Interim Report of the Joint Team/SSC Working Group on Assessment/Management Issues Related to Recruitment

September 2014

Introduction

In 2012, the Groundfish Plan Teams and Crab Plan Team appointed a working group (Robert Foy, Jim Ianelli, Diana Stram, Grant Thompson) to list and evaluate alternatives for a number of assessment and management issues related to recruitment. Phases I and II of the working group's report were structured around the list of items considered at an April 2012 workshop, and were submitted in May and September of 2012, respectively. Other working group members (Anne Hollowed, André Punt, William Stockhausen, and Farron Wallace) were added following completion of Phase II. Phase III of the report was submitted in September 2013. After reviewing the Phase III report, the Groundfish Plan Teams had suggestions for further work on items B1, B2, B4, B5, B7, and C1. The Teams' suggestions for two of these items (B1 and B7) were addressed by the working group in 2014. Because only two items are addressed, the present report is referred to as an "interim" report rather than a full Phase IV report. Revised lists of alternatives and recommendations for items B1 and B7 are shown below (the original version of Appendix D, a revised version of Appendix F, and a new Appendix G are attached).

B1: Establishing criteria for excluding individual within-regime year classes from estimates

(Text under Alternatives B1.1-B1.4 is unchanged from the Phase III report. Text under Alternative B1.5, and its associated Option, is new. A simple but quantitative evaluation of some of the alternatives listed here is contained in Appendix D. Alternative B1.5 and its associated Option are based on Appendix G.)

Alternative B1.1: Do not exclude any individual within-regime year classes from estimates.

Pro: 1) Eliminates the need to specify quantitative criteria for excluding individual year classes.
Con: 1) May include poorly estimated year classes (e.g., will stock assessment authors be required to estimate strengths of *all* year classes in the current regime, even age 0 in the current year?).

Alternative B1.2: Exclude all year classes with model-estimated CVs greater than X.

Pro: 1) Very easy to implement, where feasible. 2) Clear relationship to precision of estimated year class strengths.
Con: 1) May not be feasible, because model-estimated CVs vary greatly across assessments (for example, looking at the CVs of estimated year class strengths from 1977-2009 in the sablefish and EBS Pacific cod assessments, sablefish had only 3 year classes with a CV of less than 10% compared to 25 year classes for Pacific cod, while sablefish had 25 year classes with a CV of greater than 20% compared to 1 year class for Pacific cod).

This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the National Marine Fisheries Service and should not be construed to represent any agency determination or policy.

Alternative B1.3: Exclude all year classes with model-estimated CVs greater than a fraction X *(<1) of the CV at the first age included in the model.*

Pro: 1) Very easy to implement, where feasible. 2) Clear relationship to precision of estimated year class strengths. 3) May be more feasible than B1.2, because the *relative* CV (rather than the *absolute* CV) is the criterion.

Con: 1) May still be infeasible (i.e., if *X* is set too low).

Alternative B1.4: Exclude all year classes with model-estimated CVs greater than a fraction X *(>1) of the asymptotic CV (i.e., the limiting CV that is approached as the number of times a year class is observed becomes large).*

Pro: 1) Clear relationship to precision of estimated year class strengths. 2) Where feasible, may be more intuitive than the other approaches, because this approach explicitly focuses on using only those year classes where the estimates have truly stabilized.

Con: 1) May be infeasible, because an asymptotic CV does not always exist. 2) The most difficult alternative to implement, because the asymptotic CV may vary from year class to year class.

*Alternative B1.5 (***recommendation; see also "Option" below***): Adopt one of the models estimated in Appendix G, which were fit to a large sample of data reflecting current practices in groundfish stock assessments.*

Pro: 1) Extremely easy to implement. 2) Always feasible, unless the recommended age is higher than the largest age in the model. 3) All but one of the models includes consideration of the natural mortality rate, which was a criterion requested by the SSC. 4) Because the models were fit to data reflecting current practices, use of one of these models should minimize (to some extent) the number and magnitude of changes to current practices.

Con: 1) No necessary relationship to precision of estimated year class strengths.

*Option under Alternative B1.5 (***recommendation***): Adopt Model 7 from Appendix G:* first_age=*round((0.05/*M)+$A_{10\%}$*), where* first_age *is the first age for which the recruitment corresponding to the model's estimated abundance in the current year is included in the recruitment time series,* M *is the instantaneous natural mortality rate, and* $A_{10\%}$ *is the first age in the assessment model for which survey selectivity is at least 10%.*

Pro: 1) Has the best performance of any of the seven models developed in Appendix G in terms of AIC if data with *first_age*<$A_{10\%}$ are excluded. 2) Ranks at least in the top three of any of the seven models in terms of R^2 if data with *first_age*<$A_{10\%}$ are excluded. 3) Model 7 is the only model of the seven that assures *first_age* will always be at least as great as $A_{10\%}$, thus preventing year-classes that have not been observed significantly by the survey from being included in the time series.

Con: 1) Would require some substantial changes to current practice for those assessments that include recruitments corresponding to ages in the current year with survey selectivity < 10%.

B7: Preferred measure of central tendency in recruitment

(Text under both alternatives has changed since the Phase III report. Appendix F contains two analyses upon which some of the "pro" and "con" arguments below were based.)

Alternative B7.1: To estimate Tier 3 reference points, scale spawning per recruit by the median of the recruitment time series for the current regime.

Pro: 1) Compared to use of mean recruitment, use of the median results in less variability in long-term (but not short-term) fishing effort and catch. 2) Compared to use of mean recruitment, use of the median results in higher long-term (but not short-term) average catch. 3) Compared to use of mean recruitment, use of the median results in shorter and less variable rebuilding times.

Con: 1) Compared to use of mean recruitment, use of the median results in lower average biomass (and thus CPUE) in both the short and long term. 2) Compared to use of mean recruitment, use of the median is less likely to account appropriately for curvature in the stock-recruitment relationship. 3) Projection software (e.g., Proj, Stock Synthesis) would need to be rewritten. 4) The FMPs would have to be amended.

*Alternative B7.2 (***recommendation***): To estimate Tier 3 reference points, continue to scale spawning per recruit by the mean of the recruitment time series for the current regime.*

Pro: 1) Compared to use of median recruitment, use of the mean results in higher average biomass (and thus CPUE) in both the short and long term. 2) Compared to use of median recruitment, use of the mean is more likely to account appropriately for curvature in the stock-recruitment relationship. 3) Projection software (e.g., Proj, Stock Synthesis) would not have to be rewritten. 4) The FMPs would not have to be amended. 5) Estimates of Tier 3 reference points have been based on mean recruitment for well over a decade, with few instances in which this appeared to have been problematic.

Con: 1) Compared to use of median recruitment, use of the mean results in greater variability in long-term fishing effort and catch. 2) Compared to use of median recruitment, use of the mean results in lower long-term average catch. 3) Compared to use of median recruitment, use of the mean results in longer and more variable rebuilding times.

Appendix D: A simple analysis of the B1 alternatives

Assumptions common to all examples discussed here:

A. The observational data consist of a survey time series (of length n) of numbers at age, which, when log-transformed, are distributed normally about the true log numbers at age.
B. The time series of Q, selectivity at age, and Z at age are known.

Given the above assumptions, after n observations, the CV of a cohort's estimated initial abundance (i.e., the abundance at some age prior to the age at the first observation) is equal to sqrt($h(n)/n$), where $h(n)$ is the harmonic mean of the time series of the log-scale observation error variances. To make things even simpler, an additional assumption will be used:

C. The log-scale observation error variance is equal to the following constant function of age (t): sigma^2 = exp($a + b*t + c*t$^2).
    a. In the special case where b=c=0, the CV of the estimated initial abundance after n years is CV(n)=sqrt(exp(a)/n). Note that this value equals zero in the limit as n approaches infinity.
    b. In the special case where b≠0 and c=0, the CV of the estimated initial abundance after n years is CV(n)=sqrt(exp(a)*(exp(b)-1)/(1-exp(-b*n))). Note that this value equals zero in the limit as n approaches infinity, as in the b=c=0 case.
    c. In the general case where b≠0 and c≠0, there is no short-hand formula for the CV of the estimated initial abundance after n years . In contrast to the two previous cases, CV(n) reaches a positive asymptote (the "asymptotic CV") in the limit as n approaches infinity.

Alternatives for criteria pertaining to exclusion of the most recent within-regime year classes:

1. Exclude no year classes.
2. Exclude all year classes within the last X years.
    a. In the special case where b=c=0, the *proportional reduction* in CV relative to CV(1) will depend only on X, but the *absolute* CV will also depend on a.
    b. In the special case where b≠0 and c=0, the *proportional reduction* in CV relative to CV(1) will depend only on X and b, but the *absolute* CV will also depend on a.
    c. In the case where b≠0 and c≠0, both the *proportional reduction* in CV relative to CV(1) will depend only on X, b, and c; but the *absolute* CV will also depend on a.
3. Exclude all year classes with model-estimated CVs greater than X.
    a. In the special case where b=c=0, the number of years needed to achieve CV(n)=X and the *proportional reduction* in CV relative to CV(1) will both depend on X and a.
    b. In the special case where b≠0 and c=0, the number of years needed to achieve CV(n)=X and the *proportional reduction* in CV relative to CV(1) will both depend on X, a, and b.
    c. In the case where b≠0 and c≠0, it will be impossible to achieve CV(n)=X if X is set too low. If X is set sufficiently high, the number of years needed to achieve CV(n)=X and the *proportional reduction* in CV relative to CV(1) will both depend on X, a, b, and c.
4. Exclude all year classes with model-estimated CVs greater than a fraction X (<1) of the CV at the first age included in the model.

   a. In the special case where $b=c=0$, the number of years needed to achieve $CV(n)=X*CV(1)$ will depend only on X, but the *absolute* CV will also depend on a.
   b. In the special case where $b\neq0$ and $c=0$, the number of years needed to achieve $CV(n)=X*CV(1)$ will depend only on X and b, but the *absolute* CV will also depend on a.
   c. In the case where $b\neq0$ and $c\neq0$, it will be impossible to achieve $CV(n)=X*CV(1)$ if X is set too low. If X is set sufficiently high, the number of years needed to achieve $CV(n)=X*CV(1)$ will depend only on X, b, and c; but the *absolute* CV will also depend on a.
5. Exclude all year classes with model-estimated CVs greater than a fraction X (>1) of the asymptotic CV.
   a. In the special case where $b=c=0$, the asymptotic CV is zero, so the number of years needed to achieve $CV(n)=X*CV(\infty)$ will always be infinite.
   b. In the special case where $b\neq0$ and $c=0$, the asymptotic CV is zero, so the number of years needed to achieve $CV(n)=X*CV(\infty)$ will always be infinite.
   c. In the case where $b\neq0$ and $c\neq0$, the number of years needed to achieve $CV(n)=X*CV(\infty)$ will depend only on X, b, and c; but the *absolute* CV will also depend on a.

Note that Alternative #1 is the only one that works regardless of the values of the parameters. However, this begs the question of what to count as the "first observation." Here are some alternatives:

I. The first observation is the first age in the model. This definition could be problematic, because some models start at an age prior to the first age with data (e.g., SS always starts at age zero); conversely, an author might start the model well past the first age with data.
II. The first observation is the first age with relative abundance data for the cohort in question. This definition could be problematic if only a trivial amount of abundance data exist at the first age thus defined.
III. The first observation is the first age with *significant* relative abundance data for the cohort in question. This begs the question of what constitutes "significant." Some sub-alternatives:
   i. "Significant" means an observation error CV of less than X. This definition could be problematic if X is set so low that the definition cannot be satisfied at any reasonably low age (or, worse, not at all).
   ii. "Significant" means estimated survey selectivity greater than X in the respective age and year.

Appendix F: Two simple analyses of the B7 alternatives

Analysis #1

To compare and contrast the performance of using median and mean recruitment to compute the biomass reference points in the Tier 3 control rules, a simple simulation was conducted.

To capture the idea of occasional recruitments that are much larger or more common than would be expected from a single lognormal distribution, recruitments were drawn randomly from a weighted sum of two lognormal distributions, one of which represents "typical" recruitments and the other of which represents "extreme" recruitments.

The values of the parameters governing the recruitment distribution were as follow:

- For the "typical" distribution, $\mu = 0$, $\sigma = 0.6$
- For the "extreme" distribution, $\mu = 2$, $\sigma = 0.05$
- Proportion of time that recruitment is "typical" = 0.95

The ratio of the median "extreme" recruitment to the median "typical" recruitment is about 7.4, and the ratio of the means is about 6.2. The ratio of the *lower* end of the 95% confidence interval for the "extreme" distribution to the *upper* end of the 95% confidence interval for the "typical" distribution is about 2.1 (Figure F.1).

All fishing mortality was assumed to occur instantaneously at the start of the year, and was expressed in terms of a discrete annual exploitation rate U. Natural mortality was expressed in terms of a discrete annual rate A. A range of values (0.05, 0.10, 0.20, and 0.30) was considered for A. Selectivity and maturity were assumed to be knife-edged, with the first age of full selectivity equal to the first age of full maturity. Growth parameters were scaled so that $U_{35\%}=A$ and $B_{100\%}=1$. Values of all parameters and variables *except the median and mean of the recruitment distribution* were assumed to be known without error. The median and mean of the recruitment distribution were re-estimated in each year of the simulation, after a "burn-in" time of 30 years (i.e., no assessment was conducted during the first 30 years, so the first assessment has 30 years of recruitment data from which to estimate the median and mean).

For each value of A and each estimator, 10,000 simulations were conducted. Each simulation was initialized by assuming that the population was in equilibrium at 50% of the B_{MSY} proxy of $B_{35\%}$, where $B_{35\%}$ was scaled in terms of the respective estimator (i.e., $B_{35\%} = 0.35\times R_{med}\times SPR_{F=0}$ or $B_{35\%} = 0.35\times R_{ave}\times SPR_{F=0}$). The maximum age in the population was defined as the age at which only 0.1% of a cohort would remain in an equilibrium unfished population (where cohort size is measured at the age of recruitment), and so was different for each value of A. In each simulation, the population was projected forward for a number of years at least twice as great as the maximum age. Alternative values for the Tier 3 reference points were computed for each of the two estimators (median and mean recruitment). Catch was assumed to equal maxABC under the Tier 3 control rule in all years.

The following performance metrics were tabulated for each value of A and each of the two candidate estimators:

- Short-term (first 10 years) and long-term (last 10% of the time series) means and standard deviations of relative biomass (= biomass/$B_{40\%}$), relative exploitation (= exploitation/$U_{40\%}$), and relative catch (=catch/$C_{40\%}$); shown in Table F.1
- Five statistics pertaining to rebuilding time (to the B_{MSY} proxy, again scaled in terms of the respective estimator): upper and lower bounds of the 95% confidence interval, median, mean, and standard deviation; shown in Table F.2.

Analysis #2

Examination of some even simpler, "worst-case" scenarios may help to illustrate some differences between using the median and using the mean. Suppose, instead of the model described in Analysis #1, that recruitment can take only two values, one low and one high. If recruitment is normalized relative to the low value, the distribution is defined by two parameters: the ratio of the high recruitment to the low recruitment (*ratio_of_values*), and the probability that the high value will occur (*proportion_high*). The overall mean recruitment is then given as *overall_mean* = (1–*proportion_high*) + (*proportion_high* × *ratio_of_values*).

In this situation, the worst-case scenario in terms of bias in the estimated mean is for the stock to experience only low recruitments during the burn-in period of *nbur* years, followed by a high recruitment in the first year after the burn-in period. This bias can be derived as follows:

1. Given that recruitments are normalized to the low value, the sum of recruitments for the first *nbur* years is simply *nbur*.
2. Again, given that recruitments are normalized to the low value, the recruitment for year *nbur*+1 is simply *ratio_of_values*.
3. The sum of recruitments for years 1 through *nbur*+1 is thus *nbur*+*ratio_of_values*.
4. The mean recruitment for years 1 through *nbur*+1 is thus (*nbur*+*ratio_of_values*)/(*nbur*+1).
5. The bias is computed as the ratio of the mean recruitment for years 1 through *nbur*+1 divided by the *overall_mean*, minus 1:

$$bias_{mean} = \frac{nbur + ratio_of_values}{(nbur+1)\cdot overall_mean} - 1 \quad .$$

The worst of the worst-case scenarios occurs when there is no burn-in period at all, in which case the bias is positive and potentially large. The bias then decreases continuously with *nbur* until it equals 0 at *nbur* = $proportion_high^{-1}$–1, after which it becomes negative and reaches an asymptote at $overall_mean^{-1}$–1.

The bias in the median is independent of *nbur*, being equal simply to the lower asymptote of $bias_{mean}$.

For comparison, the eastern Bering Sea shelf bottom trawl survey now has a length of 32 years (1982-2013). In this very simple, worst-case model, the estimated bias in the mean for any assessment that includes recruitment estimates for all years in the survey time series will be negative for any value of *proportion_high* greater than about 0.03, but less negative than the estimated median.

Some examples are shown in Figure F.2.

Table F.1. Short-term and long-term (last 10% of the time series) means and standard deviations of relative biomass (= biomass/$B_{40\%}$), relative exploitation (= exploitation/$U_{40\%}$), and relative catch (=catch/$C_{40\%}$) under the two alternative estimators and a range of discrete annual mortality rates; from Analysis #1.

		Discrete mortality rate = 0.05				Discrete mortality rate = 0.10			
		Short-term		Long-term		Short-term		Long-term	
Quantity	Estimator	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Biomass	Mean	0.57	0.13	1.03	0.17	0.68	0.23	1.04	0.24
Biomass	Median	0.46	0.14	1.00	0.18	0.58	0.23	1.00	0.25
Exploitation	Mean	0.56	0.14	0.95	0.07	0.66	0.20	0.93	0.09
Exploitation	Median	0.64	0.19	1.00	0.01	0.74	0.22	1.00	0.02
Catch	Mean	0.34	0.16	0.99	0.21	0.48	0.30	0.98	0.29
Catch	Median	0.31	0.18	1.00	0.18	0.47	0.30	1.00	0.25

		Discrete mortality rate = 0.20				Discrete mortality rate = 0.30			
		Short-term		Long-term		Short-term		Long-term	
Quantity	Estimator	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Biomass	Mean	0.81	0.34	1.05	0.33	0.88	0.42	1.06	0.41
Biomass	Median	0.72	0.36	1.01	0.35	0.81	0.44	1.01	0.42
Exploitation	Mean	0.74	0.22	0.90	0.11	0.77	0.22	0.88	0.13
Exploitation	Median	0.82	0.22	0.99	0.04	0.85	0.21	0.97	0.06
Catch	Mean	0.66	0.43	0.97	0.40	0.75	0.51	0.97	0.48
Catch	Median	0.65	0.42	1.00	0.36	0.75	0.49	1.00	0.44

Table F.2. Statistics related to rebuilding time under the two alternative estimators and a range of discrete annual mortality rates, from Analysis #1.

	Mortality rate = 0.05		Mortality rate = 0.10		Mortality rate = 0.20		Mortality rate = 0.30	
Statistic	Mean	Median	Mean	Median	Mean	Median	Mean	Median
U95%CI	35	21	24	14	18	9	15	8
Mean	19	12	11	7	7	5	6	4
Median	18	11	10	7	7	5	5	4
L95%CI	7	4	3	2	2	2	2	2
St. Dev.	7.0	4.6	5.4	3.1	4.0	2.0	3.3	1.6

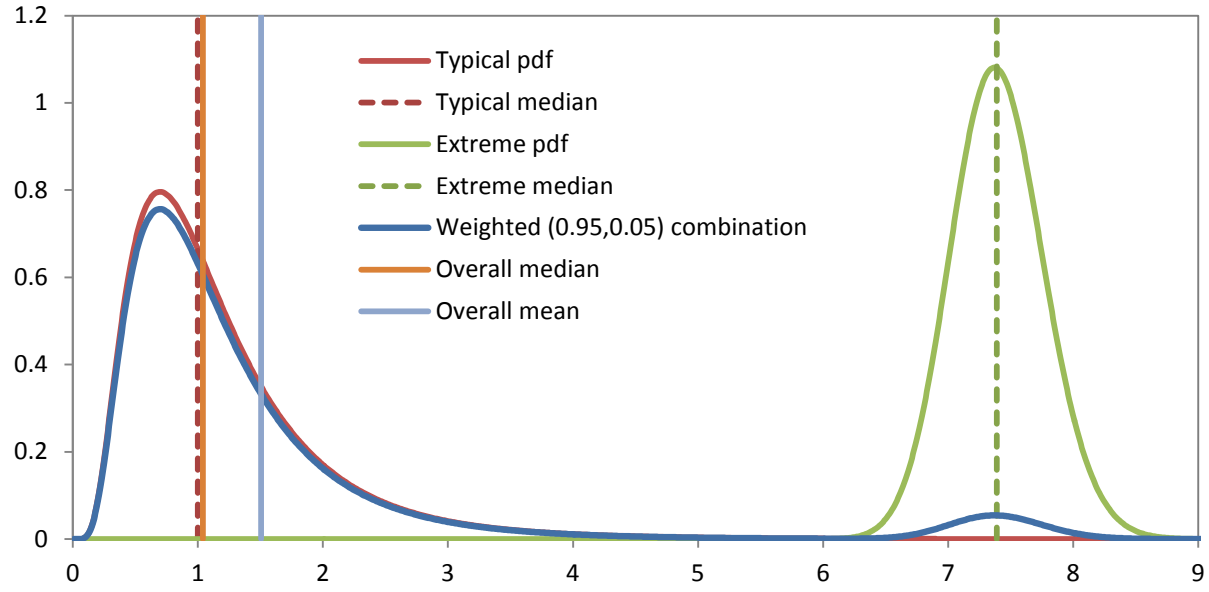

Figure F.1. Probability density function of recruitment used in Analysis #1.

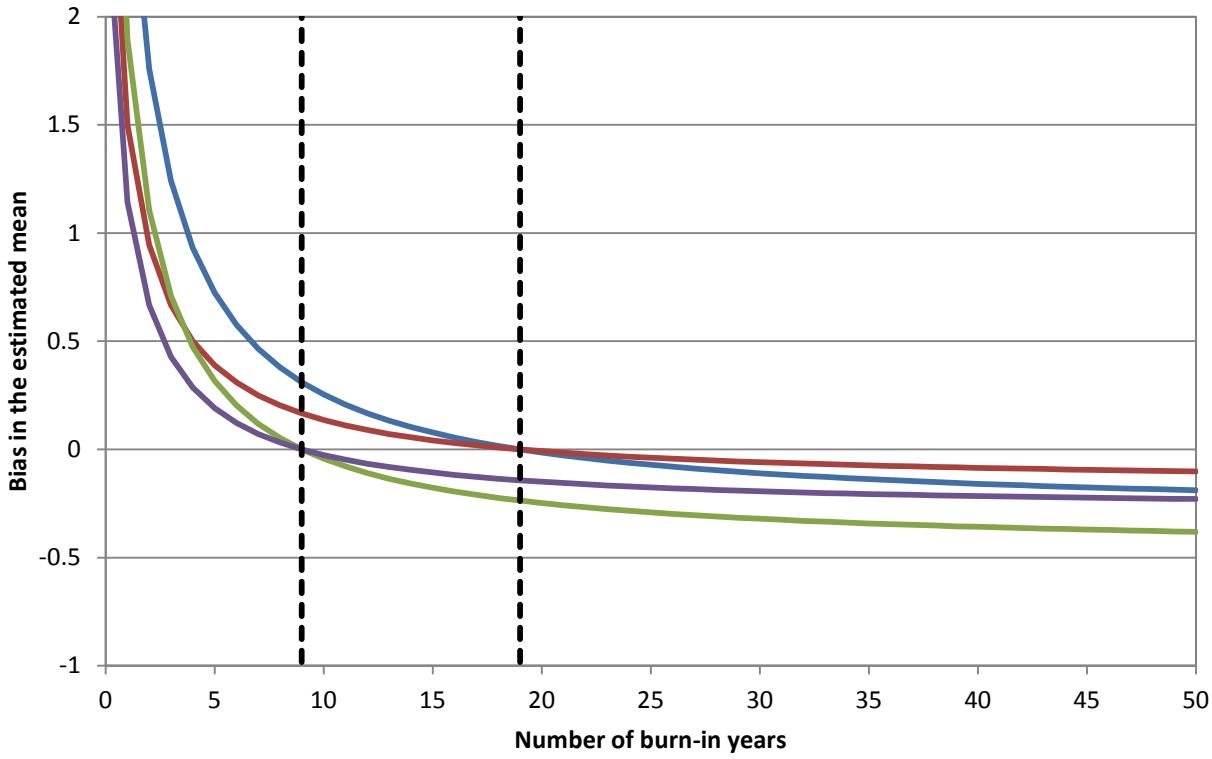

Figure F.2. Some examples of "worst-case" bias in the estimated mean from Analysis #2. Blue: *ratio_of_values*=10, *proportion_high*=0.05; red: *ratio_of_values*=5, *proportion_high*=0.05; green: *ratio_of_values*=10, *proportion_high*=0.1; purple: *ratio_of_values*=5, *proportion_high*=0.1. Vertical dashed lines indicate the value of *nbur* for which bias is zero (specific to *proportion_high*).

Appendix G: Estimating the number of recent year-classes to exclude from the time series, based on current practices

The 2013 ("Phase III") report of the Recruitment Working Group responded to an SSC request for inclusion of the natural mortality rate (*M*) in the algorithm for determining the number of recent year-classes to exclude from the recruitment time series. This response consisted of an equation (withdrawn from this year's revised draft) that used *M* and the first age at which survey selectivity was at least 10% ($A_{10\%}$) to determine the number of recent year-classes to exclude, with the idea that year classes with current age < $A_{10\%}$ should be excluded from the time series. The parameters of that equation were loosely based on a few example stock assessments. At their September 2013 meeting, the Joint BSAI and GOA Groundfish Plan Teams "recommended that the working group conduct some further analysis and that the working group consider another alternative which uses $A_{50\%}$ (age at 50% selectivity)." In other Team discussion not explicitly reflected in the minutes, interest was also expressed in basing the algorithm on a wider range of stock assessment examples.

In August 2014, assessment authors were asked to provide the following items of information for each of their age-structured assessments:

- The natural mortality rate (*M*)
- The first age at which survey selectivity reaches 10% ($A_{10\%}$)
- The first age at which survey selectivity reaches 50% ($A_{50\%}$)
- The first age for which the recruitment corresponding to the model's estimated abundance in the current year is included in the recruitment time series (*first_age*—for example, if the assessment model begins with age 1, but the estimated recruitments corresponding to ages 1 and 2 in the current year are not included in the recruitment time series, then *first_age*=3)

Authors responded by providing data for 26 assessments (15 BSAI, 10 GOA, and 1 BSAI/GOA). In cases where separate values were provided for males and females, the values were averaged. In cases where survey selectivity at the first age in the assessment model was greater than 10% (or 50%), $A_{10\%}$ (or $A_{50\%}$) was set equal to the first age in the model.

In the overall data set, three factors emerged as being potentially important for individual consideration when analyzing the data:

1. In the 2013 assessment of GOA Pacific cod, the author changed the first age for inclusion in the time series from the traditional value of 2 to 4, because she thought that she was required to do so by the Phase III report. This change produced a notable outlier in the data set, so the data were analyzed two ways: using age 2 and age 4 for this stock.
2. The Phase III report used 10% as the minimum survey selectivity that should be accepted for estimation of incoming year classes, but the Plan Teams recommended adding a survey selectivity of 50% to the analysis, so the data were analyzed both ways.
3. Many of the assessments represented in the data use model estimates of abundance at ages lower than either of the suggested selectivity cutoffs: 35% use model estimates of abundance at ages less than $A_{10\%}$, and 77% use model estimates of abundance at ages less than $A_{50\%}$. The data were analyzed two ways: including, or excluding, those assessments that use model estimates of abundance at ages lower than the suggested selectivity cutoff. For the case of $A_{10\%}$, excluding such data reduced the sample size from 26 to 17, and for the case of $A_{50\%}$, excluding such data reduced the sample size from 26 to 6.

This resulted in a factorial design of 2^3=8 different ways of looking at the data.

The following models were fit to the data (*A%* represents either $A_{10\%}$ or $A_{50\%}$, depending on which value is used in any given combination of the three factors listed above):

$$Model\,1: \quad first_age = \frac{a_1}{M + b_1}$$

$$Model\,2: \quad first_age = \exp\left(a_2 \cdot A\% + b_2\right)$$

$$Model\,3: \quad first_age = \left(\frac{a_3}{M + b_3}\right) \cdot \exp\left(c_3 \cdot A\% + d_3\right)$$

$$Model\,4: \quad first_age = \frac{a_4 \cdot A\%}{\left(M + b_4\right) \cdot \left(A\% + c_4\right)}$$

$$Model\,5: \quad first_age = a_5 + b_5 \cdot M + c_5 \cdot A\% + d_5 \cdot M \cdot A\%$$

$$Model\,6: \quad first_age = a_6 + \frac{b_6}{M} + c_6 \cdot A\% + \frac{d_6}{M} \cdot A\%$$

$$Model\,7: \quad first_age = \frac{a_7}{M} + A\%$$

All models were fit by the method of maximum likelihood, assuming a lognormal error structure. It should be noted that *first_age* is necessarily an integer, but the terms on the right-hand sides of the above equations are not restricted to integer values. Thus the analysis is necessarily only approximate, as restricting the right-hand sides to integer values would render the log-likelihood non-differentiable with respect to the parameters.

Three statistics relating to goodness of fit for each model within each combination of the three factors are shown in Table G.1: Akaike's Information Criterion (AIC); the coefficient of determination (R^2); and "Error," which describes the proportion of the data where the predicted value of *first_age* is less than *A%*. Within each combination of factors, results are sorted in decreasing order of preference for each goodness-of-fit measure:

- Focusing on AIC, Model 7 ranked as the best model whenever data with *first_age*<*A%* were excluded, Model 2 ranked as the best model when such data were not excluded and *A%* was set equal to $A_{10\%}$, and Model 4 ranked as the best Model when such data were not excluded and *A%* was set equal to $A_{50\%}$.
- Focusing on R^2, Model 6 ranked as the best model for all combinations of factors except when *A%* was set equal to $A_{50\%}$ and data were excluded whenever *first_age* was less than *A%*, in which cases Model 5 ranked as the best model. Model 7 ranked no lower than third whenever data with *first_age*<*A%* were excluded.
- Focusing on "Error," Model 7 ranked as the best model (with Error=0) whenever data with *first_age*<*A%* were excluded (just as it did when the focus was on AIC). When data with *first_age*<*A%* were not excluded, a variety of models ranked near the top, but it should be noted that the Error values for the top-ranked models in all such cases were very high (0.692-0.923).

By all three criteria, then, Model 7 does very well. The fit of Model 7 to the data with GOA Pcod *first_age*=2, *A%*=$A_{10\%}$, and data with *first_age*<*A%* excluded is shown in Figure G.1. With only one parameter, Model 7 is also easy to extend to the more realistic case where the right-hand side is constrained to take integer values. Rounding the values on the right-hand side to the nearest integer and profiling over a_7 (for the same combination of factors listed for Figure G.1) indicates that a_7=0.05 does as well as any other value of a_7.

Table G.1—Results, sorted in order of descending rank within each factor combination and measure.

Factors			AIC		R^2		Error	
Pcod age	Select.	Exclude?	Model	Value	Model	Value	Model	Value
2	0.1	0	2	48.414	6	0.431	1	0.731
2	0.1	0	1	48.653	4	0.405	2	0.731
2	0.1	0	7	48.958	3	0.352	4	0.731
2	0.1	0	4	49.063	1	0.346	6	0.769
2	0.1	0	6	50.355	5	0.299	3	0.769
2	0.1	0	5	50.921	2	0.155	5	0.808
2	0.1	0	3	51.415	7	-0.525	7	1.000
2	0.1	1	7	25.023	6	0.772	2	0.000
2	0.1	1	2	26.779	5	0.751	7	0.000
2	0.1	1	4	28.685	7	0.734	3	0.118
2	0.1	1	5	29.023	4	0.725	4	0.118
2	0.1	1	6	29.320	3	0.700	6	0.118
2	0.1	1	3	30.467	2	0.653	5	0.176
2	0.1	1	1	32.558	1	0.477	1	0.471
2	0.5	0	4	47.550	6	0.418	3	0.923
2	0.5	0	1	48.653	4	0.413	1	0.923
2	0.5	0	2	49.300	1	0.346	2	0.923
2	0.5	0	6	50.292	3	0.333	5	0.923
2	0.5	0	5	51.678	5	0.207	6	0.923
2	0.5	0	3	51.847	2	0.126	4	0.962
2	0.5	0	7	59.728	7	-7.543	7	1.000
2	0.5	1	7	-3.431	5	0.998	7	0.000
2	0.5	1	5	0.110	6	0.998	6	0.000
2	0.5	1	6	0.797	7	0.996	5	0.167
2	0.5	1	4	7.425	4	0.988	3	0.500
2	0.5	1	2	9.177	2	0.825	2	0.500
2	0.5	1	3	13.716	3	0.823	4	0.667
2	0.5	1	1	16.903	1	0.458	1	0.667
4	0.1	0	2	49.834	6	0.382	4	0.692
4	0.1	0	7	50.494	4	0.299	1	0.731
4	0.1	0	1	50.945	5	0.232	2	0.731
4	0.1	0	4	51.190	3	0.231	5	0.731
4	0.1	0	6	52.435	1	0.215	6	0.731
4	0.1	0	5	52.895	2	0.151	3	0.769
4	0.1	0	3	53.552	7	-0.615	7	1.000
4	0.1	1	7	27.664	6	0.672	2	0.000
4	0.1	1	2	29.152	7	0.630	3	0.000
4	0.1	1	4	31.335	5	0.629	4	0.000
4	0.1	1	6	31.757	4	0.603	5	0.000
4	0.1	1	5	31.987	3	0.592	6	0.000
4	0.1	1	3	33.448	2	0.572	7	0.000
4	0.1	1	1	34.831	1	0.306	1	0.529
4	0.5	0	4	49.980	6	0.322	2	0.923
4	0.5	0	2	50.648	4	0.294	1	0.923
4	0.5	0	1	50.945	3	0.235	4	0.923
4	0.5	0	6	52.671	1	0.215	5	0.923
4	0.5	0	5	53.744	5	0.130	6	0.923
4	0.5	0	3	54.081	2	0.103	3	0.923
4	0.5	0	7	60.417	7	-7.576	7	1.000
4	0.5	1	7	5.556	5	0.976	7	0.000
4	0.5	1	5	8.607	6	0.963	6	0.143
4	0.5	1	6	9.884	7	0.947	5	0.143
4	0.5	1	4	11.859	4	0.929	4	0.286
4	0.5	1	2	12.497	2	0.668	3	0.571
4	0.5	1	3	17.144	3	0.628	2	0.571
4	0.5	1	1	19.766	1	0.123	1	0.714

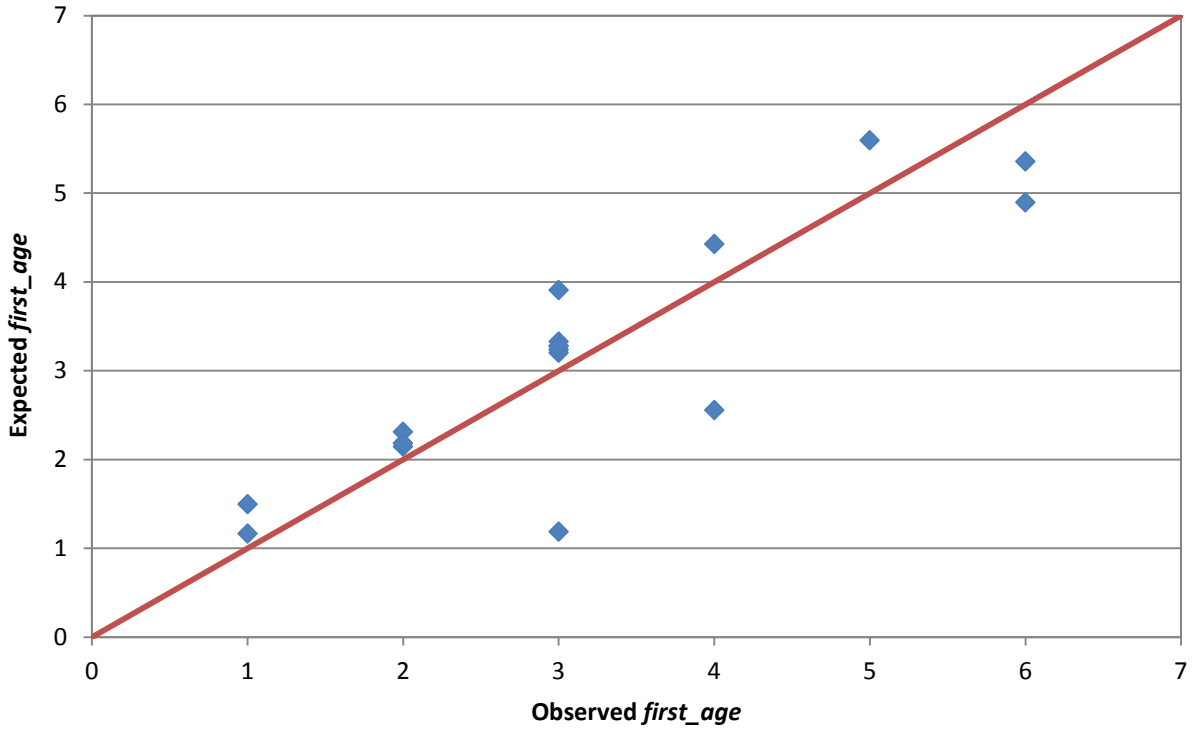

Figure G.1. Fit of Model 7 to the data (diamonds) with GOA Pcod *first_age*=2, $A\%=A_{10\%}$, and data with *first_age*<*A%* excluded. Diagonal line indicates what a perfect fit would look like.