Sampling Intensity, Aging Intensity, and its Effect on an Age-Structured Assessment of Cisco in Thunder Bay

QFC Technical Report T2018-02

DOI: 10.6084/m9.figshare.6167522

Prepared by

Nicholas C. Fisch
Quantitative Fisheries Center
Department of Fisheries and Wildlife
Michigan State University

and

James R. Bence
Quantitative Fisheries Center
Department of Fisheries and Wildlife
Michigan State University

Introduction

Herein we examine how reductions in the commercial sampling efforts of the Ontario Ministry of Natural Resources and Forestry (OMNRF) related to the Thunder Bay cisco fishery might influence the performance of a statistical catch-at-age assessment (SCAA) model for the Thunder Bay cisco stock. Specifically, we emulated a decrease in sampling or aging intensity by randomly selecting subsets of data, to see if this resulted in substantial changes in estimates or evidence of more variability and uncertainty in results from the SCAA model. This report does not fully document the assessment model, rather it documents the changes made to the Thunder Bay commercial cisco database and the performance of the SCAA in the face of those changes. These changes determine the estimated observed age composition of cisco caught in the commercial fishery in Thunder Bay, a data source that the model is fit to (among others). Readers are referred to Fisch (2018) for additional details on specifics of the SCAA.

Background

As part of a MS thesis, a SCAA model was developed for Thunder Bay cisco (Fisch 2018). This model was fit to four main data sources, one of which was the observed age composition of cisco from the commercial fishery. This observed age composition was estimated from samples of the fishery harvest residing in OMNRF's commercial cisco database. This database contains information from samples of the first 10 cisco caught in each gillnet in the commercial cisco fishery (from 1999-2015). These fish were measured to the nearest millimeter (total length and fork length), sexed, and weighed to the nearest gram. Ages were subsampled in many years based on age length key (ALK) fixed allocation bin sampling where 10 fish belonging to a 10 mm length bin of each sex and zone were selected to be aged. If less than 10 fish were in a 10 mm bin, then all fish in that bin are aged. In some years all cisco sampled from commercial gillnets were aged. Annual subsampling of cisco caught in commercial gillnets for aging is summarized in Table 1 (Columns 2-3).

For use in the SCAA model, observed age composition data for each year were estimated by pooling age data for management zones 1-4 each year and applying sex-specific ALKs. Thus removing data from the commercial database will result in different estimates of the observed commercial fishery age composition used in the SCAA. Ultimately our goal for this project was to simulate reduced amounts of biological sampling in the database and re-fit the SCAA model to determine if any decline in assessment performance is realized.

Methods

We simulated 4 different subsets of the commercial cisco database. They are as follows (in addition to their relation to the original database):

1) Instead of sampling 10 cisco per gillnet, 5 cisco per gillnet were sampled at random (out of the previous 10 sampled).

2) Instead of sampling 10 cisco per gillnet, 3 cisco per gillnet were sampled at random (out of the previous 10 sampled).

3) Instead of aging 10 cisco per zone per sex per 10mm length bin, 5 aged fish per zone per sex per 10mm length bin were randomly sampled to create an age-length key.

4) Instead of aging 10 cisco per zone per sex per 10mm length bin, 3 aged fish per zone per sex per 10mm length bin were randomly sampled to create an age-length key.

The first two scenarios focus on sampling intensity and aging intensity. While the latter two focus on aging intensity (to develop ALK). For the first two scenarios, when we reduce the number of cisco sampled per gillnet in an effort to reduce sampling intensity, some of the fish dropped from the dataset will have been aged. Thus, we are also randomly dropping aged fish (hence focusing on sampling intensity and aging intensity in some sense). Conversely, for the latter two scenarios when we drop fish from the age-length key, only the age data of those fish is excluded. Those fish are still present in the length composition, however, their age is not used in the derivation of the age-length key (hence solely focusing on aging intensity). Total number of fish sampled and aged for each subset scenario can be found in Table 1.

For each subset of the commercial cisco database, processing akin to the original SCAA was performed, i.e. aging data were pooled by management zone each year and sex-specific age-length keys were developed to estimate the observed age composition each year. Models were first run from a penalized maximum likelihood context to perform iterative reweighting of effective sample sizes (ESS) for age composition datasets (commercial age composition and fishery independent age compositions, see Fisch 2018) using method T3.4-TA1.8 of Francis (2011). Once effective sample sizes converged, the models were run in a Bayesian context. All assessment models were run for 20 million iterations, saving every 500th and burning in 2500 iterations from the final chain. Model convergence was assessed based on chains of the model estimated parameters using Geweke's diagnostic at an alpha level of 0.01. We compared the original Thunder Bay SCAA to models fit to the subsets of the database by examining changes in point estimates (medians) and estimated uncertainty through highest posterior density intervals and CVs of posterior distributions for quantities such as spawning biomass and natural mortality. We chose to focus on spawning biomass as this value may be used to calculate quotas in the future and thus is of management interest, and natural mortality as this is a parameter of ecological interest.

Results

All models converged and the resulting point estimates were highly similar to the original full model point estimates of natural mortality and spawning biomass at the end of the time series (Figures 1-3). Effective sample size for commercial fishery age composition data was 87, 61, 62, and 30 for scenarios 1-4, respectively. For context, the ESS for commercial fishery age composition data was 62 for the original SCAA model (fit using original commercial database). Thus the only scenario that provides evidence of a decrease in information content of the resulting age composition data is when only 3 fish were aged per zone and sex per 10mm bin.

Relative differences between point estimates of the final year of spawning biomass for the original SCAA compared to the model subsets were all well below 5% (Table 2). Although point estimates for spawning biomass were slightly different early in the time series, from 2011 to the end of the time series all subset scenarios converged on similar estimates to the original SCAA (Figure 2). Estimated uncertainty of spawning biomass in the final year of the assessment also did not seem to differ from the original SCAA to the subset scenarios (Figure 3). Mean CVs of the posterior distributions of spawning biomass for the full time series were all between 0.245-0.275 and CVs of the final year of spawning biomass were all within 0.205-0.225 (Table 2). Thus the subsetting procedures did not appear to influence the perceived uncertainty in estimates of spawning stock size.

Discussion

Results indicate that decreases in sampling or aging intensity identified in this report have little effect on SCAA model performance. Thus, if the sole purpose of cisco commercial harvest sampling is to inform the assessment of cisco, a reduction in biological data collection would likely have little effect on SCAA performance for cisco in Thunder Bay. Most of the effective sample sizes for the commercial fishery age composition dataset did not decrease compared to the original model, indicating that these reductions in sampling and aging intensity are not decreasing the information content of the data (Exception for subset 4 with ESS of 30). In this study we simulated decreasing the sampling intensity based on the amount of cisco sampled from each gillnet and decreasing the aging intensity based on the number of fish aged per 10mm length bin, per sex, per zone. Our aim here was to maintain the way the data were collected (e.g., X amount of fish per gillnet), and solely change the amount of data collected. That way, the same procedures could be undertaken although potentially collecting or aging fewer cisco. The reason the information content of the age composition data did not decline until aging was severely reduced is likely in large part due to the extensive data collection scheme, where cisco from every single gillnet are sampled and a robust ALK is created at fine length bins with ages spread out both temporally and spatially. Such a sampling design will tend to minimize variability when there is variation among fishing days and trips, compared to designs where a larger number of fish are sampled from a small set of fishing days and trips. In addition, even when the aging was so reduced as to lower the information content of the age compositions, critical assessment results needed for management were largely unchanged.

Acknowledgments

We would like to acknowledge the Great Lakes Fishery Commission for funding this project and the OMNRF for providing us access to their impressive dataset on Thunder Bay cisco. This work was supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research, the Michigan DNR though its support of the Partnership for Ecosystem Management at MSU, and the Quantitative Fisheries Center at MSU. This is QFC publication 2018-09.

References

Fisch, N.C. 2018. On the Assessment and Sustainability of Thunder Bay Cisco. Masters Thesis, Michigan State University, published by ProQuest.

Francis, R.I.C.C., 2011. Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci. 68, 1124–1138.

Tables

Table 1. Description of commercial cisco sampling for the original model and each subset scenario. Subsampled column (Column 2) identifies which years utilized an age-length key in the original model and column 3 displays the number of cisco sampled and aged in the original OMNRF database. Columns 4-7 of the table denote the subset scenarios described in methods. The first number in columns 3-7 represents the number of cisco sampled and the second represents the number aged (Number sampled – Number aged).

Year	Subsampled	Original Model	Subset 1	Subset 2	Subset 3	Subset 4
1999	Yes	860 – 402	439 – 215	265 – 130	860 – 218	860 – 148
2000	Yes	3241 – 594	1601 – 281	938 – 179	3241 – 268	3241 – 189
2001	Yes	1221 – 574	608 – 286	364 – 178	1221 - 304	1221 – 215
2002	Yes	1147 – 676	573 – 336	342 – 202	1147 – 312	1147 – 212
2003	Yes	1208 – 704	605 – 358	357 – 206	1208 – 308	1208 – 209
2004	Yes	1091 – 527	547 – 264	328 – 166	1091 – 268	1091 – 181
2005	Yes	661 – 280	329 – 135	192 – 74	661 – 187	661 – 130
2006	Yes	644 – 378	325 – 179	194 – 119	644 – 168	644 – 111
2007	Yes	839 – 330	419 – 165	248 – 96	839 – 206	839 – 136
2008	No	654 – 654	329 – 329	197 – 197	654 – 182	654 – 123
2009	No	638 – 637	320 – 320	192 – 192	638 – 224	638 – 149
2010	Yes	500 – 219	251 – 111	146 – 58	500 – 120	500 – 88
2011	No	563 – 562	281 – 281	169 – 169	563 – 163	563 – 107
2012	No	478 – 477	240 – 240	144 –144	478 – 165	478 – 113
2013	No	429 – 427	215 – 215	129 – 129	429 – 153	429 – 103
2014	Yes	733 – 517	367 – 256	219 – 153	733 – 206	733 – 146
2015	Yes	705 – 457	351 – 228	208 – 141	705 - 207	705 – 149

Table 2. Table of various statistics used in comparing models. Column 1 depicts relative difference for spawning biomass in 2015: (Original SCAA-Subset)/Original SCAA. Column 2 depicts the mean CV for the posterior distributions of spawning biomass over the time series. Column 3 depicts the CV of the posterior distribution of spawning biomass in 2015.

Subset	Relative Difference	Mean CV	2015 CV
Full Model	NA	0.255	0.209
1	2.61%	0.249	0.208
2	-1.52%	0.257	0.218
3	2.82%	0.273	0.215
4	-0.08%	0.268	0.223

Figures

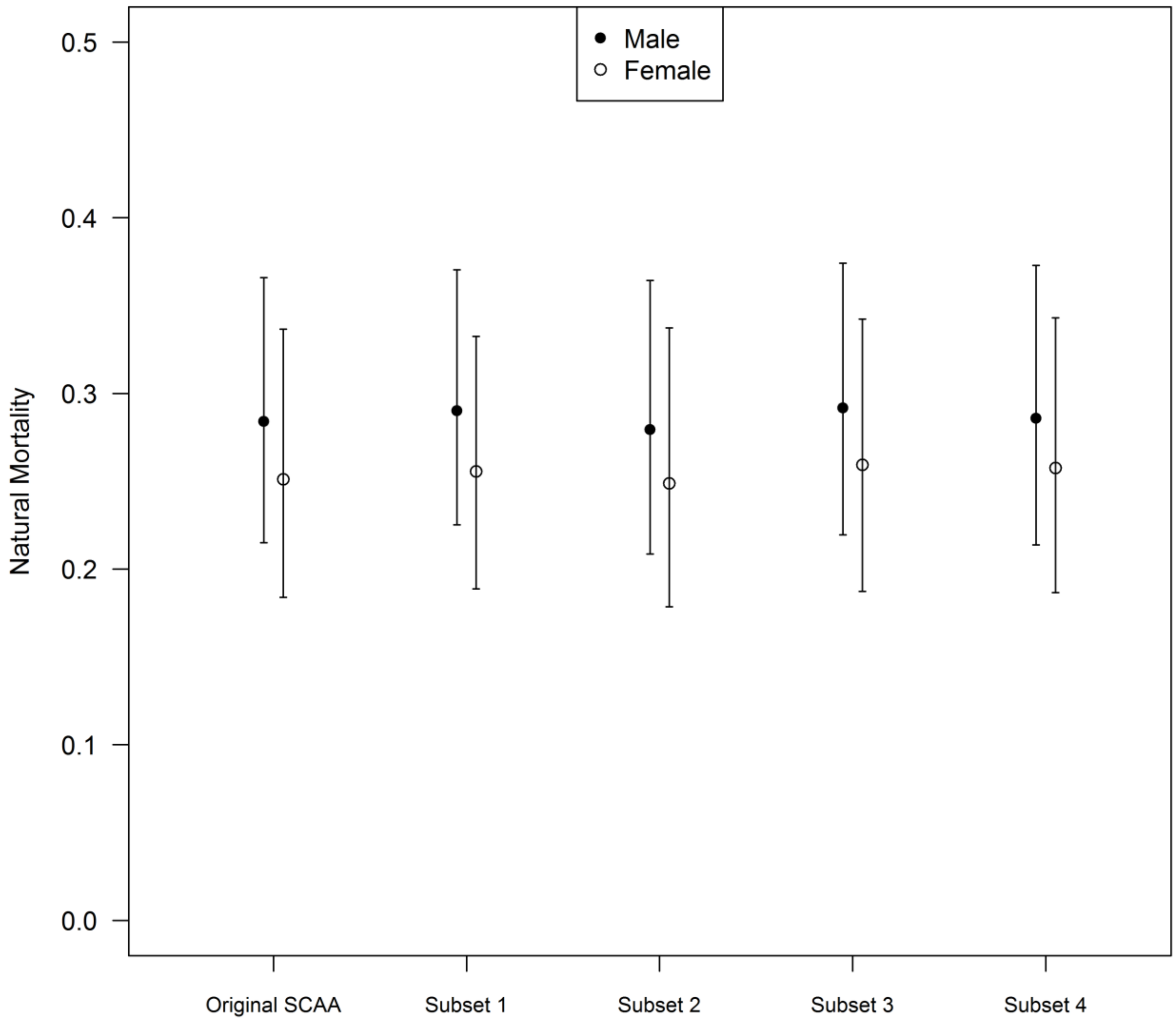

Figure 1. Natural mortality for the original SCAA and fits to the model subsets. Shown are medians of the posterior distribution and 95% highest posterior density intervals. Filled points denote estimated natural mortality for males and hollow points represent natural mortality estimated for female cisco. “Original SCAA” refers to the original stock assessment on Thunder Bay cisco using the full OMNRF commercial cisco database. “Subset 1” refers to an SCAA fit to a subset of the OMNRF commercial cisco database where 5 fish were sampled randomly per gillnet (instead of the original 10). “Subset 2” refers to an SCAA fit to a subset of the OMNRF commercial cisco database where 3 fish were sampled randomly per gillnet (instead of the original 10). “Subset 3” refers to an SCAA fit to a subset of the OMNRF commercial cisco database where an age-length key was developed from 5 aged fish in each 10mm length bin sampled randomly per zone and per sex (instead of the original 10). “Subset 4” refers to an SCAA fit to a subset of the OMNRF commercial cisco database where an age-length key was

developed from 3 aged fish in each 10mm length bin sampled randomly per zone and per sex (instead of the original 10).

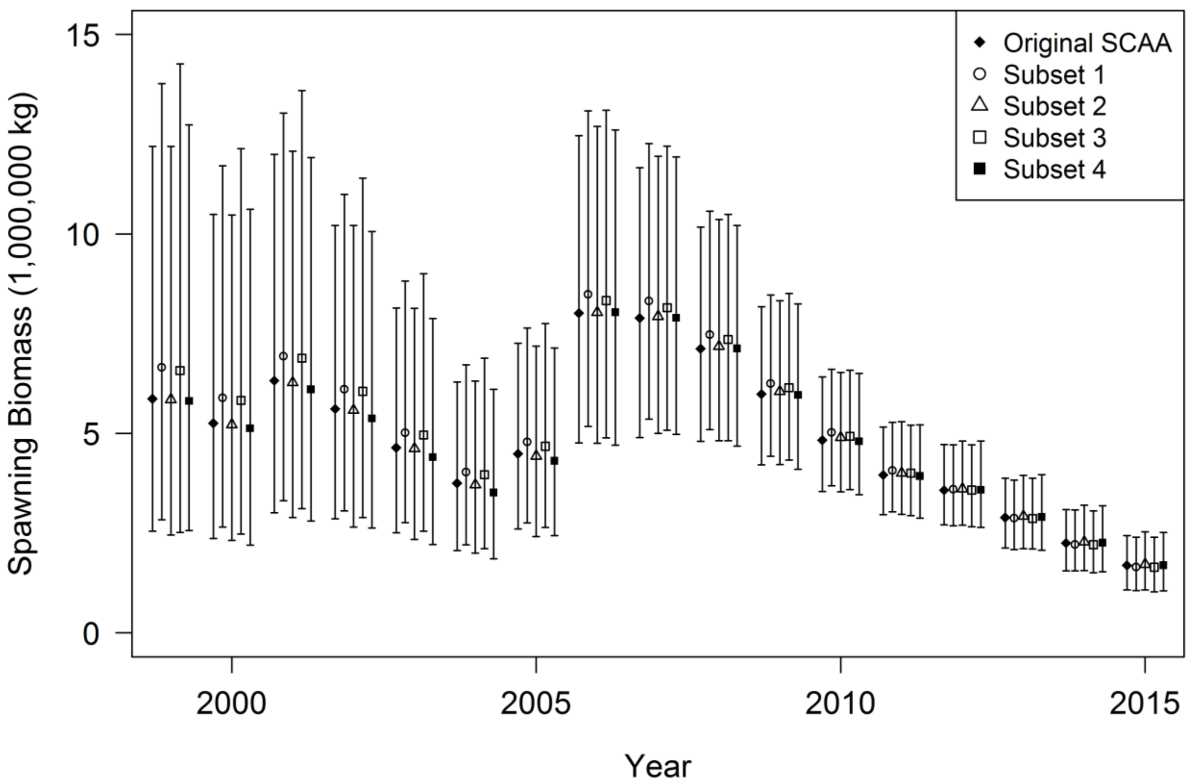

Figure 2. Spawning biomass estimates for each assessment model. Shown are medians of the posterior distribution and 95% highest posterior density intervals. Subset scenarios are as described in Figure 1.

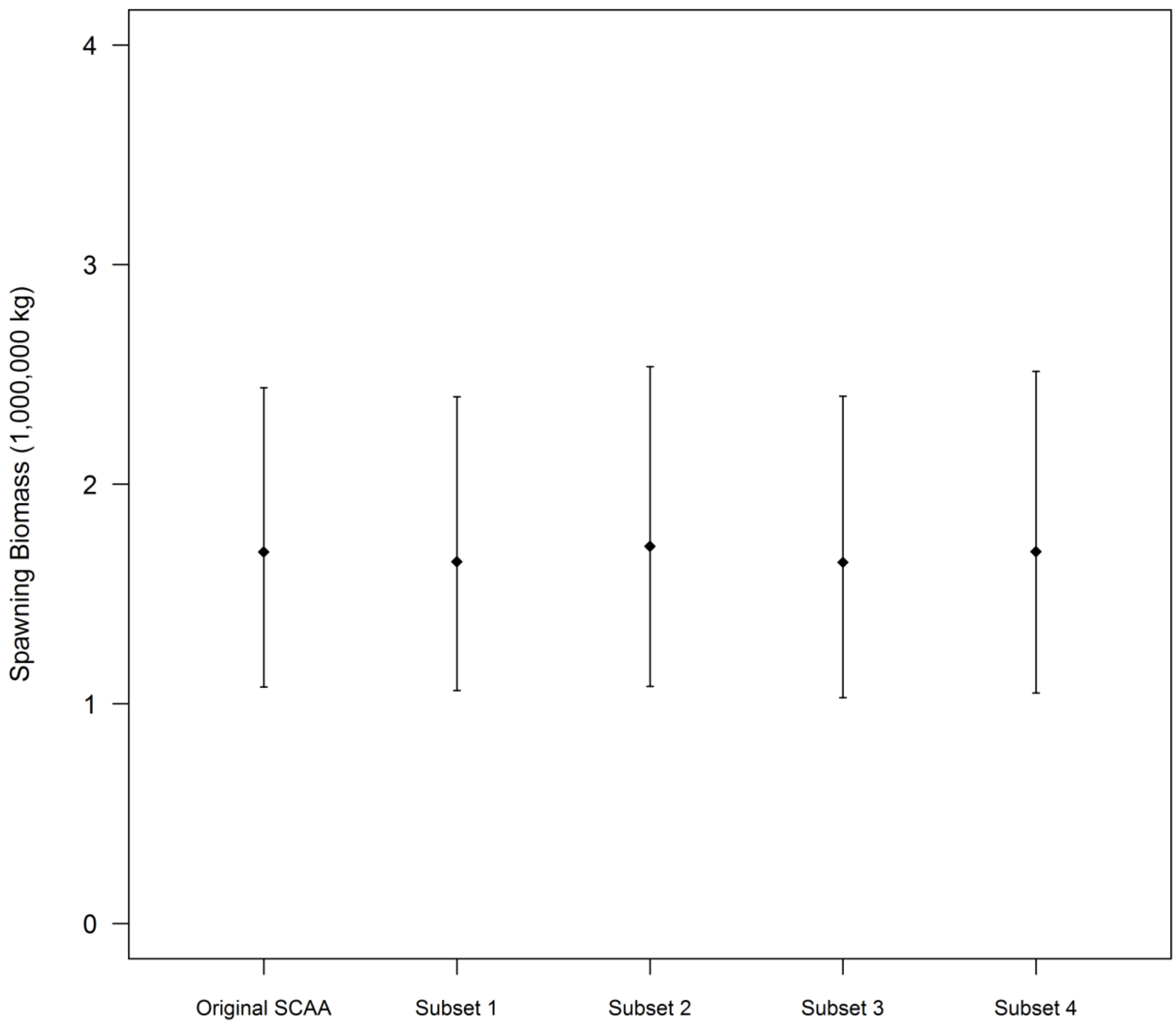

Figure 3. Estimates of spawning biomass in the final year of the assessment. Shown are medians of the posterior distribution and 95% highest posterior density intervals. X axis is defined as in Figure 1.