BSAI Groundfish Plan Team
Subcommittee on Pacific cod models

Minutes of the June 30, 2017 meeting

Introduction

Beginning with the 2010 assessment cycle, lists of models to be included in the preliminary (September) stock assessments for Pacific cod in the Eastern Bering Sea (EBS) and Aleutian Islands (AI), and usually the Gulf of Alaska (GOA), have been developed by either the Joint BSAI and GOA Groundfish Plan Teams or some subset thereof. Each year, the lists have been developed during a WebEx meeting held sometime during the first half of the year. From 2010-2013, the full Joint Teams participated in these meetings. From 2014-2016, this responsibility was delegated to a Joint Team subcommittee (Subcommittee). Developing lists of models for all three Pacific cod assessments (EBS, AI, and GOA) within a single meeting often proved to be challenging. At the request of the GOA assessment author, the GOA assessment was removed from the agenda, beginning with the 2016 meeting. Given that the GOA assessment is no longer on the agenda, development of this year's lists of models for the EBS and AI assessments was delegated to a subset of the BSAI Team (Team) membership rather than a subset of the Joint Teams membership. This year's Subcommittee meeting took place via WebEx on June 30, from 2:00-5:00 p.m. PDT. The subcommittee consisted of Dana Hanselman (chair), Alan Haynie, Allan Hicks, and Diana Stram. All members were present. Assessment author Grant Thompson provided background information, answered questions, and served as rapporteur, but did not participate in decision making. Also present were Chad See (Freezer Longline Coalition) and Paul Wilkins (Freezer Longline Coalition and Coastal Villages Regional Fund).

It was noted that, in a departure from past practice, the SSC has stated that it will no longer review the lists of models developed by the subcommittee (June 2016 SSC minutes), so the lists presented below will be considered final.

Background

A document including the following was provided to the subcommittee one week prior to the meeting:

- Links to last year's EBS/AI SAFE report chapters (2/2A), including appendices describing the preliminary assessments (2.1/2A.1) and the complete histories of assessment models (2.3/2A.3).
- Full text and summaries of last year's Subcommittee, Team, and SSC comments that are relevant to this year's assessments (26 in total, including a few duplicates).
- Descriptions of, and results from, 10 new models of the EBS stock, all but three of which include features that respond to comments from last year.
- Author's responses to selected comments from last year not addressed in the new models.
- Author's suggestions for other features to consider in this year's models.
- Draft template for use in allocating features among this year's preliminary assessment models.

The background document (with minor editing), is included as Appendix A to these minutes. In the interest of efficiency, content from that document (including numbering of model features listed therein) is not repeated in the main text of these minutes.

Although not included it in the background document, Grant's presentation on background material also addressed some issues regarding the use of empirical weight-at-age data in last year's EBS assessment. He reviewed the concerns that were expressed in last year's EBS assessment and presented some results that were discussed during last year's December SSC meeting. These are summarized in Appendix B.

Procedure, recommendations, and other conclusions

The Subcommittee's deliberations focused on amending, and filling out, Table A5 from the background document (Appendix A). The list of model features in Table A5 included all comments from last year that were not considered to be either purely procedural or duplicative, along with the 7 new features that had been included in one or more of the 10 new models developed prior to the meeting.

During the course of its deliberations, the Subcommittee developed 6 new features of its own, which were as follow:

- New1: *Allow time-varying selectivity for the fishery but not the survey*. This feature was developed as a more restrictive alternative to features Sub3 and GT6.
- New2: *Estimate survey index standard error internally ("extra SD" option in SS).*
- New3: *Use Francis weighting*. This feature was developed as a more restrictive alternative to feature BPT4.
- New4: *Give less weight to fishery comps than survey comps*, *less to sizecomps than agecomps*. This feature was developed as a substitute for features BT4, GT4, and New1.
- New5: *Use Hamel's (2015) method to develop a prior distribution for natural mo*rtality. This feature was developed as a substitute for feature SSC6.
- New6: *Report Francis weights from the terminal run if harmonic mean is used and vice-versa*.

All of the above were intended to apply to the EBS assessment only, with the exception of feature New5, which was intended to apply to both the EBS and AI assessments.

Features also considered in the deliberations included three suggested by the author in the background document (but not included among the 10 new models developed prior to the meeting), which the Subcommittee summarized as follows:

- GT8: "Do not use currently available fishery agecomp data, but do add new fishery agecomps."
- GT9: "Switch to length-based maturity or adjust age-based maturity for ageing bias."
- GT10: "Include EBS survey strata 82 and 90 (NW corner of EBS) in the data."

All of the above were intended to apply to the EBS assessment only, with the exception of feature GT9, which was intended to apply to both the EBS and AI assessments. See Appendix A for the rationale behind these features.

Not counting the standing requirement to include the current base model in the preliminary assessment, a total of 32 features were evaluated by the Subcommittee (6 from last year's Subcommittee meeting, 5 from last year's BSAI Team meetings, 5 from last year's SSC meetings, 10 from the author, and 6 from this year's Subcommittee meeting). Of the 32 features, 14 were originally intended for application to both the EBS and AI assessments, 14 were originally intended for application only to the EBS assessment, and 4 were intended for application only to the AI assessment.

The following decision was not reached until the end of the meeting, but reporting it here makes it much easier to describe the remainder of the deliberations:

The Subcommittee decided not to develop a list of models to be included in the preliminary 2017 AI assessment, thus allowing the assessment author to devote more time to the preliminary 2017 EBS assessment. The Subcommittee notes that a similar recommendation was made by both the full BSAI Team and the SSC with respect to the final 2016 assessments. The Subcommittee also notes that the author has the discretion to bring forward one or more models for the preliminary 2017 AI assessment if he so chooses. If he does not, the Subcommittee anticipates that AI Pacific cod will continue to be managed under Tier 5 in 2018.

The Subcommittee began the process of allocating features to models by prioritizing the 32 features on a scale of 0 to 2. A priority of 2 was interpreted to mean that the feature would be included in at least one model in the Subcommittee's list for this year's preliminary EBS assessment. A priority of 1 was interpreted to mean that the feature would not be included in any of the models in the Subcommittee's list for this year's preliminary EBS assessment, but the Subcommittee recommends that it be considered for inclusion in one or more future models (in either the EBS assessment, the AI assessment, or both, depending on the area of application intended in the original proposal). A priority of 0 was interpreted to mean that the feature will not be included in any of the models in the Subcommittee's list for this year's preliminary EBS assessment, and that the Subcommittee does not recommend that it be considered for inclusion in any future models. Of the 32 features, 15 were assigned a priority of 2, 7 were assigned a priority of 1 (including all 4 features that were originally intended for application only in the AI assessment), and 4 were assigned a priority of 0. The remaining 6 features were determined to be most appropriately interpreted as involving only non-model analyses, meaning that, if they are to be addressed, this should be done outside of any particular assessment model.

After prioritizing the list of features, the Subcommittee allocated the 15 features that were assigned a priority of 2 among five new models to be presented in this year's preliminary EBS assessment. The Subcommittee also chose two of the 6 non-model analyses for inclusion in this year's preliminary EBS assessment.

The results of the Subcommittee's deliberations are shown in Table 1, upon which the following pair of recommendations are based:

The Subcommittee recommends that the following models be included in this year's preliminary EBS Pacific cod assessment (note that model labels shown here are temporary placeholders; actual model labels for September will be established during the analysis, except for Model A, which corresponds to Model 16.6):

- Model A: Model 16.6 (last year's final model), after translating from SS V3.24u to V3.30.
- Model B: Same as Model A, but with the following features added:
    1. Adjust timing of the fishery and survey in SS.
    2. Do not use currently available fishery agecomp data, but do add new fishery agecomps.
    3. Switch to haul-based input sample size and catch-weighted sizecomp data.
    4. Develop a prior distribution for natural mortality based on previous estimates.
    5. Switch to age-based, flat-topped, double normal selectivity.
    6. Allow random time variability in selectivity, with σs fixed at the restricted MLEs.
- Model C: Same as Model B, but with the following features added:
    1. Use harmonic mean weighting of composition data.
    2. Allow time-varying selectivity for the fishery but not the survey.
- Model D: Same as Model B, but with the following features added:
    1. Use harmonic mean weighting of composition data.
    2. Estimate survey index standard error internally ("extra SD" option in SS).
- Model E: Same as Model B, but with the following feature added:
    1. Use Francis weighting.
- Model F: Same as Model B, but with the following feature added:
    1. Give less weight to fishery comps than survey comps, less to sizecomps than agecomps.

The Subcommittee recommends that the following non-model analyses be conducted for the preliminary 2017 EBS assessment:

- Compare σ_R to the RMSE of estimated recruitment deviations.
- Report Francis weights from the terminal run if harmonic mean is used and vice-versa.

With respect to the sets of features recommended for inclusion in Models B-F, the Subcommittee notes that several of those features satisfy various other features, which, in the interest of efficiency, are therefore not addressed explicitly in the descriptions of Models B-F. Specifically:

- Inclusion of feature GT5 ("Switch to haul-based input sample size and catch-weighted sizecomp data") in Models B-F satisfies feature SSC3 ("Set input sample size based on number of hauls sampled for length").
- Inclusion of feature GT6 ("Allow random time variability in selectivity, with σs fixed at the restricted MLEs") in Models B-F satisfies features Sub3 ("Use reasonably time-varying, double normal selectivity") and BPT5 ("Investigate the appropriate amount of time-variability in fishery selectivity").
- Inclusion of feature GT4 ("Use harmonic mean weighting of composition data") in Models C-D satisfies feature BPT4 ("Use either Francis or harmonic mean weighting").
- Inclusion of feature New3 ("Use Francis weights") in Model E satisfies feature BPT4 ("Use either Francis or harmonic mean weighting").

With respect to implementation of the above recommendations, the Subcommittee reached the following conclusions:

- For feature GT5 ("Switch to haul-based input sample size and catch-weighted sizecomp data"), the Subcommittee understands that the author will likely set initial input sample sizes equal to the number of hauls (or sets), rather than a more complicated haul-based approach such as that described by Stewart and Hamel (2014).
- For feature SSC6 ("Develop a prior distribution for natural mortality based on previous estimates"), if faced with a choice between the lognormal and normal examples given in the background document (see Appendix A), the Subcommittee prefers the lognormal.
- For feature New4 ("Give less weight to fishery comps than survey comps, less to sizecomps than agecomps"), which is used in Model F, if the Francis weightings obtained in Model E accomplish the same thing, then Model F does not need to be included. Also, the Subcommittee's preferred method for implementing feature New4 is to begin with the weightings obtained in Model E and then adjust them as little as possible subject to the constraints described by this feature.
- For feature New6 ("Report Francis weights from the terminal run if harmonic mean is used and vice-versa"), the confidence intervals surrounding the Francis weights should also be reported.

With respect to some of the features that did not receive a priority rating of 2, the Subcommittee reached the following conclusions:

- Implementing feature SSC4 ("Develop a spatially structured, combined BS/AI/GOA assessment model") would be an immense undertaking. The Subcommittee feels that this would be better suited to an exploratory analysis by a post-doctoral researcher than to incorporation in a routine assessment.
- The Subcommittee does not intend for feature Sub1 ("Allow time variability only where supported by external data") to apply to recruitment. Also, by assigning a priority of 1 to feature Sub1 and a priority of 0 to feature GT7 ("Allow deterministic time variability in selectivity, based

on environmental indices," such as was used in Models 9 and 10 in Appendix A), the Subcommittee intends to suggest that any future use of time variability in selectivity that is conditioned on external factors should be of the "time block" form, such as linking fishery selectivity to changes in fishery regulations. Alan Haynie offered to work with Chad See on developing a list of past regulatory changes that might be expected to have had an effect on fishery selectivity, to be presented to the full BSAI Team in September.

- Although the Subcommittee feels that it will not be possible to incorporate feature Sub2 ("Examine survey data from the northern Bering Sea") into this year's preliminary EBS assessment due to the fact that there is currently only a single year's worth of data in the modern NBS survey time series and the results from this year's NBS survey will likely not be ready for inclusion until after the preliminary assessment is due, it may be possible to include this feature as a non-model analysis in the final assessment.
- Although the Subcommittee also feels that features BPT3 ("Continue to compare empirical weight at age with the traditional approach") and GT10 ("Include EBS survey strata 82 and 90 (NW corner of EBS) in the data") should not be included as non-model analyses in this year's preliminary EBS assessment, it may be appropriate to include them as non-model analyses in the final assessment (note that GT10 would be appropriate to pair with Sub2, above).

Comments SSC1 and SSC12 both encourage further exploration of model averaging, but the latter comment makes such exploration conditional on the outcome of the SSC's February 2017 workshop on that subject. The minutes of that meeting conclude, "The SSC would like to see a 'test case' of how ensemble modeling works for one of our groundfish stocks." As noted in Appendix A, at least one SSC member suggested that the test case be conducted during an "off" year for an assessment that is not produced annually. **The Subcommittee concluded that the EBS Pacific cod assessment is not a good candidate for model averaging at this time.**

Comments from individual Subcommittee members

The following comments were made by individual Subcommittee members during the discussion, and do not necessarily represent the consensus of the Subcommittee:

- The following features should be included in all assessment models:
    - SSC3: Set input sample size based on number of hauls sampled for length.
    - BPT4: Use of either Francis or harmonic mean weighting of composition data.
    - New2: Estimate survey index standard error internally ("extra SD" option in SS)
- Assessments should always include "sensitivity" runs, designed to reveal the implications of possible errors in assumptions, but showing only a subset of the results required for full models.
- Sensitivity runs should not be considered as true alternatives for setting harvest specifications. Perhaps this could be ensured by presenting results of sensitivity runs in graphical form only.
- Deciding what should be included in all assessments, or all assessment models, should be a task for the AFSC rather than this Subcommittee.
- Should our process be incremental, where the features of the current base model are "tweaked" individually in the subsequent assessment?
- We should distinguish between "operational models" (those used for harvest specifications and status determinations) from "research models" (those used to explore alternative assumptions), and try to keep the operational models fairly constant over time, changing them only when the results of the research models indicate a clear advantage to doing so.
- How should we handle the situation in which we are basically indifferent between two or more model structures that have significantly different management implications?

- We need to document *why* we choose particular model configurations for inclusion in the assessment. Perhaps we should include a table describing why each feature was chosen.
- Time-varying selectivity for the fishery is easily justified, for example, as a result of changing regulations and the changing mix of gear types over time. Time-varying selectivity for the survey is less easily justified.
- If external covariates are used to justify time-varying survey selectivity, they should be factors related to the results of the survey itself (e.g., the proportion of small fish observed in the shallow stations) rather than physical environmental factors such as those used to develop Models 9 and 10 in Appendix A.
- It would be good to have a better idea of the effects, if any, of choosing the initial input sample sizes when using either Francis or harmonic mean weighting.
- Is it appropriate to reweight the input sample sizes for the individual records (years) within the various compositional data sets, rather than just reweighting the average input sample sizes for those sets?
- The prior distributions for *M* that were developed in the background materials (Appendix A) seem too broad. Perhaps some points in the data set should be down-weighted, or eliminated entirely (e.g., the estimates derived from Asian or Canadian stocks).

References

Stewart, I. J., and O. S. Hamel. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. *Can. J. Fish. Aquat. Sci.* 71: 581-588.

Hamel, O. S. 2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES Journal of Marine Science* 72(1):62-69.

Table 1. Allocation of proposals among models (A-F) to be presented in the preliminary (September) assessment for the EBS only.

Note: Model labels shown here (A-F) are temporary placeholders; actual model labels for September will be established during the analysis.

Definitions: Area = intended area of application, BPT = BSAI Plan Team, GT = Grant Thompson, New = proposal developed during this meeting, NMA = non-model analysis, No. = proposal number, Pri. = priority, SPM = starting point model, SSC = Scientific and Statistical Committee, Sub = proposal developed during May 2016 Subcommittee meeting

Row	No.	Proposal summary	Pri.	Area	SPM	A	B	C	D	E	F	NMA
0	n/a	Base model	n/a	both	16.6	x						
1	GT1	Adjust timing of the fishery and survey in SS	2	both	16.6		x	x	x	x	x	
2	GT8	Do not use currently available fishery agecomp data, but do add new fishery agecomps	2	EBS	16.6		x	x	x	x	x	
3	SSC3	Set input sample size based on number of hauls sampled for length	2	both	16.6		x	x	x	x	x	
4	GT5	Switch to haul-based input sample size and catch-weighted sizecomp data	2	both	16.6		x	x	x	x	x	
5	SSC6	Develop a prior distribution for natural mortality based on previous estimates	2	both	16.6		x	x	x	x	x	
6	GT2	Switch to age-based, flat-topped, double normal selectivity	2	both	16.6		x	x	x	x	x	
7	Sub3	Use reasonably time-varying, double normal selectivity	2	EBS	16.6		x	x	x	x	x	
8	BPT5	Investigate the appropriate amount of time-variability in fishery selectivity	2	EBS	16.6		x	x	x	x	x	
9	GT6	Allow random time variability in selectivity, with σs fixed at the restricted MLEs	2	both	16.6		x	x	x	x	x	
10	BPT4	Use either Francis or harmonic mean weighting	2	EBS	16.6			x	x	x		
11	GT4	Use harmonic mean weighting of composition data	2	both	16.6			x	x			
12	New1	Allow time-varying selectivity for the fishery but not the survey	2	EBS	16.6			x				
13	New2	Estimate survey index standard error internally ("extra SD" option in SS)	2	EBS	16.6				x			
14	New3	Use Francis weighting	2	EBS	16.6					x		
15	New4	Give less weight to fishery comps than survey comps, less to sizecomps than agecomps	2	EBS	16.6						x	
16	Sub1	Allow time variability only where supported by external data	1	both	16.6							
17	GT3	Switch to length-based, flat-topped, double normal selectivity	1	both	16.6							
18	New5	Use Hamel's (2015) method to develop a prior distribution for natural mortality	1	both	16.6							
19	BPT2	Address concerns regarding selectivity form and longline survey data	1	AI	16.6							
20	Sub4	Use either Francis or harmonic mean weighting	1	AI	16.6							
21	Sub5	Investigate alternatives to double-normal selectivity	1	AI	16.6							
22	Sub6	Investigate whether a simpler (than SS) model would be useful	1	AI	16.6							
23	SSC10	Include existing fishery age data and obtain more fishery age data	0	EBS	16.6							
24	GT7	Allow deterministic time variability in selectivity, based on environmental indices	0	both	16.6							
25	GT9	Switch to length-based maturity or adjust age-based maturity for ageing bias	0	both	16.6							
26	SSC4	Develop a spatially structured, combined BS/AI/GOA assessment model	0	both	16.6							
27	BPT6	Compare σR to the RMSE of estimated recruitment deviations	NMA	EBS	16.6							x
28	New6	Report Francis weights from the terminal run if harmonic mean is used and vice-versa	NMA	EBS	16.6							x
29	BPT3	Continue to compare empirical weight at age with the traditional approach	NMA	EBS	16.6							
30	GT10	Include EBS survey strata 82 and 90 (NW corner of EBS) in the data	NMA	EBS	16.6							
31	Sub2	Examine survey data from the northern Bering Sea	NMA	EBS	16.6							
32	SSC7	Investigate ageing bias further	NMA	both	16.6							

APPENDIX A: Background materials

Introduction

Beginning with the 2010 assessment cycle, lists of models to be included in the preliminary (September) stock assessments for Pacific cod in the Eastern Bering Sea (EBS) and Aleutian Islands (AI), and usually the Gulf of Alaska (GOA), have been developed by either the Joint BSAI and GOA Groundfish Plan Teams or some subset thereof. Each year, the lists have been developed during a WebEx meeting held sometime during the first half of the year. From 2010-2013, the full Joint Teams participated in these meetings. From 2014-2016, this responsibility was delegated to a Joint Team subcommittee (Subcommittee). Developing lists of models for all three Pacific cod assessments (EBS, AI, and GOA) within a single meeting often proved to be challenging. At the request of the GOA assessment author, the GOA assessment was removed from the agenda, beginning with the 2016 meeting. Given that the GOA assessment is no longer on the agenda, development of this year's lists of models for the EBS and AI assessments has been delegated to a subset of the BSAI Team (Team) membership rather than a subset of the Joint Teams membership. This year's Subcommittee meeting is scheduled to take place on June 30, from 2:00-5:00 p.m. PDT.

From 2010-2016, the lists of models produced by the Joint Teams or Subcommittee were reviewed by the Scientific and Statistical Committee (SSC) prior to development of the preliminary assessments. However, at its June 2016 meeting, the SSC determined that it would no longer review the lists of models.

In the past, lists of models have been developed primarily on the basis of the previous year's models and comments supplied during the past year by the Team (or Teams) and the Scientific and Statistical Committee (SSC), with little or no development of new models since the previous year's final assessments. However, at its December 2016 meeting, the SSC requested that development of at least some new models take place prior to this year's Subcommittee meeting, conditional on staff availability (see comment SSC9 below). In response, 10 new models have been developed for consideration at this year's Subcommittee, as detailed below.

In addition to the materials provided in this document, last year's EBS assessment (https://www.afsc.noaa.gov/REFM/Docs/2016/EBSpcod.pdf) and AI assessment (https://www.afsc.noaa.gov/REFM/Docs/2016/aipcod.pdf) provide important background information. The models (or model, in the case of the AI) considered in each final assessment are described in the main text, the models considered in the preliminary assessment are described in an appendix (Appendix 2.1 in the EBS assessment, beginning on page 455 of the 2016 SAFE report; Appendix 2A.1 in the AI assessment, beginning on page 582 of the 2016 SAFE report), and the complete history of all models that have been fully vetted in previous years are described in another appendix (Appendix 2.3 in the EBS assessment, beginning on page 520 of the 2016 SAFE report; Appendix 2A.3 in the AI assessment, beginning on page 633 of the 2016 SAFE report). All age-structured models since 1992 have been developed within the Stock Synthesis (SS) framework (Methot and Wetzel 2013, https://doi.org/10.1016/j.fishres.2012.10.012).

Comments from last year

Following a review of the EBS and AI Pacific cod stock assessments by the Center of Independent Experts (CIE) in February of 2016 and during the process of developing and reviewing the 2016 assessments, a large number of comments on the assessments and the assessment process were provided by the Subcommittee, Team, and SSC. Recommendations pertaining to the 2016 assessments were all addressed in those assessments. Comments related to post-2016 assessments are listed below, in chronological order (26 in total). Comments are further subdivided into those pertaining to both the EBS and AI assessments, those pertaining only to the EBS assessment, and those pertaining only to the AI assessment. In addition to listing the full text of each comment, a shorthand summary is provided for ease of reference.

Comments from the May 2016 Subcommittee meeting

During its May 2016 meeting, the Subcommittee listed several recommendations that it designated as having “medium” priority, defined as recommendations that the Subcommittee felt should be considered in either the 2017 or 2018 assessments.

EBS and AI assessments

Sub1 (originally from the 2016 review by CIE member Jean-Jacques Maguire, labeled as comment 2e.06 in the minutes of the May 2016 Subcommittee meeting): “Only those parameters where there is external information suggesting that changes are occurring should be allowed to vary, probably one at a time to avoid incorrect interpretation.” Summary: *Allow time variability only where supported by external data.*

EBS assessment only

Sub2 (originally from the December 2015 SSC minutes, labeled as comment SSC2 in the minutes of the May 2016 Subcommittee meeting): “The SSC was encouraged by the author’s explanation that dome-shaped selectivity may, in part, be explained by the possibility that some of older fish may be residing in the northern Bering Sea (NBS) at the time of the survey. This is supported by the size composition of the fish in the 2010 NBS trawl survey, which suggested that up to 40% of the fish in some larger size classes reside in this area, although the overall proportion in the NBS was small. The SSC encourages the author to further examine Pacific cod catches from trawl surveys conducted triennially by the National Marine Fisheries Service (NMFS) (1976-1991) and by the Alaska Department of Fish & Game (1996 to the present) to monitor the distribution and abundance of red king crab and demersal fish (see: Hamazaki, T., Fair, L., Watson, L., Brennan, E., 2005. Analyses of Bering Sea bottom-trawl surveys in Norton Sound: absence of regime shift effect on epifauna and demersal fish. ICES Journal of Marine Science 62, 1597-1602). While the 2010 bottom trawl survey in the NBS found relatively few Pacific cod (3% of total biomass), it is possible that the proportion of Pacific cod that are outside the standard survey area was higher in other years. A second possibility is that older Pacific cod migrate to nearshore areas to feed in the summer, making them unavailable to the survey.” Summary: *Examine survey data from the northern Bering Sea*.

Sub3 (developed by the Subcommittee during its May 2016 meeting, where it was labeled JTS5): “Use reasonably time-varying, double normal selectivity (Bering Sea only). CIE comments 2e.01 and 2e.09 suggested that some amount of time-variability in fishery selectivity is appropriate, CIE comment 2e.12 cautioned against allowing “too much” time-variability in selectivity, and CIE comment 2b.07 suggested use of the double normal selectivity function.” Summary: *Use reasonably time-varying, double normal selectivity.*

AI assessment only

Sub4 (originally from the 2016 review by CIE member Neil Klaer, labeled as comment 2a.07 in the minutes of the May 2016 Subcommittee meeting): “While there has been some recent narrowing down of agreed procedures among US west-coast stock assessors, it has also been recognized that it is not currently possible to recommend default procedures for composition and conditional age-at-length (CAAL) data. There is agreement that the Francis weighting approach is more appropriate in cases where the model is not correctly specified as it takes autocorrelation among composition data into account. It is also agreed that for a correctly specified model, the McAllister-Ianelli harmonic mean weighting method works well." Summary: *Use either Francis or harmonic mean weighting.*

Sub5 (originally from the 2016 review by CIE member Neil Klaer, labeled as comment 2b.03 in the minutes of the May 2016 Subcommittee meeting): "Further work on choice of a more appropriate selectivity function other than double-normal (or by changing the freedom of certain double-normal parameters) would probably improve the overall fit…." Summary: *Investigate alternatives to double-normal selectivity*.

Sub6 (originally from the 2016 review by CIE member Robin Cook, labeled as comment 2i.17 in the minutes of the May 2016 Subcommittee meeting): "While developing the Tier 3 model, consideration should also be given to enhancing the Tier 5 model to include a simple population model in order to obtain a little more information from the data as opposed to simply smoothing the time series." Summary: *Investigate whether a simpler (than SS) model would be useful*.

Comments from the September 2016 Team meeting

<u>EBS and AI assessments</u>

BPT1: "The Team recommends that the mid-year meetings cease unless exceptional circumstances necessitate such a meeting." Summary: *Discontinue holding annual subcommittee meetings*.

<u>AI assessment only</u>

BPT2: "It is recognized by the Plan Team that per SSC comments and the author's discretion, that the author may bring forward a better model than 16.1. The Plan Team has concerns regarding the form of the selectivity and the new data sources. We feel that these issues cannot be fully examined by November, but the Team recommends that they be addressed in the next cycle (2017)." Summary: *Address concerns regarding selectivity form and longline survey data*.

Comments from the October 2016 SSC meeting

<u>EBS and AI assessments</u>

SSC1: "The observed discrepancies among different models in these assessments are a good—if perhaps extreme—example of the model uncertainty that pervades most assessments. This uncertainty is largely ignored once a model is approved for specifications. We encourage the authors and Plan Teams to consider approaches such as multi-model inference to account for at least some of the structural uncertainty. We recommend that a working group be formed to address such approaches." Summary: *Consider multi-model inference and form a working group to do so*.

SSC2: "Regarding the mid-year model vetting process, the SSC re-iterates its recommendation from June to continue for now. The process has proven useful for the industry as an avenue to provide formal input and for the author to prioritize the range of model options to consider." Summary: *Continue to hold annual subcommittee meetings*.

SSC3: "With regard to data weighting, the SSC recommends that the authors consider computing effective sample sizes based on the number of hauls that were sampled for lengths and weights, rather than the number of individual fish." Summary: *Set input sample size based on number of hauls sampled for length*.

SSC4: “Although there is genetic evidence for stock structuring within the Pacific cod population among regions, the uncertainty in model scale for all three regions seems to suggest that some sharing of information among the three assessments might be helpful. Over the long term, authors could consider whether a joint assessment recognizing the population structuring, but simultaneously estimating key population parameters (e.g., natural mortality, catchability or others) might lend more stability and consistency of assumptions for this species.” Summary: *Develop a spatially structured, combined BS/AI/GOA assessment model*.

EBS assessment only

SSC5: “The SSC notes that, in spite of the concerns over dome-shaped survey selectivity in the survey, there are many potential mechanisms relating to the availability of larger fish to the survey gear that could result in these patterns, regardless of the efficiency of the trawl gear to capture large fish in its path. For example, in the Bering Sea the patterns could be due to larger Pacific cod being distributed in deeper waters or in the northern Bering Sea at the time of the survey. The northern Bering Sea survey planned for 2017 should provide additional information on the latter possibility.” Summary: *Examine survey data from the northern Bering Sea*. This comment is essentially redundant with comment Sub2.

Comments from the November 2016 Team meeting

EBS assessment only

BPT3: “The Team recommends comparing model predicted weight-at-age in Models 16.6 and 16.7 to the empirical weight-at-age used in Model 16.1.” Summary: *Continue to compare empirical weight at age with the traditional approach*.

BPT4: “The Team recommends weighting (tuning) composition data using the Francis method or the harmonic mean of the effective sample size (McAllister & Ianelli approach).” Summary: *Use either Francis or harmonic mean weighting.* This comment is essentially redundant with comment Sub4, except that comment Sub4 is specific to the AI assessment and this one is specific to the EBS assessment.

BPT5: “The Team believes that time-varying selectivity is important and recommends continued investigation of time-varying fishery selectivity for use in future models. In addition, the Team recommends investigating methods to determine the variance of the penalty function applied to the deviations (i.e., tuning the deviates).” Summary: *Investigate the appropriate amount of time-variability in fishery selectivity*.

BPT6: “The Team recommends comparing the estimated recruitment variability (σ_R) to the root mean squared error (RMSE) of the estimated recruitment deviations over a period of years that is well informed (i.e., when the variance of the estimated recruitment deviation is small).” Summary: *Compare σ_R to the RMSE of estimated recruitment deviations.*

AI assessment only

BPT7: “The Team recommends that the analyst propose age-structured models for consideration at the spring model specification meeting to be used in Tier 3 calculations.” Summary: *Propose age-structured models to be considered in June*.

Comments from the December 2016 SSC meeting

EBS and AI assessments

SSC6: "All three cod assessments could benefit from a formal prior on *M* based on the variety of studies referenced in each. The SSC recommends that a prior for use in all cod assessments be developed for 2017." Summary: *Develop a prior distribution for natural mortality based on previous estimates*.

SSC7: "The SSC supports the author's observation that ageing bias needs to be further investigated for cod, with results potentially applicable to all three assessments." Summary: *Investigate ageing bias further*.

SSC8: "The SSC continues to support the spring Pacific cod workshop to review and plan for model development each year, and also supports all of the technical PT recommendations for future model development." Summary: *Continue to hold annual subcommittee meetings*. This comment is essentially redundant with comment SSC2.

EBS assessment only

SSC9: "The SSC recommended discarding Model 11.5 for future analyses after one or more 16.x models incorporating time-varying selectivity in some reasonable manner (for the survey and/or fishery) are developed to take its place in this set of models. Depending on staff availability, this could be presented at the spring meeting; however, if that is not possible, it should be brought forward for the September 2017 PT meeting." Summary: *Discard Model 11.5 if other models with time-varying selectivity are developed*.

SSC10: "The SSC recommends that including existing fishery ages in the assessment and ageing additional fishery otoliths for this assessment should be priorities…." Summary: *Include existing fishery age data and obtain more fishery age data.*

SSC11: "The SSC recommends continued exploration of the treatment of weight-at-age using both internally and externally estimated values." Summary: *Continue to compare empirical weight at age with the traditional approach*. This comment is essentially redundant with comment BPT3.

SSC12: "The SSC [recommended] further considering model averaging based on the outcome of the SSC workshop during the February 2017 meeting" (term in square brackets added). Summary: *Consider model averaging if the February workshop suggests doing so*.

AI assessment only

SSC13: "The SSC also supports the PT recommendation to continue development of an age-structured model for the next assessment cycle in an effort to move the stock to Tier 3." Summary: *Continue development of an age-structured model*.

Features included in models developed in preparation for the June 2017 Subcommittee meeting

Ten new models were developed for the Eastern Bering Sea stock in preparation for the June 2017 Subcommittee meeting. They consisted of modifying Model 16.6 (last year's final model) by various combinations of the following 7 features (note that the last 4 of these are responsive to various comments listed above):

GT1: Adjust timing of the fishery and survey to conform to SS V3.30 conventions and to inform SS that the fishery sizecomp data are collected throughout the year rather than at a single point. Summary: *Adjust timing of the fishery and survey in SS.*

GT2: Switch from age-based logistic selectivity to age-based, flat-topped, double normal selectivity for both the fishery and survey. Summary: *Switch to age-based, flat-topped, double normal selectivity*.

GT3: Switch from age-based logistic selectivity to length-based, flat-topped, double normal selectivity for both the fishery and survey. Summary: *Switch to length-based, flat-topped, double normal selectivity*.

GT4: Iteratively reweight all composition data so that the harmonic mean of the effective sample size estimated by the model equals the arithmetic mean sample size specified in the data file. (Note that this feature is responsive to comment BPT4.) Summary: *Use harmonic mean weighting of composition data*.

GT5: Adjust the composition data so that the sample sizes specified in the data file correspond to the number of sampled hauls rather than being proportional to the number of sampled lengths or ages (where the proportionality coefficient for lengths depends on the year of sampling) and so that the 1991-2016 fishery size composition data from each year/week/gear/area cell are weighted proportionally to the official estimate of catch taken in that cell. (Note that this feature is responsive to comment SSC3.) Summary: *Switch to haul-based input sample size and catch-weighted sizecomp data*.

GT6: Allow random time variability in all selectivity parameters, with input standard deviations derived by Algorithm 5 in Thompson (2016), which is intended to approximate the restricted maximum likelihood estimate. (Note that this feature is responsive to comments Sub3 and BPT5.) Summary: *Allow random time variability in selectivity, with σs fixed at the restricted MLEs*.

GT7: Allow time variability in some selectivity patterns, where the variation is a parameter-specific deterministic function of a parameter-specific environmental index. (Note that this feature is responsive to comment Sub1.) Summary: *Allow deterministic time variability in selectivity, based on environmental indices*.

Models developed in preparation for the June 2017 Subcommittee meeting

Mapping features GT1-GT7 into the 10 new models

In developing models for consideration at the June 2017 Subcommittee meeting, the first task was to convert Model 16.6 (last year's final EBS model) from SS V3.24u to V3.30.04.02. All of the new models were developed using the V3.30.04.02 version of Model 16.6 as the starting point model.

Features GT1-GT7 were used in various combinations to develop 10 new models for the Eastern Bering Sea stock as follows, where the first column ("M") denotes model number (note that model numbering here does not follow the convention used in the SAFE chapter, use of which will be reserved for models that are actually vetted in either the preliminary (September) or final (December) draft):

M	Description	GT1	GT2	GT3	GT4	GT5	GT6	GT7
0	Model 16.6, converted from SS V3.24u to V3.30.04.02							
1	Like Model 0, but including feature GT1	x						
2	Like Model 1, but including feature GT2	x	x					
3	Like Model 1, but including feature GT3	x		x				
4	Like Model 2, but including feature GT4	x	x		x			
5	Like Model 1, but including feature GT5	x				x		
6	Like Model 5, but including feature GT4	x			x	x		
7	Like Model 6, but including feature GT2 (GT4 updated)	x	x		x	x		
8	Like Model 7, but including feature GT6 (GT4 updated)	x	x		x	x	x	
9	Like Model 7, but including feature GT7	x	x		x	x		x
10	Like Model 9, but with feature GT4 updated	x	x		x	x		x

In the above table, the shading in the GT4 column for Model 9 indicates that the weights applied to the various sets of composition data in that model were unchanged from those calculated using feature GT4 (harmonic mean weighting) in Model 7.

Implementation notes

"Jitter" tests were not applied to any of the new models. A reasonably small gradient and the existence of a positive definite Hessian matrix were taken as sufficient evidence of convergence for this very preliminary analysis.

<u>Model 0</u>

The objective function values from the V3.24u and V3.30.04.02 versions were identical (to the six reported digits), and the estimates of 2016 spawning biomass differed by only 1 t.

<u>Model 1</u>

Model 0 specified the timing of the fishery and the survey as occurring at mid-year, indicated by a value of 0.5 for each in the data file. This generated a pair of warnings from SS: 1) fishery timing should be set to -1, indicating that catch occurs throughout the year; and 2) timing values of 0.5 are ignored in V3.30, which requires timing to bet set in units of months (e.g., July = 7). The data file for Model 1 was adjusted accordingly.

<u>Model 2</u>

Age-based, flat-topped, double normal selectivity was implemented by fixing parameters #2, #4, and #6 of the double normal selectivity function all at a value of 10.0; fixing parameter #5 at a value of -10.0; and estimating only parameter #1 (the age at which selectivity reaches a value of 1.0) and parameter #3 (the ascending "width") internally.

<u>Model 3</u>

Length-based, flat-topped, double normal selectivity was implemented in the same manner as described above for the age-based case, with the additional consideration that the upper bound on parameter #1 had to be adjusted in light of the fact that the length data span a larger range than the age data.

Model 3 was unable to estimate the Richards growth coefficient reliably. Although an estimate of 1.01264 was obtained in one model run, the gradient was extremely large. When the Richards growth coefficient was fixed at a value of 1.0 in a subsequent run, a very small gradient resulted, so the final version of Model 3 fixed the parameter at this value.

Model 4

Four iterations resulted in the ratios of harmonic mean effective sample size to arithmetic mean input sample size being equal to unity (to two digits beyond the decimal point) for all three compositional datasets (fishery sizecomp, survey sizecomp, and survey agecomp).

Model 5

For the years 1991-2016, the numbers of hauls sampled for fishery lengths were taken from the domestic observer database. For years prior to 1990, the numbers of sampled hauls in the fishery sizecomp data were approximated by using the regression shown in Figure 2.1.13 of the 2015 EBS assessment to convert last year's Model 11.5 input fishery sample sizes into haul equivalents. For agecomp data, the number of hauls sampled for length (not age) was used, per comment SSC3. Table A1 compares input sample sizes used in Model 0 with those used in Model 5.

Figure A1 compares the 1991-2016 fishery size composition data used in Model 0 with those used in Model 5. In general, there is little difference between the two sets of sizecomp data. The effective sample sizes (treating the catch-weighted data as "true") range from 1,732 to 37,958, with a mean of 12,357.

Model 6

Four iterations resulted in the ratios of harmonic mean effective sample size to arithmetic mean input sample size being equal to unity (to two digits beyond the decimal point) for all three compositional datasets (fishery sizecomp, survey sizecomp, and survey agecomp).

Model 7

Two iterations resulted in the ratios of harmonic mean effective sample size to arithmetic mean input sample size being equal to unity (to two digits beyond the decimal point) for all three compositional datasets (fishery sizecomp, survey sizecomp, and survey agecomp).

Model 8

Model 8 involves use of Algorithm 5 in Thompson (2016) to estimate the input standard deviations for the five parameters with random effects (recruitment, fishery selectivity peak age, fishery selectivity width, survey selectivity peak age, and survey selectivity width) in an iterative fashion. For a linear-normal model, Algorithm 5 yields the restricted maximum likelihood estimate of the variance-covariance matrix of the random effects (only the square roots of the terms on the diagonal are used by SS).

Algorithm 5 was introduced in the preliminary 2015 EBS assessment (note that this is a multivariate generalization of one of the methods mentioned by Methot and Taylor (2011, https://doi.org/10.1139/f2011-092), viz., the third method listed on p. 1749), and proceeds as follows:

1. Set initial guesses for the σs.
2. Run SS.

3. Compute the covariance matrix (**V1**) of the set of *dev* vectors (e.g., element {*i*,*j*} is equal to the covariance between the subsets of the *i*th *dev* vector and the *j*th *dev* vector consisting of years that those two vectors have in common).
4. Compute the covariance matrix of the parameters (the negative inverse of the Hessian matrix).
5. Extract the part of the covariance matrix of the parameters corresponding to the *dev* vectors, using only those years common to all *dev* vectors.
6. Average the values in the matrix obtained in step 5 across years to obtain an "average" covariance matrix (**V2**).
7. Compute the vector of σs corresponding to **V1**+**V2**.
8. Return to step 2 and repeat until the σs converge.

To speed the above algorithm, the σs obtained in step 7 were sometimes substituted with values obtained by extrapolation/interpolation based on previous runs.

Reweighting of the three compositional datasets (fishery sizecomp, survey sizecomp, and survey agecomp) and estimation of the input standard deviations for the five parameters with random effects was undertaken simultaneously. Nineteen iterations resulted in all three compositional weights and all five input standard deviations converging to three digits beyond the decimal point.

The input standard deviations of the random effects for log recruitment (additive) and the four selectivity parameters (multiplicative for "peak," additive for "width") estimated by Algorithm 5 were as follow:

Recruits	Fishery peak	Fishery width	Survey peak	Survey width
0.479	0.391	0.692	0.582	1.067

The corresponding correlation matrix estimated by Algorithm 5 (but not used in SS) was:

	Recruits	Fishery peak	Fishery width	Survey peak	Survey width
Recruits	1.000	0.128	0.035	0.250	0.168
Fishery peak	0.128	1.000	0.769	0.274	0.146
Fishery width	0.035	0.769	1.000	0.213	0.005
Survey peak	0.250	0.274	0.213	1.000	0.595
Survey width	0.168	0.146	0.005	0.595	1.000

Model 9

Correlations between the time series for each of the two fishery selectivity parameters (peak and width) estimated by Model 8 were explored with respect to 12 individual covariates related to fishery characteristics, on both linear and exponential scales. The 12 individual covariates were as follow:

- Average latitude of the observed catch (measured in both numbers and weight of fish)
- Average longitude of the observed catch (measured in both numbers and weight of fish)
- Average depth of the observed catch (measured in both numbers and weight of fish)
- Timing of official catch (measured as both mean and median week of catch in weight)
- Proportion of official catch in weight taken by a gear type (trawl, pot, jig, and longline)

Correlations between the time series for each of the two fishery selectivity parameters (peak and width) estimated by Model 8 were explored with respect to 34 individual covariates related to the physical environment, on both linear and exponential scales. The 34 individual covariates were those listed under the "climate index," "atmosphere," and "ocean" categories on NOAA's "Bering Climate" website (http://www.beringclimate.noaa.gov/data/index.php; note that some of the linked datasets contain multiple time series, e.g., a time season for each season).

For each selectivity parameter, the correlations with all individual covariates in the respective fleet-specific set (fishery or survey), expressed as anomalies and also as exponentiated anomalies, were examined. If correlations greater (in absolute value) than 0.3 were found, all possible pairings of the relevant individual covariates were also examined (omitting pairs that were almost perfectly correlated; e.g., the proportions of the catch taken by trawl and longline gear have a correlation of –0.93), by adding the two anomaly time series together (adjusted by the sign of the correlation and rescaling so as to have a variance of unity). Correlations with the exponentiated sums of those anomalies were also examined. To be considered for inclusion in Model 9, a correlation had to be greater (in absolute value) that 0.5.

For the fishery selectivity peak parameter, only one covariate (proportion of catch taken by pot gear) had abs(corr)>0.3 on the linear scale, none had abs(corr)>0.3 on the exponential scale, and none had abs(corr)>0.5 on either scale, so **no covariate for the fishery selectivity peak parameter was included in Model 9**.

For the fishery selectivity width parameter, six covariates had abs(corr)>0.3 on the linear scale, and five had abs(corr)>0.3 on the exponential scale. On either scale, some of the covariates were redundant (e.g. the same quantity measured in weight or numbers of fish). In such cases, only one of the covariates was carried forward for further analysis. On the linear scale, the candidate individual covariates were: average latitude of the catch in numbers, average depth of the catch in numbers, median week of catch in weight, and the proportion of the catch in weight taken by trawl gear. On the exponential scale, the candidate individual covariates were: average depth of the catch in numbers, and the proportion of the catch in weight taken by trawl gear. The highest abs(corr) for any pairwise combination was median week of the catch in weight plus the proportion of the catch in weight taken by trawl gear, on the linear scale, which gave a correlation of 0.590. Therefore, **the sum of the median week of the catch and the proportion of the catch taken by trawl gear was used as a linear covariate for the fishery selectivity width parameter in Model 9**.

For the survey selectivity peak parameter, only two covariates (Arctic dipole (fall) and ENSO) had abs(corr)>0.3 on the linear scale, but neither of these had abs(corr)>0.5. The rescaled sign-adjusted sum (+ and –, respectively) gave abs(corr)=0.475. On the exponential scale, only two covariates (Arctic dipole (fall) and AO) had abs(corr)>0.3 on the linear scale, but neither of these had abs(corr)>0.5. The rescaled sign-adjusted sum (+ and +, respectively) gave abs(corr)=0.457. Because no individual covariate or pair of covariates gave abs(corr)>0.5, **no covariate was included for the survey selectivity peak parameter in Model 9**.

For the survey selectivity width parameter, many covariates had abs(corr)>0.3 on both the linear and exponential scales. On the linear scale, no individual covariate had abs(corr)>0.5, but on the exponential scale, Arctic dipole (winter) had abs(corr)=0.549. On the linear scale, 10 pairwise combinations gave abs(corr)>0.5, but none gave abs(corr)>0.6. On the exponential scale, 20 pairwise combinations gave abs(corr)>0.5. To narrow the field, the following 4 pairwise combinations gave abs(corr)>0.6 on the exponential scale:

1. Arctic dipole (fall) + Arctic dipole (winter) = 0.627
2. Arctic dipole (winter) + AO = 0.618

3. Arctic dipole (winter) – ENSO = 0.611
4. NPI + Siberian-Alaskan Index (Siberia) = 0.614

Pair #1 in the above seems redundant. Pair #4 is suspicious, because it seems unlikely that the "Siberia" part of the Siberian-Alaskan Index would have more impact in the EBS than either the "Alaskan" part or the combined "Siberian" and "Alaskan" parts. Pair #3 may be harder for readers of the assessment to understand because of the minus sign, and the two components are somewhat correlated (corr = –0.20), whereas the two components of pair #2 are only barely correlated (corr = 0.06). Therefore, partly by process of elimination, **the sum of the Arctic dipole (winter) and the Arctic oscillation was included as an exponential covariate for the survey selectivity width parameter in Model 9**.

Model 10

A single iteration resulted in the ratios of harmonic mean effective sample size to arithmetic mean input sample size being equal to unity (to two digits beyond the decimal point) for all three compositional datasets (fishery sizecomp, survey sizecomp, and survey agecomp).

Results

The upper part of Table A2 shows the objective function values for all 11 models, broken down by component and overall.

The lower part of Table A2 provides some context for interpreting the objective function values. Note that the objective function values for only Models 1-3 are truly comparable, because those are the only models that share both a common data file and common set of weights for the various compositional datasets (data files and sets of compositional data weights are numbered according to the model in which each was first used). Among Models 1-3, Model 2 has the lowest objective function value, beating Model 1 by 20.68 points and Model 3 by 165.93 points (note that Model 3 also has one fewer internally estimated parameter), suggesting that age-based, flat-topped, double normal selectivity outperforms both age-based logistic selectivity and length-based, flat-topped double normal selectivity (at least in the context of the features shared by those three models).

The upper part of Table A3 provides the following diagnostics for all 11 models:

- The ratio of the mean input survey standard error to the root-mean-squared error (RMSE) of the model's fit to the survey index (range 0.45-0.65)
- The ratio of the harmonic mean effective sample size to the arithmetic mean input sample size for each of the three compositional datasets
  - range 0.14-2.01 for the fishery sizecomp data
  - range 0.71-1.02 for the survey sizecomp data
  - range 0.11-1.06 for the survey agecomp data
- The ratio of σ_R to the RMSE of the estimated recruitment deviations—provided in response to comment BPT6 (range 0.90-1.22)
- Mohn's ρ (range 0.08-0.20)

The middle part of Table A3 provides some key outputs from all 11 models:

- 2016 female spawning biomass, in t (range 165,999-433,689)
- 2016 relative (to $B_{100\%}$) female spawning biomass (range 0.24-0.71)
- The relative level of spawning per recruit to which the 2016 fishing mortality rate corresponds (range 0.18-0.52)

The lower part of Table A3 repeats the "context" part of Table A2 for ease of reference.

Table A4 gives the point estimates of all non-time-varying parameters estimated internally by any of the 11 models (base values are shown for time-varying selectivity parameters in Models 8-10). Note that the Richards growth coefficient in Model 3 and σ_R in Model 8 are also shown for purposes of comparison, even though they were not estimated internally (the Richards growth coefficient in Model 3 was fixed at a value close to the value that was estimated in a version of the model that may not have converged, and σ_R in Model 8 was estimated by Algorithm 5 of Thompson (2016)). Blank cells in Table A4 indicate that the respective parameter (row) was not used in the respective model (column).

Figure A2 shows the selectivity schedules estimated by all 11 models. Figure A2a shows the time-invariant schedules for Models 0-7 and the base selectivity schedules for Models 8-10 (in Model 3, selectivity is actually estimated based on length rather than age; the age-equivalent relationship is shown in Figure A2a). Figure A2b shows the median and upper and lower 80% confidence intervals for the time-varying selectivity schedules for Models 8-10. The confidence intervals were determined empirically (i.e., by sorting the time series at each age) rather than theoretically. The age axes in Figure A2 extend only to age 7, because all selectivity schedules for Models 0-7 and all base selectivity curves for Models 8-10 reach a value of at least 95% by age 8.

Author's responses to selected comments not addressed in the new models

Comment BPT7 recommends that the author propose age-structured AI models for consideration at this year's Subcommittee meeting. AI versions of Models 0-10 are hereby proposed for consideration.

Comments SSC1 and SSC12 both encourage further exploration of model averaging, but the latter comment makes such exploration conditional on the outcome of the SSC's February 2017 workshop on that subject. The minutes of that meeting conclude, "The SSC would like to see a 'test case' of how ensemble modeling works for one of our groundfish stocks." Although not reflected in the minutes, at least one SSC member suggested that the test case be conducted during an "off" year for an assessment that is not produced annually.

Comment SSC6 requests that a prior distribution for the natural mortality rate (*M*) be developed on the basis of the previous studies referenced with respect to estimation of *M* in the Pacific cod assessments for the EBS, AI, and Gulf of Alaska (GOA). The list of previous studies in the 2016 GOA assessment (https://www.afsc.noaa.gov/REFM/Docs/2016/GOApcod.pdf) is the longest of the three, providing 15 point estimates of *M* from the EBS, GOA, British Columbia, Korea, and Japan. The lists in the 2016 EBS and AI assessments are subsets of the list in the GOA assessment. If the estimates of *M* obtained in the 2016 EBS and GOA assessments (0.36 and 0.47) are added to the list in the GOA assessment, a total of 17 estimates are available. If a normal distribution is assumed, the sample mean and standard deviation are 0.481 and 0.204, respectively (coefficient of variation = 0.424, 95% CI spans 0.082-0.881). If a lognormal distribution is assumed, the log-scale sample mean and standard deviation are –0.811 and 0.410, respectively (coefficient of variation = 0.435, 95% CI spans 0.199-0.993). Figure A3 shows the cumulative distribution function and probability density function for both the normal and lognormal cases, along with the point estimates from the 2016 EBS and GOA assessments. The mode of the normal distribution comes very close to matching the estimate from the 2016 GOA assessment, while the mode of the lognormal distribution comes very close to matching the estimate from the 2016 EBS assessment.

Comment SSC7 requests further investigation of ageing bias. An ongoing NPRB study has been examining the possibility that the stratified sampling design used for otolith collection in the past may have contributed to the observed mismatch between survey sizecomp modes and the mean lengths at age based on the age data. Both random and stratified samples were collected in 2015 and 2016. To be consistent with data collected previously, only the stratified samples from 2015 were used in the 2016 assessment (the 2016 data had not been aged yet). Based on analysis of the 2015 and 2016 samples, the stratified sampling design was abandoned this year, and from now on, only the random sampling design will be used. Either this year or in the near future, the assessment will need to account for this change.

Comment SSC10 requests that existing fishery ages be used in the EBS assessment and that obtaining additional fishery age data for this assessment should be priorities. In response, four years' worth of fishery otolith subsamples (1000 otoliths each) were requested to be aged in time for use in this year's assessments (subsamples for 2015 and 2016 were requested to be aged by July 1 and subsamples for 2013 and 2014 were requested to be aged by October 1). The subsamples for each year were chosen randomly, but in proportion to the catch taken in each week/gear/area cell. Inclusion of the few *existing* fishery age data, however, is problematic, because they were all taken during a single season by a single gear, and are therefore unlikely to be representative of the overall fishery.

Author's suggestions for other features to consider in this year's models

In the EBS models used in the 2005 and 2006 assessments, maturity was expressed as a function of length rather than age, because the version of SS available during those years did not include an option for expressing maturity as a function of age. Once age-based maturity became an option in 2007, the EBS models switched to this option on the recommendation of the author of the maturity study from which both sets of parameters (length-based and age-based) were taken (Stark 2007, http://fishbull.noaa.gov/1053/stark.pdf, and James Stark, Alaska Fisheries Science Center, *pers. commun*.), and all subsequent EBS and AI assessment models have retained this assumption. At the time when the switch to age-based maturity was made, no evidence of ageing bias had been discovered. However, all EBS assessments since 2009 have estimated an ageing bias. It may therefore be worth exploring the possibility of switching back to length-based maturity, or attempting to incorporate ageing bias into the maturity function outside of the assessment model (which would likely have to involve an iterative process).

It also might be worth exploring the possibility of including the data from the "extra" EBS shelf survey stations (strata 82 and 90, in the northwest corner of the EBS), which were added to the "standard" survey area beginning in 1987 "to investigate the distribution and abundance of snow crabs and the northern distribution of walleye pollock" (Conner et al. 2017, https://www.afsc.noaa.gov/Publications/AFSC-TM/NOAA-TM-AFSC-353.pdf). The Pacific cod biomass in strata 80 and 92 has averaged about 3% (range: 0.5% - 8.6%) of the biomass in the "standard" survey area since 1987. Including strata 80 and 92 to the 1987+ time series would probably require estimating two catchability coefficients rather than one, and perhaps would require estimating additional survey selectivity parameters as well.

Suggested procedure for the June 2017 Subcommittee meeting

In previous meetings convened to develop lists of models to be included in the preliminary Pacific cod assessments, the procedure has included reviewing proposals (comments, recommendations, features) that had been submitted prior to the meeting and developing new proposals during the meeting. Subsets of the proposals were then allocated among *up to* six models (including the previous year's final model) and some number of non-model analyses (i.e., analyses conducted outside of an assessment model). A template similar to Table A5 has helped to guide this discussion in previous years.

The following proposals (arranged by topic and area) were considered to be procedural only, and therefore not included in Table A5:

- BPT1: Discontinue holding annual subcommittee meetings

- SSC2: Continue to hold annual subcommittee meetings
- SSC8: Continue to hold annual subcommittee meetings
- SSC1: Consider multi-model inference and form a working group to do so
- SSC12: Consider model averaging if the February workshop suggests doing so
- SSC9: Discard Model 11.5 if other models with time-varying selectivity are developed (BS only)
- BPT7: Propose age-structured models to be considered in June (AI only)
- SSC13: Continue development of an age-structured model (AI only)

In Table A5, comments with identical summaries are referenced by the number assigned to the earlier comment. The following comments are essentially redundant:

- SSC11 is redundant with BPT3 (earlier)
- SSC8 is redundant with SSC2 (earlier)
- SSC5 is redundant with Sub2 (earlier)

Comment BPT4 is redundant in principle with Sub4 (earlier), but comment Sub4 is specific to the AI assessment and comment BPT4 is specific to the EBS assessment, so both are listed in Table A5.

Reference

Thompson, G. G. 2016. Specifying the standard deviations of randomly time-varying parameters in stock assessment models based on penalized likelihood: a review of some theory and methods. In revision for *Fishery Bulletin*. 49 p. Available from the author upon request (grant.thompson@noaa.gov).

Tables and Figures

Table A1. Comparison of input sample sizes used in Model 0 (“old”) and Models 5-10 (“new”).

	Fishery sizecomp		Survey sizecomp		Survey agecomp	
Year	N(old)	N(new)	N(old)	N(new)	N(old)	N(new)
1977	2	30	n/a	n/a	n/a	n/a
1978	12	160	n/a	n/a	n/a	n/a
1979	17	235	n/a	n/a	n/a	n/a
1980	15	208	n/a	n/a	n/a	n/a
1981	11	148	n/a	n/a	n/a	n/a
1982	13	187	250	313	n/a	n/a
1983	56	782	312	255	n/a	n/a
1984	138	1913	288	264	n/a	n/a
1985	204	2825	400	345	n/a	n/a
1986	178	2496	365	349	n/a	n/a
1987	339	4726	251	339	n/a	n/a
1988	105	1458	237	339	n/a	n/a
1989	70	966	237	293	n/a	n/a
1990	260	3601	134	329	n/a	n/a
1991	357	5188	171	313	n/a	n/a
1992	369	5322	228	332	n/a	n/a
1993	232	2993	247	363	n/a	n/a
1994	372	4687	330	364	204	364
1995	368	5215	218	347	163	347
1996	463	6618	222	359	203	359
1997	502	7278	218	369	205	369
1998	446	6838	227	362	181	362
1999	404	9231	277	336	246	336
2000	425	9731	298	355	246	355
2001	448	10364	469	366	263	366
2002	491	11472	290	364	248	364
2003	612	14341	293	363	361	363
2004	497	12242	257	361	284	361
2005	487	11568	268	360	365	360
2006	384	8849	288	354	371	354
2007	299	6901	304	368	412	368
2008	355	8320	308	338	346	338
2009	315	7482	396	360	403	360
2010	277	6514	179	342	369	342
2011	363	8804	492	368	358	368
2012	400	9287	310	356	372	356
2013	503	11126	443	354	405	354
2014	497	12165	426	373	349	373
2015	456	11309	458	354	244	354
2016	257	9553	407	376	n/a	n/a

Table A2. Objective function values and context for the 11 models.

Objective function	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Catch	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Equilibrium catch	0.00	0.00	0.00	0.00	0.01	0.05	0.01	0.00	0.00	0.00	0.00
Survey index	-25.21	-25.21	-21.89	-27.30	-25.13	17.08	-27.81	-26.40	-42.62	-13.28	-12.96
Sizecomp (fishery)	364.60	364.16	360.00	292.15	726.03	4447.18	605.29	586.24	576.40	546.74	559.94
Sizecomp (survey)	1008.34	1008.12	1001.80	1137.28	1016.28	1380.13	982.04	987.28	964.30	965.20	980.35
Agecomp (survey)	241.40	241.10	232.75	327.60	47.75	430.33	47.13	47.81	56.15	45.31	47.98
Recruitment	4.25	4.26	-0.91	7.97	2.18	10.15	3.54	-1.25	-13.52	-1.83	-1.62
Softpounds	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00
Deviations	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	-41.82	n/a	n/a
Total	1593.39	1592.45	1571.77	1737.70	1767.13	6284.93	1610.19	1593.70	1498.90	1542.15	1573.69

Context	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Data file	0	1	1	1	1	5	5	5	5	9	9
Set of compositional data weights	0	0	0	0	4	0	6	7	8	7	10
Relative wt. per record: fish. sizecomp	0.333	0.333	0.333	0.333	0.700	0.896	0.671	0.673	0.780	0.673	0.674
Relative wt. per record: surv. sizecomp	0.333	0.333	0.333	0.333	0.259	0.051	0.284	0.281	0.172	0.281	0.279
Relative wt. per record: surv. agecomp	0.333	0.333	0.333	0.333	0.041	0.053	0.045	0.046	0.048	0.046	0.047
Relative wt. overall: fish. sizecomp	0.412	0.412	0.412	0.412	0.738	0.924	0.711	0.713	0.815	0.713	0.714
Relative wt. overall: surv. sizecomp	0.361	0.361	0.361	0.361	0.238	0.046	0.263	0.261	0.157	0.261	0.258
Relative wt. overall: surv. agecomp	0.227	0.227	0.227	0.227	0.024	0.030	0.026	0.027	0.027	0.027	0.028
Parameter count	77	77	77	76	77	77	77	77	226	79	79

Table A3. Diagnostics, key outputs, and context for the 11 models.

Diagnostics	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Mean input survey SE / RMSE	0.58	0.58	0.56	0.58	0.57	0.45	0.58	0.57	0.65	0.53	0.53
H(eff. N)/A(input N): fish. sizecomp	1.90	1.89	1.95	2.01	1.00	0.14	1.00	1.00	1.00	1.02	1.00
H(eff. N)/A(input N): surv. sizecomp	1.01	1.01	1.02	0.91	1.00	0.71	1.00	1.00	1.00	1.02	1.00
H(eff. N)/A(input N): surv. agecomp	0.20	0.20	0.21	0.19	1.00	0.11	1.00	1.00	1.00	1.06	1.00
Ratio of σR to RMSE of rec. devs.	0.96	0.96	0.90	0.96	1.05	1.22	1.01	0.98	1.01	1.00	1.00
Mohn's ρ	0.15	0.15	0.08	0.18	0.11	0.11	0.12	0.16	0.16	0.20	0.20

Key outputs	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
2016 female spawning biomass (t)	337,454	332,825	433,689	267,138	301,312	165,999	314,601	352,823	294,975	335,954	335,082
2016 relative female spawning biomass	0.51	0.51	0.71	0.43	0.51	0.24	0.50	0.59	0.52	0.55	0.55
2016 FSPR%	0.39	0.39	0.52	0.33	0.40	0.18	0.38	0.45	0.43	0.42	0.42

Context	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Data file	0	1	1	1	1	5	5	5	5	9	9
Set of compositional data weights	0	0	0	0	4	0	6	7	8	7	10
Relative wt. per record: fish. sizecomp	0.333	0.333	0.333	0.333	0.700	0.896	0.671	0.673	0.780	0.673	0.674
Relative wt. per record: surv. sizecomp	0.333	0.333	0.333	0.333	0.259	0.051	0.284	0.281	0.172	0.281	0.279
Relative wt. per record: surv. agecomp	0.333	0.333	0.333	0.333	0.041	0.053	0.045	0.046	0.048	0.046	0.047
Relative wt. overall: fish. sizecomp	0.412	0.412	0.412	0.412	0.738	0.924	0.711	0.713	0.815	0.713	0.714
Relative wt. overall: surv. sizecomp	0.361	0.361	0.361	0.361	0.238	0.046	0.263	0.261	0.157	0.261	0.258
Relative wt. overall: surv. agecomp	0.227	0.227	0.227	0.227	0.024	0.030	0.026	0.027	0.027	0.027	0.028
Parameter count	77	77	77	76	77	77	77	77	226	79	79

Table A4. Estimates of all non-time-varying parameters used in any of the 11 models (based values for selectivity parameters in Models 8-10).

Parameter	M. 0	M. 1	M. 2	M. 3	M. 4	M. 5	M. 6	M. 7	M. 8	M. 9	M. 10
Natural mortality rate	0.363	0.361	0.440	0.324	0.412	0.320	0.380	0.426	0.479	0.414	0.413
Length at age 1 (cm)	16.401	16.401	16.415	14.161	16.425	16.510	16.312	16.348	16.390	16.356	16.357
Asymptotic length (cm)	99.387	99.406	100.247	93.137	102.475	105.753	100.946	103.753	108.910	104.110	104.126
Brody growth coefficient	0.197	0.197	0.200	0.236	0.201	0.172	0.192	0.183	0.177	0.184	0.184
Richards growth coefficient	1.050	1.050	1.013	1.000	1.000	1.151	1.073	1.086	1.049	1.075	1.076
SD of length at age 1 (cm)	3.425	3.425	3.432	3.412	3.457	3.461	3.385	3.419	3.493	3.433	3.433
SD of length at age 20 (cm)	9.717	9.716	9.405	10.389	9.310	9.662	9.999	9.422	8.445	9.206	9.204
Ageing bias at age 1 (years)	0.321	0.321	0.293	0.314	0.282	0.304	0.324	0.307	0.295	0.306	0.306
Ageing bias at age 20 (years)	0.351	0.355	0.500	0.429	0.866	0.464	0.555	0.647	1.198	0.690	0.687
ln(mean post-1976 recruitment)	13.220	13.207	13.771	12.864	13.452	12.825	13.306	13.615	13.901	13.521	13.518
Recruitment sigma (σR)	0.638	0.638	0.591	0.669	0.605	0.648	0.624	0.582	0.479	0.576	0.577
Pre-1977 recruitment offset	0.333	0.334	0.376	0.347	0.222	0.213	0.273	0.296	0.303	0.296	0.293
Initial fishing mortallity rate	0.155	0.155	0.128	0.158	0.398	1.101	0.237	0.202	0.179	0.200	0.204
Survey catchability	0.876	0.883	0.708	1.127	0.907	1.190	0.880	0.791	0.795	0.819	0.820
Fishery selectivity A50%	4.324	4.318				4.297	4.372				
Fishery selectivity A95%-A50%	1.158	1.159				1.282	1.245				
Survey selectivity A50%	1.006	1.005				0.998	1.024				
Survey selectivity A95%-A50%	0.289	0.289				0.287	0.247				
Fishery selectivity peak age/size			5.558	71.820	5.445			5.732	5.626	5.679	5.679
Fishery selectivity width			0.699	5.738	0.665			0.881	0.870	0.870	0.870
Survey selectivity peak age/size			3.025	19.993	2.660			2.646	2.736	2.535	2.535
Survey selectivity width			1.885	3.636	1.356			1.240	1.099	1.078	1.081
Fishery selectivity environ. link										0.085	0.085
Survey selectivity environ. link										0.281	0.284

Table A5. Allocation of proposals among models to be presented in the preliminary (September) assessments for both areas.

Note: Model numbers shown here are temporary placeholders; actual numbers for September will be established during the analysis.
Abbreviations: BPT = BSAI Plan Team, IAA = proposer's intended area of application, NMA = non-model analysis,
No. = proposal number, SPM = starting point model, SSC = Scientific and Statistical Committee, Sub = Subcommittee

			Bering Sea								Aleutian Islands							
No.	Proposal summary	IAA	SPM	1	2	3	4	5	6	NMA	SPM	1	2	3	4	5	6	NMA
Sub1	Allow time variability only where supported by external data	both																
SSC3	Set input sample size based on number of hauls sampled for length	both																
SSC4	Develop a spatially structured, combined BS/AI/GOA assessment model	both																
SSC6	Develop a prior distribution for natural mortality based on previous estimates	both																
SSC7	Investigate ageing bias further	both																
GT1	Adjust timing of the fishery and survey in SS	both																
GT2	Switch to age-based, flat-topped, double normal selectivity	both																
GT3	Switch to length-based, flat-topped, double normal selectivity	both																
GT4	Use harmonic mean weighting of composition data	both																
GT5	Switch to haul-based input sample size and catch-weighted sizecomp data	both																
GT6	Allow random time variability in selectivity, with σs fixed at the restricted MLEs	both																
GT7	Allow deterministic time variability in selectivity, based on environmental indices	both																
Sub2	Examine survey data from the northern Bering Sea	EBS																
Sub3	Use reasonably time-varying, double normal selectivity	EBS																
BPT3	Continue to compare empirical weight at age with the traditional approach	EBS																
BPT4	Use either Francis or harmonic mean weighting	EBS																
BPT5	Investigate the appropriate amount of time-variability in fishery selectivity	EBS																
BPT6	Compare σR to the RMSE of estimated recruitment deviations	EBS																
SSC10	Include existing fishery age data and obtain more fishery age data	EBS																
Sub4	Use either Francis or harmonic mean weighting	AI																
Sub5	Investigate alternatives to double-normal selectivity	AI																
Sub6	Investigate whether a simpler (than SS) model would be useful	AI																
BPT2	Address concerns regarding selectivity form and longline survey data	AI																

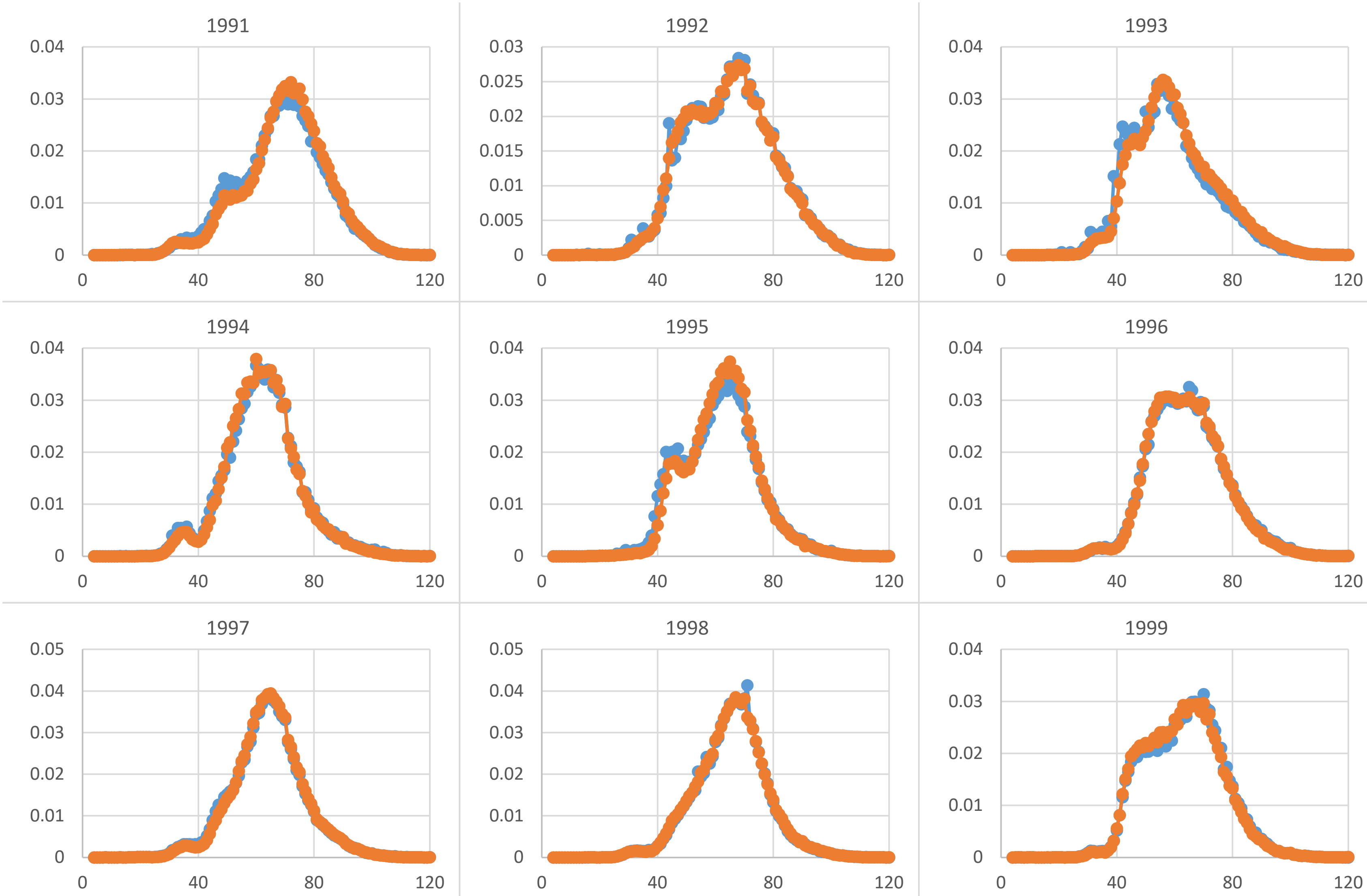


Figure A1a. Comparison of 1991-1999 sizecomp data used in last year's EBS assessment (orange) with catch-weighted sizecomp data (blue).

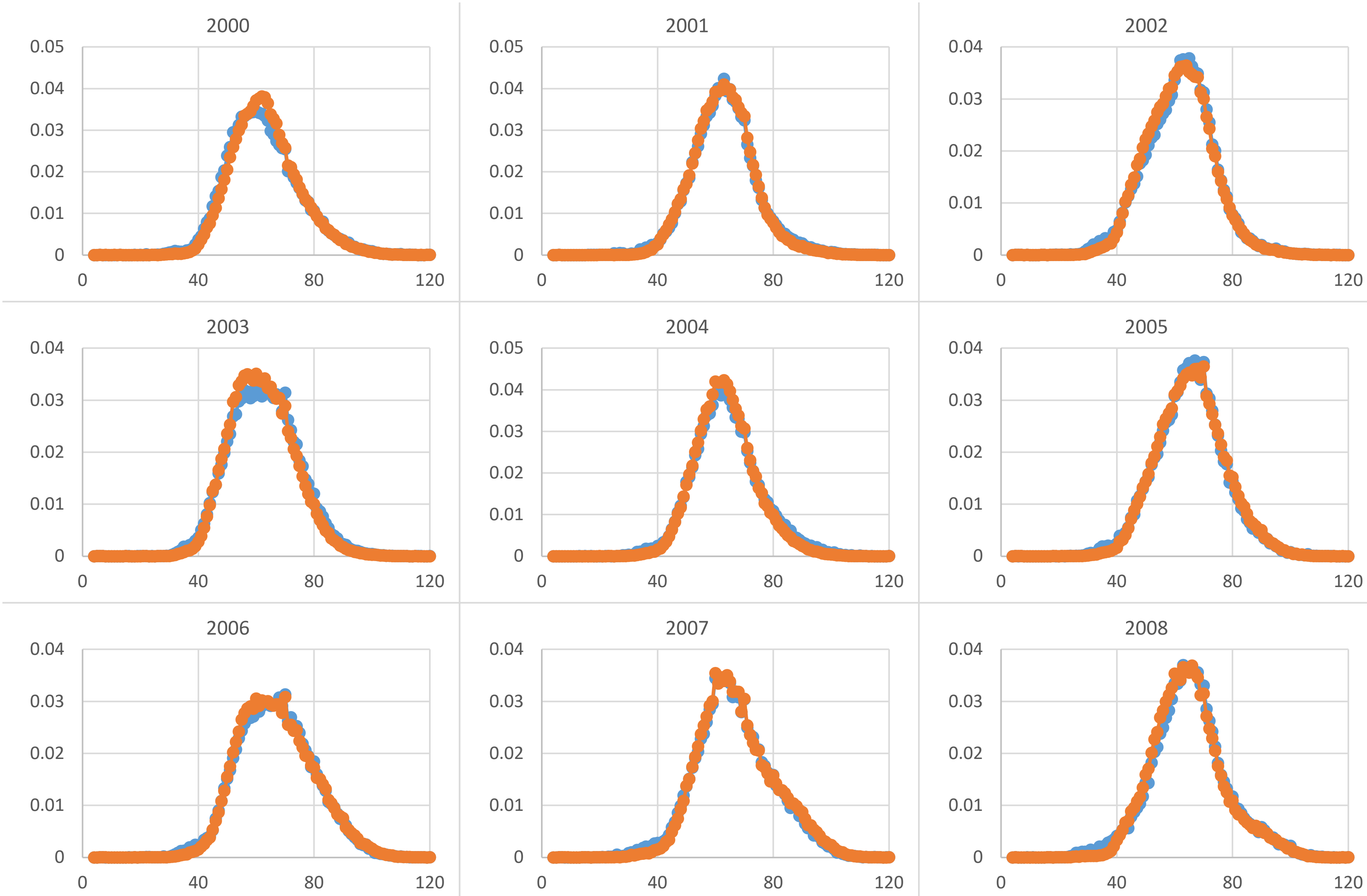


Figure A1b. Comparison of 2000-2008 sizecomp data used in last year's EBS assessment (orange) with catch-weighted sizecomp data (blue).

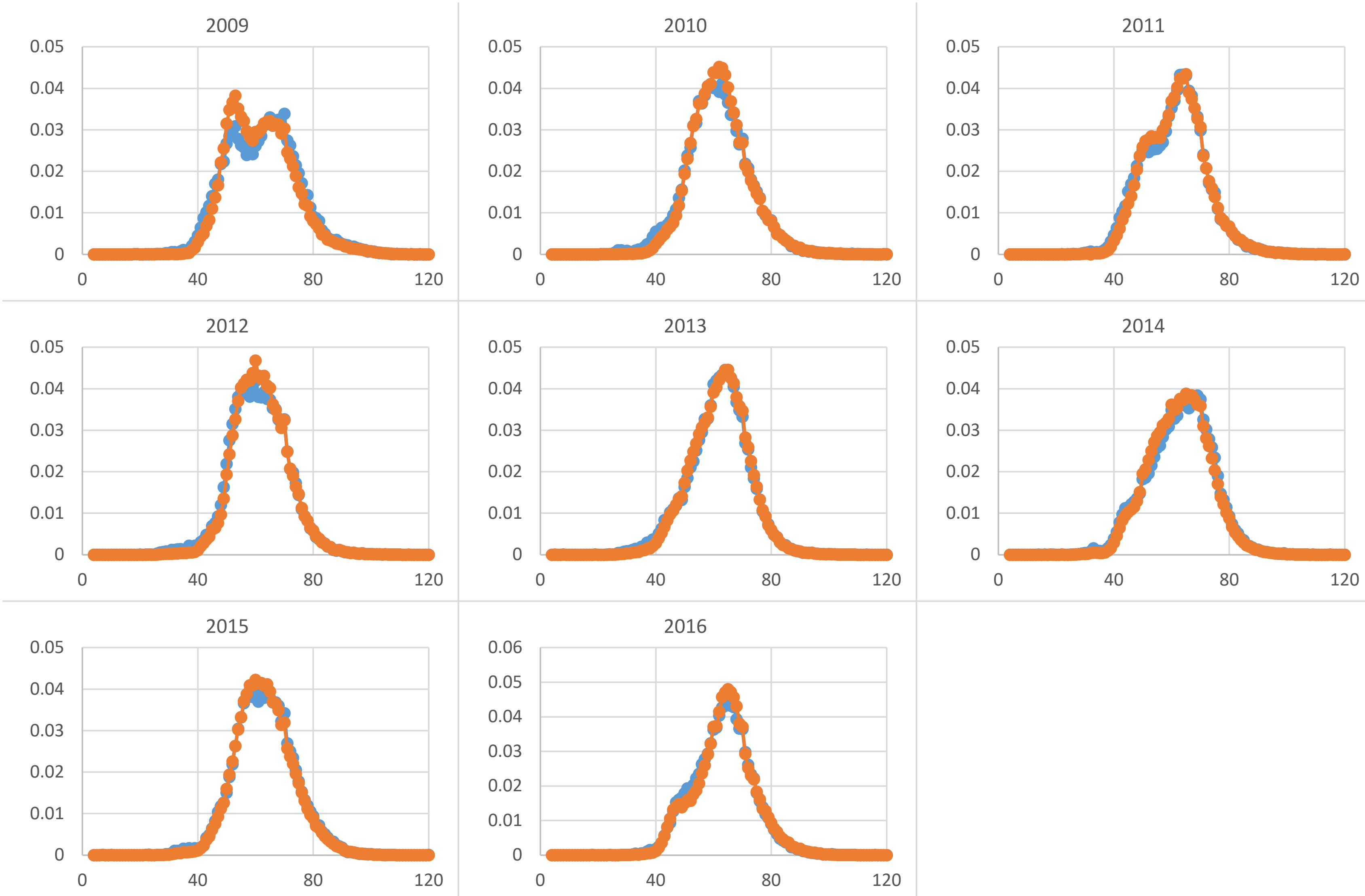
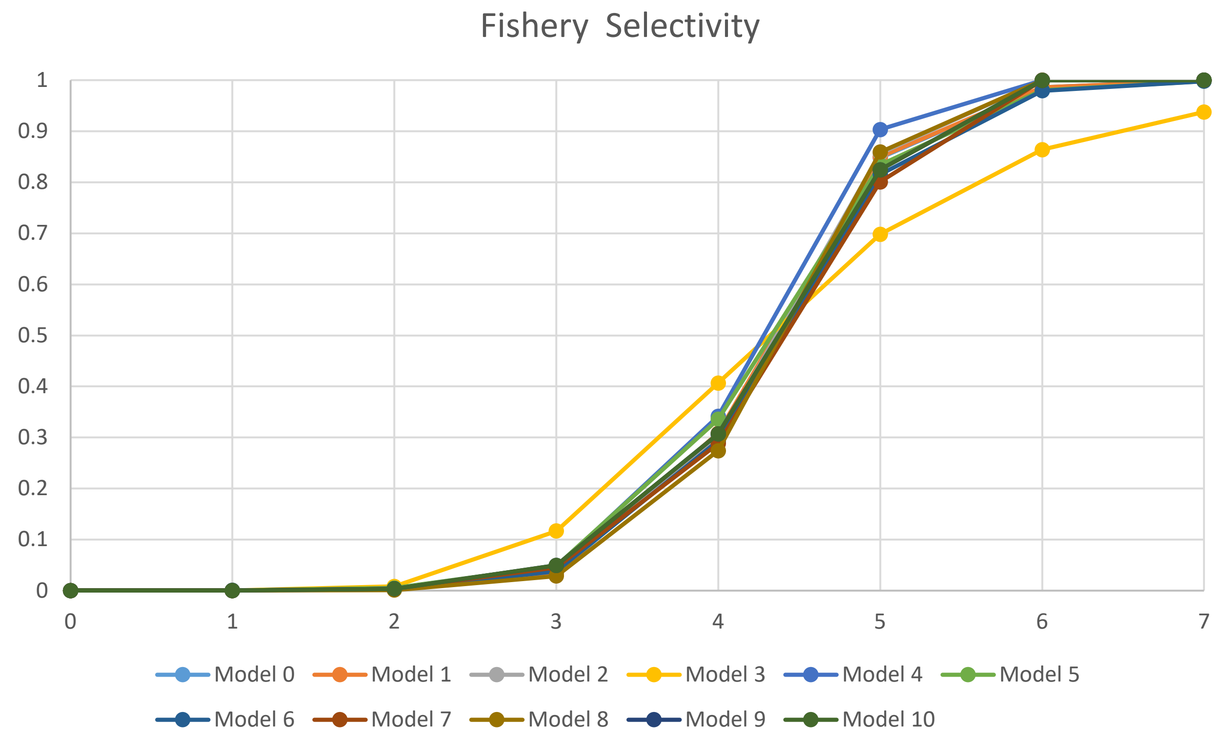


Figure A1c. Comparison of 2009-2016 sizecomp data used in last year's EBS assessment (orange) with catch-weighted sizecomp data (blue).



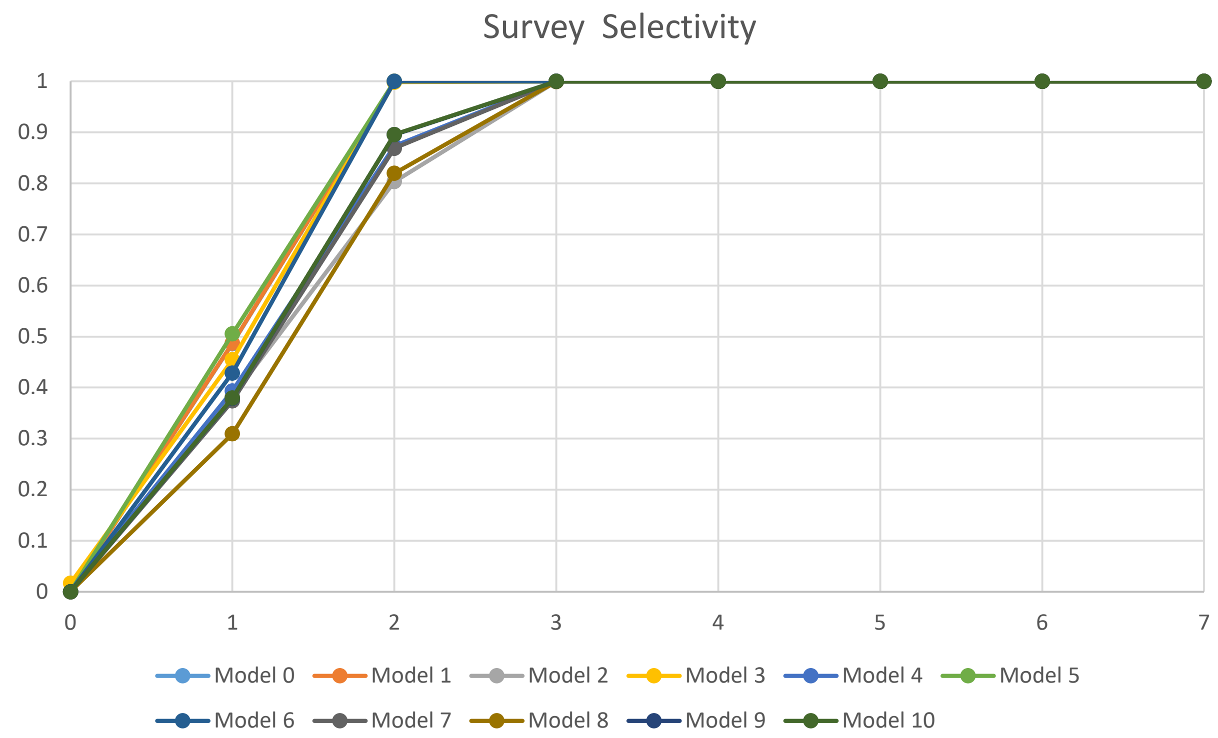


Figure A2a. Time-invariant schedules in Models 0-7 and the base selectivity schedules in Models 8-10.

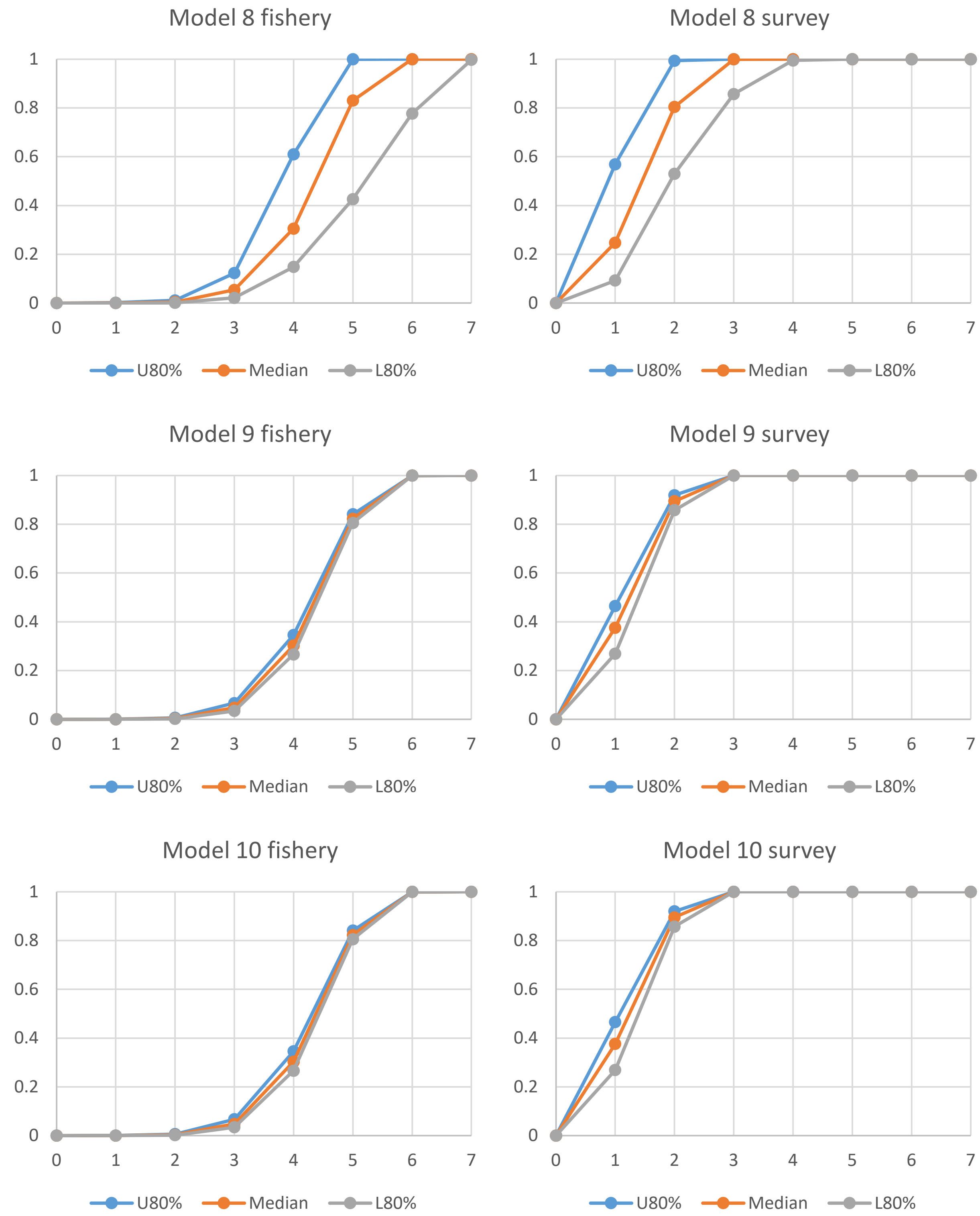


Figure A2b. Median and upper/lower 80% CIs for the time-varying selectivity schedules in Models 8-10.

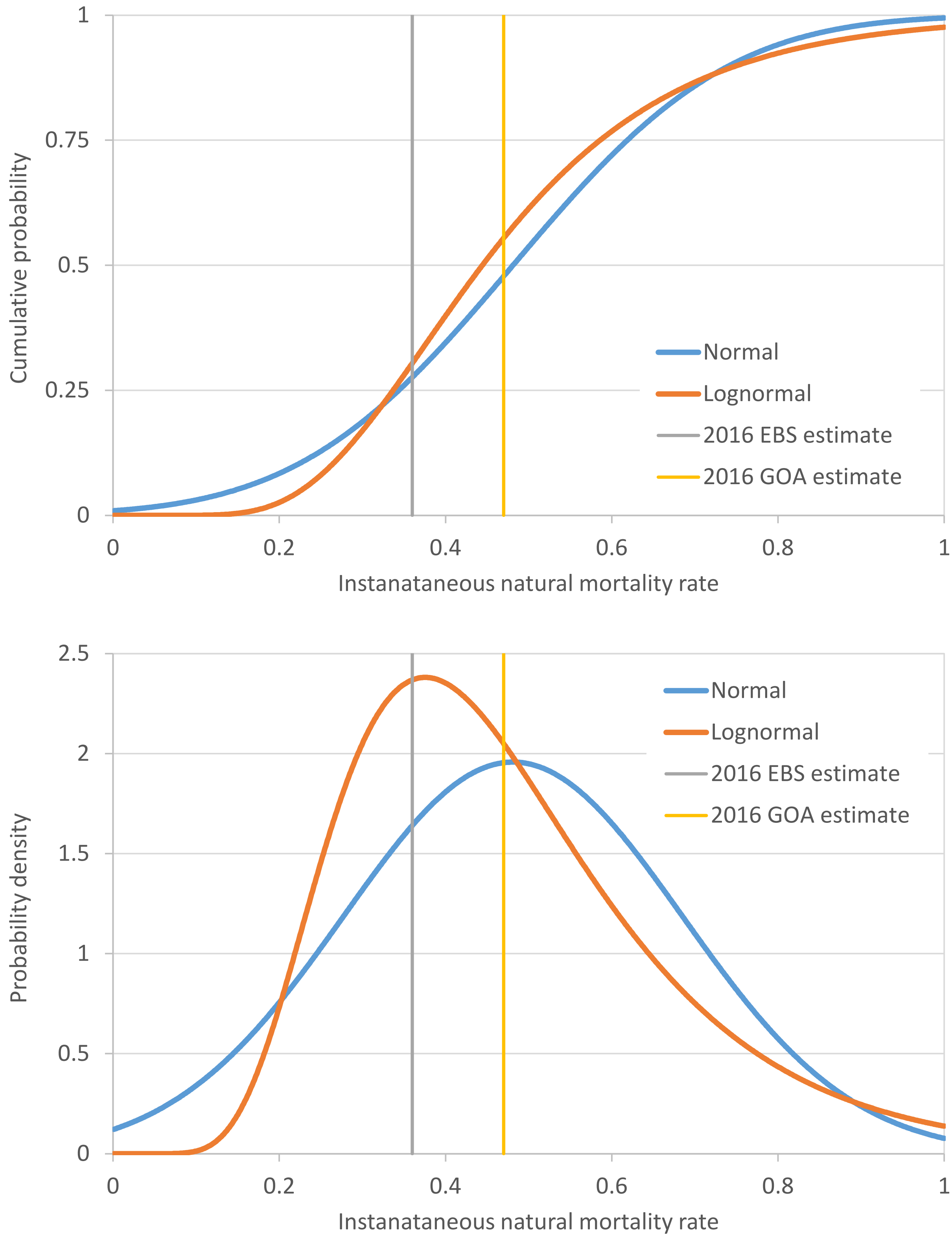


Figure A3. Possible prior distributions for the natural mortality rate.

APPENDIX B: Empirical weight at age

Use of empirical weight at age in last year's assessments

Comments BPT3 and SSC11 both encourage continued efforts to compare empirical weight at age with estimates obtained by the "traditional" approach. In the traditional approach, a constant (across time) length-at-age relationship and equations describing the between-individual variability about that relationship are estimated inside the model, and annually varying weight-at-length relationships are estimated outside the model. When combined, these sets of relationships are used by the model to convert internal estimates of population numbers at age and catch numbers at age into population biomass and catch biomass. In the empirical weight-at-age approach, annually varying weight-at-age relationships are estimated outside the model, which the model uses to convert internal estimates of population numbers at age and catch numbers at age into population biomass and catch biomass.

In the 2016 EBS assessment, three models (16.1, 16.8 and 16.9) used the empirical weight-at-age approach. The remaining models (11.5, 16.6, and 16.7) used the traditional approach (except that the weight-at-length relationship in Model 11.5 was constant across time rather than annually varying). The final 2016 EBS assessment included the following discussion of the two approaches:

> "Assuming that the estimates are accurate, the main advantage of using externally estimated weight at age is that this method integrates any changes in the length-at-age and weight-at-length relationships without having to estimate them inside the model. Disadvantages (in the context of the present assessment) include the following:
>
> 1. No smoothing was applied to the estimates, even though they exhibit a fair amount of variability, at least some of which seem implausible. For example, 10% of the within-cohort changes in weight from ages *a* to *a*+1 are negative.
> 2. Age data exist for only 18 of the 35 years in the survey time series and only 4 of the 39 years in the fishery time series. Long-term averages were used for all years with no age data.
> 3. The fishery age data come primarily from the longline fishery, and may not be representative of the overall fishery.
> 4. Because the trawl survey takes place in summer, beginning-of-year population weights at age were calculated by averaging mid-year weight(age,year) and mid-year weight(age−1,year−1), implying that weight at age changes linearly within each one-year interval.
> 5. Consistent with the last several assessments, all of the models in this year's assessment estimate a positive ageing bias (Table 2.22a), a finding which was recently confirmed by Kastelle et al. (2017) on the basis of stable isotope analysis, meaning that the empirical weights at age are likely biased downward.
>
> It may be advisable to examine more statistically sophisticated approaches for estimating weight at age outside of the assessment model, such as those that have been explored for the EBS walleye pollock assessment (Ianelli, this volume)."

The November 2016 BSAI Plan Team meeting included a discussion of the large discrepancies in the schedules of weight at age that were estimated by the two approaches, which are shown in the upper half of Table B1. In the interval between the November Team meeting and the December SSC meeting, the estimates were examined further by the assessment author, focusing primarily on the fact that the estimates used in the assessment were obtained by applying the externally estimated annual weight-at-length relationship to the externally estimated mean length at age. Because this method is known to be biased low due to the form of the relationship between weight and length (e.g., Pienaar and Ricker 1968), an alternative set of mean weights at age was derived by integrating the weight-at-length relationship across the distribution of length (assumed to be normal) for each age. The resulting estimates were much closer to the estimates obtained by the traditional approach, as shown in the lower half of Table B1. Except for age 2, the average (across years) weights at age from the empirical approach were always within 13% of the average obtained by the traditional approach.

Reference

Pienaar, L. V., and W. E. Ricker. 1968. Estimating mean weight from length statistics. *J. Fish. Res. Bd. Canada* 25(12):2743-2747.

Table B1. Comparison of model weights at age to two implementations of empirical weight at age.

Ratio of begin-year weights (model to empirical), with empirical weights based on mean length at age.

Year	2	3	4	5	6	7	8	9	10	11	12	13	14
1999	2.26	2.20	2.33	1.86	1.56	1.41	1.30	1.19	0.94	0.97	0.94	0.96	1.01
2000	2.21	2.71	2.29	2.42	1.98	1.47	1.43	1.21	1.23	0.97	0.98	0.98	1.03
2001	2.07	2.71	2.41	2.11	2.04	1.89	1.30	1.27	1.16	1.16	1.08	1.05	1.10
2002	1.90	2.44	2.56	2.08	1.88	1.75	1.72	1.25	1.04	1.19	1.05	1.08	1.14
2003	1.89	2.01	2.29	2.09	1.74	1.60	1.51	1.41	1.21	0.98	1.18	1.12	1.18
2004	1.87	2.44	1.97	1.95	1.70	1.49	1.39	1.23	1.11	1.13	1.01	1.13	1.18
2005	2.47	2.24	2.28	1.88	1.72	1.49	1.36	1.31	1.07	1.07	1.10	1.09	1.15
2006	1.73	2.80	2.28	2.05	1.72	1.49	1.24	1.15	1.14	0.95	1.01	1.03	1.09
2007	1.94	2.05	2.18	1.93	1.73	1.47	1.25	1.11	0.99	1.04	0.96	1.02	1.07
2008	2.41	2.16	1.80	1.68	1.56	1.40	1.20	1.07	1.02	0.84	0.97	0.98	1.03
2009	1.89	2.05	1.87	1.58	1.50	1.47	1.34	1.19	1.11	1.09	0.94	1.06	1.11
2010	2.12	1.88	1.71	1.68	1.49	1.44	1.29	1.26	1.08	1.05	1.00	0.98	1.03
2011	1.55	1.82	1.58	1.41	1.46	1.37	1.35	1.09	1.19	1.05	1.01	0.98	1.04
2012	1.67	1.63	1.60	1.44	1.29	1.31	1.21	1.21	0.94	1.05	1.02	1.00	1.05
2013	2.30	1.66	1.62	1.44	1.42	1.26	1.31	1.17	1.13	0.99	0.99	1.03	1.08
2014	1.53	1.79	1.57	1.54	1.35	1.40	1.20	1.21	1.12	1.00	1.12	1.09	1.15
2015	1.58	1.80	1.71	1.57	1.47	1.33	1.29	1.13	1.07	1.04	0.96	1.12	1.18
Mean:	1.96	2.14	2.00	1.81	1.62	1.47	1.34	1.20	1.09	1.03	1.02	1.04	1.09

Ratio of begin-year weights (model to empirical), with empirical weights integrated over distribution.

Year	2	3	4	5	6	7	8	9	10	11	12	13	14
1999	0.98	1.20	1.41	1.15	0.88	0.90	0.91	0.89	0.76	0.97	0.96	1.08	1.16
2000	0.74	1.16	1.24	1.22	1.15	0.94	1.09	1.05	1.08	1.18	1.08	1.08	1.26
2001	0.67	1.37	1.26	1.17	1.04	0.97	0.85	0.89	1.06	0.92	1.03	0.92	1.05
2002	0.81	1.18	1.57	1.20	1.13	1.02	1.01	1.01	0.84	1.10	0.94	0.97	1.01
2003	0.50	1.00	1.33	1.22	1.06	0.97	0.89	0.98	1.07	1.00	0.98	1.37	
2004	0.63	0.97	1.02	1.04	0.92	0.86	0.86	0.83	0.90	1.02	0.95	0.98	
2005	0.68	1.01	1.18	1.11	1.02	0.89	0.79	0.92	0.90	1.00	1.06	0.92	1.01
2006	0.56	1.21	1.27	1.27	1.17	0.91	0.81	0.74	0.89	0.88	0.95	0.97	
2007	0.72	0.91	1.18	1.09	1.13	0.94	0.80	0.75	0.73	0.96	0.98	1.04	
2008	0.87	0.89	0.90	0.99	0.89	0.80	0.76	0.71	0.72	0.74	0.92		
2009	0.66	0.90	0.87	0.95	1.02	0.96	0.86	0.81	0.79	0.90	0.95	1.13	
2010	0.76	0.80	0.83	0.89	0.95	1.00	0.89	0.92	0.80	0.93	0.87		
2011	0.56	0.85	0.78	0.83	0.91	0.89	0.91	0.79	0.84	0.91	1.05	0.83	
2012	0.67	0.76	0.92	0.87	0.88	0.92	0.82	0.86	0.79	0.83	0.95	0.93	
2013	0.75	0.74	0.90	0.86	0.92	0.86	0.93	0.87	0.91	0.95	0.87	1.00	
2014	0.55	0.71	0.75	0.88	0.82	0.91	0.87	0.89	0.95	0.89	1.13		0.95
2015	0.48	0.76	0.79	0.83	0.83	0.83	0.88	1.01	0.83	0.88	0.94	1.00	
Mean:	0.68	0.97	1.07	1.03	0.98	0.92	0.88	0.88	0.87	0.95	0.98	1.02	1.08