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

C S A S | ***S C C S***

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

Proceedings Series 2011/075 | **Compte rendu 2011/075**

Pacific Region | **Région du Pacifique**

Proceedings of the Regional Advisory Meeting on the biological sampling of Pacific Herring, and factors influencing the variability in Pacific Herring egg layers and considerations to stock assessment

Compte rendu de la réunion sur les avis scientifique régionale sur l'échantillonnage biologique du hareng du Pacifique et facteurs ayant une incidence sur la variabilité des couches d'œufs de hareng et considérations relatives à l'évaluation des stocks

January 18-20, 2011
Nanaimo, BC

Du18 au 20 janvier 2011
Nanaimo, C.B.

Sean MacConnachie,
Chairperson

Sean MacConnachie,
président de réunion

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

May 2012 | **Mai 2012**

Foreword

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

Avant-propos

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

Proceedings Series 2011/075

Pacific Region

Proceedings of the Regional Advisory Meeting on the biological sampling of Pacific Herring, and factors influencing the variability in Pacific Herring egg layers and considerations to stock assessment

January 18-20, 2011
Nanaimo, BC

Sean MacConnachie,
Chairperson

Compte rendu 2011/075

Région du Pacifique

Compte rendu de la réunion sur les avis scientifique régionale sur l'échantillonnage biologique du hareng du Pacifique et facteurs ayant une incidence sur la variabilité des couches d'œufs de hareng et considérations relatives à l'évaluation des stocks

Du 18 au 20 janvier 2011
Nanaimo, C.B.

Sean MacConnachie,
président de réunion

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

May 2012 **Mai 2012**

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

Published and available free from:
Une publication gratuite de :

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

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

CSAS-SCCS@DFO-MPO.GC.CA

Correct citation for this publication:

DFO. 2012. Proceedings of the Regional Advisory Meeting on the biological sampling of Pacific Herring, and factors influencing the variability in Pacific Herring egg layers and considerations to stock assessment; January 18-20, 2011. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2011/075.

TABLE OF CONTENTS

SUMMARY

Given reductions in resources for data collection following the Larocque court decision, concern has been expressed about the adequacy of ongoing data collection programs. Consequently, there is ongoing interest in exploring the cost-benefits and tradeoffs of varying spatial and temporal sampling coverage versus the precision of parameter estimates. Conclusions and recommendations from several past reviews have also identified this need. These types of evaluations are required in order recognize whether datasets in the time series can be used to distinguish different biological characteristics between regional stock groupings. Despite area closures in three of the five major regional assessment areas, the biological sampling and spawn survey programs operate annually. However, there is some concern about the adequacy of the biological sampling program in providing information on fish size and age composition of major herring stocks for stock assessment analysis and modelling.

Although herring spawn data have been collected for over 50 years, a detailed study of factors that influence the number of egg layers deposited has not been conducted. A better understanding of factors governing the density of spawn deposition will provide valuable information on herring reproduction. The assumption that low numbers of egg layers in a specific spawning site is symptomatic of a low spawning biomass should be investigated since this view may be implied by stock assessment sampling and modelling.

Mortality of eggs during spawning has been examined and documented at most major spawning sites, both in the Atlantic and Pacific Ocean. The loss of eggs during spawning has special relevance to Pacific herring stock assessments that rely on a quantitative index of herring spawn as a key component for annual assessments. A particular concern is that as the financial and logistical support for spawn surveys has diminished the timing of the surveys may be relatively later than during earlier surveys. Relatively later assessments of spawning by SCUBA surveys could result in an under-estimate of spawn, hence an underestimate of the spawning biomass. The potential scale of such possible under-estimates is uncertain, but even a relatively small daily loss (~2%) would result in a total loss of over 25% during a 14-day incubation period. A daily loss of 5% would result in total reduction of more than 50% during the same period. Also, there are other uncertainties that affect the estimates of spawn survival, including density dependent survival of eggs, with survival to hatching being lower in very high densities, etc. It is plausible that this is a concern for areas of the BC coast where spawn has tended to concentrate in fewer areas in recent years.

SOMMAIRE

Étant donné la diminution des ressources affectées à la collecte de données à la suite de la décision judiciaire *Laroque*, certaines personnes s'inquiètent de la possible insuffisance des programmes actuels de collecte de données. En conséquence, on remarque un intérêt soutenu pour l'examen des coûts/avantages et des contreparties que présente la couverture spatiale et temporelle variable lors d'échantillonnages par rapport à la précision des estimations de paramètres. Les conclusions et les recommandations tirées de plusieurs examens menés par le passé soulignent aussi cette situation. Il est essentiel de mener ce genre d'évaluations afin de déterminer quels ensembles de données dans la série chronologique peuvent servir à distinguer les différentes caractéristiques biologiques parmi des regroupements régionaux de stocks. Malgré des fermetures de zones dans trois des cinq principales aires d'évaluation régionales, les programmes d'échantillonnage biologique et de relevés sur le frai se déroulent encore chaque année. Toutefois, certaines personnes se préoccupent de l'insuffisance possible du programme d'échantillonnage biologique à fournir des renseignements sur la composition selon l'âge et la taille des principaux stocks de harengs aux fins d'évaluation et de modélisation des stocks.

Même si des données sur le frai du hareng sont recueillies depuis plus de 50 ans, aucune étude détaillée sur les facteurs ayant une incidence sur le nombre de couches d'œufs déposées n'a encore eu lieu. Une meilleure compréhension des facteurs régissant la densité des œufs déposés permettra d'obtenir de précieux renseignements sur le frai du hareng. On doit vérifier l'hypothèse selon laquelle le faible nombre de couches d'œufs dans une frayère en particulier refléterait une faible biomasse du stock reproducteur, car cette hypothèse pourrait être fondée sur l'échantillonnage et la modélisation des stocks.

On a étudié et consigné la mortalité des œufs durant le frai pour la plupart des principales frayères, dans les océans Atlantique et Pacifique. La perte d'œufs pendant le frai revêt une importance particulière dans le cadre des évaluations des stocks de harengs du Pacifique qui s'appuient sur un indice quantitatif du nombre d'œufs de hareng en tant que principal composant pour les évaluations annuelles. On doit particulièrement se préoccuper du fait que, comme le soutien financier et logistique lié aux relevés sur le frai a diminué, la période des relevés pourrait débuter sensiblement plus tard que pour les relevés précédents. Des évaluations du frai en plongée effectuées sensiblement plus tard pourraient entraîner une sous-estimation du frai, et donc une sous-estimation de la biomasse du stock reproducteur. L'ampleur potentielle de telles sous-estimations demeure incertaine, mais même une perte quotidienne plutôt faible (~ 2 %) se traduirait par une perte totale de plus de 25 % pour une période d'incubation de 14 jours. Une perte quotidienne de 5 % se traduirait par une perte totale de plus de 50 % pour la même période. En outre, d'autres incertitudes peuvent avoir une incidence sur les estimations de la survie des œufs, y compris la survie des œufs en fonction de la densité, la survie des jeunes lors de l'éclosion étant plus faible si la densité est très élevée, etc. Cela est sans doute préoccupant pour les zones côtières de la Colombie-Britannique, où les harengs tendent à se regrouper pour frayer dans un plus petit nombre de zonesau cours des dernières années.

INTRODUCTION

The Centre for Science Advice – Pacific (CSAP) review was held January 18, 2011 at the Pacific Biological Station in Nanaimo, British Columbia. External participants from industry and First Nations were in attendance at the meeting. American colleagues also participated via Webinar. The Chair (Sean MacConnachie) opened the meeting by welcoming participants, reviewing the agenda and reviewing to the terms of reference.

The following working papers were reviewed:

- "Biological sampling of BC herring: Analysis of sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations" by Vivian Haist
- "Factors influencing the variability in Pacific herring egg layers and considerations to stock assessment" by Doug Hay.

REVIEWS

WORKING PAPER #1:

"Biological sampling of BC herring: Analysis of sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations." by Vivian Haist.

**Paper was accepted subject to revisions*.*

A number of points of clarification were raised including a question about the mismatch between reported dates of the fisheries and the resulting landings and biological samples. It was suggested that there may be a code in the database to identify and separate out test fishery catches that occurred outside the fishing dates. There was a question about the maturity of 2 year old herring and whether they came from the test fishery or the commercial fishery and it was noted that most were in the test fishery catches. A question was raised about adjusting the gonadosomatic Index (GSI) according to the condition of the fish since the maximum GSI may increase as condition increases. It was agreed that condition could be a co-variate with the GSI to forecast spawning date. Clarification of the interpretation of effective sample size was raised and the question of whether reducing the sampling period from the present to a shorter 14 day window would bias the sampling precision. It was noted that although fewer samples might be obtained the effective sample size should be similar.

Review #1

A nice overview of the age composition data was presented. An additional piece that could be added is whether there is an objective way to weight age composition data as an outcome from these analyses since advice is needed on how different weights would affect the stock assessment. The reviewer questioned whether more weight should be placed on age composition in assessment. The response was to use EffN as initial weighting. The modeled number of aged fish versus the true number fish aged plateaus at ~500-600. Therefore, should the EffN be used in the assessment as a relative weight? The response was that yes, this has been effectively done in the past. It was noted that down weighting of the data was related to

observed retrospective pattern in the past. There is a need to reconcile information in age composition data and cohort strength with the trend data in the spawn deposition data.

There was a discussion around the maintenance of precision in the age composition data in the expectation of reduced sampling window of 14 days. It was noted that there would likely be fewer days that a charter boat would be out there sampling; it didn't make sense to reduce the spatial component. A question was raised about the need to set up a proper random sampling design in the test fishery. It was noted that the charter days are determined in advance of season and may be extended depending on when fish spawn; and locations of sampling are somewhat ad-hoc (locations are where the fish are located) but the focus of test fishery changes as season progresses; at the beginning, the test fishery is looking widely but as the fishery approaches the test boats concentrate effort in the likely fishing area.

There was discussion about the minor difference in mean age between the commercial and test fisheries. It was noted that this could be due to the timing of the fisheries since older fish spawn earlier and the fishery could be impacting the age structure selectively. It was also noted that although mixed ages approach the spawning grounds early in the season it appears that the youngest fish separate out and move out of the area.

A question was raised about how one would deal with ageing error and how would it affect the number of samples to collect. It was noted that it was important to understand the extent and direction of the ageing bias rather than collect more samples.

A question was raised regarding the basis for determining sample size? If a criterion is set for a specified age frequency CV will we be able to detect a weak year-class? The response was the one should be able to detect very weak year-classes even with moderate CVs in the age composition data.

There was an extensive discussion of the rules used for sampling in the test fishery and whether they are adequate to ensure representative sampling. It was noted that the program does try to sample the fish throughout the area. However the commencement of the pool system in 1998 changed the focus of the fishery. The current tendency is to look more broadly than during the previous periods of the test fishery. There was a suggestion that this results in more samples earlier in the season and then later, but also greater sampling closer to the spawn timing.

A discussion of the possible rules for weighting the sampling data followed. It was noted that it would be difficult to weight samples after the fact. Suggestion that one needs to collect comments on each sample as it is collected. It was also suggested that samples collected within three days of the start of a fishery would not be considered as test fish sampling.

Review #2

The reviewer felt that a clearer description of the link between the current analyses, past work and needs for future research would be useful. The justification for the application of the cluster analysis as a tool for the analysis could be strengthened. The author noted that it was used primarily because of the multinomial assumption in the age composition data. The definition of good and moderate in the assessment of clusters should be improved. Figure 5 was difficult to interpret and would benefit from a clearer description as well as adding the number of the samples per cluster. A discussion of the sample collection occurred in the review and focussed on how the test fishery samples were collected compared to the fishery samples. It was noted that the time series extended from 1971-2010 but some uncertainty existed about samples prior

to 1980 because the vessels were focussed on obtaining their test payment rather than random sampling. There was also a suggestion that reduction fishery samples may be biased because of uncertainty about where the catch and biosamples were collected. A comment was made that the objective of the test fishery program for scientific data collection should be reviewed.

Clarification of the sex ratio analysis and how it might apply to the stock assessment was requested. Discussion occurred about whether the spawning population or spawning aggregations were the target of biosampling. Questions arising during discussion included: How does one deal with the sample weighting issue especially when the sampling may not continue through the entire spawning period? What is the target of the sampling and what do the late samples represent in terms of the spawning population that is targeted? Is it possible to relate biosamples to individual spawning events?

The utility of the GSI analysis for forecasting spawning time and consequently for sample weighting was discussed at length. It was agreed that it looked promising but required some additional analysis.

There was a discussion of whether there is any indicator of smaller scale population structure. It was noted that data isn't available to address the question of whether sufficient data are available to provide advice on a smaller spatial scale. It was suggested that data for areas 15/16 should be reviewed. Areas 132 and 135 appear distinct from the cluster analysis. It was also noted that genetic differences have been identified between spawning waves from European studies. It was noted that there were a number of data inconsistencies in the database (spawn timing dates versus survey dates) that made it difficult to address some of the research questions and that additional work was needed to clean up some of the database errors prior to further analysis.

Further discussion of the target population and whether age 2 fish are representative of the population. Are they resident fish or schools that haven't left the area in which case they should not be included in the age structure. Should we only be sampling the fished population? However, we are modelling the entire population and collecting sample s of both the test program and fishery sampling.

A question was raised about the accuracy of the ageing information and the need for validation. It was noted that the ageing laboratory will be re-ageing some samples from the herring scale archives to assess historic ageing consistency.

RECOMMENDATIONS

1. Establish a subcommittee to develop and test different options for weighting scenarios e.g. length of spawn by section.

2. Amend working paper to include conclusions of the work and results of modifications form the Chih paper and apply the corrected GSI equation for maximal roe weight.

3. Catch database issues need to be addressed to permit investigation of data weighting issues.

WORKING PAPER #2:

"Factors influencing the variability in Pacific herring egg layers and considerations to stock assessment" by Doug Hay

**Paper was accepted subject to revisions*.*

A participant asked if 0.01 layers, which is set to reflect trace layers, represents 39,000 eggs per square meter, as suggested in Appendix 7. The Authors felt that egg density was likely less. There was a comment about whether proportions of egg layers can actually be detected, as implied in the manuscript. The Authors responded that the original purpose of trace egg layers was to be a qualitative description of where the limits of a spawn occurred. It was suggested that the effects of trace layers on spawn deposition estimates be evaluated.

It was noted that data from 1984 onward is used in paper. One of the participants stated that diver surveys began in 1988 and 1984-87 data were based on research surveys which were likely more methodical.

Reviewer's comments

The first reviewer thought that paper was an excellent summary of information. This reviewer wondered if apparent trend in increase in trace layers and reduction in egg layers is a consequence of variations in spawning biology or due to variations in methodology. It was suggested that biology may not be known well enough to eliminate a biological explanation. The Reviewer wondered why contract divers were not contacted to ask if methodology has changed. One Author responded that a quality control system was in place to try to minimize variations in diver methodology. The Reviewer thought that there should be a definition of a spawning event to help interpret the figures. In addition, the definition of fish density was unclear. The Reviewer suggested that scale be included in the description to help interpretation. The Reviewer felt that there was no support for density of spawning herring controlling density or layers of eggs and wondered if the decrease in spawn area and egg layers might suggest more of a biological rather than methodological explanation. One Author suggested that cumulative length and area wouldn't change and suggested that egg layer changes are independent. The Reviewer was concerned if the trend in spawning density (Fig. 6) was a function of suspect data before 1990, because densities seem to be unrealistically high. An Author responded that the figure intended to show an inconsistency, that is, spawn index in last five years is different than earlier years. The Reviewer noted that it is possible that the earlier data are biased because surface surveys would underestimate spawn widths and therefore spawn areas; if these are excluded then the effects over the last five years becomes less dramatic. The Reviewer noted that Fig. 7 shows that cumulative spawn is dome-shaped and suggested that this could be a function of a higher proportion of surface surveys earlier in the time series which would bias area estimates downward. One Author noted that uncorrected surface survey data could have been used in the analysis which could have influenced the results as suggested by the Reviewer. The Reviewer noted that survey duration appears to be consistent over time. There is an implication that the detection of trace layers could be affected by size of spawn with larger spawns being sampled less comprehensively. The Reviewer suggested that using a number of quadrats that is proportional to transect length may decrease the impact of trace layers and thought that this could be an issue for BC herring surveys because the number of transects is fixed. One Author responded that the current BC spawn survey methodology should be immune to this bias because the divers start surveying at the deep end of a spawn.

Reviewer 2 felt that the paper was valuable and a good first step for investigating the variability in egg layers over time. The Reviewer expected there to be an increase in trace layers with a decrease in the proportion of more dense egg layers, which wasn't reflected in Appendix 4, Fig. 4. An Author responded that the egg layer categories are relative so increases in trace layers do not necessarily reflect changes in all denser egg layers. (Doug referred to the data as being relative (proportions?); however, the caption says that frequencies, which should be independent, are plotted.) The Reviewer felt that assigning a quantity to trace egg layers of 0.1 rather than 0.01 layers was arbitrary and suggested that the effect of this could be dealt with by simulations designed to evaluate the effect of including the trace layers. It was noted that predation was addressed in the manuscript but there were no data on predator abundance trends presented which could have been used to evaluate the effect of predation. The Reviewer recommended that the Authors test for statistically significant changes in metrics and test the qualitative descriptions; there were instances (eg. Table 1 and Appendix Fig. 2 and 3) where qualitative conclusions are inconsistent with quantitative information. The Reviewer suggested that standard error bars be included in the figures. The survey indices are important and there appears to be trends in them that may affect the spawn trends. The Reviewer suggested in duration of surveys may result in increased frequency of trace layers. One Author noted that there is an error in the figure (Appendix 5, Fig. 2) that show a recent increase in spawn survey duration.

Discussion – One participant asked if Table 1 includes spawns that were not surveyed. One Author responded that un-surveyed spawns are excluded. This participant suggested that there is more recent focus on trace layers because spawning has declined.

The Participants noted that Appendix 4. Fig. 3 showed a decline in trace layers with increases in spawning stock biomass. One participant suggested that if this was reflecting methodology then the linear relationship would break down.

RECOMMENDATIONS

The paper will be accepted after revisions are made to reflect the Reviewers' and Participants' comments. It was recommended that:

1) there should be a modeling and statistical evaluation of the effect of trace layers on spawn deposition estimates;

2) the variability in diver assessment of all egg layers should be investigated, and diver egg deposition estimates should be compared with counts of eggs from associated samples, and;

3) other descriptions of spawn abundance should be explored.

ACKNOWLEDGEMENTS

Thank you to Jake Schweigert and Ron Tanasichuck for being the rapporteurs during the review. Thank you to Janeane MacGillivray for coordinating all of the documents, organization of the meeting and managing the webinar. Thank you to Marilyn Joyce for the advice on managing the meeting.

APPENDIX A: AGENDA

PACIFIC HERRING

Regional Advisory Process

Centre for Science Advice Pacific
January 19-20, 2011
Nanaimo, British Columbia
Chairperson: Sean MacConnachie

Wednesday, January 19, 2011

Working Paper to be reviewed:

Biological sampling of BC herring: Analysis of sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations by Vivian Haist. 9:00	Introductions	Sean MacConnachie
9:10	Review Agenda & Housekeeping	Sean MacConnachie
9:20	CSAS Overview & Procedures	Sean MacConnachie
9:40	Review of Terms of Reference as pertains to biological sampling research document	Sean MacConnachie & RAP Participants
9: 45	Presentation of Working Paper	ian Haist
10:30	***Break***	
10:50	Questions of Clarification	AP Participants
11:15	Presentation of Reviews & Authors' Responses	viewers & Author(s)
12:00	***Lunch Break***	
1:00	Discussion	AP Participants
2:30	*Break*	
2:50	Building Agreement on Conclusions, Recommendations, Advice and Future Work	AP Participants
4:30	***Adjournment***	

Thursday, January 20, 2011

Working Paper to be reviewed:

Factors influencing the variability in Pacific herring egg layers and considerations to stock assessment *by Doug Hay et al* 9:00	Introductions	Sean MacConnachie
9:10	Review Agenda & Housekeeping	Sean MacConnachie
9: 30	Review of Terms of Reference as pertains to investigating variability in herring egg layers	an MacConnachie & RAP Participants
9:45	Presentation of Working Paper	Doug Hay
10:30	***Break***	
10:50	Questions of Clarification	ıP Participants
11:00	Presentation of Reviews & Authors' Responses	viewers & Authors
12:00	***Lunch Break***	
1:00	Discussion	ıP Participants
2:30	***Break***	
2:50	Building Agreement on Conclusions, Recommendations, Advice and Future Work	ıP Participants
4:30	***Adjournment***	

APPENDIX B: ATTENDEES

First Name	Last Name	Affiliation	Jan 19	Jan 20
Jennifer	Boldt	DFO	X	X
Kristen	Daniel	DFO	X	X
Marilyn	Joyce	DFO	X	X
Sean	MacConnachie	DFO	X	X
Bruce	McCarter	DFO	X	X
Lisa	Mijacika	DFO	X	
Brenda	Spence	DFO	X	
Ron	Tanasichuk	DFO	X	X
Randy	Webb	DFO	X	X
Jake	Schweigert	DFO	X	X
Jaclyn	Cleary	DFO	X	
Charles	Fort	DFO		X
Peter	Midgley	DFO	X	X
Matt	Thompson	DFO		X
Dennis	Chalmers	Province of BC	X	X
William	Gladstone	Heiltsuk First Nation	X	X
Lorena	Hamer	HCRS	X	X
Doug	Hay	DFO Scientist Emeritus	X	X
Bill	Wilson	Aboriginal Vessel Owners	X	
Earl	Newman	Heiltsuk First Nation	X	X
Ashleen	Benson	SFU	X	
Vivian	Haist	Haist Consulting	X	X
Steve	Martell	UBC Fisheries	X	X
Sherri	Dressel	NOAA	X	X
Sharon	Jeffery	Haida Fisheries	X	
Ed	Safarik	HCRS	X	X

APPENDIX C: TERMS OF REFERENCE

Terms of Reference

Assessment of Pacific Sardine, biological sampling of Pacific Herring, and factors influencing the variability in Pacific Herring egg layers and considerations to stock assessment

Pacific Regional Advisory Process

January 18-20, 2011
Nanaimo, British Columbia

Chairperson: Sean MacConnachie

Background

The *Centre for Science Advice Pacific (CSAP) Pelagics Standing Committee*, along with additional invited participants as required, meet to review information related to assessing pelagic fish stocks. These reviews are based on specific questions outlined in formal *Requests for Science Information and Advice*. A Regional Advisory Process (RAP) is being planned to review three Research Documents and one Science Advisory Report pertaining to Pacific sardine or Pacific herring in British Columbia waters.

Working Paper 1: Assessment of Pacific sardine in British Columbia waters, with an emphasis on seasonal abundance and migration estimates

Context

Following a RAP conducted in April of 2009, harvest advice in British Columbia (BC) has been based on a 3 year running average of the most recent BC migration rate estimates, a current estimate of the coastwide adult stock biomass (from the US assessment) and applying the US harvest rate (15% in recent years). Sardine seasonal migration into BC has been estimated from west coast Vancouver Island (WCVI) midsummer surface trawl surveys. Seasonal abundance and migration estimates based on WCVI trawl survey data have been considered to be minimum estimates because they don't include other regions of the province where sardines also occur midsummer. Following a meeting with industry in June of 2010, an agreement was made to investigate the feasibility of developing a harvest strategy based entirely on the results of the annual west coast of Vancouver Island trawl survey. A need to investigate the possibility of assessing sardine abundance in other unsurveyed areas of the coast was also identified.

Objectives

1.1 To review the feasibility of developing a seasonal estimate of sardine biomass in the Canadian zone from the summer research trawl survey data, independent of the US stock assessment.

1.2 To identify issues (e.g. biological and technical, including information gaps) associated with changing from the methodology adopted in April 2009.

1.3 To provide advice on seasonal abundance and/or migration to apply to harvest control rules and setting quotas for the 2011/2012 fishing season

Working Paper 2: Biological sampling of BC herring: Analysis of sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations

Context

Given reductions in resources for data collection following the Larocque court decision, concern has been expressed about the adequacy of ongoing data collection programs. Consequently, there is ongoing interest in exploring the cost-benefits and tradeoffs of varying spatial and temporal sampling coverage versus the precision of parameter estimates. Conclusions and recommendations from several past RAPs have identified this need. Furthermore, these types of evaluations are required in order recognize whether datasets in the time series can be used to distinguish different biological characteristics between regional stock groupings. Despite area closures in 3 of the 5 major regional assessment areas, the biological sampling and spawn survey programs operate annually. However, there is some concern about the adequacy of the biological sampling program in providing information on fish size and age composition of major herring stocks for stock assessment analysis and modelling.

Objectives

2.1 To review an investigation which explores effects of varying spatial and temporal sampling coverage to adequately characterize fish size and age structure of Pacific Herring stocks in the major assessment areas.

2.2 To determine if existing data can be used to evaluate whether the accuracy and precision of estimates of biological characteristics has changed over time.

2.3 To determine if data trends of biological characteristics are indicative of similarities or differences between stocks in some areas (e.g., Central Coast (CC) subareas 6, 7, and 8).

Working Paper 3: Factors influencing the variability in Pacific herring egg layers and considerations to stock assessment

Context

Although herring spawn data have been collected for over 50 years, a detailed study of factors that influence the number of egg layers deposited has not been conducted. A better understanding of factors governing the density of spawn deposition will provide valuable information on herring reproduction. The assumption that low numbers of egg layers in a specific spawning site is symptomatic of a low spawning biomass should be investigated since this view may be implied by stock assessment sampling and modelling.

Mortality of eggs during spawning has been examined and documented at most major spawning cites, both in the Atlantic and Pacific Ocean. The loss of eggs during spawning has special relevance to Pacific herring stock assessments that rely on a quantitative index of herring spawn as a key component for annual assessments. A particular concern is that as the financial and logistical support for spawn surveys has diminished the timing of the surveys may be relatively later than during earlier surveys. Relatively later assessments of spawning by SCUBA surveys could result in an under-estimate of spawn, hence an underestimate of the spawning biomass. The potential scale of such possible under-estimates is uncertain, but even a relatively small daily loss (~2%) would result in a total loss of over 25% during a 14-day incubation period. A daily loss of 5% would result in total reduction of more than 50% during the same period. Also, there are other uncertainties that affect the estimates of spawn survival, including density

dependent survival of eggs, with survival to hatching being lower in very high densities, etc. It is plausible that this is a concern for areas of the BC coast where spawn has tended to concentrate in fewer areas in recent years.

Objectives

3.1 To review information related to physiological, ecological and behavioural controls that affect Pacific herring spawning behaviour and factors associated with the estimation of the number of egg layers.

3.2 To review information related to factors that affect the estimation of variability in egg layers.

Expected Publications

- CSAS Proceedings
- CSAS Science Advisory Report (1), based on Pacific sardine assessment
- CSAS Research Documents (3)

Participation

DFO Science Branch
DFO Fisheries and Aquatic Management Branch
Commercial and recreational fishing Interests
First Nations organizations
Non-government organizations
Academia

References Cited

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

DFO. 2009. Proceedings of the Pacific Scientific Advice Review Committee (PSARC) meeting for the assessment of scientific information to estimate Pacific sardine seasonal migration into Canadian waters. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2009/034

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

Schweigert, J., McFarlane, G., Hodes, V. 2010. Pacific sardine (*Sardinops sagax*) biomass and migration rates in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/088.

APPENDIX D: DETAILED REVIEWS

Review of: Biological sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations.

by: Steven Martell
January 15, 2011

Summary

This working paper presents a comprehensive overview of the biological sampling program from the BC herring Sn-roe and test fisheries and is a continuation of a previous review that was presented to CSAP back in September 2011. This review specifically examines data from the Sn-roe and test fisheries between 1971 and 2010 and attempts to answer eight specific questions that are geared towards improving the underlying stock assessment model and or highlighting specific issues to be considered in an operating model for future management strategy evaluations. A series of different analyses were performed to address these questions including: cluster analysis on age-composition data, comparisons of sex ratios in catch by period and growth rates, use of linear models to predict relative GSI, and a "power analysis" to determine the effective sample size.

The specific research questions are paraphrased below and my interpreted answers to these questions are given parenthetically.
1. Is there evidence for spatial structure, or sub-stocks, within each of the Stock Assessment Regions (SARs)? [Yes.]
2. Do Sn-roe and test fisheries sample different populations? [Appear to sample the same population, however, mean age in the Sn-roe fishery is slightly higher than the test fishery.]
3. Are there sex specific differences in growth and selectivity? [No significant difference in size-at-age, gillnet roe fisheries are highly selective for female fish up to age-6, Sn-roe fishery highly selective foe age-2 males.]
4. Can GSI be used to associate herring with spawning events? [A relative GSI measure holds some promise to potentially associate samples with specific spawning events.]
5. Is the assumption that test fishery samples are representative of spawning populations reasonable? [Yes and No. In some years test fishing occurs just prior to, during, or immediately following spawning, but in some years test fishing occurs 20+ days before the onset of spawning.]
6. Is there an objective basis for weighting the Sn-roe and test fishery age composition samples? [No. There are inconsistencies in the catch database and the timing of the test fishery.]
7. Has the precision of the age-composition data changed over time? [No. In recent years there is no indication that the precision of age-composition has changed.]
8. How would a decrease in sampling affect precision and accuracy of age-composition data? [For age-proportions greater than 5%, estimates of c.v.s would generally be less than 0.25 in the Sn-roe and gillnet fishers with 6 to 12 samples per fishery.]

Overall, I think this working paper provides a nice overview of the bio-sampling programs and how historical age-composition data have been used in stock assessments. I particularly like the cluster analysis within each of the SARs, as I think this highlights some potential problems with weighting of age-composition data. I feel that the paper fell a bit short on recommendations to deal with the noted shortcomings, namely how assign weights to the age-composition in stock assessment models.

CSAP Questions
• Is the purpose of the working paper clearly stated? Yes.
• Are the data and methods adequate to support the conclusions? No real conclusions were actually made. The methods appear to be adequate to support the cluster analysis, with an apparently low probability of a Type II error.
• Are the data and methods explained in sufficient detail to properly evaluate the conclusions? I struggled a bit to understand the effective sample size calculation, specifically exactly how values of _ were obtained. I certainty could not repeat it.
• Are the recommendations provided in a form useful to a fisheries manager? There are no specific recommendations provided other than various comments, for example, of separating the test fishery samples from the Sn-roe fishery, and sexually explicit dynamics due to the differences in sex ratios in the landings, especially the gill net fishery.
• Does the advice reflect the uncertainty in the data, analysis or process? No real management advice is given.
• Can you suggest additional areas of research that are needed to improve our assessment abilities? Many of these same questions should be vetted in an MSE framework, as specific issues regarding effective sample size could be addressed, and objective ways to weight the various data used in stock assessment models could be addressed from a fisheries management perspective.

Clarifications (in no particular order)
• I'm note sure I understand how the _ term is estimated in calculating the Effective sample size EffN.
• I'm confused by Figure 19. The actual number of samples collected, is this the same as number of different samples taken from each of the fisheries. Is this the same as N in

•

$$\frac{N \sum_{i=1}^{8} \sum_{j=2} p_{ij}(1-p_{ij})}{\_n_i} = \frac{N}{2} \quad 8 \sum_{j=2} (s_j)_2?$$

You state in the 2nd paragraph in section 6, that there is no objective basis for weighting individual samples so in the following analyses all samples are treated as simple random samples. If I understand this correctly and the estimates of effective sample sizes in Figure 19 (which are much larger than N) should the stock assessment models place more weight on the age composition data than it currently is? Or should the actual weights be reduced because there appears to be spatial structure within each of the SARs?
• I believe that Pennington and Vølstad (1994) found similar results with reduced samples sizes, but maintaining precise estimates required sampling more locations (hauls). Does the reduction to 14 days maintain the same spatial coverage as the existing program?
• In section 6.2.2 I do not understand how the EffN equation is modified, and the last sentence of the 3rd paragraph confuses me. "For the full sampling coverage, 7% of fisheries had effective Ns less that 100". Do you mean the actual test fishery data, or the combined Sn-roe and test fishery data? I'll assume its only the Test fishery.

References

Pennington, M. and Vølstad, J. (1994). Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. Biometrics, 50(3):725–732.

CSAP Review

Biological sampling of BC herring: Analysis of sampling requirements for characterizing age structure and other biological characteristics of fisheries and spawning populations.

Author: Vivian Haist
Reviewer: Jaclyn Cleary, DFO

Date: January 16, 2011

Summary

This Working Paper is the second paper addressing a Request for Science Advice developed in spring 2010. The first paper arising from this Request for Science Advice, Review of the biological sampling program for BC herring (Cleary and Benson (in prep.); presented to CSAP September 2010), presents an initial evaluation of the biosample data, examining trends in the timing of the biosampling and spawns survey programs.

The Rationale for Science Advice questions the adequacy of the biological sampling program for providing information on fish size and age composition of the major herring stocks for stock assessment analysis and modelling, particularly in areas that are closed to fishing.
The Request identifies the following questions:

1. What is the optimal spatial and temporal sampling coverage required to adequately characterize fish size and age structure of Pacific herring stocks in the major assessment areas?
2. Do the existing data indicate whether the precision of estimates of biological characteristics has changed over time?
3. Are there differences in characteristics that might suggest separate biological stocks in some areas, e.g.., Central Coast (CC) subareas 6, 7, and 8?

Following on these questions, the Terms of Reference specified three objectives for this working paper:

1. To investigate the effects of varying spatial and temporal sampling coverage to adequately characterize fish size and age structure of Pacific herring stocks in the major assessment areas.
2. To determine if existing data can be used to evaluate whether the accuracy and precision of estimates of biological characteristics has changed over time.
3. To determine if patterns in biological characteristics are indicative of similarities or differences between stocks in some areas (e.g., Central Coast (CC) subareas 6, 7, and 8).

A number of additional research questions were also explored:

- Is there evidence for spatial or temporal structure (sub-stocks) within herring stock assessment regions (SARs)?
- Do Sn-roe and test fisheries sample different populations?
- Are there sex-related differences in biological characteristics that should be captured in the stock assessment?
- Can gonadosomatic indices (GSI) be used to associate herring samples with spawning events?
- Is the assumption that test fishery samples are representative of spawning populations (in some years) reasonable?

- Is there an objective basis for weighting roe fishery and test fishery samples in developing age-compositions for stock assessments?
- Has the precision of age compositions changed over time?
- How would a decrease in sampling effect the precision (and accuracy) of age-compositions?

Given the scope of analysis required to adequately address these questions, the author notes that this analyses can not provide definitive answers but is a first step toward identifying issues and developing methods to address the questions.

General comments:
My overall impression of this paper is that it presents a variety of interesting statistical analyses that are thorough and generally well described by the author. However, in order to evaluate the significance of this analysis for management, discussion needs to examine links between results presented herein and the quality of information provided by the sampling program. Given the immediacy of funding restrictions and ongoing changes in the bio-sampling program, it is imperative to begin to focus our analyses on management-specific questions. As such, I feel that impact of this paper could be greatly improved with a more focused scope. For example, highlight the cluster analysis and sampling weighting work, and tie the discussion back to the specific questions identified in the TOR and the impacts for management.
Specific comments (in order of appearance):
1) Cluster analysis
-Include justification as to why this method was chosen. What does cluster analysis assume about the population?
-Include discussion of interpreting sampling effects vs. population effects. I.e., the test fishery samples only “major spawning aggregations”. Given this sampling bias, what does the appearance of “clusters” tell about the population?
• Pg 9, Section 2.1.3
-Include criteria for good, moderate, poor
• Pg 12, Figure 5
-The information being presented by the left-side of these figures is not clear.
• Pg 11, para 4
-“The clusters may reflect spatial structure…”: evidence to support this statement is not clear here. Please elaborate.
• Pg 12, para 2
-Author observes higher probabilities for sections in the northern SARs than the southern SARs. However, there are far fewer sections in the northern SARs than southern SARs, also, sections are highly variable in size (both within and among SARs). These factors should be considered when discussing the results.
• Pg 13, para 1
-Section 135 (SoG): high degree of consistency in the age compositions within section – Is this evidence of “distinction from other sections” or lack of geographic proximity from other sections used in the cluster analysis?
2) Time-series: 1972-2010
-Are there not problems with the accuracy of Section and SAR information for biosamples collected prior to 1980? Would this not impact the calculation of probabilities (that a sample from section *i* is in the same cluster as a sample from section *j*)? (Pg 12)
3) Sn-roe vs. test samples
• Pg 14
-The analysis supports the hypothesis that the sn-roe fishery is selective for older (larger) fish. **For discussion:** What are the potential implications for stock assessment of combining biosamples from sn-roe and test, given this observation?

• Pg 14, para 6; Pg 15, Figure 6
-The author suggests there is a temporal trend in mean age between sn-roe and test fishery, suggesting there may be a trend in the roe fishery becoming more selective. While this is an interesting discussion point, I don't think Figure 6 is enough to support this statement. Please include (or suggest) follow up analyses to explore this trend.
For discussion: How can this information be used to inform gear selectivity?
4) Sex ratios
• Pg 17-18
-Identify the type of sex-related differences in the data that would justify the need to move to a 2-sex model. Examples in other fisheries?
-Results are not linked back to statements in the 1st paragraph: Are there sex-related differences in biological characteristics that should be captured in the stock assessment?
-Why include length ratios by age (Figure 9) when the assessment uses weight-at-age?
5) Relative GSI
• Pg 19 (last para), Figures 11-16; Pg 26, para 2
"…hence samples can be assumed to reflect the spawning populations" – Conclusions from Sept 2010 CSAP meeting were that the test fishery does not sample spawning population in each SAR, only the 'major spawning aggregates'.
-The majority of discussion points from Section 4 relate to comparing biosampling dates with recorded spawn deposition – similar to material presented in Cleary and Benson. Comparisons are much easier to visualize when sampling dates are overlaid on spawning window (Cleary and Benson, Fig 6).
-Specify which years show low GSI at end of sampling period. Does this indicate spawning will occur at a later date or does it support the hypothesis that herring can reabsorb eggs? If the former is more plausible, does this then support the notion that "spawn was missed"?
6) Predicting relative GSI with LMs
-I am not convinced Section 4.1 is necessary. I recommend removing it from the paper.
7) Sample weighting (Pg 29-31)
-Author states the weighting section is included to promote discussion… Should also include suggestions as to what type of additional data should be collected in order to stratify biosamples by catch weight. Eg, we do have access to skipper estimates of volume fish for each test set (and number of associated test samples). Similarly, we can get offload estimates from packers (and number of associated catch biosamples).

Minor edits
Pg 4, para. 3
-Include brief description of workshop (motivation, participants), including dates (June 2010).
-Clarify recommendations: separating data sources and 2-sex model were identified as issues to be explored (i.e., 'should be considered', not 'should be done').
Pg 14, para 2
-"The cluster analysis was run on the combined…" Shouldn't this read independent?
Pg 16, Table 6
-Presentation of results in text (Pg 15) are clear; however, Table 6 is a little confusing. Suggest eliminating table.
Pg 16, Table 7 (heading)
-replace quartile with quantile to be consistent with text.

Concluding remarks
Throughout this review I considered the following six questions:

1. Is the purpose of the working paper clearly stated? [yes]
2. Are the data and methods adequate to support the conclusions? [interesting methods, no overall conclusions]
3. Are the data and methods explained in sufficient detail to properly evaluate the conclusions? [methods well explained, conclusions not linked to management advice]
4. Are the recommendations provided in a form useful to a fisheries manager? [propose 'next steps' linked to provision of management advice]
5. Does the advice reflect the uncertainty in the data, analysis or process? [as above]
6. Can you suggest additional areas of research that are needed to improve our assessment abilities? [development of MSE framework to further this work, i.e., evaluate necessity of test fishery; consider value to stock assessment of fishery-independent biosample data; data requirements for finer-scale management advice; explore cost-benefit/ tradeoffs of varying spatial and temporal sampling coverage vs. precision of parameter estimates]

The initial Request questions the adequacy of the biosample data for providing information on fish size and age composition for stock assessment and modelling.
This paper presents statistical analyses around fish size and age composition however findings are not discussed in terms of stock assessment and modelling. I find that this paper would benefit from a more focused consideration of management-specific questions, including:

- The issue of "what" the test fishery program is sampling is still unresolved. This paper provides some evidence to suggest the test fishery samples the "spawning population" in some years/SARs, however not in others. Identify additional research that will help to clarify this issue.
- Discuss recommendations for evaluating the usefulness of the test fishery data for stock assessment. How do we evaluate the usefulness of this program for stock assessment? For example, with reductions in funding (post-Larocque) would we be better served spending additional funds on improving the spawn survey program and using only biosample data from the catch, OR should we consider developing a fishery-independent source of biological data? Given reductions in funding, how valuable is the biosample data for the provision of science advice (biomass estimates, biological reference points), as compared to the spawn index.
- Clearly identify next steps/ future work that will bring us closer to providing recommendations to fisheries management.

Review of: CSAP Working Paper 2011/P49 January 3, 2011 DRAFT
Investigating changes in herring spawn intensity (layers)
Authors: D.E., Hay, C. Fort, J.F. Schweigert, L. Hamer, and P.B. McCarter

Reviewer: Jennifer Boldt

Thank-you for the opportunity to review the working paper titled “Investigating changes in herring spawn intensity (layers)”. This was an interesting and valuable summary of survey and herring spawn data and will be very useful in guiding further research on this subject. Included below are written comments on the authors’ methodology, interpretations, and recommendations. The comments below are intended to be as constructive as possible and include answers to the following six questions:

- Is the purpose of the working paper clearly stated?
- Are the data and methods adequate to support the conclusions?
- Are the data and methods explained in sufficient detail to properly evaluate the conclusions?
- Are the recommendations provided in a form useful to a fisheries manager?
- Does the advice reflect the uncertainty in the data, analysis or process?
- Can you suggest additional areas of research that are needed to improve our assessment abilities?

A. General comments

1. This was an interesting summary of survey and herring spawn data. The authors did a good job at summarizing a lot of information and the Appendices were helpful.
2. The objectives were clearly presented.
3. The data and methods used to assess the objectives were good, but could use some improvement, additions, and/or clarifications (see comments below).
3. Most of the data presentation was clear (exceptions noted below) and of sufficient detail (exceptions noted below) to properly evaluate the authors’ conclusions regarding those particular data sets. In some cases, it appeared that some data were not presented, or perhaps the results were confusing (see below).
4. Most of the recommendations were provided in a form that would be useful for guiding further research on this subject (see below).
5. Uncertainty was not directly addressed in this paper and improvements in this may be helpful (see comments below).
6. Additional suggestions are included below.

B. Specific comments

1. The authors attribute the change in the mean egg layers to the observed changes in the ‘trace’ category, but do not address the fact that the 1-2 and 2-3+ layer categories also changed. There was a decrease in the 1-2 and 2-3+ layer categories over time and this likely affected the mean egg layers. If there was a decrease in egg layers (due to causes other than methodological sampling protocols), could it be expected that there would be a decrease in the 1-2 and 2-3+ layers and an increase in ‘trace’ layers? Could this indicate that the changes in mean egg layers were due to biological factors instead of or in addition to methodological changes in the survey? This should be addressed prior to ‘publication’ of this paper. The results and abstract sections of the paper attribute the change to survey methodological changes, but the synopsis and review section acknowledge there may be biological causes (i.e., size at age). There seems to be a slight mismatch between information presented in the results and abstract compared to conclusions presented in the synopsis and review.
2. The authors’ recommendation to use 0.1 as the ‘trace’ category rather than 0.01 seems equally arbitrary. A better recommendation may be, as the authors suggest, to examine the effect of setting trace layers to zero as well as other values (simulations run on an expected

range of values?) and examine the effects on the spawn index. In addition, issues in point #1 (above; i.e., changes in 1-2 and 2-3+ layers) should be addressed.
3. The authors discount widespread increases in predation. There is evidence that marine mammal populations (i.e., including whales) have increased in recent years. The authors' conclusions would be better supported by a brief summary of current predator population trends.
4. It would be helpful to include statistical tests to see if there have been significant changes in metrics over time rather than qualitative observations of trends. This could be included in the Appendix 3 Tables 1-5. In addition it would be helpful if the authors included error bars on graphs, where possible (e.g., Figure 5?).
5. Were the data for the duration of the interval between the spawning date and the subsequent survey date presented? See Results, Variation in survey effort, 4th sentence. Table 5 presents qualitative descriptions of trends and refers to Figure A5F3, however, Figure A5F3 shows diver time at stations. Figure A5F2 shows diver survey dates, but it is unclear if this is what the authors are referring to when discussing interval between the spawning date and the survey date. Also, in addition to plotting diver survey dates, the authors could plot the duration, as it looks like there was an increasing trend in duration of surveys (starting earlier and ending later). This should be examined for trends.
6. note: The authors quantified survey effort in various ways. Another metric to consider may be the number of transects made per recorded spawning event (?).
7. Results, Temporal variation in region-metrics versus event-metrics, 2nd paragraph, last sentence (comparing and contrasting the event-metrics to the region-metrics): Actually, the percent of metrics with no change is similar between event- (20%) and region-metrics (17%). The difference between event- and region-metrics is the % of metrics that showed an increase or decrease: a higher percent of event-metrics (30%) showed an increase compared to region-metrics (7%).
8. Table 4: Sometimes having a qualitative description of trends can be very useful. It was unclear however, which trends were significant. For example, Table 4, row 1 indicates a decreasing trend from ~2.5 to <2 mean layers. The data presented in Appendix 2 Figure 6, however, does not appear to show a significant trend, but that cannot be assessed with the information presented (see point #4 above). In addition, it appears that the qualitative descriptions are based on the linear regressions, which do not appear to fit the data in some cases (e.g., mean layers). The authors did state that where the 26-year trend is not clear, the trend is based on recent years. This becomes confusing to follow in the results. This table was not as useful as it could have been, given these considerations.
9. Table 5. Same concerns as for Table 4. Also, what does "no pattern expanding within" mean? (Table 5, row 3). See point # 4 above for additional comments.
10. Results, Shoal/school structure and size versus SSB- potential implications, 4th paragraph, beginning with "Recent evidence from analysis…": the threshold density is cited as 0.2 fish/m^2, but how does this compare to what is seen in BC (i.e., Figure 6)? Or are they comparable? Could the spawning density decline with declining SSB to some threshold value? If the number of shoals is determined by SSB (Croft et al. 2003), and if there are multiple shoals that spawn in the same area, how does that affect egg density?
11. Synopsis and review, Density-dependent mortality- suffocation and predator swamping: It is not clear how this section pertains to the results in this paper.
12. Introduction, 4th paragraph: This paragraph should be re-worded. Why do the authors state that it is "especially important *in the present situation*" that scientists not automatically dismiss the possibility that field observations are incorrect regarding herring biomass? It's not clear what the point is here when discussing different perspectives on the relative abundance of herring.

The authors' conclude in the synopsis and review section that it is "…clear that BC herring stocks have declined significantly in the last 5 years."

13. Appendix, Page 3: Regarding tall vegetation and the records that exceed 100%, the authors state: "It may be a concern about whether the incidence or frequency of the high (>100) estimates is consistent over time." Why is this? What is the basis for this? Is this pertinent to the objectives of this paper? No information is provided to support this statement. How have the incidence or frequency estimates changed over time?

14. Appendix 3, Figure 3: Are the R^2 or p-values for this metric presented somewhere? The abstract states that the spawn area increased for some areas which seems to be the case for two areas.

C. Editorial comments

1. Introduction, 1st paragraph, 4th last sentence: "…expect egg deposition density…" there are some missing words in this sentence.

2. Introduction, 3rd paragraph, last sentence: change "indicates" to "indicate"

3. Introduction, 4th paragraph, 3rd sentence: need a year for the Walters and Hilborn reference.

4. Introduction, 5th paragraph, 3rd sentence: insert "their" prior to "geographical distributions"

5. Indices of spawn and stock assessments and basic assumptions, 1st paragraph, 1st sentence: delete "of" prior to "the number of eggs produced…."

6. Methods, Quantifying spawn survey effort in space and time – an overview and definitions, 1st paragraph, last sentence: "The distance between the outermost…" Is this the outermost margin of spawn?

7. Analysis and test, Temporal and spatial trends in abundance: region-based versus event-based metrics, 1st paragraph, 1st sentence: is the cumulative spawn the linear length of all spawning beds/events? How is the area calculated?

8. Analysis and test, Temporal and spatial trends in abundance: region-based versus event-based metrics, 2nd paragraph, 1st sentence: is a spawn event the same as a spawning bed?

9. Analysis and test, Frequency of trace layers, 2nd sentence, regarding "0.01": it would be helpful at this point to define this as 0.01-0.49 and explain it. The "0.01" is used here and in the first part of the pertinent results that refers to Table 4. Table 4, row 3 shows this as 0.01-0.49 and it isn't explained until after it is referred to in the results, so it was confusing.

10. Analysis and test, Quantifying survey effort, 1st paragraph, 2nd sentence: change "data" to "date" in the phrase "…the duration of the interval between the spawning data and the subsequent survey date…"

11. Analysis and test, Quantifying survey effort, 1st paragraph, 2nd sentence: change "recoding" to "recording"

12. Analysis and test, Quantifying survey effort, 1st paragraph, 3rd sentence: change "later" to "latter"

13. Results and Discussion, Temporal variation in region-metrics versus event-metrics, 1st paragraph, last sentence: The meaning of this sentence is not clear (i.e., "…that weight the trends of each year equally.")

14. Results and Discussion, Variation of layers in time and space, 5th paragraph, 3rd sentence: "(See also Appendix Table Figure3)". Should this be changed to Appendix 4, Figure 3?

15. Synopsis and review, decreases in spawn layers and density of spawning fish, 3rd paragraph beginning with "With respect to explanation 1…", 3rd sentence: insert "of" so the phrase is " Therefore if we assume that two of the three key spawn parameters….".
16. Synopsis and review, Effects of changes in size-at-age and age composition effects on layers, 2nd paragraph, 3rd sentence: change "in" to "are", so the phrase is: "…demographic changes in the population are more likely.."
17. Synopsis and Review, Limitations of spawning habitat, 2nd sentence: delete "herring", so the phrase is "If the density of spawning…"

Review of "Investigating changes in herring spawn intensity (layers)"
Sherri Dressel
Alaska Department of Fish and Game
Commercial Fisheries Division
January 17, 2011

The authors of "Investigating changes in herring spawn intensity (layers)" have done an excellent job compiling and summarizing a large amount of information on herring spawn on the coast of British Columbia and potential factors that affect it. The paper provides a valuable contribution to the interpretation of herring egg deposition patterns as well as a critical re-evaluation of survey methods that is necessary for any agency conducting surveys over long periods of time. Overall, I agree with the authors conclusions and am excited to see the work that was done in order to reach them. I have some suggestions of additional things to consider which could adjust conclusions somewhat, but likely won't change them greatly, as well as additional suggestions for edits that I hope will provide additional strength to the paper. I have included responses to the six main questions below and then have included additional questions and smaller suggested edits at the end.

Is the purpose of the working paper clearly stated?

The authors clearly state two objectives for the paper on page 5 in the section "Objectives for this paper". However, in the first paragraph of the introduction, the authors list questions addressed in the paper which confuse things. It would be helpful if the questions in the introduction were taken out of question form and integrated into the text so it is clear that the objectives of the paper are those listed in the "Objectives for this paper" section.

It would have been very helpful to have heard the information presented in the "Shoal/school structure and size versus SSB – potential implications" section at the beginning of the paper. In fact, hearing a review of what has been written on school, cluster and population structure, as well as the fact that the density of herring within a school is believed to remain static regardless of SSB, would help set up the paper and provide a basis for why the authors are making the comparisons that they do. If this information is provided at the beginning of the paper, the single objective of the paper could be the second objective listed on page 5.

Related to the purpose of the paper is the assumption stated in the first sentence of the abstract, "the density of spawning herring controls the density or layers of eggs". While it is stated as obvious, I believe it is actually in question and should not be assumed. I don't believe that geographic distribution of eggs on the substrate/vegetation necessarily reflects the spatial distribution of fish in a school at a snapshot in time. I do believe that the density of eggs at a particular spot approximately reflects the biomass of fish that spawned on that spot, but do not believe it provides a complete picture of the formation of fish in the water since eggs can be laid in the same spot by different fish over time. For instance, the density of eggs in a location may result from multiple schools of herring spawning over time in the same location. Because I do

not believe that the density of eggs on the spawning grounds necessarily maps the density and location of fish in a school, I do not agree that the density of eggs is controlled by the density of spawning herring (how closely packed fish are in a school). This apparent disagreement with the authors affects interpretations throughout the paper. If, however, the term "density" was used by the authors with a different meaning (if it didn't refer to how closely packed fish are in the water), then it is possible that this difference may be addressed sufficiently by adding a definition in the paper of what the term "density" means.

Given the objective to review information about the factors that affect the estimation of egg layers and the variability in those estimates, the observation that "[c]hanges in criteria used to assess spawn density have occurred in the past (Hay and McCarter 2009)" from page 27 could be a strong support for why the authors are looking for potential changes in survey methodology. It may be useful to include this earlier in the paper and it would definitely be helpful to describe what these changes were.

One of the initial questions in this paper is whether changes in survey methods have affected estimates of SSB. The authors are faced with the difficult task of using the process of elimination to determine whether changes in methods are likely to have occurred. A question which may be worth addressing in a sentence in the introduction is why this process of elimination was necessary and why wasn't it possible to ask the contracted divers if methods have changed?

Are the data and methods adequate to support the conclusions?

Yes, except for a few suggestions noted below, the data and methods are adequate to support the conclusions in this paper.

The authors have done an exceptional job of presenting and eliminating a multitude of explanations other than methodological changes that could account for changes in the number of egg layers. However, they seem certain that they have addressed all of the possibilities when I'm not convinced it is ever possible to know so. Therefore, I believe in some cases more cautious language should be used to allow for the chance that not every possibility has been exhausted.

One possible explanation for the increase in egg density that was not addressed was whether the increase in egg density could have been due to more schools spawning at a site. For instance, one school spawning after another in the same location could result in an increased density of eggs, even if the density of fish in schools remains constant. This question might be addressed by looking to see if the duration of spawning in an event has changed over time. If this occurs, it could explain why the mean length and width of spawn events remain constant, but the mean egg layers follow changes in SSB.

The patterns of density of spawning herring (kg/square m) in Figure 6 made me question whether the methods prior to 1990 could have differed from those after, causing inconsistent calculations of cumulative spawning area over time. The authors indicate in Appendix 2, Figure 5 that spawn records in early years were categorized differently than later years. Could this have affected estimates of cumulative spawning area in Figure 7 or could other methods have changed prior to 1990 affecting cumulative spawning area? The reason I ask is that the extremely high densities of spawning herring in Figure 6 seem like they may be unreasonable in the early years. If the first five years of data are left off, cumulative spawning area (Figure 7) would decrease over time in Haida Gwaii, Central Coast, and West Coast Vancouver Island and densities (Figure 6) in Haida Gwaii, Prince Rupert, Strait of Georgia, and West Coast Vancouver Island would all be in the range of 2-6 km/square m. In the Central coast, for which spawn

widths are half of the other regions (likely indicating steep shorelines), densities would range from 2-10 km/square m if the first five years were left off. So if the pre-1990 data was in question and excluded, the decrease in density addressed as a concern in this paper would be much more moderate and the change in area for over half the regions would decrease as density decreased. While the conclusion to check survey methods would still be valid and useful, the summary of the data presented in the paper could change considerably.

The authors employ an unconventional use of correlation analysis and the rational for why such a use is appropriate should be made clear or another analysis should be used. While it is common to test a correlation between variables that are not independent, such as whether height and weight of individual organisms are correlated, testing a correlation between two variables where one is directly calculated from the other raises concern. I found the p-values helpful, but I wondered whether assumptions were violated in the construction of the test and whether there was a reason why the p-values shouldn't be trusted. Additional explanation and justification would be helpful to address these questions. It is likely that if the p-values aren't appropriate to calculate, the correlations could still be presented without associated p-values.

On page 27, the authors describe size at age as gradually declining for several decades, whereas the decrease in egg layers has been rapid. There is no data included to support this conclusion.

I agree with the authors that an increased number of "trace layers" estimates can affect the overall estimate of eggs and, depending on the corresponding change in area, that the magnitude and direction of change is unknown. As a diver, I also agree for logistic reasons that it may be more likely to count "every last egg" when populations are in low abundance, whereas some areas with "trace layers" may not be counted when populations are in high abundance. Given how egg abundance is calculated, such a change in methods can affect estimates and should be investigated. However, there doesn't appear to be support in the paper for the suggestion that the impact of using 0.01 layers as the arbitrary value is great, why using 0.1 layers would be better, or that using a more realistic estimate of trace layers would mitigate any error associated with the addition of additional low spawning areas (p.28). In order to know these conclusions, it seems that explanatory data summaries or simulations would be necessary. Without these, it seems that the greatest statement that can be made is a suggestion that it be examined to see its impact, without making preliminary conclusions.

Are the data and methods explained in sufficient detail to properly evaluate the conclusions?

More detail is needed in the section "Analysis and tests". The authors describe what they are going to compare, but not how they are going to do so (such as what statistical tests, at what level of alpha are they going to conclude significance, and what statistical software packages are used). For instance, the authors refer to spatially-based analyses, but don't explain to which ones they are referring and how the analyses will be conducted. Similarly, they refer to event-based analyses, analyses of egg layers, examinations of variation in diver survey effort, and variation in types of vegetation, all of which need more description.

As a reader that was not extensively familiar with the spawn survey methods used in BC, explanations of terminology and methods were important for me. The authors have done a good job of familiarizing an outside reader with the BC process, but there were a couple of additional things that would have been helpful. The first is how is an event/record defined? Is it a contiguous patch of spawn on the shoreline? If there is a small gap in the geographical extent of spawn, are two the patches recorded as separate events? How big does the gap need to be to record it as a separate event? Is there a time factor, such that all spawn that occurred at the

same time regardless of location in a region is called an event? If two spawns occur with a separation in time, but happen in the same geographical location (the second spawn occurring on top of substrate that was previously spawned upon) is it one event or two? Without knowing this, I had difficulty understanding and evaluating some of the comparisons that were made.

In some cases, a description of what was said in a reference would have been helpful in addition to a citation. For instance, including the equation referred to in the sentence on page 8 ("Egg deposition for each sampling quadrat is estimated from the predictive equation described in the 1989 assessment (Haist and Schweigert 1990, Schweigert 1993)"), in Appendix 1b, and in Appendix 7 would have been helpful. Similarly, describing what "changes in criteria [were] used to assess spawn density" occurred in the past (Hay and McCarter 2009) would be helpful, as would including a summary of how the spawn index is calculated (p. 12) so the reader would know it is a function of cumulative length(?), width, and layers.

Throughout the paper, the term "biological interactions" was vague enough to be confusing. For example, a more descriptive term would make the following sentence on page 14 more readily understandable: "if two metrics of spawn events (mean length and width) are determined by the biological interactions within the spawn event, it is odd that mean egg layers appear to be under the influence of the SSB".

Are the recommendations provided in a form useful to a fisheries manager?

Yes, authors laid out the recommendations clearly and succinctly and put them in helpful context.

Does the advice reflect the uncertainty in the data, analysis or process?

Many conclusions made by the authors reflected the uncertainty in the data well. For example, the examination of fish size on the abundance of eggs laid was effective as showing why it could not explain the magnitude of increase in egg density. Some conclusions were made with more certainty than I believe is possible, however, given the complex topic. For instance, conclusions were made regarding the density of fish based on the distribution of eggs, even though fish movement cannot be accounted for directly based on eggs alone.

Can you suggest additional areas of research that are needed to improve our assessment abilities?

Because divers' estimates of egg layers are converted to an absolute estimate of abundance/biomass, diver calibration studies may be helpful to evaluate and adjust divers' estimates. In southeast Alaska, diver calibration samples are taken each year to compare divers' estimates of egg abundance with abundance estimates made by subsampling eggs in the lab. Results show that most divers underestimate the number of eggs in a quadrat and the degree of underestimate increases as the egg abundance increases. Results also show that over- and underestimates vary by diver and by kelp type. In southeast Alaska, calibration estimates are used to adjust divers' estimates of eggs for inclusion in the spawn deposition index. If divers in BC differ in their estimation of egg layers, calibrations could be a helpful way to adjust estimates.

When considering changes to survey, using a number of quadrats per transect that is proportional to transect length may provide a sampling method that is less sensitive to the effect of increases or decreases in the number of trace observations. I am not certain that this would remedy the situation, but it could be an option worth investigating.

Additional questions and suggestions:

- Although the authors' point remains unchanged, it doesn't appear that the calculations for mean density in the second example shown in Appendix 7 follow the computation described in Appendix 1b or the equation from a source cited in Appendix 1b (Schweigert 1993). Appendix 1b indicates that average egg density is estimated as the weighted mean of the means of a series of quadrats where the weighting is based on the length of each quadrat. When I interpret that, I get: =400. Also, the calculations show estimates of SSB that aren't referred to in the Appendix legend or in Appendix 1b.
- Is "BC" in appendix graphs the sum of the other areas? Why does it occur in some figures, but not others?

- What does confound mean (p.5, paragraph 1)?

- Does "searching for spawn" mean aerial surveys, skiff surveys, or foot surveys (p.5, paragraph 1)?

- Does "measure spawn" mean dive surveys (p.5, paragraph 1)?

- Does "mis-match between" mean not enough (p.5 paragraph 2)?

- Does "evaluate spawn" mean estimate abundance through dive surveys (p.5, paragraph 2)?

- I wasn't able to understand the following sentence (p.5, paragraph 4): "If egg density is controlled mainly by density-dependent processes operating within spawn events, there may not necessarily be any relationship between total abundance (SSB) and the unique properties of spawn events, especially egg layers". Does it mean that the length, width, and density of events may be the same, but there may be more or fewer of them?

- Does "approximate margins of spawn" mean the length or the width (p.7, paragraph 4)?

- P.7 paragraph 4 - Does "outermost" mean the outermost extent of spawn or the outermost quadrat (or are those the same)? Similarly, is the "point closest to the shore" the point closest to the shore with spawn or the quadrat closest to the shore with spawn?

- On page 19, what does "depth zone close to chart datum" mean?

- Why do the explanations in the last paragraph of p.21 necessarily need to be mutually exclusive?

- Is the conceptual model referred to on page 23 a new one created by the authors or an existing one from the literature?

- If the predation of small benthic organisms removes a consistent amount of eggs each day, doesn't this mean there will be a consistent loss of eggs, not a proportional loss of eggs? (p.26)

- On page 27, the authors say "In recent years, there tends to be clusters of negative residuals [in Cleary et al 2010]". What is recent (e.g. since 1988 or since 2008)?

- I'm not convinced that if the density of spawning herring changes as a function of biomass than this means spawning habitat is limiting (p.27). Additional arguments may be necessary to eliminate other options.

- In Table 2:

    - o Add what bold indicates (does it indicate correlations that are not significant and the α=0.05 level?)

- In Table 4:

    - o For Central Coast mean layers estimated from regions, what does decreasing from ~2.5 to <3 mean?

- In Table 5:

    - o The timeframe being summarized should be clarified. For instance, for Haida Gwaii the diver observation time is increasing since 2004, but constant overall. A second example is that the number of transects in the Central coast decreased since 2000, but they were increasing prior to that.

- In Table 6:

    - o The summaries in Table 6 appear to be for the recent few years, while the graphs that are being summarized present data over a longer period of time. The difference is confusing and may seem selective unless it is clarified in the legend.

    - o What does bold font indicate?

    - o Why is flat kelp highlighted for Central Coast but not Haida Gwaii?

    - o Why is leafy algae for West Coast Vancouver Island not highlighted?

    - o Why is string algae for Strait of Georgia highlighted?

    - o In the legend, what does the term “simple” in “simple observations” mean?

- Appendix 5 figure 4 - If the number of transects were zero in the first few years for several areas, how were transect widths, areas, and egg layers determined? Could this affect the cumulative spawning area calculations in Figure 7, explaining the dome-shaped trends and exacerbating the density decreases estimated in Figure 6?

18. Synopsis and Review, Limitations of spawning habitat, 3rd sentence: “where herring have spawned on nearly 25 of the total…” Should this be 25 %?

19. Synopsis and Review, A computational issue relative to ‘trace’ layers?, 8th sentence, beginning with “In this circumstance, the decrease the mean layer…”: insert “in”, so the phrase is: “In this circumstance, the decrease in the mean layer…”

20. Summary and Recommendations, 4th paragraph, 1st sentence: insert “an”, so the phrase is: “….diver surveys could lead to an increase in exposure time…”

21. Appendix 3, Figures 2-7: the figure captions should be changed to refer to a Figure not a Table and Figure: e.g., Appendix Table 3, Figure 2.

22. Appendix 3, Figure 7: no linear regression line is shown.

23. Appendix 6, Figure 2: label the panel (a) and (b) in the figure and identify panel (a) in the caption. Also, what are the error bars showing?

24. Appendix 6, Figure 3a: what lighthouses were used?

25. Appendix 7, caption, 2nd last sentence: insert “by”, so the phrase is: “When the mean density is multiplied by total spawn area…”
26. Figure 5 caption: “…mean number of egg layers from spawn events is compared to the SSB by linear regression…” This graph does not include SSB.