality (M) and fishing mortality (F). The values of M for each age bin are drawn from previous peer reviewed literature but with additional annual variability that is modeled through random effects that are correlated across ages and years to produce a box-type covariance function. The Baranov catch equation is used to calculate total deaths resulting from F (Baranov 1945). F is modelled as stochastic process with random effects, similar to the M process, with fixed correlation between years (functions like a Random Walk). Total catch for each year is estimated through a censored likelihood approach. The reported catch is considered a lower bound. Different catch scenarios, with relatively lower or higher upper bounds may be tested and applied depending on managers' confidence in reported catch or supporting data. The censored likelihood calculates the probability that the variable x (in this case catch) falls within the given bounds. This generates a very flat probability distribution, which has been adjusted to have a heavier tail. Fishing mortality for young ages (2-4) is treated differently because catch of these young cod is probably only accidental. When an NCAM model is set up, a first run is conducted to fit catch between the given bounds, which provides starting values for the analysis run.

NCAM INPUTS

Presented by: Noel Cadigan and Christoph Konrad

This model incorporates a diverse range of data types, including total catch estimate, catch age composition, DFO RV Survey indices, inshore acoustic biomass estimates, inshore Sentinel Survey indices, and tagging data.

Catch Age Composition

The most recently reported catch age compositional data are presented in a 2009 Research Document (DFO 2009). Availability of the component data used to estimate catch at age and reporting of the methods are limited. Consequently, in NCAM, age compositions are simply modeled as a logistic normal multinomial (LNM; Francis 2014) distribution. This distribution requires non-zero age compositions; therefore, zero catches for some ages present a problem. As a solution, zeroes are replaced with a minimum value (0.5) and total catch is rescaled to match the raw data sum. At present, no sample size effect is included, though some participants agreed it may be a valuable addition. As discussed before, an estimated 40% of total catch is taken in recreational fisheries in recent years. Participants raised concern that the age compositions drawn from commercial data may not apply equally to the recreational portion of the catch.

DFO Survey Data

The offshore DFO RV Fall Survey provides model data for Northern cod ages 2-14, between 1983-2014. The Sentinel Survey provides inshore gillnet data for Northern cod ages 3-10, between 1995-2014

The DFO survey Q appears to have changed during the 1995-2009 period. This change likely occurred due to a change in the overwintering distribution of stock, mainly in Smith Sound, specifically, a group of older fish that may have moved from the offshore to the inshore. This hypothesis is based on work by George Rose, which indicates Northern cod overwintering in Smith Sound in 1995 (Rose 2003). Participants pointed out that this assumption may not be completely valid, as little or no sampling was conducted prior to 1995; researchers did not anticipate the large aggregation of cod in Smith Sound, however there is no evidence of their absence. In general, there was extensive discussion on the plausibility and/or evidence for the q-change approach implemented in NCAM (Cadigan, 2015). After consideration of multiple alternate hypotheses, it was agreed that the approach proposed be accepted.

Over the time series, gear has changed three times. Catchability conversions have been calculated across gear types for Northern cod, however participants discussed at length whether the stock conditions at the time of conversion reflect the present stock. Specifically, there may not have been enough cod in the large size groups at the time of Campelen conversion study. Conversion factor studies included Northern cod up to 85 cm. Survey results indicate that the stock is now experiencing an increase in the abundance of cod $\geq$ 85 cm. The return of those large fish introduces a source of uncertainty in the conversion factors for this dataset.

An unpublished review by Brodie and Walsh in 2013 provided a detailed summary of all the analysis on conversion factors and tow length (including 5, 10, 15 minute tows). They found no significant difference for catch rate of different length groups, including sizes beyond 85 cm and well into 110 cm +. At present, there seems to be little evidence to indicate a bias in the conversion factors, however meeting participants felt it prudent to recommend an update to the original conversion studies under current environmental and stock conditions.

Inshore Acoustic Biomass

As discussed above, a portion of the stock was hypothesized to overwinter in Smith Sound for a portion of the moratorium period (approximately 1995 to 2009), and was thus unavailable to the autumn DFO RV Survey. The size of the portion of the stock overwintering in Smith Sound is based on the acoustic biomass estimate and the age composition. Complementary Campelen trawls were used to provide the age compositions for the acoustic biomass estimates. Catch efficiency parameters are fixed, and incorporate the fact that there seem to have been very few juvenile cod in Smith Sound during this period; NCAM assumes that the acoustic signal records 70% of age 2 cod present, 90% of age 3 cod, and 100% of older age classes based on trawl sampling information. These constant values were used for the Campelen trawl age-composition data (available for 1995, 1998, 1999-2003). The underlying assumption is that selectivity of the acoustic signal and the sampling trawl are the same, which has not yet been confirmed.

Offshore Acoustic Biomass

Given that the 2015 offshore acoustic biomass estimate was not available at the time of this meeting, a hypothetical value of 500 kt spawning aggregation was considered in order to gauge utility and influence on the model. In this case, the acoustic estimate is considered as the lower bound of Spawning Stock Biomass (SSB) for the whole stock. If the model predicted SSB is large then the acoustic biomass estimate will not add any new information. However, if the model predicts a relatively small SSB, then adding the acoustic biomass estimate will increase the SSB estimated by the model. The sensitivity run with this hypothetical value included (modelled using a censored likelihood) was poorly fitted and the 95% confidence interval for the estimated SSB was fully below the assumed acoustic biomass value. Based on this hypothetical value, reviewers and the model author agreed that the offshore acoustic biomass could not be reasonably fit by NCAM.

Sentinel Survey

The Sentinel Survey Q is based on the dome shaped age selectivity characteristic of gillnet fisheries. Parameters are freely estimated for each age bin. However, the Sentinel Survey does not seem to be reflecting general trends in the stock over the full time period 1995-2015. The model author suggested that changes in the size of offshore stock components and migration patterns impacted survey Q. This discrepancy is addressed through the incorporation of catchability year effects, based on a Random Walk that begins at zero. This approach allows for a long term trend in Q, which the author expects to vary according to the fraction of the offshore stock available to the survey each year.

Tag Data

Available data includes roughly 238,000 Northern cod tag deployments, and over 36,000 recaptures between 1954-2015 in NAFO Divs. 2J, 3K and 3L. Tagging was not conducted every year, and deployments peaked between 1980-83. There are no discernable patterns in recovery rate by region, however at tagging depths greater than 400 m recovery is relatively low. A declining trend in the fraction of tags recovered is persistent across geographical regions and depth strata since 2007.

Of the total tag data, 130,000 T-bar tags deployed since the mid-1980s were used to estimate NCAM. All other tag types are excluded to reduce confounding variables related to tag loss and reporting rates. Age of the tagged fish was calculated based on recorded length and a relatively small sample (n = 155) of tagged fish that were caught and their otoliths returned for age determination. This is the first time this analysis has been conducted on these data and the authors recommend some review for accuracy. For example, the calculated age tables include a very small number of 40 cm fish at age 2; this is unlikely, as tagging is conducted only for fish > 45 cm, generally age 4 and older. Minor discrepancies in the age calculations may be the result of mixing commercial otolith ages with tagged fish lengths. At the youngest ages, commercial fisheries may preferentially select faster growing fish, and, additionally, the otolith data include very few individuals in the older age classes (e.g. only three otoliths at age 12). The 155 otoliths included as reference data for the NCAM age calculations were chosen because they belonged to traceable cohorts. Any additional otolith data would be subject to similar standards of comparability, or separation by gear type (i.e. otoliths should be retrieved from fish caught by the same gear type as tagged fish).

Tag Return Model

The input data for the tag return model was limited to experiments with > 70 tagged fish and age classes with > 10 individuals. The ages of tagged fish were estimated from length (described above) and age at capture was inferred based on time at liberty.

Commercial catches exert an unknown harvest rate on a population of unknown size. Tag returns share the unknown harvest rate, but the population size is known with great precision (i.e. number of tag deployments). This information can then be applied to back-calculate a harvest rate. Following tag deployments, a pulse fishery would catch a fraction of tagged individuals immediately in year one, some would be lost to natural mortality, and a fraction survives until the year two fishery. Exploitation rate is derived from the return rate from year one. Tag return in year two provides information on harvest rate and natural mortality.

The value of these tag data depend on reporting rate, which is easily estimated following the 1997 introduction of high reward tags and double tagging studies to estimate rates of tag loss. Prior to 1997, reporting rates are accounted for within the model as the proportion of tags that are potentially available to be recovered (see Myers et al. 1996). The consequence of estimating this by experiment means that the magnitude of tag-returns is not informative on mortality rates; only the rate of decline of recaptures with years-at-liberty is informative. In NCAM, tag loss is treated as another source of "natural" mortality in the tagged population. NCAM uses Kirkwood's model directly, which assumes that per capita tag loss is fish dependent, and is coded as a random effect. The author generated a non-parametric tag loss model as well but the Kirkwood model resulted in the best fit.

NCAM INITIAL FINDINGS[1]

Presented by: Noel Cadigan

The range in the 2014 estimates of total mortality rate (Z) from the various NCAM formulations 0.16-0.28, of which fishing mortality (F) was only a small portion: 0.04. Increasing catch bounds does not have much effect on the predicted exploitation rate, rather this serves only to rescale biomass. Previous research has indicated that fishing mortality was increasing at the time of the moratorium (Myers et al. 1996), however NCAM suggests that the level of F during this period was much smaller than the level of M. Based on these results, Northern cod appears to be a M-driven stock, rather than a F-driven stock as previously thought. This distinction does not diminish the importance of sustainable harvest advice, but provides critical information for stock assessment, planning, and predictions of future stock performance. If management is meant to provide a buffer against variations in natural mortality, then this will alter catch advice.

Catch multiplier projections were used to provide short-term management advice. The geometric mean recruitment of the 3 latest assessment years was used in the 5 year projections. For this presentation, NCAM considered three scenarios:

1. status quo catch;
2. 50% increase in catch, and
3. 50% decrease in catch.

At present, F is so low, that these three scenarios produce relatively similar outcomes. The author strongly recommended restricting decision-making to projections with a CV below 30% which effectively limits the time-period of reliable projections to 1-3 years.

Model Fit

Total model predicted catch varied within the catch bounds provided. The model fit the commercial age compositions very well, with some discrepancies in the young age class (years two and three). Survey fit was separated in pre- and post-moratorium. Before the moratorium, survey results are too variable year-to-year for the model to fit really well. NCAM tracks the decline of the Northern cod stock well, but does not extend that fit into the recent survey

[1] Refer to Appendix 9 for NCAM Post-meeting analyses

increases. There is a conflict between DFO RV survey and catch-at-age compositions. The increase in RV recruitment is not evident in catch-at-age compositions.

NCAM results showed very good fit to the aggregated tagging data. The model fit relatively well to the inshore acoustic survey in Smith Sound, however the addition of the tagging data (new since the model was originally published; Cadigan 2016) seems to have slightly reduced model fit to Smith Sound biomass and age compositions. Notably, the model could not fit to the hypothetical offshore acoustic biomass value.

DISCUSSION

The author, external reviewers and meeting participants agreed that NCAM displayed very good fit overall, and yielded results that could form the basis of the next assessment of this stock.

The likelihood approach to parameter estimation and weighting was discussed at length, with emphasis on potential differences between data types. Survey data was entered into the model as mean values for thousands of tows, while tagging data was entered more directly (and thus, it was entered in greater abundance). This raised the concern that a subtle misspecification of the model to the tagging data may be amplified by the sheer scale of the dataset. An external reviewer considered this question and recommended testing the tag data for internal correlation, but emphasized that the likelihood inferences are robust. It was agreed among participants that the model's ability to accurately model the tagged portion of the Northern cod population does not necessarily equate with accurately modeling of the total population. For this reason, NCAM is built to understand tagging data in clusters that do not experience the same F as the broader population. However, when F was modeled for tagged and untagged portions of the stock separately, very little difference was found. The fishery distribution is not homogenous, but F did not appear to vary significantly between tagging experiments or between tagged and untagged cod.

The precision of the model predictions were briefly discussed. Despite high model flexibility and built-in process error, NCAM delivered a relatively precise SSB estimate (CV of 10%). Participants and author agreed that this should be investigated; other parameters yielded much higher CVs (e.g. the CVs on F and M were 30-40%).

On a more theoretical level, it was suggested that characterization of past stock performance and projection of future conditions are two distinct problems that may not be best answered by a single model. For instance, there are additional processes and management strategies, like risk aversion, that influence how assessments are projected and applied to future scenarios. This was carefully considered by meeting participants, however discussion returned to the view that the only measurable success for a model is its ability to reproduce observations. All other criteria, regardless of perceived value, are limited to some level of speculation. The utility of simulations was also briefly debated. It was suggested that a stock assessment would expect simulation testing of the operating model. In response, one reviewer argued that simulation testing only serves to identify a model that successfully predicts a simulation, not necessarily a model that successfully predicts reality. From this, many fundamental questions were drawn: How do we estimate parameters and which data do we want to consider? It appears that much of the available data is very informative for the model. Depending on how the model is structured, the result may be a model that estimates the parameters for which we have more data very well, whether or not those particular parameters are most important to managers. For example, much emphasis is placed on predicting catch, despite findings that suggest this stock is driven by natural mortality. The study and inclusion of ecosystem processes that may increase understanding and prediction of M may be more useful than additional data related to catch. Many participants were concerned that although a highly flexible model fits any data well, it may lack predictive power. These challenges remain central to model formulation, are not issues restricted to modelling of Northern cod dynamics and their resolution was far beyond the scope of this meeting.

A recommendation was made for future NCAM iterations to expand the time series earlier than 1983, to better capture the stock history. Participants also suggested that more process error may be included to compensate for migration, an important consideration for Northern cod, where stock fidelity is not well known.

STATISTICAL CATCH AT AGE MODEL (SCA)

The Statistical Catch at Age model (SCA), developed and presented by Sean Cox, represents a more ‘conventional approach’ to stock assessment, based on total catch, proportion at age, DFO RV Survey data and biomass indices.

INTRODUCTION

This model is similar to the model applied to assessments of Atlantic Cod in the Gulf of St. Lawrence, Atlantic Halibut (fit to proportions at length instead of proportion at age), and several Pacific and US stocks, including Pacific Hake, Halibut and Sablefish. Unlike NCAM, process errors are counted as model parameters and they are treated as fixed effects. In total the model included roughly 300 parameters, however it is driven by the leading 4-6.

Natural mortality (M) is modeled as a Random Walk in age blocks. All three age-class priors for M are set to 0.2 with a CV of 25%. The SCA model is sensitive to the priors on initial M, so these values are important. Over time, M changes according to the Random Walk: the previous M plus the normal mean zero deviation. Fishing mortality (F) is separated into age effects and time effect. The values of F per age class are determined by first predicting the catch (according to the Baranov catch equation) and then the total catch over all ages:

F= [selectivity at age] + [fully selected fishing mortality over time]

This calculation assumes that observed catch is normally distributed around predicted catch (with 10% error). The model for recruitment is unique for this stock. In most assessments recruitment requires some prior on the parameter because the stock is not depleted to the point where the model curve can draw a prediction. However, the Northern cod stock is depleted to the point (< 20%) that the model for recruitment performs well.

The weight at age and mature at age parameters are drawn from empirical values reported in the literature, however this assumes that weight at age is consistent pre- and post- stock collapse. Like F, the DFO RV survey is also divided into time and age effects, with CV set at 0.2. The model assumes that three quarters of the year passes before the survey takes place.

SCA INITIAL FINDINGS[2]

Presented by Sean Cox

SCA model results indicate the stock declined significantly before the moratorium. This is attributed to a peak in M, although it is possible that a portion of M is counting unreported catch. Like NCAM, in recent years the model results are below RV Survey stock estimates. This lag could be the model accounting for unreported catch in M, thereby maintaining a slightly inflated M that does not allow the model to track recent gains.

[2] Refer to Appendix 10 for SCA Post-meeting analyses

In general, SCA achieves a noisy, yet reasonable fit to the RV Survey age compositions (CV of 0.53) with some variability in year-to-year fit. This could be the result of an aging error or variability in sampling. Specifically, the SCA model consistently overestimates the RV results for the age 4 class. The fit is even noisier for the Commercial age compositions (CV of 0.62) and consistently overestimates the age 5 class, although this trend disappears in recent year predictions. The age 3 recruitment estimate performs very well (CV of 0.15).

SCA results indicate that productivity was stable following the moratorium, and increased in the early 2000s. SCA estimates unfished spawning stock biomass of Northern cod at about 10 million tonnes, with a large confidence interval. The limit reference point currently adopted for the stock (average SSB from 1983-1989) was estimated at 967,000 t which places BLIM at 9.6% of the unfished biomass. The current biomass estimate of 213,000 t (23% of BLIM) is broadly consistent with NCAM findings. Preliminary retrospective analyses identified no trends.

The SCA model assumes that catch is measured with some error, and that the reported catch is unbiased. However, under some circumstances one or neither of these assumptions may be true, and may lead to unreported catch being incorrectly accounted for as natural mortality. SCA further assumes process error and observation error variance are known (these are treated as fixed values in the model). Finally, the model ignores temporal correlation in observation errors, which may arise from changes in stock availability to the RV Survey and the commercial fishery.

DISCUSSION

In general, participants agreed that the SCA model provides useful information for an assessment, and displayed good or reasonable fits to the observed data (with noted some variation between datasets, discussed above).

Although the purpose of the meeting was to review both models, without necessarily comparing the two directly, this was a natural progression of the discussion along this line. NCAM represents a complex attempt to integrate ecological features for a specific stock, while SCA aims fit the data in the most parsimonious way. Participants found it very useful to review the two models beside each other (Table 2). NCAM incorporates as much data as possible and its flexibility allows for all the hypotheses we have about this stock to be represented (e.g. movement of fish in Smith Sound). The complicated NCAM model benefits from comparison with a standard stock assessment model and vice versa.

Both models employ the same age composition data, and authors attributed the difference in performance and fit to NCAM's use of tagging data. There may also be some discrepancy in how SCA treats catch at age. SCA predicts higher F values but low M by age 6, indicating that the fishery is concentrated on the older fish or that there are fewer older fish in the population. As a result, losses are entirely attributed to F for the older fish, but losses of younger fish are attributed to M. The unobserved catch may be unreported catches of fish and/or discards, perhaps due to high-grading. In addition, NCAM moves a portion of the stock out of the survey and into Smith Sound; SCA can only cope with a decline in survey by increasing mortality. Accounting for these differences, the underlying pattern of M appears somewhat consistent between the models.

This line of thinking resulted in a detailed discussion about the Smith Sound aggregations. If the Smith Sound acoustic biomass surveys are accepted, then the abundance of cod in that area grew to a peak over time (i.e. this does not represent a single large scale migration event). From this statement, three hypotheses were considered:

1. a portion of the stock moved inshore in the winter of 1995, staying there and building until 2009 when a mass emigration occurred. However, fish only seem to appear in Smith Sound at age 4, which suggests that;
2. a group of Northern cod moves into the Smith Sound area in 1995, and age 4 fish from offshore continued to migrate into the sound in the following years or;
3. the overwintering Smith Sound cod represent an inshore population that was always present and did particularly well following the moratorium.

In all hypotheses, the cod move offshore again around 2009, when they are no longer observed by the inshore survey, despite a lack of evidence for high catch or high natural mortality. Hypothesis 2 appears to fit well to the inshore and offshore data. Abundance of individuals over age 5 declined sharply in Divs. 2J and 3K at the same time as the appearance of Northern cod was documented in Smith Sound in 1995. This connection is somewhat confounded by the gear change that also occurred at this time; the DFO RV Surveys converted to the Campelen trawl, which catches fish of the small size classes more efficiently than previous gear type (Engel trawl).

Table 2. Summary of comparison between NCAM and SCA models.

Model	NCAM	SCA
Time Period	1983-2014	1959-2014
Data included	• DFO RV Survey-at-ages 2-14, 1983-2014 • Catch Proportions at age, 1983-2014 ages 2-14 • Total Catch (as lower bound of annual removals) • Tag release / recapture data from 1983-2014 • Smith Sound Acoustic Biomass index, 1995-2009 • Age composition from SS Acoustic index • Sentinel GN Catch rate at age, 1995-2014 • '*Exploratory run' which included a hypothetical offshore acoustic biomass estimate of 500 Kt for 2015*	• Total Survey Abundance • Survey Proportions at age • Total Catch (*one run as unbiased; second run considering annual estimates of misreporting*) • Catch Proportions at age
Model Class	Integrated Assessment Model	Statistical Catch at Age
Model Structure	• Recruitment as two independent random values (pre vs post collapse) • Process error in M (auto-correlated in ages & years) • Auto-correlated random walks in F • Separable q for Sentinel GN • Allows an age/ time-varying survey catchability component to 'move' between survey area and Smith Sound over 1995-2009 (AR1) • Random Deviations on tagging F in year of release	• Beverton-Holt SR model Recruitment Deviations • Three M random walks (age groups 1-3/4&5/6+) correlated across year • Age separable F w/ asymptotic selectivity • Survey Index w/ asymptotic age-selectivity
Projections	Yes	Not yet, could do projection: F or catch based or closed loop via link to mseR [for longer term projections involving harvest control rules]

The extent of the modeled time-series provided another point of contrast between the two models. NCAM begins in 1983, while SCA extends back to 1959. The consideration of catch reporting accuracy changes over time. References to possible under-reporting of catch refer mainly to the 1990s onward. Before that, reported catch is derived from NAFO statistics without any supporting information or literature to suggest the catches were under-reported, or to what scale. It was briefly suggested that the SCA model be restricted to the data from 1983 onward to facilitate comparison to NCAM, however it was quickly noted that the purpose of this meeting was not to adjust one model to match the other, but to judge each model for its utility to a

Northern cod stock assessment. In particular, several participants were seriously concerned about restricting any model to the post 1980 time frame. Many agreed that it is not prudent to limit assessment to recent years while ignoring the history of the stock pre-decline. This creates a new baseline that ignores full stock capacity, particularly the much higher stock levels of the 1960s.

The following recommendations were made for the SCA model:

1. Adjust the gillnet selectivity function from flat-topped to dome shaped, which may greatly effect catch estimates post-moratorium.
2. Consider catch parameter as a lower boundary on an estimate and include recreational catch as much as possible.

Following this discussion, additional runs were conducted to address suggested edits to the SCA model. To examine variation in M, it was held as a constant pre-1983, based around 0.5 on ages 1-3, 0.3 on age 4, 0.2 for 5 and 6 (follows NCAM M rates). The change in M yields a very different fit; in particular, the updated model has much better fit to observed values in the late 1980s. Additionally, M spikes much higher (to 3.5 instead of 1.2). The final spawning biomass is about the same in the original and follow up model runs, but the changed M reduces residuals around the time of the moratorium. However, changes to M did not alter the behaviour of the model pre-moratorium; large residuals are still predicted at age 4. The large M at age 4 appears to be masking something else, potentially the inshore/offshore movement of cod (see Appendix 10).

BIO-ENERGETICS-ALLOMETRIC MODEL

Relative Contributions of Seals, Fisheries and Food Availability on the Lack of Recovery and Dynamics of the Northern Cod Stock

Presented by: Alejandro Buren

This model seeks to identify patterns of interaction between Northern cod and the surrounding ecosystem in order to assess the relative contributions of fisheries removals, predation by harp seals, and food availability (as indexed by capelin) on the lack of recovery and dynamics of the Northern cod stock. It can help in better understanding varying natural mortality rates. It is a structurally simple model meant to complement the stock assessment models presented at this meeting. At present, this model is not able to generate projections into the future.

The bio-energetic allometric model considered acoustic capelin biomass estimates, DFO RV Survey Northern cod biomass estimates, total fisheries catch, and consumption of cod by harp seals (based on three harp seal diet composition scenarios). The model was fit to Northern cod biomass estimates over the time period 1985-2007. Since capelin are not surveyed every year, some interpolation was conducted to fill gaps in the dataset. These analyses indicate that cod is a small proportion of the harp seals diet. Biomass dynamics were best explained by a combination of fisheries removals and capelin availability, whereas seal consumption was found not to be an important driver of the Northern cod stock. This work demonstrates a mechanistic link between cod and capelin that may illuminate natural mortality for future stock assessments. Cod stomach contents data are available for a limited number of years, however available data indicate that capelin are a high fraction of cod diet. Nutrient analysis also indicates that capelin are a very high quality food, they are a significantly more efficient meal for cod than alternative forage prey, like shrimp. The link between capelin and cod may be a function of condition. When few capelin are present, the cod have to turn to sub-optimal prey leading to poor or skip spawning events.

The information presented by in the bio-energetic model was of great interest to meeting participants. Consensus was quickly reached that ecological considerations are an important addition to any stock assessment. However, it was not immediately clear how the results presented in this study could be incorporated into one of the previously reviewed stock assessment models. Capelin biomass data is inconsistently collected, but if modeled, it may provide an additional input to NCAM or SCA. Participants also suggested using body condition directly; Spring condition is recorded, and may eliminate the need to treat capelin as a kind of condition proxy. Reviewers cautioned that if the capelin model cannot be restructured to make projections, this information will not be applicable to the SCA, NCAM or similar assessment models. Model authors accepted this feedback with plans to update the work.

MSY REFERENCE POINTS FOR NORTHERN COD

Stochastic MSY Reference Points for Northern Cod, including Variable Natural Mortality Rates and a Spatial Stock-Recruit Model

Presented by: Noel Cadigan

Development of reliable catch advice is the main goal of the stock assessment process, yet current catch levels remain uncertain. As a solution, models are used to estimate catch and generate catch multiplier projections. These models are related to very specific regulations to manage target species so that they move out of a predefined critical zone of low abundance. Although some fishing is still allowed in critical zone to allow flexibility to socioeconomic needs, the emphasis of Maximum Sustainable Yield (MSY) management is to rebuild the stock. The term Maximum Sustainable Yield has become synonymous with fisheries management, yet it remains poorly defined. “Maximum” is misleading; the real goal is constrained maximization. The stock should be optimized within some limits that do not risk damage to itself or other species. “Yield” may refer to biomass, profit, social benefit, or some combination thereof.

The calculation of meaningful MSY reference points requires long term projections with accurate growth and maturation rates. The stochastic M process from NCAM may be applied, along with growth, maturation and some ecosystem considerations to account for variable M. In summary, the recruit per spawner plot for Northern cod is predicted to decline until reaching a biomass 500 kt threshold, after which recruit per spawner increases. The area occupied by the stock is simultaneously expected to expand. A spatial Beverton Holt model was paired with a model of area occupied vs SSB increase, demonstrating distinctly non Beverton Holt patterns in the stock recruitment. Specifically, this model hybrid shows a non-monotone response in recruit per spawner. The resulting long-term forecast for Northern cod suggests that the next 30 years will be highly variable,. The common feature for this stock is highly variable M, both past and future.

An immediate limitation to this work is that the dome shaped fisheries selection model will not last; as the stock increases, gillnets will be replaced by trawlers. Furthermore, while a 200,000 t catch may be sustainable in the extended term, if managers choose to control harvest at a constant F, it must be accepted that fishery success (and thus, profits) will vary year-to-year. If the aim of managers is to control harvest to buffer against natural variability, it will require a different strategy.

Further research may be pursued to extend this model back to 1959 (as with NCAM), to develop a spatial meta-population model, and to examine possible shifts in maturity/weight as the stock rebuilds.

An interesting feature that was noted is that by modeling a single species M, this work reflects similar results as multi-species MSYs. Questions were also raised about strategies to reduce the risk of overfishing – the MSY work presented here is framed purely by yield optimization. If the focus is shifted to minimizing risk (down-side management), substantial change to the research will be required. It was suggested that the most useful question to managers is not for a clear definition of MSY, but to explain the purpose of the number that would be generated, so that yield, risk management or some combination may define more relevant reference points.

Participants suggested that further research may also include a study of density dependent stock performance to identify threshold values.

MANAGEMENT STRATEGY EVALUATION

Presented by: Sean Cox

Stock assessment models today are meant to estimate both historical and more recent stock status however, the key parameters are not constant. For example, growth rates vary, making stock status very difficult to forecast. If we acknowledge the limitations of forecasting, researchers and managers need to have a good plan for reacting to changes in the stock. If we have another sharp natural mortality event, are we ready? Will we be able to catch it in time? If the sole management objective is to maximize average yield, scientific research of the stock can lead all decision making. However, other goals like social and economic stability of the fishery rely on a different management structure. Management Strategy Evaluation (MSE) employs simulations to compare datasets, analysis methods and management efforts with the goal of identifying the best performing management strategy.

Many generic reference points are suggested in the literature (e.g. 20% of B0), however there are no generic fisheries and yield rarely resembles the expected curve. In the 1990s when the Precautionary Approach (PA) was introduced, B_{LIM} was created as a threshold for avoidance of stock collapse. However, for this goal harvest should be restricted before the stock reaches B_{LIM}. Furthermore, if managers want to avoid conditions wherein fisheries are opened and closed from year to year, fishing effort should be reduced above B_{LIM} and before B_{MSY}.

Resilient management will consider a broader range of risks (high natural mortality events, or high variability in natural mortality); the idea is not meant to predict the future, but to assist decision making. Several different management strategies were presented, comparing and contrasting management strategies across Canadian stocks. For example, for Pacific herring stocks, it has been suggested that harvest be limited to half of F_{MSY} as a conservative reference point, modeling results showed that the current DFO method was more robust due to large variation in natural mortality. For Atlantic halibut stocks, Harvest Control Rules (HCRs) such as discounted release of larger fish or size limits are modeled with reasonable success. Atlantic halibut could be considered as a template for how to proceed with Cod assessment and HCRs.

DISCUSSION

Participants noted that the discussion of management strategy evaluation is not limited to this meeting; managers and stakeholders will contribute crucial input to this topic in the lead up to and during the upcoming Northern cod stock assessment. The NCAM and SCA models are amenable to an MSE, and the author suggested that this could be conducted with the participation of stock assessment researchers, managers and stakeholders.

The main challenge is the high uncertainty regarding catch; identifying a sustainable total allowable catch (TAC) will be a challenge given lack of estimates and lack of control on the recreational fishery. Several operating models could be constructed offering different scenarios on historical catch, to compare their effects on the results. This suggestion was well received by meeting participants. Although the participants did not feel that direct catch advice would be required from an MSE immediately, it may be advantageous to begin MSE analyses in advance simply to improve understanding of both the stock and the method. However, any decision to proceed with a Northern cod MSE would require a decision process beyond the scope of the meeting. Beyond the biological and ecological performance of the stock, management strategy development would require discussion and clarification of social and economic goals for the fishery.

It was noted that a significant amount of work would be required to generate the MSE operating models, and that this measure is neither appropriate or practical to conduct a full MSE within the timeline of the next assessment in March 2016. However, there was agreement that MSE simulations are both possible and relevant in the context of the NCAM and SCA models presented here. Discussion of a Northern cod MSE may continue in the upcoming assessment and may be of particular interest to Northern cod recovery working groups.

CONCLUSIONS AND RECOMMENDATIONS

Full reviewer reports can be found in Appendices 5-8. The following summarizes reviewer presentations and discussion as the Framework Meeting concluded.

The Northern cod stock presents a difficult assessment challenge for several reasons, the stock ranges over a vast area, exhibits complex stock structure and undertakes large-scale seasonal migrations, a high proportion of catch is unreported and there appear to be significant changes in productivity and spatial distribution over time.

REVIEW OF DATA SOURCES

In general, reviewers agreed that a detailed review of each dataset was difficult to achieve, based solely on the data summaries. Insufficient information was provided to fully evaluate the utility of many of these data sources in the verbal presentations.

General concerns that reviewers raised across data sources included:

1. consistency of methods over time (or ability to account for changes in model);
2. calculation of uncertainty estimates (not presented for every dataset);
3. spatial resolution and extent of the datasets (spatial context would help with the consideration of sources of variance in the model).

Several of the data sources are subject to changes in methods, most notably the gear changes in the DFO RV Survey, however perceived inconsistencies in the acoustic biomass survey were also discussed at length. It is crucial that methodological changes are examined and well understood to assure comparability of data over time.

DFO RV Survey

Reviewers highlighted a number of limitations, including missed strata, variation of survey timing change in availability of the stock in the mid-1990s to mid-2000s. However, the survey appears to track cohorts well and it represents a key data input for any stock assessment model.

In particular, reviewers indicated that change in stock availability to the survey is worth investigation. There is some uncertainty regarding the previous comparative fishing experiments done in the 1990s, as large cod may not have been readily available. As cod rebuild, selectivity assumptions of survey gear should be investigated experimentally.

DFO Tag Data

The Floy tagging studies were designed mostly to estimate fishery exploitation rates issue. The high reward tagging was used to estimate tag reporting rates among four regions, but in this assessment focused on tag reporting rates in one region, namely the inshore of Divs. 2J3KL. Tagging data can also provide valuable spatial information. Analysis of data from satellite tags, rather than Floy tags, is another interesting approach and a good opportunity to introduce students and researchers to non-standard modelling for assessment.

The meeting recommended further study of tag reporting rate to determine the extent of a declining trend in tag reporting and, if necessary, to develop a solution (e.g. increase value of high and low reward tags). Tag reward values ($10 and $100) have not been increased since 1997 and increasing the reward values may encourage harvesters to return tags.

Acoustic Biomass Survey

It is strongly recommended by the reviewers and participants of this meeting that acoustic biomass data be considered at an independent peer-review meeting, including recognized assessment and acoustics experts. The aim of an acoustic biomass peer review process would be to provide advice on the use of offshore acoustic biomass survey data in a stock assessment model for Northern cod. However, the conclusions of this meeting will also be relevant to the use of Smith Sound acoustic biomass data, which is currently considered in the NCAM model presented at this meeting.

Sentinel Survey

The Sentinel Survey is inconsistent with the DFO RV Survey in mid to late 1990s. This discrepancy may be related to changes in availability and/or design of the sentinel survey in terms of number of sites sampled and the duration of coverage by the survey. Thus the Sentinel Survey may not provide a consistent index for a stock assessment model. NCAM addressed this by discounting the inter-annual trend, however it is still possible that the age compositions are not representative.

Commercial Catch At Age Data

The commercial catch data are treated as standard data sources in most assessments, however it cannot be ignored that these data are limited by unreported landings, discarding/high-grading, and unrecorded recreational catch. A significant portion of total catch is unseen (estimated 40% of total catch is recreational harvest since 2006), and there is no evidence that age compositions are similar to the reported commercial catch. This should be investigated, and may require both management and research efforts to examine and report recreational catch.

Commercial Logbook Data

Reviewers further cautioned that CPUE from harvester logbooks is not a measure of population abundance and should not be applied to models as such.

REVIEW OF PRESENTED MODELS

Two models were reviewed: the NCAM and SCA model. NCAM was custom built for Northern cod and with inputs from a wide number of data sources. As presented, the meeting participants are confident that NCAM could be applied to the 2016 Northern cod stock assessment. The SCA model would also be acceptable for use in this assessment if concerns with the residual patterns were addressed. A bioenergetics model was also presented that provides a possible explanation for trends in natural mortality estimated in NCAM and SCA. A bio-energetics model was presented to provide insight into variation of M in the Northern cod stock.

NCAM Review

NCAM is based on the state-space model framework, is specifically tailored to this stock, and facilitates examination of ecological hypotheses. The flexibility of NCAM allows a lot of room to add other data (e.g. fish harvester questionnaire). Reviewers noted that NCAM achieves a good fit for a diverse array of data sources, while maintaining largely objective self-weighting. In particular, NCAM is notable for the inclusion of tagging data which provide information on F and M, and facilitates distinction between the two.

Reviewers agreed that sensitivity analyses should be presented for NCAM assessment during the March 2016 meeting; some preliminary sensitivity work was presented and appeared to be very robust. Retrospective analyses should also be investigated to test model performance. Retrospectives and sensitivity analyses would be a valuable exercise – models can be run very quickly now, and the definition of model robustness has been transformed.

The historical extent of the model was discussed by several reviewers. It may be useful to investigate extending NCAM further prior to 1983, to include earlier stock capacity. Conversely, one reviewer questioned whether the conditions of the stock pre-collapse and during the moratorium accurately reflect current rebuilding. If this is the case, due to changes in growth, recruitment and/or survey availability, it may be more prudent to model a subset of time-series. Spatial-temporal analysis may be useful here too, particularly because this area is more data-rich.

NCAM suggests extreme variability in M, however there was little support provided beyond better model fit. One reviewer suggested further work to identify mechanisms and elaborate on consequences for management. Specifically, NCAM results indicate that this stock is M driven, and that the collapse is closely linked to increases in M. However, modeled M has a number of other spikes over time, without full explanation. This has important implications for the projections – if M is highly variable, and prone to increase rapidly without warning, that may radically change the catch advice.

Reviewers agreed that the NCAM model would be a reliable tool for assessing status of 2J3KL Northern cod. Although discussion during this meeting indicated that projections should be limited to 3 years, one reviewer suggests that it may be appropriate for 5 year projections after some modification to the way that M is modelled (i.e. set average M at a value near the long-term average of estimated M).

SCA Review

The SCA Model represents a standard easy-to-use approach focused on a few of the most important data sources, and includes stock-recruitment relationship which is important as the stock rebuilds. Limited flexibility in the model means that flexible components (e.g. M) will vary to account for other features, potentially making the model vulnerable to misspecification. The original model assumes catch estimates are unbiased; a follow up for this meeting was reformulated to consider reported catch as a minimum estimate (see Appendix 10). Further investigation on the strength of these assumptions is important to understanding how to apply the model to a stock assessment. Unlike NCAM, which moves a portion of the offshore stock into Smith Sound, SCA attributes non-stationarity in the 1995-2007 period to elevated M, though it could also be attributed to a change in availability of older age groups to the survey.

Although some problematic residual patterns were identified in the SCA results, particularly in the 4 and 5 age classes, this model delivers broadly similar patters as NCAM. Reported residuals should be investigated. As with NCAM, reviewers agreed that retrospective and sensitivity analyses (e.g. sensitivity to fixed selectivity time blocks) are important next steps for this model. One reviewer recommended that issues with fit to the age-composition data be resolved before applying SCA to stock assessment.

Review of Bio-Energetic-Allometric Model

The bio-energetic model presents promising work that may be included as an index, or proxy of M. The iteration of the model presented here requires some updates to the current year before it is applied to a stock assessment. Reviewers wondered whether previous capelin declines have influenced cod abundance in the same ways suggested by this model. Regime shifts were detected in the 1920s and in 1991 that include climate, oceanographic and biological changes. Changes at the time of the collapse are not limited to cod; it was an ecosystem shift. Prior capelin collapse may not have influenced cod in the same way, due to differences in fishing pressure and other environmental factors.

COMMENTS ON THE FRAMEWORK REVIEW PROCESS

Reviewers noted that Northern cod represent a very interesting assessment situation and they were glad to be involved. Organizers were commended for including an excellent range of backgrounds and experience on the review panel and at the meeting, which provided stimulating discussions covering a wide range of topics and hypotheses. It was noted that the more timely provision of relevant documents and information on the datasets (e.g. presentation of example plots, indication of whether cohorts can be tracked) would have facilitated a better review. Future meetings would also benefit from broader presentations, for example better understanding of ecosystem function would certainly help in interpreting model output better.

Reviewers also suggested that future meetings begin with a brief presentation on what has been done in the past and what aspects of stock assessment could be most improved. A strategic discussion of the primary goals of the stock assessment model would also help frame the review and guide discussion.

REFERENCES CITED

Baranov, F. I. 1945. On the Question of the Biological Basis of Fisheries: On the Question of the Dynamics of the Fishing Industry. Indiana University.

Cadigan, N. 2013. Update of SURBA+ for 2J3KL cod. Can. Sci. Advis. Sec. Res. Doc. 2013/054. P. 1-19.

Cadigan, N.G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 10.1139/cjfas-2015-0047.

DFO. 2007. Stock Assessment of Northern (2J3KL) cod in 2007. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2007/018.

DFO. 2009. Stock Assessment of Northern (2J3KL) cod in 2009. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2009/009.

DFO. 2013. Stock Assessment of Northern (2J3KL) Cod in 2013. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2013/014.

Francis, C. 2014. Movement of Atlantic cod around the British Isles: implications for finer scale stock management. Journal of Applied Ecology. 51(6):1564.

Gavaris, S., and C. A. Gavaris. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Can. Spec. Pub. Fish. Aquat. Sci. 66: 178-182.

Kirkwood, G.P., 1981. Generalized models for the estimation of rates of tag shedding by southern bluefin tuna (*Thunnus maccoyii).* Journal du Conseil. 39(3): 256-260.

Lilly, G.R., 2008. The decline, recovery, and collapse of Atlantic cod (*Gadus morhua*) off Labrador and eastern Newfoundland. *In* Resiliency of gadid stocks to fishing and climate change. Edited by G. Kruse, K. Drinkwater, JN Ianelli, JS Link, DL Stram, V. Wespestad, and D. Woodby. Alaska Sea Grant Program. University of Alaska, Fairbanks, Alaska. pp.67-88.

Myers, R.A., Barrowman, N.J., Hoenig, J.M. and Qu, Z. 1996. The collapse of cod in Eastern Canada: the evidence from tagging data. ICES Journal of Marine Science: Journal du Conseil. 53(3): 629-640.

Román, E., C. González-Iglesias and D. González-Troncoso. 2015a. Results for the Spanish survey in the NAFO Regulatory area of Division 3L for the period 2003-2014. NAFO SCR Doc. 15/019.

Román, E., A. Armesto and D. González-Troncoso. 2015b. Results for the Atlantic cod, roughhead grenadier, redfish, thorny skate and black dogfish of the Spanish survey in NAFO Div. 3L for the period 2003-2014. NAFO SCR Doc. 15/020.

Rose, G. A. 2003. Monitoring coastal northern cod: towards an optimal survey of Smith Sound, Newfoundland. ICES Journal of Marine Science: Journal du Conseil. 60(3): 453-462.

Rose, G. A, and Rowe, S. 2015. Northern cod comeback. Canadian Journal of Fisheries and Aquatic Sciences. 72(12): 1789-179

Stansbury, D. E. 1997. Conversion factors for cod from comparative fishing trials for Engel 145 otter trawl and the Campelen 1800 shrimp trawl used on research vessels. NAFO SCR Doc. 97, 73.

APPENDIX 1: TERMS OF REFERENCE

NORTHERN COD FRAMEWORK REVIEW MEETING

Regional Peer Review – Newfoundland and Labrador Region

November 30 - December 4, 2015
St. John's, NL

Chairperson: Don Power

Context

- The framework review meeting will focus on methodology for estimating the population size of the Northern Cod (NAFO Divs. 2J3KL) stock.
- The most recent (2013) stock assessment utilized a survey-based model (SURBA) to determine stock status. This method relies only on DFO survey data, generating absolute estimates of total mortality rates (Z) and relative estimates of stock size. A drawback to this approach is that it cannot directly provide catch-based management advice. A Framework Review meeting on Northern Cod will review multiple models of population dynamics, and also discuss the utility of various data sets available for assessing this stock. For example, Cadigan (2015) describes a state-space population dynamics model which estimates stock size of Northern Cod by integrating input data from several sources (survey data, mark-recapture data and commercial fisheries data).
- The goal of the Framework will be to determine which data sets and models will be considered at the March 2016 Regional Peer Review (RPR) as a basis for providing management advice. No provision of management advice will occur during the framework meeting, rather management advice will be provided through the 2016 RPR.

Objectives

The Framework meeting will address the following:

1. Evaluate the utility of data sources that may provide information on the stock status of Northern Cod. If possible, examine the uncertainty in the data and the methods used to estimate measures of uncertainty.

In addition, for each model considered:

2. Is the structure of the model and the estimation techniques employed sufficient (i.e. robust and reliable) to estimate the stock size of Northern Cod?
3. Are proposed projection methods for evaluating future catch options sufficiently robust and reliable? Are any additions or modifications to methods or inputs required before providing advice to Fisheries Managers?
4. Recommend priority short- and medium-term research areas to improve data sources, assessment model formulation and estimation, and projection methods.
5. If reference points have not been estimated, provide direction for future work towards estimating Precautionary Approach reference points (both biomass and removal references).
6. Acoustic surveys have recently been conducted by the Centre for Fisheries Ecosystems Research at Marine Institute of Memorial University. (Data from these surveys are not

currently available.) Consider extensions to the model structure that can incorporate these data as an index of stock size.

7. If time permits, discuss the extent to which the method (and underlying computer software) is amenable to conducting Management Strategy Evaluation (MSE).

In order to address the Terms of Reference and enable across-method comparison for each method considered, the following results should be available for review two weeks prior to the beginning of the meeting:

- Measures of Fit, including graphical display of residuals,
- All parameter estimates and/or random effects, with measures of uncertainty,
- Estimates of Total Abundance, Recruitment (at age 3), Biomass (Ages 3 and older), Mature Biomass, Fishing Mortality (averaged over both ages 4-6 and ages 7-9), Natural Mortality (if applicable), and Catchability Parameters, and
- Retrospective analysis (five years or greater).

External experts will participate and provide independent peer-review of the Framework Review. Identified reviewers will be required to provide a reviewer report on the final day of the meeting.

Expected Publications

- Proceedings (which shall include reports from the reviewers)

Participation[3]

- Fisheries and Oceans Canada (DFO) - Science, NL Region
- Fisheries and Oceans Canada (DFO) - Fisheries Management and Policy & Economics, NL Region[4]
- Memorial University of Newfoundland and Labrador
- Fisheries Scientist, Fish Food and Allied Workers Union
- Other invited Regional, National and International Scientific Experts, including Academia
- Invited reviewers:

  Dr. Joanna Mills-Flemming, Associate Professor of Statistics, Dalhousie University (Halifax, Canada)

  Dr. Anders Nielsen, Senior Research Scientist, DTU Aqua (Copenhagen, Denmark)

  Dr. Chris Legault, Supervisory Research Fishery Biologist, NMFS (Woods Hole, USA)

  Dr. Doug Swain, Research Scientist, Fisheries and Oceans Canada (Moncton, Canada)

[3] Considering the technical nature of the review, participation will be limited to individuals with the necessary scientific expertise to provide review of population models and/or related datasets. Consistent with the participation guidelines for all Canadian Science Advisory Secretariat (CSAS) processes, attendance is by invitation only. Questions on participation can be directed to: Dale Richards, Coordinator, Centre for Science Advice, NL Region.

[4] Representatives of Fisheries Management and the Policy & Economics Branches of DFO will be asked to participate as needed, to provide specific presentations and/or answer questions of clarification.

References

Cadigan, N.G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 10.1139/cjfas-2015-0047.

APPENDIX 2: AGENDA

Northern Cod Framework Meeting – Draft Agenda

Memorial Room - Northwest Atlantic Fisheries Centre - St. John's, NL

November 30 – December 4, 2015

Chairperson: Don Power (DFO Science)

November 30, 2015

Time	Activity	Presenter
0900	• Opening remarks, Welcome from DFO, Round of introductions	Chair + Ben Davis
N/A	• Review of Data Sources: Questionnaires (Q1-Q4) ~15 mins/source)	Various
0920	• Commercial Catch • Commercial Catch-at-age • Recreational Fishery Removals • DFO RV Surveys	Brian Healey/FAM Danny Ings FAM/CP Brian Healey
1045	• Offshore Acoustic Survey • Smith Sound Acoustic Survey • DFO Tagging/Telemetry • Sentinel Survey • Inshore Mobile Gear Survey • Harvester Logbooks	George Rose Sherrylynn Rowe Rose/Rowe Brian Healey Brian Healey Don Power Brian Healey
1315	• Inshore Harvester Questionnaire • Beach Seine Surveys • Satellite Tagging Studies • Spanish RV Survey – 3L outside EEZ	Erin Carruthers Bob Gregory Sherrylynn Rowe Don Power
N/A	• Review of Data Sources: Questionnaires (Q5-Q8) ~15 mins/source)	Various
1415	• Commercial Catch • Commercial Catch-at-age • Recreational Fishery Removals • DFO RV Surveys	Brian Healey/FAM Danny Ings FAM/CP Brian Healey
1530	• Offshore Acoustic Survey • Smith Sound Acoustic Survey • DFO Tagging/Telemetry • Sentinel Survey • Inshore Mobile Gear Survey • Harvester Logbooks • Inshore Harvester Questionnaire • Beach Seine Surveys • Satellite Tagging Studies • Spanish RV Survey – 3L outside EEZ	George Rose Sherrylynn Rowe Rose/Rowe Brian Healey Brian Healey Don Power Brian Healey Erin Carruthers Bob Gregory Sherrylynn Rowe Don Power

December 1, 2015

Time	Activity	Presenter
0900	• Modelling: NCAM + Discussions	Noel Cadigan
1045	• Modelling: NCAM + Discussions	Noel Cadigan
1330	• Continued discussion on NCAM or any issues to date	Various

December 2, 2015

Time	Activity	Presenter
0900	• Modelling: Statistical Catch-Age Model	Sean Cox
1045	• Modelling: Statistical Catch-Age Model	Sean Cox
1330	• Continued discussion on S C-A model or any issues to date	Various

December 3, 2015

Time	Activity	Presenter
0900	• Discussions on Way Forward - Framework Decisions • Reviewers comments to date	Various
1045	• Continued discussion on any issues to date	Various
1330	• Bioenergetics−allometric model: relative contributions of seals, fisheries and food availability on the lack of recovery and dynamics of the northern cod stock	Alejandro Buren
1445	• Review of Terms of Reference • Continued discussion on any issues to date	Various

December 4, 2015

Time	Activity	Presenter
0900	• Reviewers Report - Highlights • Recommendations - Group	Various
1045	• Reviewers Report – Highlights • Recommendations - Group	Various
1330	• Continued discussion on any issues to date	Various

APPENDIX 3: PARTICIPANT LIST

Name	Affiliation
Don Power	DFO Science-NL Region
James Meade	DFO Science-NL Region
Emilie Novaczek	MUN
Anders Nielsen	DTU-Aqua
Chris Legault	NMFS
Steve Walsh	DFO Science-NL Region
Kris Vascotto	GEAC
George Lilly	DFO Science-NL Region
Geoff Evans	DFO Science-NL Region
Ben Davis	DFO Science-NL Region
Robin Anderson	DFO Science-NL Region
Mariano Koen-Alonso	DFO Science-NL Region
Robert Gregory	DFO Science-NL Region
Peter Shelton	DFO Science-NL Region
Erika Parill	DFO Science-NL Region
Sherrylynn Rowe	MUN-CFER
Fred Phelan	DFO-SSB-NL Region
Rick Rideout	DFO Science-NL Region
Leslie Norman	DFO-SSB-NL Region
Christoph Konrad	MUN-CFER
John Brattey	DFO Science-NL Region
Corey Morris	DFO Science-NL Region
Erin Carruthers	FFAW
Noel Cadigan	CFER-MUN
Brian Healey	DFO Science-NL Region
Doug Swain	DFO Science-Gulf Region
Sean Cox	Simon Fraser University
Eugene Lee	DFO Science-NL Region
Danny Ings	DFO Science-NL Region
Alejandro Buren	DFO Science-NL Region
Paul Glavine	DFO-SSB-NL Region
Patricia Williams	DFO-RMAF-NL Region
Joanna Mills-Flemming	Dalhousie University

APPENDIX 4: DATA QUESTIONNAIRES

DFO RESEARCH VESSEL SURVEY

The DFO RV multi-species survey covers the entire continental shelf off Newfoundland and Labrador; it is the longest standing time-series of data for Northern cod. The survey has changed over time, but in general, the autumn survey covers Divs. 2J3KL (the entire known stock area) and the spring survey covers Div. 3L.

1. How are data collected? [Detailed description with methods, sample sizes, standardized design or description of sampling protocol]

DFO in the Newfoundland region has been conducting scientific surveys of all or part of the stock area of Northern cod since 1949. However, it was not until 1971 that these surveys became standardized (prior to this they were line-transect design). In the spring of 1971, the first random, stratified bottom trawl surveys (with depth as the major stratifying variable) were carried out on the Grand Banks (Divs. 3LNO) to provide abundance and biomass indices (and other biological sampling) of groundfish stocks. The methods were approved by the International Commission of Northwest Atlantic Fisheries (ICNAF) and the Northwest Atlantic Fisheries Organization (NAFO). Documentation with regards to the development of these surveys is contained in the following reports: Grosslein and Pinhorn 1971 and Grosslein 1971. In 1977, autumn surveys were started on Div. 2J, expanded to 3K in 1978, 3L in 1981 and 3NO in 1990). The time series for Northern cod begins in 1983 (all of Divs. 2J3KL was covered). Over the time series, there have been three gear changes. Prior to 1983, a Yankee gear was used on chartered vessels; after 1984-95 (autumn) an Engel 145 high lift otter trawl was used on the vessels *Gadus Atlantica* and *Wilfred Templeman* and since 1995 (autumn) and 1996 (spring), the Campelen 1800 trawl (a shrimp trawl) has been used on either the *Teleost*, *Templeman* or the *Alfred G. Needler* (sister ship with the Templeman). Comparative fishing experiments were carried out between the *Gadus* (Engel) and the *Teleost* (Campelen) and conversions factors exist for Atlantic cod, among other species. Comparative fishing experiments were also carried out between the *Templeman* (Campelen) and the *Needler* (Engel) and conversion factors developed between these as well for Atlantic cod.

The survey relies on a stratified random design, with stations allocated in proportion to stratum surface area. However, a minimum of two fishing sites are allocated per stratum. Fishing is done on a 24-hour basis. Prior to 1995, a standard fishing tow (or set) was 30 minutes at a speed of 3.5 knots; since then it is 15 minutes at a speed of 3.0 knots. The trawl has attached net-monitoring (SCANMAR) equipment which ensures the ability to evaluate the consistency of gear deployment in real time during the tow.

Over time, and especially in recent years, there have been three sources of uncertainty: gaps in coverage (missed strata, reduced numbers of sets); changes in timing (synopticity; survey coverage extended later, coverage of some strata/Divisions often spread out over a much longer period than planned); and vessel effects (few direct comparisons of the 3 vessels used, using vessels in areas where they have little or no coverage in previous years).

For details on the complete description of the survey methods and standardized design, as well as the information gaps, please see "History of Annual Multi-Species Trawl Surveys on the Atlantic Coast of Canada" and the following papers:

Brodie, W. 1996. A description of the 1995 fall groundfish survey in Div. 2J3KLNO. NAFO SCR Doc. 96/27, Serial No. N2700, 7 pp.

Brodie, W. 2005. A description of the fall multispecies surveys in SA2+ Divisions 3KLMNO from 1995-2004. NAFO SCR Doc. 05/8, Serial No. N5083, 21 pp.

Brodie, W. and D. Stansbury. 2007. A Brief Description of Canadian Multispecies Surveys in SA2+ Divisions 3KLMNO from 1995-2006. NAFO SCR Doc. 07/ 18, Ser. No. N5366, 24 pp.

Grosslein, M. D. and A. T. Pinhorn. 1971. Progress in development of a co-ordinated survey program in the ICNAF Area. Int. Comm. Northw. Atlant. Fish. Res. Doc. 71/128, Ser. No. 2634 (mimeo).

Grosslein, M. D. 1971. Some observations on accuracy of abundance indices derived from research vessel surveys. Int. Comm. Northw. Atlant. Fish. Redbook 1971 (III): 249-266.

Pinhorn, A. T. 1971. Objectives and characteristics of existing and proposed groundfish surveys by the Fisheries Research Board of Canada Biological Station, St. John's, Newfoundland. Int. Comm. Northw. Atlant. Fish. Res. Bull. 71/36, Ser. No. 2527 (mimeo).

Power, D., Healey, B. P. and Ings, D. W. 2015. Performance and description of Canadian multi-species bottom trawl surveys in NAFO subarea 2 + Divisions 3KLMNO, with emphasis on 2012-2014. NAFO SCR Doc. 15/022, 25 pp.

Stansbury, D. E. 1997. Conversion factors for cod from Comparative Fishing Trials for Engel 145 Otter Trawl and the Campelen 1800 Shrimp Trawl used on Research Vessels. NAFO SCR Doc. 97/73. Serial No. N2907, 10 pp.

Warren, W.G. 1997. Report on the comparative fishing trial between the *Gadus Atlantica* and *Teleost*. NAFO Sci. Coun. Studies 29: 81-92.

Warren, W.G., Brodie, W., Stansbury, D., Walsh, S., Morgan, J., and Orr, D. 1997. Analysis of the 1996 comparative fishing trial between the Alfred Needler with the Engel 145 trawl and the Wilfred Templeman with the Campelen 1800 trawl. NAFO SCR Doc. No. 68, Serial No. N2902, 12 pp.

For a description on how the data from surveys is used for Divs. 2J3KL cod assessments, please see:

Brattey, J., Cadigan, N. G., Dwyer, K., Healey, B. P., Morgan, M. J., Murphy, E. F., Maddock Parsons, D. and Power, D. 2010. Assessment of the cod (*Gadus morhua*) stock in NAFO Divisions 2J+3KL in 2010. CSAS Res Doc. 2010/103, 108 pp. (particularly pages 7 & 8)

Information on gear used in survey:

McCallum, B.R., and Walsh, S.J. 1996. Groundfish survey trawls used at the Northwest Atlantic Fisheries Centre, 1971 to present. NAFO Sci. Coun. Studies 29: 93-104.

2. Has the design changed over time? [For example, changes in the execution of the DFO survey and coverage deficiencies will be documented]

The design of the survey has changed over time as more accurate hydrographic charts became available, as well as expanding to inshore strata and deeper offshore strata. Currently, the maximum depth in the autumn survey is 1,500 m while only 731 m in the spring. Northern cod assessments take into account only "index strata". These are: strata in the range of 100-500 m in Divs. 2J3K and 55-366 m in Div. 3L. (The difference across NAFO Divisions exists as Div. 3L was depth-stratified much earlier than Divs. 2J3K, and fathoms were used as units of depth). In some years deep water sets were not carried out due to operational or weather constraints. In recent years, there have been reductions in the numbers of sets – refer to Power et al. 2015 for details on numbers of sets by division and survey timing. It is not known how the reduction in sets affects estimates or associated estimates of precision. See the following paper(s) for details.

Brattey, J., Cadigan, N. G., Dwyer, K., Healey, B. P., Morgan, M. J., Murphy, E. F., Maddock Parsons, D. and Power, D. 2010. Assessment of the cod (*Gadus morhua*) stock in NAFO Divisions 2J+3KL in 2010. CSAS Res Doc. 2010/103, 108 pp. (particularly pages 7 & 8)

Brodie, W. 1996. A description of the 1995 fall groundfish survey in Div. 2J3KLNO. NAFO SCR Doc. 96/27, Serial No. N2700, 7 pp.

Brodie, W. 2005. A description of the fall multispecies surveys in SA2+ Divisions 3KLMNO from 1995-2004. NAFO SCR Doc. 05/8, Serial No. N5083, 21 pp

Power, D., Healey, B. P. and Ings, D. W. 2015. Performance and description of Canadian multi-species bottom trawl surveys in NAFO subarea 2 + Divisions 3KLMNO, with emphasis on 2012-2014. NAFO SCR Doc. 15/022, 25 pp.

3. How are the data analysed? [Detailed description of the statistical methods employed etc.]

Estimates of means or total abundance (and biomass), along with their associated variances are calculated using the standard formulae for stratified random designs (Cochran 1977; Smith 1990). Details of statistical analysis of the DFO survey are found in:

Doubleday, W. G. 1981. Manual on Groundfish Surveys in the Northwest Atlantic. NAFO Sci. Coun. Studies, 2: 7-55

Smith, S. J., and G. D. Somerton. 1981. STRAP: A user-oriented computer analysis system for groundfish research trawl survey data. Can. Tech. Rep. Fish. Aquat. Sci. 1030: iv + 66 p.

4. Are there estimates of uncertainty? [How are these constructed?]

Standard error estimates are obtained by repeated sampling within survey strata. Confidence intervals are produced on the basis of the Student's t-distribution (Cochran 1977) but these CIs may not be considered the appropriate measure of error due to the fact that catches may have skewed frequency distributions and small sample sizes within strata can cause extremely long CIs or negative lower limits (Smith 1997).

Cochran, W.G. 1977. Sampling techniques. 3rd ed. John Wiley, New York.

Doubleday, W. G. 1981. Manual on Groundfish Surveys in the Northwest Atlantic. NAFO Sci. Coun. Studies, 2: 7-55

Smith, S. J., and G. D. Somerton. 1981. STRAP: A user-oriented computer analysis system for groundfish research trawl survey data. Can. Tech. Rep. Fish. Aquat. Sci. 1030: iv + 66 p.

Smith, S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. Can. J. Fish. Aquat. Sci. 54: 616-630

5. Does this index provide an index of stock size?

Yes, it is a standardized, fishery-independent measure of biomass and abundance. It was designed originally to survey cod (and other commercial groundfish stocks). It is now a multispecies survey that monitors a vast array of species but still retains the capability to survey commercial groundfish stocks.

6. How does this data source contribute to assessing stock status?

It is a major part of the assessment of this stock. The analytical estimates from the 2013 assessment were based solely on information from this survey. Furthermore, biological samples collected during surveys yield data on fish condition, weight at age, maturity at age, and food/feeding.

7. Are the analysis methods (including uncertainty) appropriate?

In general, yes; however, there are criticisms with respect to the usage of the normal or Student's *t*-distribution for approximation of the sampling distribution as a measure of uncertainty (Cadigan 2011; Smith 1997).

Cadigan, N.G. 2011. Confidence intervals for trawlable abundance from stratified-random bottom trawl surveys. Can. J. Fish. Aquat. Sci. 68: 781–794.

Smith, S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. Can. J. Fish. Aquat. Sci. 54: 616-63.0

8. Suggestions for improvements/research recommendations?

Reviewers highlighted a number of limitations, including missed strata, variation of survey timing change in availability of the stock in the mid-1990s to mid-2000s. However, the survey appears to track cohorts well and it represents a key data input for any stock assessment model.

In particular, reviewers indicated that change in stock availability to the survey is worth investigation. Previous studies when large cod may not have been readily available; as cod rebuild, selectivity assumptions of survey gear should be investigated experimentally.

INPUT DATA QUESTIONNAIRE – EU-SPAIN TRAWL SURVEY OF 3L

1. How are the data collected?

Beginning in 2003, EU-Spain has conducted a stratified random survey annually (except 2005) to sample within NAFOs Regulatory Area of Div. 3L during July- August. A total of 24 strata were sampled in the Flemish Pass area from 2006 onward. During 2003, only 17 strata were sampled and the 2005 survey was cancelled due to logistical problems. A minimum of two tows are conducted per strata using a Campelen trawl with the allocation of sets proportional to stratum size. Each tow is 30 minutes duration at a mean speed of three knots. Tows are conducted daily between 6:00 and 22:00. Details can be found in Román et al. 2015 a and b.

2. Has the design changed over time?

No.

3. How are the data analyzed?

Mean biomass and abundance per haul, stratified mean biomass and abundance per haul and annual variance values are presented along with distributional data.

4. Are there estimates of uncertainty?

Annual variance values are calculated.

5. Does this data series provide an index of stock size?

No. Only a small portion of the stock area is within the scope of the survey and there is no evidence that significant concentrations of the stock are sampled by the survey.

6. How does this data source contribute to assessing stock status?

These data have not been used directly in assessing stock status.

7. Are the analysis methods (including uncertainty) appropriate?

N/A

8. Suggestions for improvements / research recommendations?

N/A

Key References:

Román, E., C. González-Iglesias and D. González-Troncoso, 2015a. Results for the Spanish survey in the NAFO Regulatory area of Division 3L for the period 2003-2014. NAFO SCR Doc. 15/019.

Román, E., A. Armesto and D. González-Troncoso, 2015b. Results for the Atlantic cod, roughhead grenadier, redfish, thorny skate and black dogfish of the Spanish survey in NAFO Div. 3L for the period 2003-2014. NAFO SCR Doc. 15/020.

INPUT DATA QUESTIONNAIRE – COD TAGGING DATABASE

There is an extensive electronic database of cod tagging information based on releases of several hundred thousand tagged cod extending back to about 1948 (pre-1993 data more difficult to access, 1995 onward readily available using web-based application) . These data have been used previously to investigate stock structure and migration patterns, management boundaries, and measure exploitation rates.

1. How are the data collected?

Various tags have been used (internal, external disc, and Floy). Historical data (to 1993) are summarized in Taggart et al. 1995. Can Tech. Rep. Fish. Aquat. Sci. 2042. Three main papers that use these data:

- Myers et al. 1997 CJFAS 54: 1-12;
- Myers et al. ICES J. Mar Sci 53: 629-640;
- Barrowman and Myers 1996. Biometrics 52:1410-1416.

Since 1995 only external Floy tags at base of first dorsal fin. Tag has unique #, return address, and reward value (since 1997). Program extensively advertised. Cod >=45 cm only. For post moratorium period total tagged=68,000. Tagging is opportunistic but attempts to spread tagging widely across stock area (mostly inshore since 1995, offshore only in 2008 and 2015. Sample size per year variable, minimum/year ~1,000; max about 5,600 per year in post-moratorium period. Double tagging used to estimate tag loss rate; high-reward method to estimate tag reporting rate (annually since 1997). Field expts (including acoustic telemetry) to estimate initial tagging mortality.

- Main papers dealing with tagging mortality, tag loss rates and tag reporting rates:
- Brattey and Cadigan 2004 Fish. Res. 66: 223-233;
- Cadigan and Brattey 2006. Can. J. Fish. Aquat. Sci. 63: 1944-1958.

2. Has the design changed over time?

Different tag types used pre-moratorium. Focus on inshore since moratorium (no offshore fishery), more widespread tagging (and recoveries) pre-moratorium. No reporting rates pre-moratorium.

3. How are the data analyzed?

Several approaches. See references above and various CSAS Research Documents (not complete listing):

- Brattey, J., and Healey, B. P. 2003. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/032.
- Brattey, J., and Healey, B. P. 2005. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/047.
- Brattey, J., and Healey, B. P. 2007. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/027.
- Cadigan, N., and J. Brattey. 2003. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/037.
- Brattey, J., B. P. Healey, and D. P. Porter. 2008. DFO Can. Sci. Advis. Sec. Res. Doc. 2008/047.
- Brattey, J. 2013. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/104.

4. Are there estimates of uncertainty?

For tag loss rates and reporting rate estimates.

5. Does this data series provide an index of stock size?

No.

6. How does this data source contribute to assessing stock status?

Provides information on stock structure and migration patterns that can be useful in understanding trends in indices. Can provide estimates of exploitation / could infer stock sizes if catch information is reliable.

7. Are the analysis methods (including uncertainty) appropriate?

Yes; see references.

8. Suggestions for improvements / research recommendations?

The meeting recommended a study of tag reporting rate to determine the extent of a declining trend in tag reporting and, if necessary, to develop a solution (e.g. increase value of high reward tags).

INPUT DATA QUESTIONNAIRE – SATELLITE TAGGING

1. How are the data collected?

Tags were attached to large cod (smallest 85 cm, most >95 cm) with a purpose designed backback and set to pop after one year. Not all tags make the full year and pop early and there are some failures. Where a tag can be located and returned (ours by fishermen) two minute data resolution over the full year ispossible. For a tag which transmits but is not returned a maximum of hourly data is transmitted (transmissions are random over the full span of the data so that if battery power fails prior to full transmission as appropriate a sample as possible is returned - this varies but most tags give up >50% of their hourly data and some all of it).

Tagging was done in May of 2012-15. Numbers of tags below:

- 2J (10 in 2015)
- Notre Dame Channel 3K (4 in 2015)
- Bonavista Corridor (3KL) (40 in 2014, 32 in 2013, 4 in 2012)

Data are taken from tag initially processed by MTI (manufacturer). Data are still coming in - preliminary reporting could be made available by March 2016. Tags provide data applicable to survey estimates - in particular where fish were when surveys-fisheries are being conducted, and data on how this relates to stock structure and rebuilding.

2. Has the design changed over time?

No. Other than locations of tagging.

3. How are the data analyzed?

Not yet completed. Temperature and depth straightforward. Latitude and longitude requires filtering and an algorithm being developed.

4. Are there estimates of uncertainty?

For latitude and longitude possibly. Not really applicable to temp and depth as sensors are highly accurate.

5. Does this data series provide an index of stock size?

No.

6. How does this data source contribute to assessing stock status?

At present, it is not incorporated, however this data may provide spatial context, migration data, and information on stock fidelity in the future.

7. Are the analysis methods (including uncertainty) appropriate?

N/A (too preliminary).

8. Suggestions for improvements / research recommendations?

N/A (too preliminary).

INPUT DATA QUESTIONNAIRE – ACOUSTIC SURVEY OF COD BIOMASS IN OFFSHORE - CFER

1. How are the data collected?

Acoustic survey methods in late winter and spring have been applied to cod distributed in the Bonavista Corridor region since 1990 (several publications based on these data are in the primarily and DFO-CSAS literature, beginning with Rose 1993 with the most recent one Rose and Rowe 2015). The acoustic equipment, vessels, and survey designs have changed over these 25 years, but the backscatter measures are thought to be consistent and comparable, and the scaling to a biological measure has been consistently applied. All raw data used were 38 kHz backscatter from calibrated scientific echosounders and all were supported by directed trawling to identify species, size and other biological properties of the identified fish signal. Additional frequencies are available for the most recent years and these data are being explored for additional information of species and size. A"dead zone" estimate is made from a linear expansion of the backscatter in the bottom 5 m (ranges from <10-20% of signal). Some details of this are in earlier publications and in Rose and Rowe (2015).

Despite the changes over the years, it is thought unlikely that they could cause anything like the changes in backscatter measured which span orders of magnitude. Nevertheless, in describing the recent increased abundance of cod perhaps the most important of these data were collected since 2012 when there were no changes in equipment, timing, vessel and only minor changes in survey coverage in this area (the 2015 survey covered much more ground to the north).

With the exception of 2012, which was a survey of opportunity with only two large fishing sets - the primary goal being tagging - the surveys from 2013-15 have dozens of directed sampling sets covering the full survey region. Most sets were done with the Campelen trawl but a few with the GOV trawl - it is feasible that there is a trawl bias in the size composition data which is being investigated.

The standard survey design is to run latitudinal transects across the shelf spanning 500 to approximately 250 m depth range each 5 sq. n. mi. (in some instances 10 sq. n. mi.). Over most all of this region cod are distributed at this time of year in single species aggregations of varying density, hence species ID is not an issue. Capelin are also routinely assessed during these surveys.

2. Has the design changed over time?

Yes, but only minor changes since 2012. Especially during the late 1990s and early 2000s the fish were so few and densities near uniformly low that sampling intensity was reduced considerably.

3. How are the data analyzed?

Echoview software is used for echogram visualization and editing. The bottom is identified as the various species based on visual recognition, trawling results and target strength correspondence. These data are then integrated by 1 sq. n. mi. bins. Several methods have been employed to estimate mean density and CIs. For the full series (most recently published in Rose and Rowe 2015) up to 2014, a simple randomizing bootstrap methods was used on 1 sq. n. mi. estimates of cod density (the basic sampling unit based on backscatter scaled by mean target strength based on the length based models of Rose and Porter 1996 and Rose 2009). This was done 1,000 times with n equal to the number of basic sampling units and from that the estimate of mean and 95% CIs were derived.

We have also used geostatistics (see Mello and Rose 2008, 2009 Res Docs). This is likely a preferred method but earlier surveys could not be handled this way so for consistency this was not used in recent analyses and publications. In addition, there have been issues (technical) with the degree and scale of spatial autocorrelation on which this method is dependent, and these are currently being investigated.

4. Are there estimates of uncertainty?

Yes, in all cases. See point 3.

5. Does this data series provide an index of stock size?

Unknown. See Question 8.

6. How does this data source contribute to assessing stock status?

May provide either an index or an absolute estimate (pending review) of spawning stock abundance and biomass by age within the region surveyed. See Question 8.

7. Are the analysis methods (including uncertainty) appropriate?

Unknown. See Question 8.

8. Suggestions for improvements / research recommendations?

The Framework Meeting was convened to review input data sets and stock assessment models. Other data series considered by the group were well understood and have been debated extensively during previous assessment meetings. After discussion on questions 1-4 for this data source, several questions were raised on the technical details of both the data collection and analysis methods. It was not possible to draw a conclusion during the meeting, in part because the Principal Investigator was unable to attend.

Meeting participants had only limited acoustic experience and there was discussion on whether some of the questions that had been raised were pivotal ones, and equally likely, there could be important issues that were not discussed.

Given the potential importance of this data set to provide valuable information to understanding recent changes observed in Northern cod, the Framework meeting **recommended** that the acoustic biomass data be considered at an independent peer-review meeting, including recognized assessment and acoustics experts.

INPUT DATA QUESTIONNAIRE – ACOUSTIC SURVEY OF COD BIOMASS IN SMITH SOUND

1. How are the data collected?

All acoustic data used either a 38 kHz EK500 or EK60 echosounder (except 1995) and a blocked survey design (see Rose 2003). All sounders were calibrated using standard tungsten-carbide calibration sphere methods. Acoustic integration was done in Echoview software. Up until 2003 sampling was done using short trawl sets on CCGS Teleost (a few on CCGS Shamook). After 2003 no suitable method was available. Catches in all cases were more or less 100% cod and species ID is not considered an issue. Data on these elements have been published in Rose 2003 and Rose et al. 2009.

2. Has the design changed over time?

Design has not changed. The only change has been different timing of the survey, and different vessels, but all were done in the winter-spring period when the cod aggregation or what was left of it was present (in some of the later years multiple surveys were conducted over the winter-spring period).

3. How are the data analyzed?

Standard echo integration, scaled by a target strength-to-weight (TS/kg) relationship - this is not the best way to do this but for these data it is the simplest and given the lack of reliable size data in the later years the only feasible method.

4. Are there estimates of uncertainty?

Yes. They are based on sampling error of the echo integration values within each survey block.

5. Does this data series provide an index of stock size?

No, but it may provide an estimate or index of a portion of the stock unavailable to the DFO RV Survey.

6. How does this data source contribute to assessing stock status?

This data is employed in the NCAM model to account for some non-stationarity in the stock.

7. Are the analysis methods (including uncertainty) appropriate?

See #8 of previous data questions (Offshore Acoustic Survey).

8. Suggestions for improvements / research recommendations?

See #8 of previous data questions (Offshore Acoustic Survey).

INPUT DATA QUESTIONNAIRE - BEACH SEINE SURVEYS

1. How are the data collected?

Age 0 & 1 year old juvenile Atlantic cod have been enumerated and measured at 12 sites in Newman Sound, 1995-2015. The same sites have been sampled every two weeks from July to November each year throughout the period (see #2). A total of 120 seine samples are taken each year in a typical year. Fish samples are collected using a 25 m demersal beach seine, deployed from a 6 m boat at a distance of 55 m from the shore. The seine sampled the lowest 2 m of the water column. The same methods have been used throughout the entire 20 year period of the program.

All fish collected are identified, enumerated, and measured. Juvenile cod were assigned to tentative age groups based on previously established age-length relationships in Newfoundland waters in late autumn. These estimates are refined by examination of length frequency trajectories through each season.

2. Has the design changed over time?

The methods have not changed appreciably over the entire 20 years of the program, with some early exceptions. In 1995, only six sites were sampled and only from September to November. In 1996, nine of the eventual 12 sites were sampled. In all subsequent years, the sampling frequency and effort have been maintained.

In some years, additional sampling has been undertaken to affirm the prevalence of certain coast-wide dynamics relevant to interpretation of age 0 & 1 catch data in Newman Sound. These efforts confirmed the broad relevance of the data. In the first ten years of the program, effort was expended to analytically tie this survey into the spatially broader "Fleming Survey (St. Mary's Bay to Notre Dame Bay – 1959-64, 1992-97, & 2001).

3. How are the data analyzed?

The catch data are analyzed using descriptive statistics and linear regression techniques among and within cohorts between capture years, within and among recruitment datasets (e.g. Fleming and inshore RV survey data).

4. Are there estimates of uncertainty?

Descriptive statistics include simple metrics of standard error of means, 95% confidence intervals and similar metrics.

5. Does this data series provide an index of stock size?

No. It is a recruitment index.

6. How does this data source contribute to assessing stock status?

Measures age 0-1, and captures age 1 reliably. Has been used most reliably in the stock assessment to bridge the gap in young age classes.

7. Are the analysis methods (including uncertainty) appropriate?

Methods were not reviewed here, but general consensus is that they are appropriate and acceptable to the meeting participants.

8. Suggestions for improvements / research recommendations?

It should be noted that this relationship will get proportionally weaker as stock rebuilds due to spillover from primary nursery habitats. Should be analyzed as the rebuilding trend continues.

Aside from saturation issues, will this still provide a recruitment index? May not index the stock as a whole in the future (area occupied is shifting and expanding). This could be analyzed with other recruitment indices like a Fleming survey, which offers a wider approach.

Does this relate to density dependent recruit per spawner trends?

INPUT DATA QUESTIONNAIRE – INSHORE MOBILE GEAR SURVEY 2J3KL 2006-2011

1. How are the data collected?

From 2006 to 2011, DFO in cooperation with the FFAW under the Fisheries Science Collaborative Program, conducted a stratified-random survey of the coastal/nearshore area of 3KL and part of 2J (as far north as Black Tickle 53°30'N). A total of 150 sets were planned. Coverage was apportioned by stratum size separately for two areas. The first covering 32 coastal strata (95 sets in depths <50 m; 3,800 sq. n. mi.) and 15 perimeter strata on the seaward side of the coastal strata (55 sets in depths from 51 m – 200 m; 9,100 sq. n. mi.)

Five different 59' - 64' vessels conducted the work from July/Aug 2006-11 utilizing a Balloon Star 300 otter trawl with a 40 mm liner (same gear as 4RS3Pn and 4T sentinel mobile gear surveys). Sets targeted a 30 minute tow at a speed of 2.5 knots (10 minutes acceptable).

A minimum of two tows were conducted per strata.

2. Has the design changed over time?

Yes. Original coastal strata were developed quickly based on dated nearshore hydrographic information. Prior to the 2009 survey, a re-stratification was undertaken based on more recent hydrographic charts. The area in the coastal strata there was a net increased from 3,800 to 4,200 sq. n. mi and the area in the perimeter strata decreased from 9,100 to 8,800 sq. n. mi

3. How are the data analyzed?

Stratified mean number and weight per tow as well as by length group and by age and associated variances and 95% CI's for each area, and, swept area biomass and abundance produced for overall areas as well as by three zoned areas within 2J3KL (Inshore North, Central and Southern areas).

4. Are there estimates of uncertainty?

Yes. (see #3 above).

5. Does this data series provide an index of stock size?

Overall this survey has shown high variability and does not appear to be tracking cohorts and has therefore not proven to be a good index of changes in the stock.

6. How does this data source contribute to assessing stock status?

Results may be strongly influenced by annual changes in the magnitude and timing of movement of migratory cod from the offshore. Its contribution remains unknown.

7. Are the analysis methods (including uncertainty) appropriate?

N/A

8. Suggestions for improvements / research recommendations?

This survey was discontinued after the 2011 survey.

INPUT DATA QUESTIONNAIRE – SENTINEL SURVEY

Sentinel Survey projects were formally announced by the Minister of DFO in October, 1994. The Sentinel Surveys in the NL Region of DFO are an extension of the index fishermen's project from the Northern Cod Science Project Program with modifications to allow for science activities achievable only under a fishing moratorium.

The Sentinel Survey had the following objectives:

1. To develop a catch rate series for use in resource assessments.
2. To incorporate the knowledge of inshore fish harvesters in the resource assessment process.
3. To describe the temporal-spatial distribution of Atlantic Cod in the inshore area over a number of years through, for example, the use of catch rate information, tagging studies, bycatch information and participant' observations.
4. To gather length frequencies, sex and maturity data and sample ages for use in resource assessment.
5. To establish a long-term physical oceanographic and environmental monitoring program of the inshore areas.
6. To provide a source of biological material for other researchers. For example, tissue for genetic, physiological and toxicological analyses, cod stomachs for food and feeding studies, and by-catch information.

An in-depth review of the Sentinel Surveys across Atlantic Canada was conducted in November, 2001 (DFO 2002). The intent of the review was to:

1. Document the various sentinel projects, and how they contribute to provision of advice;
2. Consider changes to project design as necessary to better meet the needs of stock assessment and the financial impact of any such revisions; and
3. Identify any challenges in implementing such changes while maintaining the utility of various time-series data produced by the program to date.

1. How are the data collected?

Repeated sampling in designated inshore areas. Intent of program was to continue inshore fishing at traditional times in an area established through traditional knowledge. Additionally, these areas are mostly shore-ward of the DFO RV survey. Various fishing gears were used at different sites, including: Gillnets, (mesh-sizes of both 5.5 inches and 3.25 inches have been used), Linetrawl (i.e. longlines), cod traps, and handlines. The smaller mesh gillnet is deployed attached to the commercial-sized 5.5 inch mesh, and the intent was to capture smaller cod to provide data on cod recruitment Cod traps were phased out over time, and were not been used in the Sentinel Survey after 2002. Hand lines were used mostly in conjunction with nets or trawls as a means of acquiring cod for tagging purposes or determining presence of cod when nets were not catching fish.

A fixed (control) location on the fishing grounds was established for each site and will remain fixed for the duration of the project. Each fishing day, up to half of the gear was set at the control site. The remainder of the gear (experimental) was set at one or two other locations on the fishing grounds at the discretion of the crew. The location of each fishing set, and the time of the set and the soak time for the gear were recorded. Other environmental observations were noted, including wind direction and speed, percent cloud cover, tide conditions, presence of invertebrates (bait) and other fish species in the area, marine mammals, sea birds and any other variables which might have influenced fishing behavior.

Catches from control and experimental gear were measured for length and sex (i.e. 100% measured). Otoliths were sampled on a fish length-stratified basis to provide a means to study age and growth. Selected sites were instructed to collect a length-stratified sample of up to 100 frozen fish on a biweekly basis for detailed biological sampling by DFO technical staff at the Northwest Atlantic Fisheries Centre. Weight analysis measurements were taken on these samples

2. Has the design changed over time?

The number of sites sampled by the Sentinel Survey has generally been declining since its inception. Initially, the program comprised 66 sites. By 2004, this had been reduced to 58, with a further reduction to 43 by 2008. In 2012, 48 sites were occupied.

In addition to variation in the number of sites sampled, the annual sampling effort – measured by number of weeks allocated to each site – has also declined. In 1995, sampling was conducted over a 15 week period, while over 1996-98, 12 weeks were allocated. In 1999, a minimum of eight weeks were allocated, though many sites were allotted extra time. Each enterprise involved surveyed for 10 weeks in most recent years.

3. How are the data analyzed?

Sentinel observations are subjected to data screening to remove outliers (e.g. extended soak time if gear could not be retrieved at normal interval due to inclement weather). Screening criteria were proposed by an informal WG struck at the 1999 Zonal assessment meeting in Rimouski, Quebec, and these have been modified only minimally since that time. Catch per unit effort is computed as number of fish per net for gillnets (both mesh sizes), and number of fish per 1,000 hooks for linetrawl.

Sentinel catches are aged using the aging information collected at selected sites distributed across the inshore of Divs. 2J3KL. Distinct age-length keys were compiled within Unit Area (e.g. 3Ki, 3Lh, etc.) / year / quarter-of-year cells. The age range included in further analysis for each gear type are as follows:

- Linetrawl – Ages 3-9;
- 5.5 inch Gillnet – Ages 3-10; and
- 3.25 Inch Gillnet – Ages 2-10.

To standardize the CPUE data, a generalized linear model was applied to the catch and effort data for each gear and survey method. Spatial, temporal and age effects are incorporated, as month-within-site and age-within-year parameters are estimated. A Poisson response distribution (with log link function) was assumed, and the effort (amount of gear) used as offset variable. Given the properties of the Poisson distribution, data were aggregated prior to modeling, within each:

- Type/Gear/NAFO/Division/Location/Month/Year/Age cell, where "type" refers to the control vs experimental site at each location.
- Least-squares means of the year –age effects are the primary output of interest. Annual CPUE is constructed by summing the age-specific estimates in a given year.

4. Are there estimates of uncertainty?

Yes, estimates of variation are the 95% Confidence Intervals for the LS means.

5. Does this data series provide an index of stock size?

The Sentinel Survey provides valuable information about availability of commercial sized fish to fishing gear set a long (north-south) but narrow strip of the inshore. The portion of the stock available to the passive gears employed by this survey has quite likely varied over the duration of the project.

From the period of collapse in the early 1990s until the mid-2000s, there were very few adult cod in the offshore. Additionally, the only 'large' over-wintering aggregation known to exist over that period was resident in the inshore (Smith Sound, Trinity Bay) and these fish were clearly available to Sentinel gears throughout the summer and fall. However, as noted in the 2013 stock assessment, more recent observations in (from both tagging and acoustic telemetry), have indicated that the inshore-offshore migrations observed prior to the collapse of this stock has resumed. Increased numbers of cod are apparently over-wintering in the offshore, and some of these migrants would be available to the Sentinel survey. However, it is unknown what fraction of the over-wintering fish are residing in the offshore year-round, and would therefore be unavailable to the Sentinel survey. These dynamics have most certainly affected catch rates at different Sentinel sites along the coast to differing degree.

Various index and stewardship fisheries have been conducted over the duration of the Sentinel survey, and fishery timing has (at least in part) overlapped with the period of Sentinel survey activity. Competition effects between Sentinel survey gear and other passive gears would impact the catch rates, and these impacts are likely variable over the time-series given changes to management regulations in these fisheries.

Variations in local oceanographic conditions and cod interactions with both predators and prey impact the fixed-gear catchability by an unknown amount. These differences would be captured in the estimated year effects, but distort the potential of the survey providing and index.

The 2001 workshop on the Sentinel program (DFO 2002) noted: *'An observed strength of the Groundfish Sentinel Program is that it is capable of sampling the stock at otherwise inaccessible locations and at times not normally covered by other fishery-independent assessment tools, particularly research vessel surveys. However, it was recognized that the full utility of abundance monitoring based on sentinel activities cannot be known until its responsiveness is tested as stocks undergo periods of change more pronounced than has generally been the case to date.'*

Given the above, it would seem that the Sentinel survey does not provide an index of stock size, but possibly provides an index of relative changes in the abundance of cod available to (very) near-shore fisheries.

6. How does this data source contribute to assessing stock status?

The sentinel in recent period shows a trend that is consistent with other information (RV), Cohort strength signals also consistent with RV. Sentinel trends for north, central, south agree well with spatial trend information from RV for recent period.

Sentinel data has improved understanding of changes in stock structure through the post-moratorium period. Cohort tracking has illuminated differences between age structure inshore and offshore in previous years. The age composition information from the Sentinel is valuable to stock assessment models (e.g. used in NCAM).

7. Are the analysis methods (including uncertainty) appropriate?

Methods are appropriate if there are no strong interactions (see 8)

8. Suggestions for improvements / research recommendations?

What was the impact of decline in number of sites? How did this change the spatial coverage of the survey? What was the impact in the restriction of time range of the sentinel?

Improvements in modelling:

- Consider using all data, not just fixed sites
- Extend maximum age-range included in modelling exercise?

Key References:

DFO. 2002. Workshop on the Groundfish Sentinel Program. CSAS Proceedings Document 2002/003.

Maddock Parsons, D. 2014. Update of Sentinel Survey Results in NAFO Divisions 2J3KL for 1995-2012. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/014. v + 32 p.

Maddock Parsons, D., and Stead, R. 2009. Overview of Sentinel Surveys in NAFO Divisions 2J3KL and subdivision 3Ps 1995-2008. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/092. iv + 36 p.

Maddock Parsons, D., R. Stead and D.E. Stansbury. 2000. Sentinel surveys 1995-1999: catch per unit effort in NAFO divisions 2J3KL. DFO Can. Sci. Advis. Sec. Res. Doc. 2000/102.

Stansbury, D.E., D. Maddock Parsons, and P.A. Shelton. 2000. An age-disaggregate index from the sentinel program for cod in NAFO Divisions 2J3KL. DFO Can. Sci. Advis. Sec. Res. Doc. 2000/090.

INPUT DATA QUESTIONNAIRE – COMMERCIAL CATCH AT AGE SAMPLING

1. How are the data collected?

Fish length measurements and otoliths are collected by independent observers onboard fishing vessels or by DFO Science Branch personnel through the port sampling program. All fish are unsexed. The sampling design for port sampling is based on cells composed of quarter, gear type, and unit area. Currently, minimum length sampling is set at 200 per sample. Minimum requirements for otolith collections are five per 1 cm length group per cell. Design for observer sampling is based on cells composed of quarters, gear type, and NAFO Division. Minimum length measurements are 250 per set with one set monitored per day in directed fisheries and as many sets as possible for bycatch fisheries. Two otoliths per length group per week are required for directed fisheries and one otolith per length group per trip is asked for bycatch fisheries.

During some years, conservation and patrol officers also collect length data from individuals encountered at sea or dockside and these data plus aging data from the commercial sampling or research trawl survey are used for catch at age analysis of the recreational/food fishery.

2. Has the design changed over time?

Yes. Originally, the sampling design was based on 3 cm length groups, which changed to 1 cm groups in 1993. Port sampling had minimum sampling of 10 otoliths per 3 cm length group in cells composed of quarter and unit area. Most of the sampling was conducted at fish processing plants in the 1980s, with 300-400 fish measured per set and an attempt was made to increase sampling for larger catches (Stevenson 1983).

With the short duration of the fishery during recent years, DFO staff indicate it is challenging to collect spatially representative samples. Further, retention of fish of preferred lengths for personal use prior to any potential sampling activity also reduces the number of random samples available for measurements.

3. How are the data analyzed?

Data from both the commercial fishery and the sentinel survey are combined into one analysis of catch at age. Length measures are available from almost every fish collected during the sentinel survey, but the sampling frequency is much lower for the commercial fishery. Length frequencies are calculated by cells composed of source (sentinel, commercial), month, gear and unit area and assigned to the corresponding landings. When length frequencies are not available for a particular cell, the catch is added as a bump-up factor for that gears total catch. Quarterly age length keys are applied to length frequency data to age the catch and derive proportions at age. Keys from another source (i.e. Sentinel surveys), neighbouring unit areas or a combination of gears are used when aging data are unavailable for a cell. In the event that no appropriate aging data are available, catch from that cell would typically be age-disaggregated using the age distribution from a larger cell (e.g. all gillnet catch in first half of year).

The number at age is calculated from the total catch divided by the average weight at age and multiplied by proportions at age. Further details are available in Gavaris and Gavaris (1983).

4. Are there estimates of uncertainty?

Yes. A standard error and C.V. is provided for each age class and a sum of products is calculated.

5. Does this data series provide an index of stock size?

No.

6. How does this data source contribute to assessing stock status?

Average weight at age estimates generated by the catch at age analysis were used previously in assessment models (i.e. ADAPT) for Northern cod to derive stock biomass. The current (i.e. 2013) assessment model (SURBA), uses average weight at age data from the autumn survey to predict stock biomass at the beginning of the year (Cadigan 2013). Weight at age data from commercial and sentinel fishing are available as an informal check on these estimates.

7. Are the analysis methods (including uncertainty) appropriate?

Yes.

8. Suggestions for improvements / research recommendations?

The commercial catch data are treated as standard data sources in most assessments, however it cannot be ignored that these data are limited by unreported landings, discarding/high-grading, and unrecorded recreational catch. A significant portion of total catch is unseen (estimated 40% of total catch is recreational harvest since 2006), and there is no evidence that age compositions are similar to the reported commercial catch. This should be investigated, and may require both management and research efforts to examine and report recreational catch.

Key References:

Cadigan, N. 2013. Update of SURBA+ for 2J3KL cod. Can. Sci. Advis. Sec. Res. Doc. 2013/054. p. 1-19.

Gavaris, S., and C. A. Gavaris. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. Can. Spec. Pub. Fish. Aquat. Sci 66: 178-182.

Stevenson, S. C. 1983. A review of sampling of commercial groundfish catches in Newfoundland. Can. Spec. Pub. Fish. Aquat. Sci 66: 29-38.

INPUT DATA QUESTIONNAIRE – STEWARDSHIP CATCH

Accurate information on fishery removals is imperative for both stock assessment and eco-certification purposes, and both are connected to ensuring long-term sustainability of the resource. Although catch data obviously do not provide an index of stock status, they are an essential element in population dynamics models which attempt to estimate stock size. This intent of what follows is to document some of the issues with respect to estimates of removals for both commercial and recreational fisheries.

Commercial catches are estimated by DFO Statistics Branch. Since the mid-1990s, an independent Dockside Monitoring Program has been in place for commercial index and stewardship cod fisheries. Total cod landings are estimated using a combination of Dockside Monitor verified weigh-off of catch, self-declared weigh-off of catch, information from Fisheries Officers, and commercial logbook data. A ‘Dockside Monitor’ is an agent of a company independent of the fish buyer, whose role is to witness all weigh-off of catch and to attest that the weight recorded is valid. Self-declared weigh-offs apply to a relatively low fraction of the landing events, and occur in isolated areas with low numbers of landing events, and/or when monitors are unavailable in ports with moderate or high numbers of landing events. For any self-declared weigh-off to occur, prior to offloading any catch, a harvester must phone the monitoring company, and provide an estimate of catch on board the vessel. If no monitor is available to witness catch off-loading, the monitoring company provides the harvester with a unique authorization number (without this, off-loading of fish is not permitted; violators are subject to prosecution under the *Fisheries Act*). Fisheries Officers inspect landing events of both monitored and unmonitored (with authorization) off-loads. Statistics Branch provides estimates of removals by landing event for quota monitoring and stock assessment purposes, but at present there is no available documentation on the precision of these removals.

Various issues have been raised by different industry stakeholders over time regarding potential biases of the official landing statistics, and some of these are as follows:

- Abuse of Authorization Number Process (catch offloaded without monitor)
- Accuracy of weigh-offs
- Abuses of fish designated as ‘Personal use’ (i.e. not sold commercially)
- High-grading for selected sizes when there was a price differential across size categories
- Dumping, discarding at sea
- Misreporting by species and/or stock area (more of an issue around the time of moratorium)
- Monitor Avoidance / Trading info on monitor location

The extent to which any of these issues are, or continue be, problematic with respect to estimating catch accurately is largely undocumented, and no new information is available for review at this meeting. However, recent discussion has been initiated across multiple sectors of DFO and the intent is to seek to identify whether or not each of these issues are affecting the accuracy of *current* cod landings statistics. (Additionally, a more fulsome review of what issues have been raised will be conducted.) Further, in cases for which there is insufficient information to come to a conclusion on whether or not they are impacting the validity of the removals, suggestions will be provided on what information or process is required to come to conclusions for each case.

INPUT DATA QUESTIONNAIRE – RECREATIONAL FISHERY

At present there are no estimates available of the amount of cod removed by Recreational fishery. Effort is controlled through various management measures.

Subsistence fisheries have been ongoing in the Northern cod stock area for centuries, and these fisheries were unregulated prior to the 1992 commercial fishing moratorium. Since 1992, the Food / Recreational groundfish fishery has had various management regulations, and the following details have been provided by Resource Management and Aboriginal Fisheries Branch:

- 1993-1995: No recreational fishery
- 1996 to 2000: Recreational fishing over 1-2 weekends only, with a 10 fish daily bag limit per individual, maximum of 50 fish per boat. (No fishery in 1997.)
- 2001 and 2002: Pilot Recreational Licence regime, 8-9 week season, retention tags required with a 10 fish daily bag limit. In 2001, 30 tags were available per licence, reduced to 15 tags per license in 2002.
- 2003 – 2005: No recreational fishery.
- 2006 to Present: No licence required. Daily bag limit of 5 fish per person, up to a maximum of 15 fish per boat. Season length has been approximately 30 days in recent years.

Since the inception of the (post-1992) recreational groundfish fishery, various processes have been employed to estimate the recreational groundfish harvest. For example, estimates of removals have been generated based upon license-return data when licenses were required. In other years, estimated removals were generated by Conservation and Protection staff using effort estimates (number of vessels and individual participating combined with at-sea / dockside data on the numbers of fish per participant and/or vessel). For the 2007 season, two estimates of recreational catches were available – one from C&P, and a second from a telephone survey conducted by an independent consulting firm contracted by DFO. Estimates from each source were considerably different (371 t vs ~2,000 t) and although reasons for these discrepancies were explored, neither was invalidated. No direct estimates of recreational groundfish landings have been available since 2008.

In recent assessments, recreational groundfish removals have been inferred based upon comparison of tag returns from recreational vs commercial fishers, and applying the ratio of these values to the commercial landings to generate the approximate level of removals from recreational fishing.

INPUT DATA QUESTIONNAIRE – 2J3KL INSHORE FISH HARVESTER QUESTIONNAIRE

The cod sentinel survey, which began in 1994, was designed to develop a reliable catch rate series, to describe cod distribution in the inshore, to collect length frequency data and biological material, and to incorporate the knowledge of inshore fishermen into the process of resource assessment (Davis et al. 1996). During the early years of the survey, results were reviewed with harvesters at regional meetings. Harvesters were asked to comment on abundance trends, catch rates and condition of cod, and signs of recruitment. For example, comments from the post-sentinel meetings indicated that catch rates were low and cod were small in Div. 2J during the 1998 fishery (Jarvis 1999). In contrast to the consistent message from Div. 2J, reports from Div. 3K harvesters differed among communities with some reporting excellent catch rates and fish in good condition. However, communities on the Northern Peninsula reported very poor catch rates in 1998. Div. 3L harvesters reported good catch rates at 17 of the 20 communities where regional meetings were held (Jarvis 1999).

The 2J3KL fish harvester telephone survey was initiated in the mid-2000s and replaced regional meetings as a means of gathering information on catch rates, condition, recruitment, and the status of the fishery. The 2014 survey questions are included below.

1. How are the data collected?

The telephone survey follows a stratified random design using the 2008 Groundfish License Holders list. The call list for each NAFO division is based on a randomized order of license holders. Each person doing the survey is handed a randomized list and a standard introduction to the survey.

135 license holders were surveyed in 2014. 36 license holders were surveyed in Div. 2J, 32 in 3K, and 67 in 3L. Target response rates have been 10-15% of license holders within each NAFO division.

2. Has the design changed over time?

The telephone survey design has not changed since regional meetings were replaced by the survey.

However, questions about tagged fish were added in 2007. Responses to these questions provide information on the tag reporting rate from each NAFO division.

3. How are the data analyzed?

Data are summarized and presented at stock assessments. To date, there have been no additional analysis beyond regional summaries.

4. Are there estimates of uncertainty?

To date, no estimates of uncertainty have been presented. Variance among different vessel classes or regions could be compared.

Background references:

Davis, B. + 29 co-authors, 1996. The 1995 Sentinel Survey in NAFO Subdivision 3Ps. Internal DFO and FFAW Report.

Jarvis, H. 1999. 2J3KL: 1998 Peer Review of Sentinel Results & Fish Harvesters' Comments on Index Fishery Results. Internal FFAW Report.

2014 Telephone Survey Questionnaire Script

I am calling from the FFAW. In preparation for the upcoming science meeting to assess 2J3K3L Northern cod stock, we are collecting information from fish harvesters. We strongly believe that this stock cannot be accurately assessed without the knowledge you have gained from being on the water.

We would like to ask you a few questions about Northern cod, it will only take a few minutes, would you like to participate?

1. How old are you?
2. How many years have you fished for cod?
3. Did you fish for cod in 2013?
4. What is the size range of the vessel?
   a) under 35ft
   b) 35 to 45ft
   c) 45 to 65 ft
5. Compare the abundance of cod in your area during 2013 with that of the late 1980's. Would you say that in 2013, cod were:
   a) Less abundant
   b) About the same
   c) More abundant
6. Compare the abundance of cod in your area during 2013 with that of 2012. Would you say that in 2013, cod were:
   a) Less abundant
   b) About the same
   c) More abundant
7. The size range of cod in your area during 2013 appeared to be:
   a) Mostly under 20 inches in length
   b) Mostly 20 inches and over
   c) An even Mixture of ALL sizes
8. Compare the overall size of cod in 2013 with 2012. In 2013 cod, overall, were:
   a) Smaller
   b) Same
   c) Larger
9. The condition or health of cod in your area during 2013 appeared to be:
   a) Poor
   b) Average
   c) Good

10. With consideration given to the traditional time cod migrated to your area, would you say the migration in 2013 was:
    a) Earlier
    b) About the same time
    c) Later
11. With consideration given to the traditional time cod migrated out of your area, would you say the migration in 2013 was:
    a) Earlier
    b) About the same time
    c) Later
12. In the inshore, cod have historically been found in specific areas commonly referred to as "traditional fishing grounds". During 2013, cod were found:
    a) Only in one location
    b) Average distribution
    c) Throughout the area
13. Baitfish are important to the recovery of our cod stocks. For species cod feed on it is important to know what is happening to their abundance level.

A. Capelin abundance is:

- At a low level and declining
- At a low level but increasing
- Average abundance & not changing
- Good abundance but decreasing
- Good abundance and increasing

B. Squid abundance is:

- At a low level and declining
- At a low level but increasing
- Average abundance & not changing
- Good abundance but decreasing
- Good abundance and increasing

C. Herring abundance is:

- At a low level and declining
- At a low level but increasing
- Average abundance & not changing
- Good abundance but decreasing
- Good abundance and increasing

D. Mackerel abundance is:

- At a low level and declining
- At a low level but increasing
- Average abundance & not changing
- Good abundance but decreasing
- Good abundance and increasing

14. In 2013, did you catch one or more cod with tag(s)? Yes/No. If yes, did you return the tag? Yes/No

15. Since the late 1980's are there more seals in your area: Yes/No/not sure

16. In your area what effect do seals have on cod?

    a) High
    b) Moderate
    c) Low
    d) No Impact on Cod

17. Based on your observation what was in the cod bellies in 2013:

    a) Sandlance
    b) Capelin
    c) Herring
    d) Mackerel
    e) Squid
    f) Small Crabs
    g) Other

Survey Respondents are also asked if they have additional comments.

INPUT DATA QUESTIONNAIRE – COMMERCIAL LOGBOOKS, <35' FLEET

A new "science logbook" was introduced to record catch and effort data for vessels <35 ft. in the re-opened fishery in 1998. Prior to the 1992 moratorium, the only data for vessels <35 ft. came from purchase slips, which provided limited information on catch and no information on effort.

1. How are the data collected?

Harvesters fill in the daily catch record, providing information on estimated weight caught, location fished, type and amount of gear used, and by-catch details. At te end of the fishing season, harvesters are required to return the logbooks to DFO.

2. Has the design changed over time?

Yes. Over time, additional data fields were added to the logbook (e.g. bycatch of SARA-listed species, fields for recapture detail of tagged cod). These logbooks were provided by DFO for many years. However, now, as with all other logbook programs, these are industry funded.

3. How are the data analyzed?

Median commercial gillnet catch rates were calculated separately for each unit area for years when the directed inshore cod fishery was open. There were insufficient data to produce a time series for other gear types (i.e. line-trawl or hand-line) and there was no directed fishery for cod during 2003-05.

4. Are there estimates of uncertainty?

Yes. For each cell, the 5^{th} and 95^{th} percentiles are used to illustrate the variability around the annual medians.

5. Does this data series provide an index of stock size?

No.

6. How does this data source contribute to assessing stock status?

Provides information on fishery performance, can be used to compare to Sentinel Survey results as well as the inshore fisher phone survey conducted by FFAW.

7. Are the analysis methods (including uncertainty) appropriate?

Reviewers further cautioned that CPUE from harvester logbooks is not a measure of population abundance and should not be applied to models as such.

8. Suggestions for improvements / research recommendations?

N/A

Key References:

Brattey, J., Cadigan, N.G., Dwyer, K., Healey, B.P., Morgan, M.J., Murphy, E.F., Maddock Parsons, D., and Power, D. 2011. Assessment of the cod (*Gadus morhua*) stock in NAFO Divisions 2J+3KL in 2010. DFO Can. Sci. Advis. Sec. Res. Doc. 2010/103. viii + 108 p.

DFO. 2013. Stock Assessment of Northern (2J3KL) Cod in 2013. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2013/014.

APPENDIX 5: REVIEWER REPORT 1

BACKGROUND

An independent review panel was formed to evaluate the data and modeling approaches for Northern (Divs. 2J3KL) cod. The panel members were:

- Dr. Joanna Mills-Flemming, Dalhousie University (Halifax, Canada)
- Dr. Anders Nielsen, DTU Aqua (Copenhagen, Denmark)
- Dr. Chris Legault, NMFS (Woods Hole, USA)
- Dr. Doug Swain, Fisheries and Oceans Canada (Moncton, Canada)

Others were invited to the meeting to provide working papers, information, or discussion regarding Northern cod. The meeting was held 30 November through 4 December, 2015 in St. John's, NL. The guidance provided to the panel regarding our reports was "Each reviewer should provide a description of any strengths or weaknesses of the construction of/analysis of the various data sets, population models presented, and/or the Framework Review process."

This report has sections for the three major aspects reviewed during the meeting: data sets, the Northern Cod Assessment Model (NCAM), and the statistical catch-at-age model (SCA). There are two additional sections on the bioenergetics model presented during the meeting and feedback on the process used for the review. As a reviewer, I have more comments that could be considered negative than positive, but this only reflects my role to provide feedback that can be used to improve the scientific basis for assessing and managing this stock. It is not my intent to denigrate any of the excellent work presented, but instead to provide suggestions for improvement.

DATA SETS

Summary sheets were provided for thirteen data sources for possible use in the Northern cod assessment. These summaries provided the reviewers a broad overview of the data in a consistent manner, which was quite helpful during discussions of the modeling. However, these overviews were necessarily brief to allow review of the modeling. Thus, as a reviewer I cannot provide much guidance on their use or recommend any changes to the way the data are collected.

Some general guidance I can provide is that it is desirable to include as many sources of information as possible within the modeling framework. This allows for direct comparison of the influence of different data sources and identifies the types of flexibility that must be incorporated into the model to allow for a self-consistent explanation of the sources. This is preferred to the practice of only including a small subset of data in the model and then relying on individuals to create stories that relate the un-modeled information to the model results. The inclusion of as many data sources as possible within the model also allows testing of hypotheses regarding the relative importance of different data types or model formulations.

There are some general considerations when considering whether a time series of data can be used as a tuning index in a stock assessment model. One issue is the representativeness of the population that the time series will be compared to in the model. For example, is the spatial extent of the data sufficient to adequately represent the entire population, can the age or length composition of the data be adequately modeled, and do the data link directly or indirectly to the population are questions that should be examined for each time series. Another important component of a time series is the consistency of the methods over time. If the data collection methods differ radically, then the model will be forced to include variability in the catchability parameter, and thereby reduce the influence of the time series as a tuning index in the model. Uncertainty estimates should also be examined because they can be useful in determining how to treat time series with large spikes. Models should be more influenced by observations with low uncertainty than similar observations with high uncertainty. Finally, attempts should be made to model any known sources of bias in the data. This is especially important in catch series when under-reporting or missing components of the catch are suspected or known.

The main trawl survey for Northern cod is currently assumed to have a flat-topped selectivity pattern based on experiments conducted in the 1990s. During the period of the experiments, there were not many large cod available to be caught. As the Northern cod population rebuilds, consideration should be given to repeating some of these important gear experiments to directly test this important modeling assumption. This is especially true if the fishing gear or vessel changes.

The harvester logbooks can be used to estimate CPUE for different fleets. The resulting CPUE estimates should not be used as indicators of change in the population (i.e. a tuning index in the model) because CPUE responds to many factors other than the population (e.g. other fishing opportunities, fishery regulations, the price of gas, and ex-vessel price of cod). These data can provide important information regarding when and where the fleets are operating and their catch rates and so it may be useful to collect the information.

It was not clear from the data summary whether the off-shore acoustic data could be used as a time series of population abundance due to questions about the area surveyed over time. It was also not clear whether the 2015 estimate could be used as an absolute biomass estimate in the stock assessment due to questions about the estimation process used to derive the biomass estimate. Given the apparent discrepancy between the models examined during the review and the acoustic biomass estimate, I support the meeting’s recommendation to have an independent peer review of the acoustic data.

NORTHERN COD ASSESSMENT MODEL (NCAM)

This model was developed by Noel Cadigan specifically for Northern cod accounting for many features specific to this stock. I was impressed with the completeness of the modeling, with even tagging information being incorporated to improve estimation of fishing mortality rates over time. The use of Template Model Builder (TMB) allowed a high degree of objective weighting of the different data sources, with most of the weights determined by the model’s ability to explain the annual changes in the observations. The model is highly flexible and I commend the author for his ability to develop the model in the relatively new TMB framework and his willingness to explore additional formulations during the review meeting.

My main concern with the NCAM results is the unexplained high precision for the estimates of spawning stock biomass given the high degree of flexibility in the modeling framework and lower precision for many other parameters such as M and F. This may just be a result of the current low fishing mortality rate and the fact that annual SSB combines many cohorts, but a CV less than 10% seems to me to be too low given the much higher uncertainty in M and F and the high degree of flexibility in the modeling framework. The reason for this low CV should be found and put forward to avoid similar confusion before this model is used for management advice.

A second major concern is how the natural mortality rate is modeled, especially the implications in the projections. Conceptually, the approach used to model M is fine with M deviations estimated from a “mean” value. However, in practice, the “mean” value is not at all representative of the central tendency of the estimates, with almost all the estimated M values being above the “mean.” In the projections the M values will default back to the “mean” which is not an appropriate measure of central tendency for the time series estimated. This results in the projections increasing rapidly due to M being much lower than it had been previously. My suggestion is to set or estimate the “mean” values so that they better reflect the central tendency of the M estimates over the time series. The author noted during this meeting that this is a relatively easy fix and would be done before the model was used for management advice. The performance of this change, and the model in general, could be evaluated through the use of retrospective forecasting.

The complexity of NCAM is both a strength and a weakness. It is able to incorporate local knowledge about the fish and fishery in customized code. This makes testing hypotheses related to Northern cod easy and straightforward. However, it does require a high level of expertise to run the program. Additionally, customized code can have errors especially when written in a hurry. A demonstration of this was provided during the meeting when a short snippet of code was shown on screen and the author noted an error. This section of code had just been written and not fully tested, but demonstrates the potential for problems if changes to the program are made quickly and not sufficiently tested.

I was surprised to see similar types of data treated differently within the model. For example, the catch data were treated as a total amount with age compositions while some of the survey data were treated as age-specific indices. The author noted some reasons for these differences, but at other times noted that he would do things differently if he had to do them again. So it was not entirely clear if the reviewed formulation is the most appropriate to use for management advice. This issue could easily be addressed through sensitivity analyses to determine whether or not the different treatments result in vastly different estimates. Given the strong signals in the data, I do not expect this to be the case, but it would provide more comfort to demonstrate this through sensitivity analyses than simply make this assumption.

In addition to the sensitivity analysis, there are a few places where NCAM could be modified to better reflect the Northern cod situation. The precision of the fishery age composition, as determined by the variable σ_P, could be estimated separately for the pre and post moratorium years to reflect the different proportion of catch actually observed. In the pre moratorium years the precision would be expected to be higher because a greater proportion of the total catch had observations relative to the post moratorium years when unreported landings and recreational catch, which are not observed, constitute a greater proportion of the total catch. The reporting rate estimated for the tagging data could have its associated uncertainty included in NCAM instead of using just the point estimate as is done currently. This is particularly important due to the demonstrated instability of the reporting rate estimates, which changed noticeably with the addition of one year of data. The conversion of the tagged fish length to age might be better done within NCAM than outside to better represent the cohorts associated with given lengths as opposed to the simple age slicing approach used currently. If it is too unwieldy to incorporate this estimation within NCAM, then a stochastic age slicing process might perform better than the deterministic age slicing currently used.

Despite the NCAM estimation with relatively high precision, I remain concerned about the extreme variability in the natural mortality rate. I suspect that the M variable may be accounting for a number of non-modeled features such as movement of the fish out of the region, unaccounted fishery removals beyond those already modeled, and additional changes in fishery selectivity beyond those modeled. This may just reflect my personal bias though. While M almost certainly does change from year to year in reality, the amount and rate of change estimated in NCAM is well beyond what I am accustomed to seeing without a strong hypothesis to explain it. The large sudden spike in M in the early 1990s that caused the recent population crash is particularly concerning, but there are additional spikes in more recent years that do not have an obvious explanation. Just producing a better fitting model is not a sufficient reason to estimate a wildly changing M, in my opinion. I suggest considering a modeling approach that either links the changes in M to a hypothesized cause, such as capelin abundance, and tests the significance of this relationship over time, or else assume a constant (or nearly so) M and look for other places in the model to account for the changes in the population.

One topic for possible future research is the inability of the TMB approach to account for different levels of uncertainty associated with values within a given time series of data. For example, a survey could have a few years that are quite different from the rest in terms of precision due to incomplete sampling, a single extremely large catch, or highly unusual weather conditions. In the TMB approach, the additional uncertainty in these years is ignored and each year in the time series is treated the same. This has the potential to produce biased results. I provide this topic as future research because there is not an obvious solution other than using subjective emphasis factors, something that the TMB approach tries to avoid. I do not expect this to be an issue for Northern cod specifically, but rather a general research topic.

In my opinion, the modeling approach used in NCAM could be used to provide management advice for Northern cod once the issues noted above are addressed. The approach has a sound scientific basis and provides reasonable estimates for parameters.

STATISTICAL CATCH-AT-AGE MODEL (SCA)

This model was developed by Sean Cox using the basic data of catch and the main bottom trawl survey. It was intended to be a more standard and easier to understand model than NCAM. Despite its relative simplicity, it was still able to incorporate a number of features specific to Northern cod and produced estimates of spawning stock biomass that were quite similar to those from NCAM. It had the added advantage of starting earlier in the history of the fishery and incorporating a stock-recruitment relationship, the latter will be especially important as the stock rebuilds. I commend the author for responding to questions and concerns during the meeting with additional analyses, I found these quite helpful in understanding why there were initially differences between NCAM and SCA results.

The major concern I had with this model is the lack of fit to the fishery and survey age composition data. This problem needs to be resolved before the model can be used to provide management advice. It is not clear to me what is causing this lack of fit. Perhaps the selectivity at age function is too stiff and an age-specific estimation would improve the fits. I'm sure there are many ways to make the model more flexible to produce fits to the data that do not have the large blocks of residuals. These should be explored to see if different approaches result in different estimates of important variables such as spawning stock biomass.

It would be helpful to extend the model to include additional sources of data, particularly additional tuning indices. I would not expect the tagging data to be included in this type of model. As the model was extended during the meeting, it was suggested that some of the parameters became highly correlated (e.g. M on young fish and recruitment). Correlated parameters are always an issue with complex stock assessment models and should be considered more thoroughly before the model is used for management advice to ensure that there are not misinterpretations of results (e.g. changes in recruitment interpreted as changes in M due to high correlations between parameter estimates). Additional sensitivity analyses should also be conducted to explore assumptions such as the survey CV of 0.2, which seems quite tight to me. Given the strength of the signal in the data, I do not expect any radical departures to be found in the sensitivity analyses, but conducting them would clearly demonstrate the stability of the model to these subjective decisions. As with NCAM, I remain concerned about the estimate of M in the early 1990s that shows a large sudden spike.

In my opinion, the modeling approach used in SCA could be used to provide management advice for Northern cod once the issues noted above are addressed. The approach has a sound scientific basis and provides reasonable estimates for parameters.

BIOENERGETIC MODEL

This model was not as thoroughly reviewed as the two stock assessment models. It was presented to provide a hypothesis to explain the high natural mortality rate estimated in both stock assessment models in the early 1990s when the stock decreased rapidly. The bioenergetics model related abundance of Northern cod to capelin, which also decreased rapidly in this region at the same time as cod. A number of lines of evidence were presented to support this hypothesis, including changes in cod condition factor, weight of cod stomachs, and cod diet. However, the model presented ended in 2007 and needs to be updated (in a more modern computer language) to see if the hypothesized relationship continues to hold. A question was raised during the meeting whether capelin abundance had declined suddenly in years prior to those considered in the bioenergetics model and it there was any correlated change in cod at that time. This should be explored to provide additional support to the hypothesis because while cod and capelin abundance is strongly correlated in the time period examined, this correlation could be due to many factors other than direct causation.

PROCESS

The late delivery of many working papers, not until during the meeting for some, made review of the material more difficult than if they had been available prior to the meeting. However, the willingness of presenters and others in the room to discuss issues in detail during the meeting in a frank and open manner allowed me to become comfortable with the stock assessment results.

I found the presentation of two quite different modeling approaches to be highly beneficial. The ability to compare and contrast model estimates improved my comfort that both models were performing well. The strong signals in the data certainly influenced the similarities in the model results, but it was still useful to see that both a new, highly complex model and a more standard model explain the observations in a similar way. I commend the authors of both models for their willingness to openly discuss their results in a non-confrontational manner and explore alternative formulations to help the review panel understand the processes and assumptions of both models.

As a pragmatic issue, as opposed to a scientific issue, I remain concerned that both modeling approaches were presented by authors outside DFO. The lack of DFO authored models at the review demonstrated a capacity issue that needs to be addressed urgently.

I was impressed with the range of experience and backgrounds of the people invited to the review meeting. The discussions were kept on topic and covered a wide range of viewpoints in a professional setting. In my opinion, separating the provision of management advice from the scientific discussions was a major advantage of this meeting and will result in better scientific advice than if this meeting had to produce management advice as well as review the scientific basis for the advice. I thank the meeting organizers for inviting me to participate in this review and hope that my input has been helpful.

APPENDIX 6: REVIEWER REPORT 2

Thank you for your invitation to participate in the Framework Review meeting convened by Fisheries and Oceans Canada (DFO) to review methodology for estimating the population size of the Northern cod (NAFO Divs. 2J3KL) stock. I apologize for being unable to attend these meetings in person. Participating via WebEx proved to be difficult as I missed substantial portions of the discussion due to both the poor quality of the audio component and the complete absence of video capability. Notwithstanding these limitations I learned a great deal during the meeting and would certainly be open to working with your organization in future.

In terms of the Framework Review itself, I believe that external reviewers (like myself) would benefit tremendously from an initial session providing a historical perspective on the stock itself, as well as expert opinion(s) on the quality of previous stock assessments and the goals (with priorities indicated) going forward. While some of this information was indeed made available, it was difficult to absorb during what was a very tight timeline, particularly for the purposes of making useful recommendations. However, I must commend both the organizers and all participants for their constructive yet frank appraisal of the various assessment methods. It is indeed clear that such work is *extremely important* given that the Northern cod stock has supported a substantial fishery across a vast area for centuries. In fact I would strongly recommend that such an exercise be undertaken each and every year.

One of the mandates of the Framework Review was to consider the input data available. The limited information provided in this regard both before and during the meeting made it very difficult to assess these datasets. I was particularly interested in learning more about the spatial resolution of the various datasets (some nicely prepared visualizations would have been very useful here) as well as understanding the unique features of acoustic data suggested to be useful for stock assessment. I fully believe that incorporation of relevant spatial information would prove useful for interpreting results as well as evaluating and enhancing the various assessment methods proposed.

The Framework Review was also asked to consider potential assessment methods for integrating datasets to provide estimates of absolute stock size. These approaches are in contrast to recent assessments of the Northern cod stock that have primarily focused on survey-based estimates. Of the three assessment methods discussed, the NCAM proposed by Dr. Noel Cadigan showed the most promise for true integration. It was also the assessment method with which I am most familiar. I would encourage Dr. Noel Cadigan to ensure that his code is well documented and that his state-space model is described in a way that is more standard in the mainstream statistics literature (clear references to the observation and process equations for example). Though this effort may seem unnecessary at first, I see great value in the NCAM and feel that this would help with both external review and knowledge transfer going forward as I am certain others will wish to attempt similar models. I would also like to see a full sensitivity analysis performed. With the advent of the TMB package the computation time required to run this type of model has reduced dramatically such that a number of ‘competing’ models can be run and compared. Doing so allows one to fully investigate the impact of the various underlying model assumptions, the need for robust enhancements, as well as to more formally quantify the bias and uncertainty associated with both resulting estimates and predictions.

A Statistical Catch-Age Model was discussed during Day 2 of the meeting. Unfortunately, I was unable to follow much of the discussion due primarily to some trouble with WebEx and I also didn’t receive relevant background materials in time to review them thoroughly. As such, I don’t feel comfortable in making any formal comments on this modelling approach.

While I found the Bioenergetics-Allometric Model of Dr. Alejandro Buren nicely presented I believe this model is still in need of refinement. I think it will first be important to ascertain whether the relationship between the cod and capelin populations is causal or simply indicative of a correlation.

Finally, during the Framework Review I did raise the issue of whether the underlying dynamics of the Northern cod have changed so substantially through time that the search for a single assessment method that spans both the collapse and the ensuing recovery (and all relevant datasets) is even realistic. Perhaps it is simply too ambitious to imagine that this could be done and instead a model averaging type approach may be entirely more suitable. I think this is definitely worth contemplation now that extremely fast computational tools are available.

Again, thank you for this interesting opportunity.

APPENDIX 7: REVIEWER REPORT 3

EVALUATION OF POTENTIAL DATA SOURCES

The meeting was asked to evaluate the utility of various data sources that may provide information on the status of Northern cod. Descriptive summaries of these data sources were available to the meeting. However, insufficient detail was provided to fully evaluate these data sources. Nonetheless, many of these data sources have been subject to intensive reviews and discussion in the past. Given the lack of detailed information, my comments are limited to a few key data sources.

Offshore acoustic survey

These data could potentially provide important information on stock status. However, in my view, a detailed review of these data by appropriate experts would be needed to determine whether and how these data could be incorporated in the assessment. Key questions from my perspective include:

1. How well are the species and size compostion of acoustic aggregations characterized by trawling? What were the criteria used to determine when trawl samples were obtained? How frequent and variable were the trawl samples? What is the evidence that the species and size composition of trawl samples reflected that of fish higher in the water column?
2. How were densities in the acoustic “deadzone” determined? What assumptions were used to make this determination? What evidence is there to validate these assumptions.
3. How are acoustic signals converted to biomass estimates?
4. Should biomass estimates be used as relative or absolute measures of biomass? (Information reported to the meeting indicated that acoustic biomass is often used to provide relative indices when used in stock assessments. Model estimates of catchability can be considerably greater than one.)

Given the lack of detail provided to the meeting on the offshore acoustic data, and the lack of acoustic expertise among the meeting participants and reviewers, my view is that it is essential that these data be subject to a detailed, independent review by stock assessment and acoustic experts before being used to assess stock status of Northern cod.

Offshore RV survey

These data are a key assessment input. A number of issues were raised with respect to this data source:

1. Missed strata: Not all strata have been covered in all years. However, the proportion of missing “index” strata used to assess cod status was reported to be very low except in 2004. A solution would be to omit the 2004 RV indices, as was done in the updated NCAM model (Cadigan 2015).
2. Changes in survey timing and possibly in the timing and extent of offshore migration in the fall may affect the availability of cod to the survey. This issue appears to be important in 1995-2009, when a reduced proportion of the stock appeared to overwinter offshore. This can be accommodated in the population modelling (Cadigan 2015).

Sentinel Gillnet Index

Time trends in age-aggregated indices were inconsistent between the RV and sentinel indices. This may be related to changes in cod availability to and/or survey design of the sentinel fishery. For example, there has been a substantial decline in the number of sites fished and the duration of fishing in the sentinel program. As a consequence, standardized catch rates in the sentinel program may be biased in recent years relative to those in earlier years. This could be investigated by examining the strength of site x year and month x year interactions in the sentinel data. If these interactions are strong, then the sentinel index may not provide a consistent index for use in population modelling.

DFO tagging data

These data are an important assessment input. They provide:

1. information on fishing and natural mortality if rates of tagging mortality, tag loss and tag reporting are estimated, as is the case here.
2. information on the relative importance of recreational and commercial fisheries (if reporting rates are reliably known for both fisheries).
3. information on migration patterns.

POPULATION MODELLING

Northern Cod Assessment Model (NCAM)

This model addresses major difficulties in assessing the Divs. 2J3KL cod stock:

1. changes in natural mortality,
2. unreported catch, which is known to be high in some periods, and
3. a non-stationarity in the RV survey data in the 1995-2009 period. This non-stationarity is addressed by hypothesizing that it reflects a decline in the proportion of the stock overwintering offshore during this period. This proposed decline in availability to the RV survey coincided with the occurrence of an aggregation of cod overwintering inshore in Smith Sound. The abundance of “offshore” fish (i.e., fish that normally overwintered offshore) overwintering inshore was estimated based on acoustic estimates of annual abundance of the Smith Sound aggregation. Age-composition of these fish was estimated from trawl samples (not available after 2003).

A strength of the model is that it integrates data from a variety of sources, including standard population modelling inputs (e.g. survey indices, fishery catch and age composition) as well as a large tagging database and acoustic data (e.g. on the Smith Sound overwintering aggregation). Incorporating the tagging data may be a particularly important feature of this model, providing important information on fishing (F) and natural (M) mortality. This additional information may help to distinguish between unreported catch and M (e.g. previous modelling work on estimating unreported catch used assumed values of M [Bousquet et al. 2010]), and may contribute to the model’s ability to estimate the variance of process error in M.

Another strength of this model is that it is, to a large degree, “self-weighting”. One feature that contributes to this is the use of the continuation-ratio logit to model proportions at age in the catch. This approach does not require specification of “effective sample samples”, which can have a large influence on the results. Another important feature in this regard is its ability to estimate the process error variance for M.

I agree that zero values for survey indices can provide information on abundance that should be retained for the model. When stock size is low, these zero values may represent low abundance values below the detection limit. I think that the approach used in NCAM to deal with these zero values is a good one (and hope to employ it in future modelling).

While more self-weighting than most stock assessment models (that I'm familiar with), some subjectivity persists (which I think is often necessary). For example, fixed values are used for the measurement error variance for catch, upper (and sometimes, lower) multipliers for the bounds on total catch, and age adjustments to the variance of the catch age compositions. Similarly, catchability (*q*) to the RV survey was constrained to be equal for ages 6-14, and assumed values for q were used for the trawl used to obtain age compositions of the Smith Sound fish. Using fixied values for these catchability parameters could be avoided by fitting selectivity functions to the data. Flat-topped selectivity could be fit using only two parameters, thus requiring no increase in the number of model parameters.

Sensitivity analyses could be used to examine the robustness of results to the fixed values used in the modelling. One sensitivity analysis was presented, using different values for the upper and lower bounds on catch. This analysis demonstrated that the stock biomass depended on the values used for these bounds, but that stock status relative to the Blim reference point was robust to changes in these values.

This model provides a good fit to the data. There is some "blocking" of residuals in the RV survey age compositions, but the residual patterns are not extreme. In my view, this model would be a good tool to estimate the stock status of Northern cod. Estimated stock size would be dependent on the bounds selected for catch, but stock status (relative to Blim appears to be robust and reliable regardless of the values chosen for these bounds. Model results are influenced by the assumption that the overwintering cod aggregation in Smith Sound in 1995-2009 represents cod that would normally overwinter in the offshore where they would be available to the RV survey, and model structure would need to be revised if a different hypothesis were preferred. Results are influenced by the acoustic biomass estimates for Smith Sound, which have not been adequately evaluated to my knowledge. Age composition data for this inshore aggregation is limited. This model also provides a reasonable approach for incorporating offshore acoustic estimates of cod biomass, if incorporation of these estimates is desired.

NCAM also provides a good tool for stock projections, given some modification. Currently, *M* is modelled based on *M* deviates from initial values in 1983 (*M*=0.5, 0.3 or 0.2 depending on age). *M* deviates are assumed to have a mean of 0. However, M deviates are almost all positive. This will have undesirable effects in projections (i.e. *M* gradually declining to the fixed "initial" values rather than to the longterm mean). To address this problem, *M* should be modelled based on deviates from the longterm average.

Statistical Catch at Age (SCA)

SCA is a commonly used assessment model. The data inputs were an age-aggregated RV index, total fishery catch and age-compositions of the fishery and RV catches. Unlike NCAM, this model did not incorporate tagging data. The initial formulation also did not incorporate bias in the catch data (i.e. allowing true catch to exceed reported catch), but a second model formulation did allow for this. Also unlike NCAM, the model did not incorporate the hypothesis that a portion of the cod that normally overwintered offshore instead overwintered in Smith Sound in the 1995-2009 period. Instead, the model attempted to account for an apparent loss of older fish in the offshore during this period through elevated *M*. The model was run estimating *M* since 1959 or estimating *M* since 1983 with earlier values of *M* assumed to be equal to the 1983 values. There is little information available to the model to estimate *M* prior to 1983 (the earliest year with survey data), and in my view the preferred formulation estimates *M* beginning in 1983. Despite the differences between the SCA and NCAM models, the two models estimated similar levels and trends in cod biomass and similar status relative to Blim (based on the SCA model which estimated *M* beginning in 1983).

A strength of the SCA model is that it provides a long-term historical view of this stock, with biomass estimates extending back to 1959 (compared to 1983 for NCAM). It is important to maintain this long-term perspective in order to avoid a “shifting baseline” in our view of the potential productivity of this stock.

Like NCAM, this model incorporates process error in *M*, an important feature in modelling the dynamics of this stock. The modelling of this process error differed somewhat between the two models. In NCAM, *M* for age *a* in year *y* was modeled as deviations from the initial value assumed for *M* of age *a* in 1983. In the SCA model, *M* for age *a* in year *y* was modeled as deviations from *M* for age *a* in year *y*-1. Instead of using fixed values for $M_{a,1983}$ (as was done in NCAM), *M* in 1983 was estimated with a prior provided (*M* was assumed to equal the 1983 values in earlier years). In both models, *M* deviations were assumed to have a mean value of 0. While estimated deviations likely had means near 0 in the SCA model (this result was not shown to my recollection), the means were considerably greater than 0 (with almost all values positive) in NCAD. This issue could be avoided in NCAD by modelling process error in *M* as deviations from the longterm average. This issue is not expected to affect NCAM estimates up to the present but will affect projections.

The SCA model also included a Beverton-Holt stock-recruit model with process error. This could be a useful feature for projections as stock size increases. While NCAM lacked this feature, treating recruitments as uncorrelated lognormal random variables, NCAM could presumably be modified to incorporate a stock-recruit model if desired.

Like NCAM, the SCA used a self-weighting approach for calculated the age composition likelihoods.

Unlike NCAM, the SCA model uses fixed values for the variances of observation error in the RV survey and process error in *M*. In this regard, there is greater subjectivity in the SCA modelling. However, observation error variance is normally estimated in SCA (and VPA) models, and I expect that this will also be the case for this SCA model once it has progressed further from its current preliminary stage. On the other hand, estimating appropriate values for process error variances can be problematic. Values estimated to maximize the log likelihood can be too small to eliminate severe residual and/or retrospective patterns, symptoms of non-stationarity in the observation or process models. In this case, I would argue that it is preferable to use fixed values that are large enough to account for the non-stationarity (e.g. Swain and Benoît 2015). The ability of the NCAM model to estimate an appropriate value for process error variance may reflect the large amount of information on *F* and *M* provided to this model by the tagging data.

While the model fit to the age-aggregated survey index was good, there were strong residual patterns in the survey and fishery age compositions at ages 4-6, at least in early formulations. This problem should be addressed before using this model as an assessment tool for 2J3KL cod. This problem may be related to a decline in availability of older fish to the RV survey in 1995-2009 (when these missing fish are thought to have overwintered inshore). The SCA model interprets this decline in older fish as a consequence of elevated *M*. This may create an inconsistency between the survey and fishery age compositions because the fish unavailable to the offshore RV survey in late fall are available to the inshore fishery in summer. A possible solution would be to incorporate an age-dependent change in availability to the RV survey in 1995-2009, like in the NCAM model.

While no projection was presented for this model, the SCA model could be used to make projections.

Conclusions

1. The NCAM model was custom built for the 2J3KL cod stock, including many features specific to this stock and addressing assessment difficulties posed by this stock. This model integrates a wide range of input data, and is in my view a powerful tool for assessing the status of 2J3KL cod. It is however based on a particular hypothesis for the cause of low abundance of older cod in the RV survey catches from the mid-1990s to the mid-2000s (i.e. overwintering inshore by cod normally migrating offshore in autumn). This is a plus if support for this hypothesis is strong. Also a minor modification to how non-stationarity in *M* is modelled should be made before using this model for projections.
2. The SCA model is a commonly used model adapted for use on Northern cod. It provides a long-term perspective on this stock, extending back to 1959. It requires less expertise to run than the NCAM model. On the other hand, it integrates less data (e.g. no tagging data) and does not incorporate the hypothesis that availability of older cod to the RV survey was reduced when a high proportion of these fish overwintered inshore in 1995-2009. While this model shows promise as an alternate assessment model for 2J3KL cod, some lack of fit issues need to be addressed before this model should be used for this purpose.
3. Incorporating time-varying *M* is a key feature of both models which is essential for the assessment of this stock.
4. Despite their different structure and complexity, both models reach similar conclusions regarding biomass trends and status relative to B_{lim}. This strengthens the confidence in the robustness and reliability of these conclusions.

ECOSYSTEM PERSPECTIVE

Both models indicated wide variation in *M* of 2J3KL cod over the past 30 years. While not strictly necessary to assess stock status relative to reference points, understanding causes of variation in *M* is important for understanding sources of variability in productivity and for informing management decisions. A bioenergetic model was presented linking variation in cod productivity to the availability of a key prey, capelin. This provides a possible cause of the variation in *M* estimated by the assessment models, and is useful information for presentation at stock assessment meetings. Ideally, a suite of hypotheses for causes of elevated *M* should be examined and their relative support weighed. Other hypotheses that could be examined include the unusually cold ocean conditions in the late 1980s and early 1990s and misspecification of unreported catch as *M* in the assessment models.

REFERENCES

Bousquet , N., Cadigan, N., Duchesne, T., and Rivest, L.P. 2010. Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Can. J. Fish. Aquat. Sci. 67: 1247−1261

Cadigan, N. 2016. Updates to a Northern Cod (*Gadus morhua*) State-Space Integrated Assessment Model. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/022. v + 58 p.

Swain, D.P. and Benoît, H.P. 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Mar. Ecol. Progr. Ser., 519, 165-182.

APPENDIX 8: REVIEWER REPORT 4

EXECUTIVE SUMMARY

The framework meeting was not based on a collected assessment report. Instead the assessment meeting was based on data summary sheets and working papers. Some documents only became available during the meeting, which was not optimal, but the authors did a great job of presenting everything in a very understandable way. Two good assessment model candidates were presented — each with strengths and weaknesses, as outlined in the report. The model NCAM is tailored to include many of the data sources available, and to allow options specific to the Northern cod stock. The SCA model is closer to a standard model used for cod stocks elsewhere, and only include standard data sources. The results from the two models were overall fairly similar, which strengthens confidence in both models. The framework meeting was a pleasure to attend with many great discussions about interesting modelling approaches.

DESCRIPTION OF THE REVIEWER'S ROLE

This reviewer has independently read the assessment documentation, traveled and participated actively in the framework meeting, identified key issues in the assessment, and independently authored this review report.

FINDINGS FOR EACH TERM OF REFERENCE

1. **Evaluate the utility of data sources that may provide information on the stock status of Northern cod. If possible, examine the uncertainty in the data and the methods used to estimate measures of uncertainty.**

The framework meeting was setup to focus on assessment models. A short overview of the potential data sources was given in the form of: history, collection and processing method, changes over time, and for some sources, an evaluation of their utility to inform assessment models about the stock. The presented data sources are:

Table 3. The available data sources. X indicate that the data source is used in the assessment model mentioned above. The X in parenthesis indicate that the acoustic offshore data is only included in a hypothetical test run

Data source	Used in NCAM	Used in SCA
Commercial catch at age	X	X
DFO RV survey	X	X
Recreational Fishery	-	-
Acoustic offshore	(X)	-
Acoustic Smith Sound, Trinity Bay	X	-
Satellite tags	-	-
Tagging database	X	
Telemetry database	-	-
Sentinel survey	X	-
Spanish survey	-	-
Inshore mobile	-	-
Beach seine survey	-	-
Inshore harvest questionnaire	-	-
Commercial logbook	-	-

The data source presentation was helpful and started discussions on how these data sources could, or could not, be used to inform the assessments, and how to avoid misinterpretation of the different sources. Actual data in form of figures or tables were not presented.

A key input for many analytic assessment models is the total catch, and the age composition of the total catches. The discussion revealed that there are unusually many difficulties in obtaining and trusting this data. The difficulty comes from:

1. There is a potential for misreporting. The commercial catches are - to some extent - self-reported. There are landing sites which are not monitored, and fishermen could get information (e.g. via phones to avoid monitoring).
2. Discarding and high-grading at sea of the catch could also be an issue.
3. Missing recreational catches. Recreational catches could be a large fraction of the total removals, and no estimates of those are available. Returns of conventional tags suggest that the recreational catch could be as much as 40% of the total catch.

The DFO RV survey is the main standardized fishery independent index of stock size. It has been reviewed, and in 2013 the assessment was based solely on data from this survey.

A much debated data source at the framework meeting was the acoustic estimates (Offshore and in Smith Sound, Trinity Bay). Acoustic estimates can potentially provide an estimate of the absolute spawning stock abundance within the survey region. The acoustic estimates have not been subjected to a formal review procedure, and the framework meeting recommended that the acoustic estimated are evaluated at an independent review meeting. Such a review would carefully evaluate to what extent the acoustic estimates are describing the actual spawning stock biomass, and potential sources for biased estimates to be aware of.

The conventional tagging data can provide valuable spatial information, as well as some indication of the fraction of recreational catches relative to commercial catches. Even if this fraction is not directly used in the assessment models it can still help validate and justify the model assumptions.

The data from the satellite tags need to be further processed before the data can enter the assessment models directly. Analyzing satellite and Vemco tags is challenging, but also a great opportunity. At DTU-Aqua we have used these geolocation problems to introduce students and young researchers to non- standard modelling. It is much simpler than a full assessment model, but much the same toolbox is needed.

The Fish Harvester Questionnaire is data from a telephone survey. It is not currently used directly in the assessment. An example question is "Compare the abundance of cod in your area during 2013 with that of 2012". Consider how this information could be utilized in an assessment context. Let X_i be the response from the *i'th* harvester (1 if more cod more abundant, 0 if not). Then this could be included in an assessment model as:

$$X_i \sim Bin(1, p), \text{ where: } \text{logit}(p) = \alpha + \beta \Delta \text{SSB}_{12-13}$$

This would reflect that the harvesters would be more likely to answer "more" if underlying abundance had increased a lot. This is just mentioned as an example.

4. Is the structure of the model and the estimation techniques employed sufficient (i.e. robust and reliable) to estimate the stock size of Northern cod?

Noel Cadigan's Northern Cod Assessment model (NCAM)

NCAM is a state-space assessment model, which is a model type used to assess most commercially dominating stocks in ICES. The model type has many benefits compared to the other commonly used alternatives, which are deterministic and full parametric assessment models. The main benefits are that state-space assessment models allow:

1. flexible time- varying processes to be described with very few model parameters;
2. statistically rigorous quantification of uncertainties; and
3. within model framework for prediction.

NCAM was tailored specifically to match the data available for Northern cod. It already includes more data than any other model presented for Northern cod. In addition the state-space framework is very flexible, so it should be reasonably easy to include more plausible data sources. The model is (mostly) self-weighting, so the different data sources are weighted to the degree they provide information by an objective criteria. This makes the model a good candidate to extend by including new data, because - simply put - if useless data is added the model will simply down weight the data and not make the results worse.

The model is not a "canned" software package, so using and extending this model will require some expertise in state-space assessment modelling, and in the basic software package in which the model was developed (Template Model Builder, which is a tool similar to the more known AD Model Builder). Noel Cadigan has agreed to run the model for a few years forward. It should be feasible to setup a configurable user interface for the model (e.g. as is done for other models at http://www.stockassessment.org) if there is a need for it.

A key feature of NCAM is the censored observation likelihood used for the total catch. Instead of assuming that the reported catch is an observation of the true catch it is assumed that the true catch is above the reported catch but below a set upper limit (e.g. 2 times reported catch). In addition to the interval width a variance parameter is set, which determines the tolerance for catches outside the specified limits. This is a reasonable approach as the catch data is likely a lower limit, because there is potentially both under reporting and discarding and because the recreational catches are not included. A couple of comments to this approach.

The choice of interval width and variance parameter is subjective and will influence the results by restricting the estimate of true catch and by changing the relative contribution of catches versus other data sources in the model fitting criteria. It is important to carry out sensitivity runs with regards to these settings.

The model uses random effects, which requires integration. The integration is carried out via the Laplace approximation. The Laplace approximation is accurate for linear Gaussian models, but useful for a much wider class of models. The distribution for the censored catches is non-standard and it should be investigated if the Laplace approximation is valid approximation here. It is suggested to set up a simpler example with a censored likelihood and validate the Laplace approximation against an MCMC or importance approach.

An alternative approach to the censored likelihood could be the following. Instead of modelling the total catch per year and age compositions separately, the observed catch at age C_{oay} could be used. A yearly scaling parameter S_y could be estimated such that

$$\log C_{oay} \sim N(\log(S_y C_{ay}), \sigma^2)$$

If desired the S_y parameter could be bounded between 1 and 2. This approach should give a similar effect, but possibly be easier to solve in terms of the Laplace approximation, because the scaling parameter is moved to "outer problem" and hence fixed when the Laplace approximation is performed.

The current censored model assumes that the age distribution is the same in the unobserved part of the catch, as in the monitored part of the catch. If a large fraction of the true catch is not monitored it is likely that the age distribution of the true total catch is different than the monitored age distribution. The alternative approach outlined above would possibly allow for separate scaling factors for groups of age classes, which would soften the assumption about same age distribution.

The sensitivity analysis demonstrated (e.g. of interval width) looked convincing as quantities of interest changed very little, but further sensitivity analysis would be useful. Things to consider investigating could be:

1. The fixed censor variance describing the tolerance for catches outside the censor interval (expected not to be important).
2. The fixed correlation parameters.
3. The independence assumptions of the I-at-age.
4. The assumptions about catchabilities.
5. The fixed ratios between variance parameters (e.g. 2:1:3). When looking at these sensitivities it is of interest to look at the relevant estimates, but also at the estimated uncertainties.

The model has evolved a lot since the 2015 paper (Cadigan and Marshall 2015). It now uses more tagging data, and has increased in complexity, so it would be relevant to update the self-simulation test to ensure that the model parameters are still identifiable, and that no simple mistakes have occurred.

Residuals are important to diagnose if the model is describing the data well, but residuals are more tricky in state-space models compared to conventional purely parametric models. The standard residuals based on observations minus the prediction based on all data does not give independent residuals. The authors presented the standard residuals, which is what is normally done in assessment context. To get independent residuals two approaches are available: The one-step-ahead approach (Harvey 1989), and the single-joint-sample approach (Waagepetersen 2006). Both methods are a bit non-standard in assessment context, so outlined briefly here. Details and code examples can be obtained from this reviewer.

One-step-ahead residuals: For each observation y_i, $i = 2 \ldots n$

- Use model to predict yi ONLY from observations $y_1 \ldots y_{i-1}$
- Use residual:

$$\sum_{i|i-1}^{-\frac{1}{2}}(yi - \hat{y}_{i|i-1})$$

Single joint sample: For given model parameters θ

- If (Y,U) is distributed according to joint pdf.$\sim L(y, u)$
- Observed y is then a sample from marginal distribution with pdf. $\sim \int L(y,u)du$
- Generate one sample $u*$ from conditional distribution of $U|Y = y$
- Then the set $(y, u*)$ is a sample from joint distribution of (Y,U)
- Assumed distribution of $u*$ can be validated by standard tests

There are some retrospective patterns. It is not uncommon in assessment models, and is often caused by some parameter that is kept time-invariant in the model, but is time varying in the true system. It would be useful, but likely also difficult, to identify the cause of this retrospective pattern.

Since NCAM uses so many different data types it would be possible to run the so called leave-one-out diagnostics, where a number of runs are conducted where one of the data sources is left out completely (e.g. a run without the 'Acoustic Smith Sound, Trinity Bay' data, and a run without the 'Sentinel survey', and so on). When comparing each of these runs to the run with all data it is possible to get a sense of how much, and in what direction, each data source is pulling the important estimates.

The structure of NCAM appear sufficient and the estimation techniques employed appear sufficient (i.e. robust and reliable) to estimate the stock size of Northern cod. It should be investigated if the Laplace approximation is appropriate for this model.

Sean Cox's Statistical Catch at Age model (SCA)

An alternative Statistical Catch at Age model was presented for the Northern cod stock. The model uses the total commercial catch, its age composition, the RV survey catch, and its age composition. The model is closer to standard models used for cod stocks elsewhere. The SCA model is quite different from the NCAM model both in the data used, but also in the way the model is set up. It strengthens confidence in both approaches to see that the overall results were similar.

SCA is able to use data all the way back to 1959, which is useful in providing a longer historic time series. The model formulation has a number of subjectively fixed model parameters (e.g. recruitment deviance penalty, natural mortality random walk variances, initial N deviation, variance of catch, variance of survey biomass). This is not unusual in assessment models, but this model has a high level of such fixed parameters.

Validating how well the assessment model is describing the observations is important, and especially so when some of the model parameters are fixed. A key plot for doing so is the standardized residual plot, where the difference between observed and model predicted, divided by the estimated or assumed standard deviation, is plotted. The residual plots presented for the age compositions are not optimal for validating the model fit. For the older ages (above 10 years old) all the plotted symbols are so small that it is impossible to see if the residuals are positive or negative. At the meeting it was confirmed that these were supposed to be standardized residuals, but that the standardization was not correct. It is important get these standardized residuals corrected, because it is not possible to evaluate the fit to the observations, and especially the assumed variance structure without it.

For the younger ages the residual plot shows a systematic pattern with regards to age, which could indicate that either the age-specific mortality, or the shape of the fishing selectivity is too restrictive.

Fixing model parameters (instead of estimating them) may influence the results, and it is important to investigate these effects, but since it is mainly variance parameters that are fixed the greatest effects are expected to be seen in the uncertainties of the quantities of interest (e.g. SSB, F, and R). Uncertainty estimates are obtained by propagating the assumed (or estimated) uncertainties of the observations - through the estimation routine - to the final estimates of interest. It is therefore especially important to be able to validate the assumed variances (residual plot), and to consider the uncertainty estimates also when carrying out sensitivity analysis.

The residual pattern suggested that a more flexible selectivity pattern should be investigated, but also the year blocking could be looked at. One suggestion would be to use a finer-scaled blocking in the periods where the fishing mortality is changing a lot.

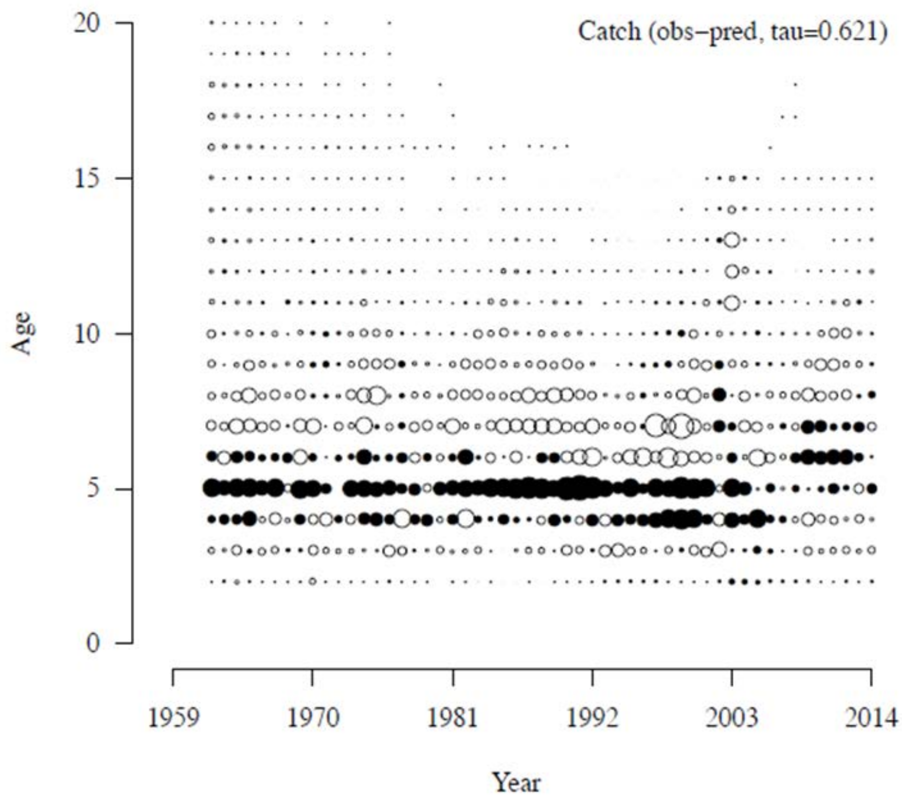

Figure 1. Example SCA residual plot.

At the framework meeting some sensitivity analysis were shown. One of them showed a decoupling of the natural mortality (M) of ages 4 and 5. The result was a bit unexpected, as one of the M-at-age series showed a trend nothing like its neighboring age classes. This reviewer suspects that using only catch and the DFO RV survey, as SCA does, leaves SCA with very weak information to estimate M from. In the period around the moratorium there should higher contrast, and hence more information about M, but outside this period very little. This means that the estimated M pattern relies more on the process assumptions, the fixed variances, and the coupling between ages. Two comments to this:

1. When deciding to couple or decouple random walks with regards to ages it is does not have to be an on/off decision. Another option is to setup the age-specific processes as correlated processes, which allows to scale - via a single correlation parameter - from independent (decoupled) to any degree of correlation including correlation of one (coupled). The optimal correlation is rarely zero or one.
2. Self-simulation test is important to see if a simulated true M-process can in fact be recovered from this amount of data.

The structure of SCA appear sufficient and the estimation techniques employed appear sufficient (i.e. robust and reliable) to estimate the stock size of Northern cod. The residual issue should be solved and the model should pass the self-simulation test.

Bio energetic model

The bio energetic model presented is not an assessment alternative to estimate the stock size of Northern cod. It did however demonstrate a plausible link between the capelin stock and the cod stock. The availability of capelin could potentially be correlated with the natural mortality of cod. It would be interesting to investigate if this connection could be useful in providing more accurate predictions for the cod stock.

3. **Are proposed projection methods for evaluating future catch options sufficiently robust and reliable? Are any additions or modifications to methods or inputs required before providing advice to Fisheries Managers?**

The proposed projection methods for evaluating future catch options, were fairly standard running of different catch scenarios, and hence judged to be sufficiently robust and reliable.

As the stock rebuilds the NCAM model should properly consider replacing the two level recruitment function with a proper stock-recruitment relationship. In projecting from both NCAM and SCA it should be carefully considered what level (or distribution) of natural mortality (M) is used projecting forward. This is not an easy question, as great changes in M has been seen in the past.

4. **Recommend priority short- and medium-term research areas to improve data sources, assessment model formulation and estimation, and projection methods.**

A lot of research topics are listed above, most of which are medium-term. The list below is what this reviewer considers short-term:

- The validity of the Laplace approximation should be checked in the NCAM.
- The correctly standardized residuals should be produced from the SCA model.
- The SCA model should be self-simulation tested.

5. **If reference points have not been estimated, provide direction for future work towards estimating Precautionary Approach reference points (both biomass and removal references).**

The framework meeting was focused assessment models. The main reference point used was the DFO PA biomass limit reference point B_{lim}, which is defined as average SSB from 1983-89.

Both NCAM and SCA produce all the standard output from an assessment model, so all standard reference points can be calculated from the output of these models (including MSY reference points). MSY reference points depend on the estimated stock-recruitment relationship, the selectivity, the natural mortality, and how these things are projected into the future. Noel Cadigan presented stochastic MSY calculation to illustrate the uncertainty related, and to show the effect of using a spatial stock-recruitment relationship.

6. **Acoustic surveys have recently been conducted by the Centre for Fisheries Ecosystems Research at Marine Institute of Memorial University. (Data from these surveys are not currently available.) Consider extensions to the model structure that can incorporate these data as an index of stock size.**

Both SCA and NCAM are sufficiently flexible to incorporate data such as an index of stock size. It was demonstrated in the NCAM model by extending the model to include a hypothetical stock size estimate.

The bigger question is how to correctly interpret the observations from these acoustic studies. The framework meeting did not include independent acoustic experts, so the recommendation from the framework meeting was that a separate review meeting focused on the acoustic observations should be held before these data are used in the assessment models.

7. If time permits, discuss the extent to which the method (and underlying computer software) is amenable to conducting Management Strategy Evaluation (MSE).

An impressive MSE package was presented by Sean Cox. It gave a good impression of what could be done. Both NCAM and SCA could potentially be used in a MSE framework. Using SCA would be simpler than using NCAM, because the data setup is simpler and because the model estimation time is faster.

References

N.G. Cadigan, C.T. Marshall. 2015. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 72 (999): 1-13

A. C. Harvey 1989. Forecasting, structural time series models and the Kalman filter. Cambridge University Press.

R. Waagepetersen 2006. A simulation-based goodness-of-fit test for random efects in generalized linear mixed models. Scandinavian journal of statistics, 33(4):721-731.

APPENDIX 9: NCAM POST-MEETING ANALYSES

A small error was discovered in NCAM, related to how much M was applied to tagged fish in the year they were released. Fixing this error had little impact on the NB NCAM results (see Cadigan 2015). Results for two other model formulations were not reviewed during the framework meeting but are provided here completeness.

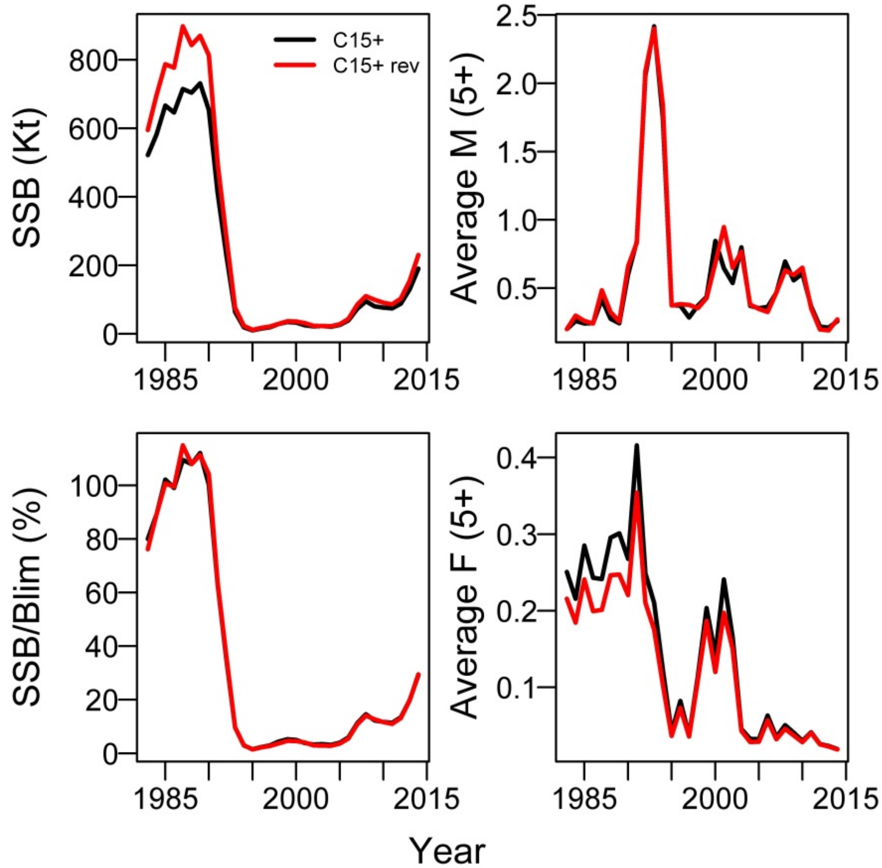

Figure 2. A comparison of mature biomass (top-left panel), stock status relative to B_{lim} (bottom-left panel), average M (top-right panel), and average F (bottom-right panel) during 1993-2014 for the update (C15) NCAM formulation and the revision (rev) with a change in how M was modelled for tagged fish in their year of release.

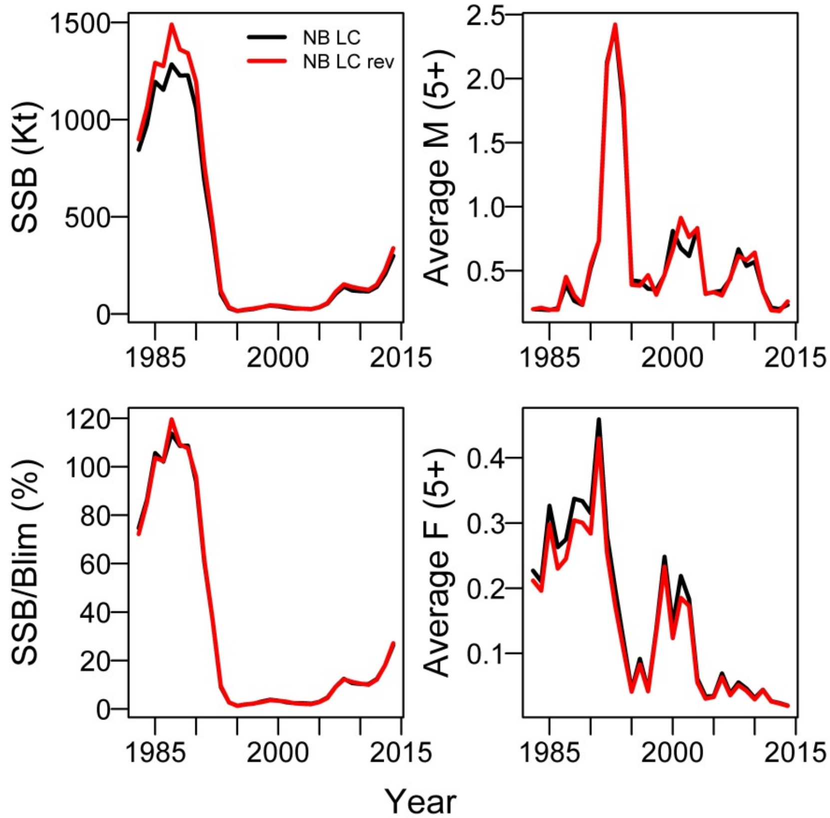

Figure 3. A comparison of mature biomass (top-left panel), stock status relative to B_{lim} (bottom-left panel), average M (top-right panel), and average F (bottom-right panel) during 1993-2014 for the NB NCAM formulation with higher bounds (LC) on total catch weight and the revision (rev) with a change in how M was modelled for tagged fish in their year of release.

APPENDIX 10: SCA POST-MEETING ANALYSES

Reviewer comments mainly focused on the lack-of-fit of the SCA model to the age-proportion data, especially the consistent over-estimation of age-5 and age-4 proportions in the fishery and RV data, respectively. The age-proportion likelihood for the SCA model was revised to reduce the influence of very small age-proportions that were common in the catch-at-age matrices. For each year, I used a threshold proportion of 0.005 as the smallest proportion allowed in the likelihood. This was implemented by (i) for each year (t) identifying the youngest age (a') for which p(a,t) < threshold; (ii) summing the proportions for all ages > a' for year t; (iii) saving the maximum age a'; (iv) using a' as the plus group age in year t to compute the model predicted age proportions in Eq SCA.15. Maximum ages ranged from 8-15 and 6-10 for the fishery and RV data, respectively.

The revised treatment of the age-proportions improved the fit to the age-composition data substantially and reduced the residual error variance (see standard errors in Figs. 5 and 6 legends). As requested by the reviewers, the strong age-dependency in residuals for both data sets was removed.

Although the effects on current biomass estimates were small, the effect on historical estimates and current status were larger, as illustrated in the following plots.

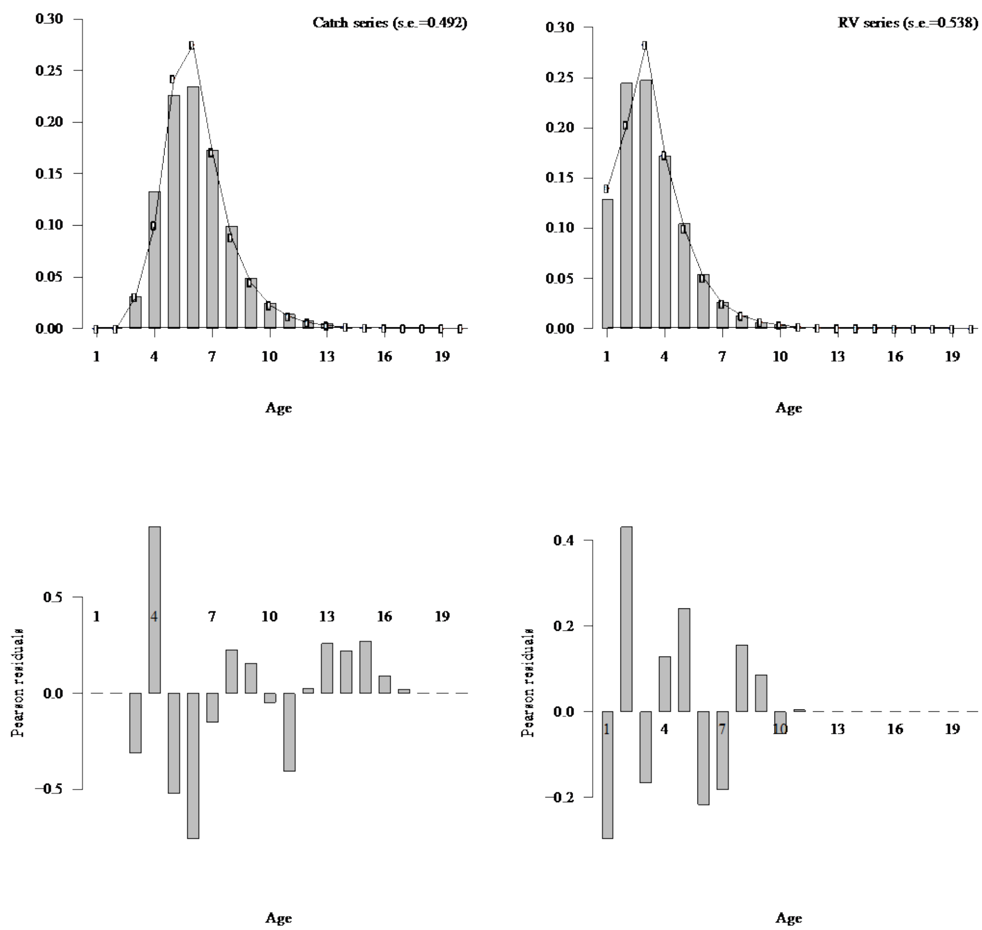

Figure 4. Residual plots for aggregated age-proportions for the revised SCA model.

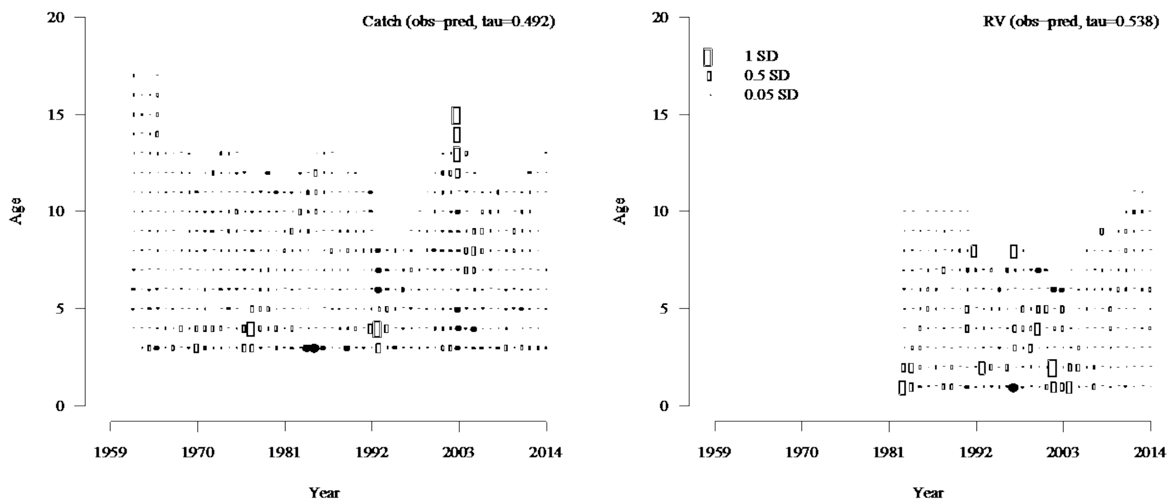

Figure 5: Residual plots for dis-aggregated age-proportions for the revised SCA model. The blank areas indicate the areas in which observed age-proportions were less than 0.005.

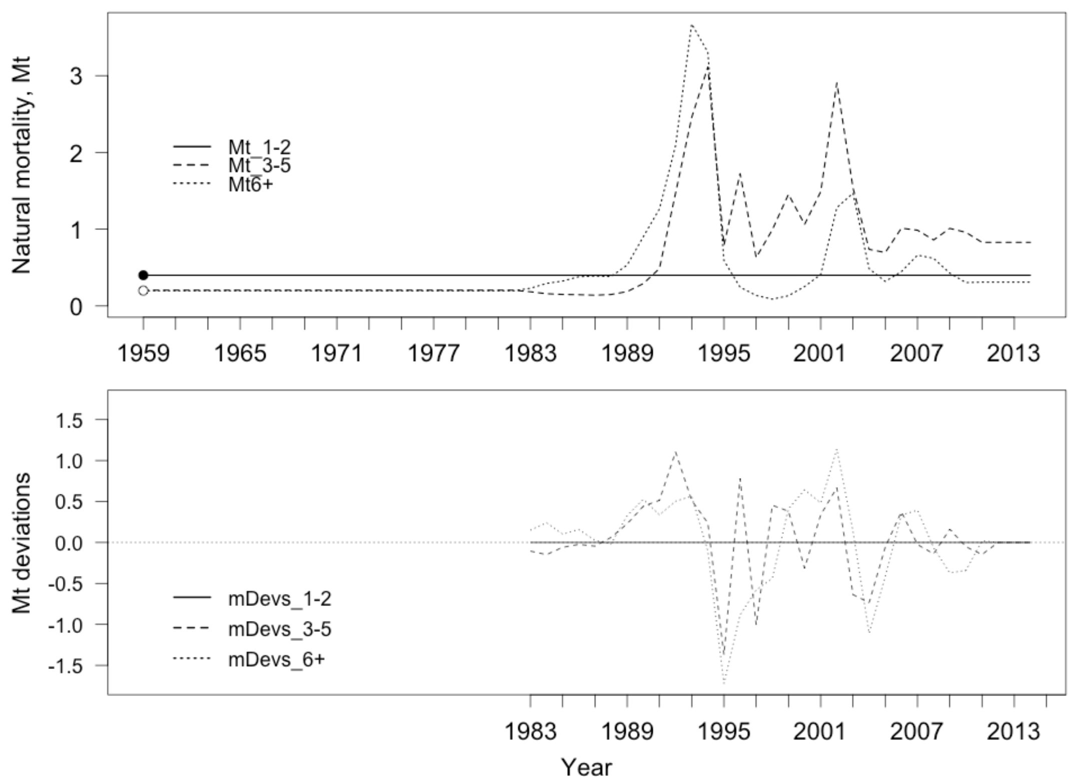

Figure 6. Revised estimates of natural mortality.

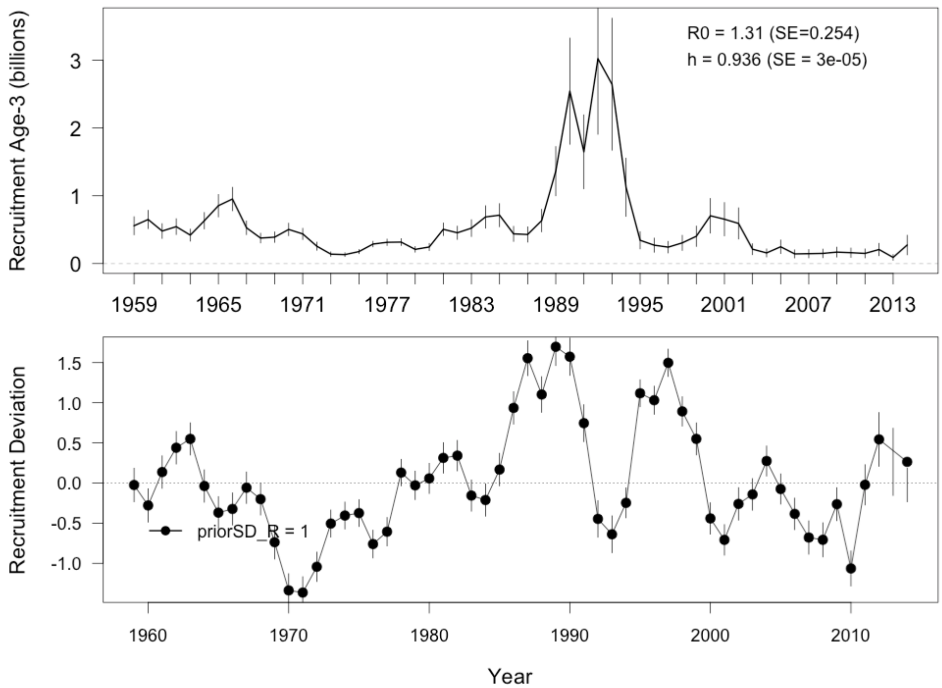

Figure 7. Revised estimates of recruitment.

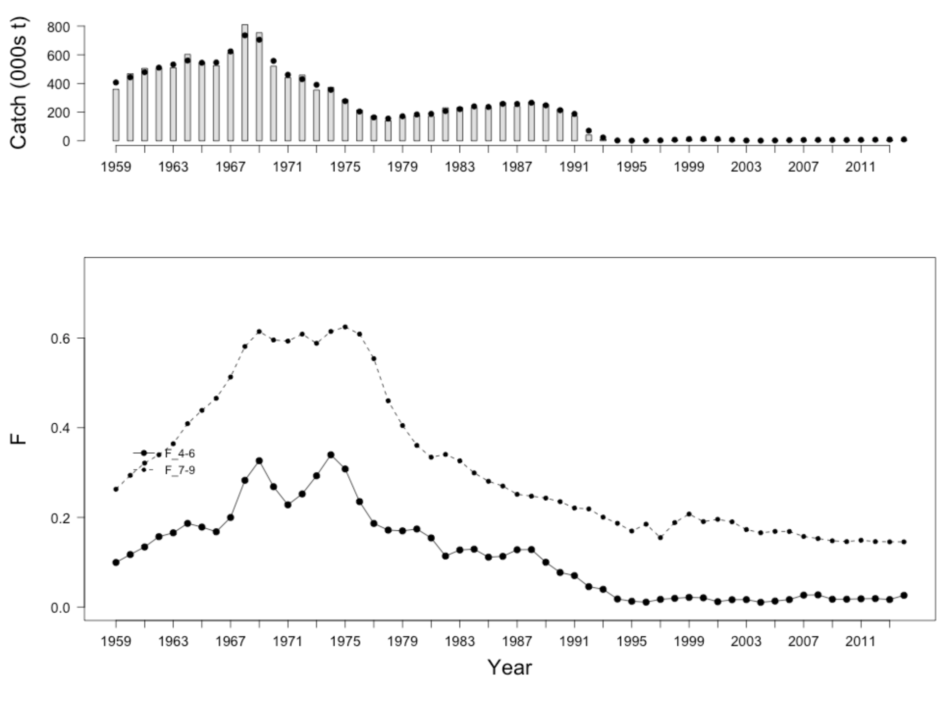

Figure 8. Revised estimates of catch.

Figure 9. Revised estimates of biomass.

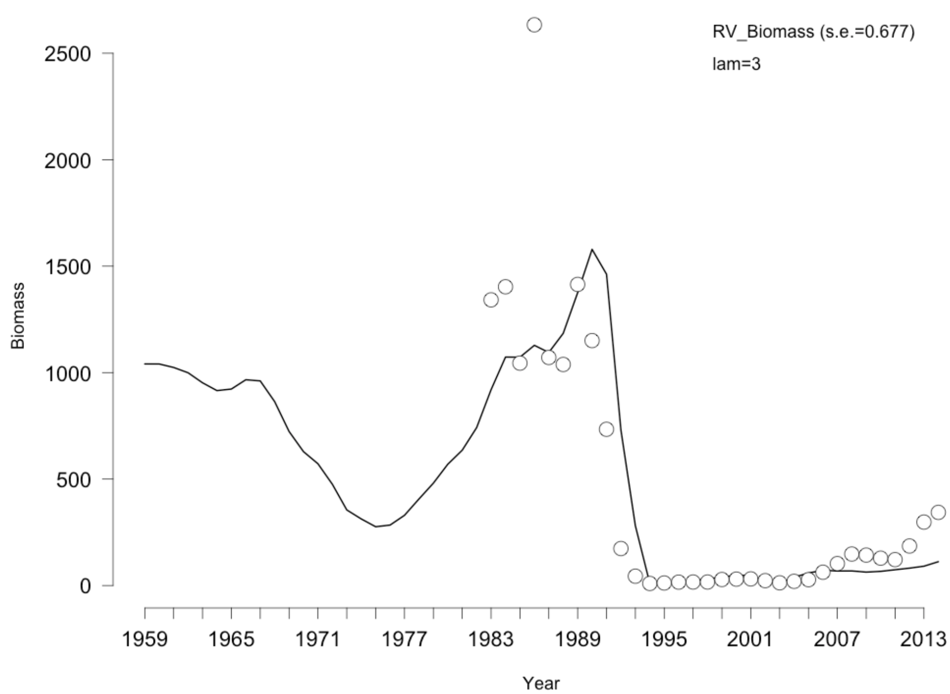

Figure 10. Revised estimates of RV Biomass

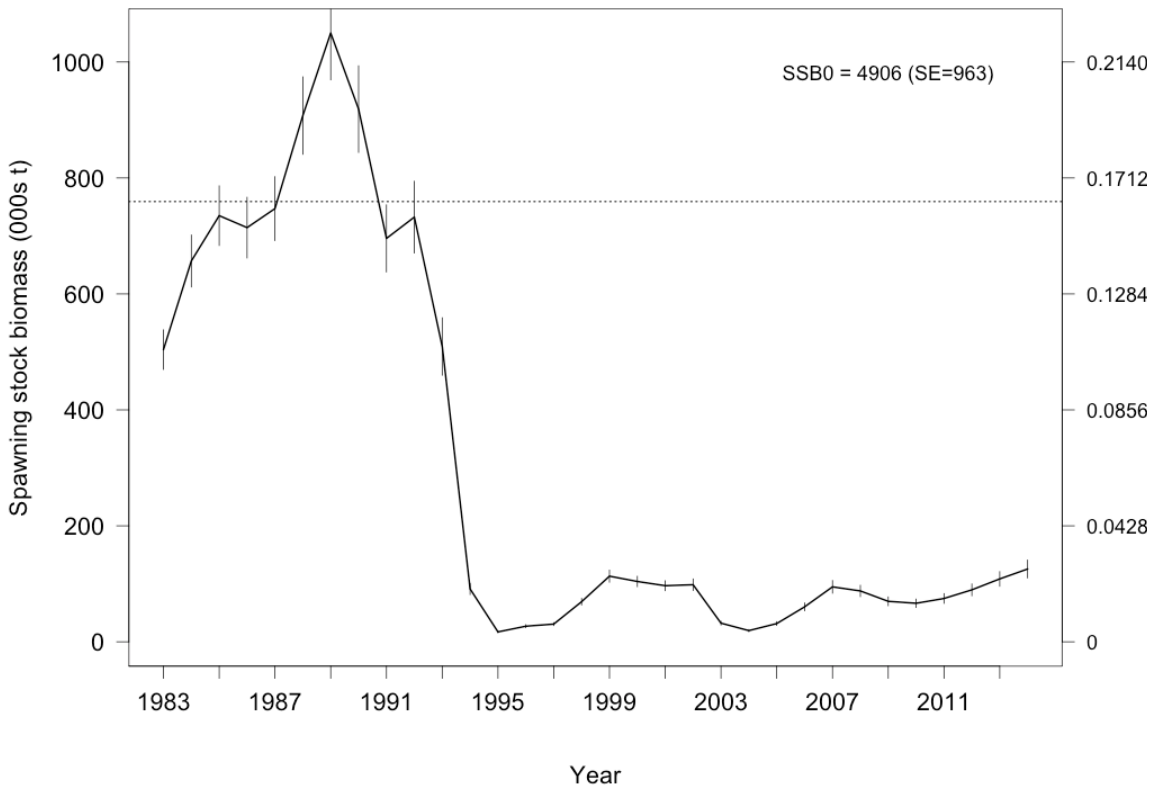

Figure 11. Revised estimates of SSB.

Distribution of random walk residuals

The second revision requested by reviewers was to show whether the random-walk natural mortality deviations followed a normal distribution as implied by the prior specification: $N\left(0,\sigma^{M}_{\delta,a}\right)$.

This test (for ages 3-5 estimates) suggests that the natural mortality deviation estimates are more dispersed compared to the prior (Fig. 13).

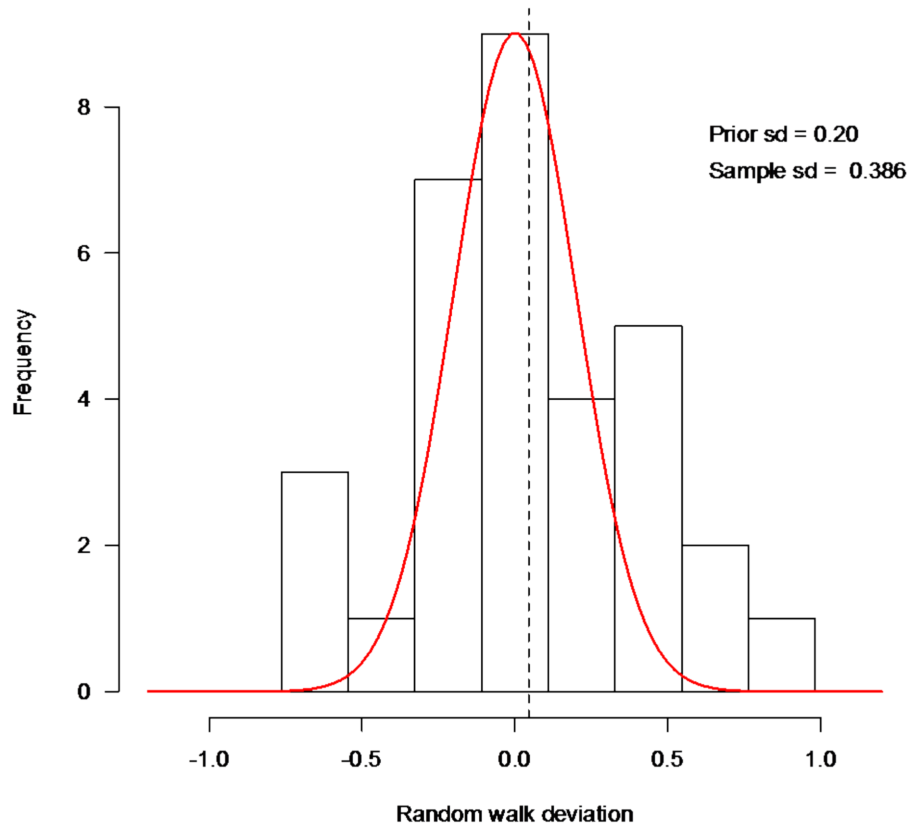

Figure 12. Distribution of the estimated random-walk deviations from the SCA model (bars) compared to the normal prior $N\left(0,\sigma_{\delta,a}^{M}\right)$*(red line).*

Self-test simulations

Self-test simulations were intended for the original submission, but were not because of time requirements to develop the base model. Reviewers, therefore, suggested that self-test simulation be included in revisions. The SSB simulations (Fig. 14) suggest that the SCA model does a reasonably good job estimating SSB since the moratorium (1994+), but is generally biased before that. Bias is not unexpected for the 1959-82 period, because there are no fishery-independent survey abundance or age-composition data for that period. Bias for the period 1983-93 could partially result from the model compensating for pre-1994 behaviour. The stock collapse around 1994 seems to break the bias pattern.

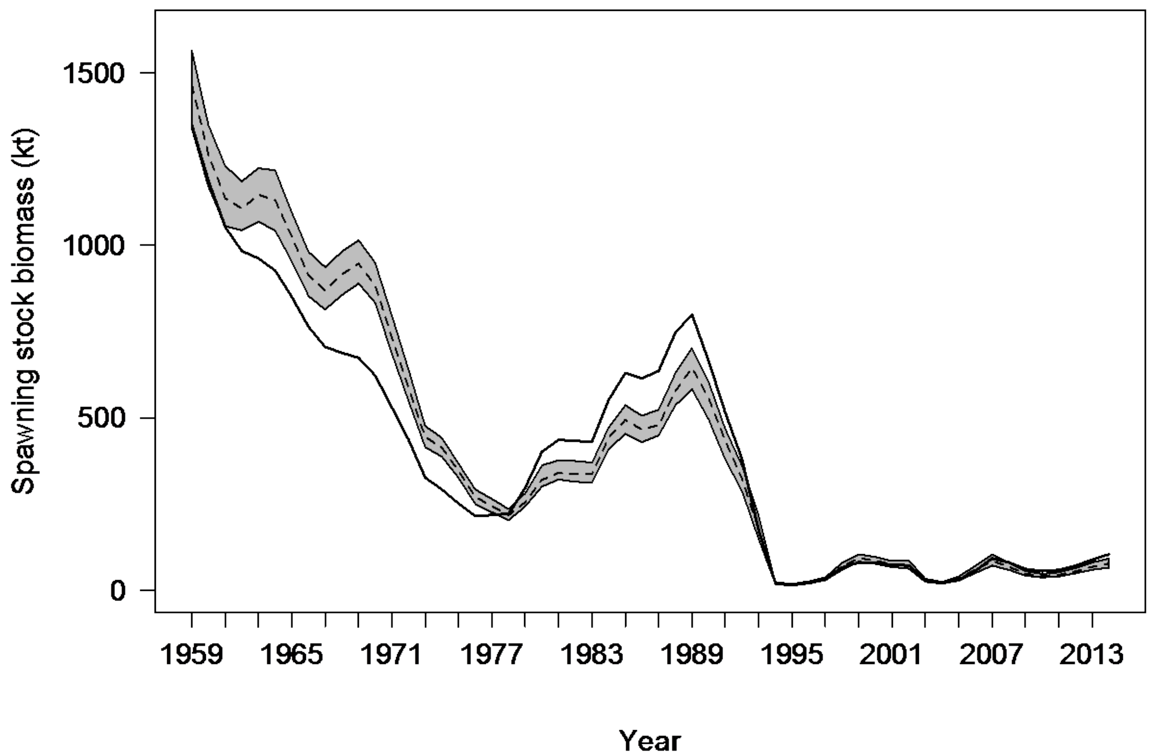

Figure 13. Self-test simulation output for SSB (1959-2014) comparing the MLE estimates (thick black line) with the distribution of estimates obtained via parametric bootstrapping from the estimated index and age-composition sampling distributions.

The natural mortality simulations (Fig. 15) suggest that the SCA estimates are possibly unbiased except for a brief period around 1982-90. Reasons underlying this bias are probability similar to SSB (i.e., weak data), although natural mortality appears less sensitive.

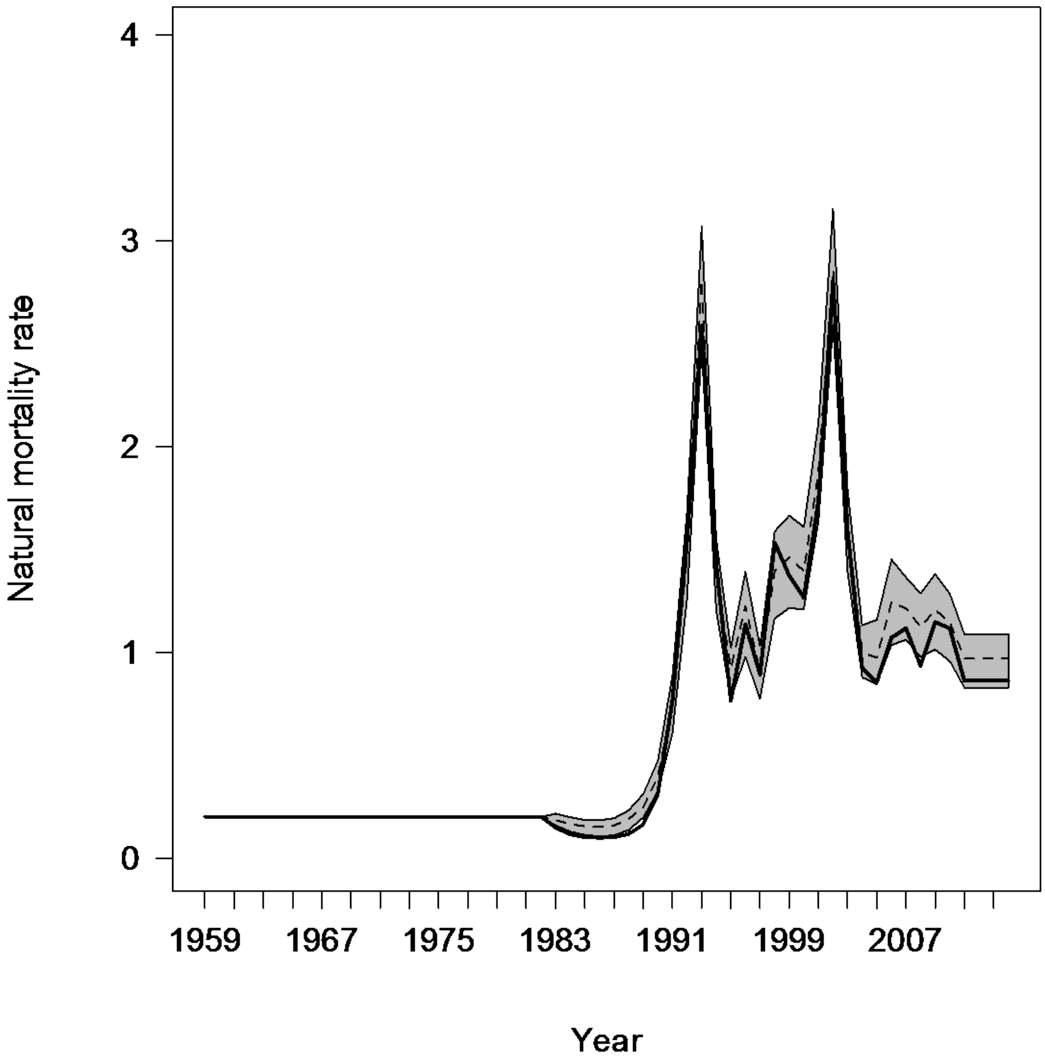

Figure 14. Self-test simulation output for natural mortality of ages 3-5 (1959-2014) comparing the MLE estimates (thick black line) with the distribution of estimates obtained via parametric bootstrapping from the estimated index and age-composition sampling distributions. Gray shaded area encompasses the central 90% of 100 simulated estimates.

Table 4. Summary of reviewer suggestions and responses.

Status	Reviewer comment/suggestion	Response
done	1. Check whether the random-walk properties of the estimated random-walk deviations in M follow the normal prior distribution	The estimated deviations are more dispersed than under a N (0,0.2^2) distribution in the prior.
-	2. Check correlations in M deviations among the age groups	-
done	3. Reduce residual pattern in age-comp fits to both fishery and survey – try alternative selectivity forms	The residual pattern has been resolved. The residual pattern was the result of using too many age-classes in the original data. With 20 age-classes, there were several low proportions each year and the age-comp likelihood was very sensitive to these.
done	4. Include self-test simulations	Self-test simulations suggest reasonable model behaviour given the data. The model is not ideal for the pre-1983 data and reasonable good since then. For future work, the SCA should probably fit 1983-2014 data and use 10 age-classes rather than 20. This would improve model performance and probably also ease of use by DFO if necessary.
done	5. Improve fit to ages 4-5 age-comps	See above
done	6. Run self-test simulations	See above
done	7. Use a constant M 1959-1983	This revision was made to the SCA and helped to improve overall model performance