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

August 2013

Introduction

In 2012, the Groundfish Plan Teams and Crab Plan Team ("GPTs" and "CPT," respectively) appointed a working group (Robert Foy, James Ianelli, Diana Stram, and Grant Thompson) to list and evaluate alternatives for a number of assessment and management issues related to recruitment. The working group's first activity was a workshop held at the AFSC Seattle laboratory during April 2012. The workshop was intended to address a long-standing request from the BSAI GPT for analysis of recruitment-related issues such as: which cohorts to include in estimation of reference points, how to estimate parameters related to recruitment (including parameters of a stock-recruitment relationship), and how to determine the reliability of the F_{MSY} probability density function. The workshop was also intended to satisfy the following SSC request (from the February 2012 minutes):

> *"The SSC supports the previous recommendation of the Groundfish PT ... to hold a workshop to develop guidelines on how to address environmental changes in the SR relationship into biological reference points and how to model environmental forcing in stock projection models.... The SSC believes it would be useful to have members from both the Groundfish and Crab Plan Teams present, because the issues are common to both groups."*

The workshop agenda, a list of modifications to the agenda that occurred during the workshop itself, a list of references, and a list of participants are attached in Appendix A. The workshop initiated discussion of existing and proposed approaches and provided ideas for further analysis of the ten workshop topics:

A. Identification of regime shifts, either for an ecosystem or some subunit thereof
    1. Current policy on identification of regime shifts
    2. Possible improvements to current policy, including consideration of risk
B. Estimation of parameters (average recruitment, stock-recruitment relationships, σ_R)
    1. Establishing criteria for excluding individual within-regime year classes from estimates
    2. Use of "conditioned" stock-recruitment parameters (e.g., $F_{MSY}=F_{35\%}$, $B_{MSY}=B_{35\%}$)
    3. Specification of priors, including hierarchical Bayes and other meta-analytic approaches
    4. Alternatives for setting/estimating σ_R
    5. Determining "reliability" of the F_{MSY} pdf
    6. Other issues involving the stock-recruitment relationship
C. Forecasting environmental variability
    1. Best practices for incorporating environmental forcing in stock assessments
    2. How knowledge of environmental forcing changes perceptions of reference points

Phase I of the working group report was completed just before the May 2012 meeting of the CPT. The Phase I report was created on such a short timeline because guidance on four of the ten workshop topics was deemed essential for the May 2012 CPT meeting to be successful. These four topics were A1, A2, B1, and C2. The Phase I report contained a listing of alternatives for these four topics, qualitative analysis for each of those alternatives and quantitative analysis for some, and a provisional recommendation for each of the four topics. The Phase I report was reviewed by the CPT at its May 2012 meeting and by the SSC at its June 2012 meeting. The CPT and SSC made the following comments on the Phase I report:

From the May 2012 CPT minutes:

> *"The CPT recommended that the default assumption for recruitment is to start with the full time series and use the alternatives listed in A2.2 – A2.6 (or other) to recommend a modification to the default timeframe. The team noted the necessity of consistency across stocks in how the set of recruitments is evaluated, and that all authors should look at several ways to detect breakpoints in productivity. Once a breakpoint has been identified, some plausible biological explanation or rationale should also be provided to support the identified change in productivity. The team stressed the need for transparency in how the breakpoint years are selected when defining reference point, and that the same software should be employed by all authors. The software would include all of the main approaches raised in the report and discussed by the team. André and Steve will pursue software for use by authors prior to the September assessments. The software will include the core methods to be used across all assessments."*

From the June 2012 SSC minutes:

> *"The SSC views the April workshop a great success…. The SSC agrees that the recommendations made in the Phase 1 report should be viewed as preliminary until the report is finalized and it receives review by both the Crab and Groundfish Plan Teams. The SSC notes that environmental forcing need not express itself through regime shifts and urges researchers to also consider environmental events and relationships. The SSC requests thorough documentation of the breakpoint analysis and software, including assumptions and statistical methodology or modeling. The SSC would also like to see some discussion of how workshop recommendations affect determination of virgin (or unfished) biomass. The SSC also suggests that life history, length frequency distribution, and ecosystem considerations could be useful in refining recommendations about analyzing SRRs. The SSC suggested that the Plan Teams should consider life history when selecting the years to exclude from the time series. The SSC anticipates that a deliberative process will be needed to finalize recommendations and so does not expect all recommendations to be implemented until 2013. The SSC looks forward to the final workshop report."*

A Phase II report, completed in August 2012, was originally intended to address all ten of the workshop topics in a similar manner. However, the available time proved insufficient to accomplish this task. Instead, the Phase II report included only a slightly modified version of the Phase I report (for example, a revision of one of the alternatives under topic B2, intended to address the SSC concern about considering life history when selecting years to exclude from the time series) and a listing of alternatives with provisional recommendations—but no analysis—for the six topics not covered in the Phase I report.

The GPTs and SSC reviewed the Phase II report at their September and October 2012 meetings, respectively. The SSC did not comment on the Phase II report. The GPTs had the following comment:

> *"For topic C1 (best practices for including environmental forcing in assessments), the provisional recommendation specifies the use of log-linear models because: 1) this is a mathematically convenient functional form; 2) this is the functional form that is typically used in such analyses; and 3) the 'true' functional forms underlying the relationships between environmental variables and recruitment have not been identified for any BSAI or GOA groundfish stocks. One Team member suggested modifying this alternative, or adding a new*

> *alternative, in which the assessment authors first attempt to determine the 'true' functional forms, then use log-linear only in those cases where the attempt is unsuccessful."*

The BSAI GPT included the following recommendation in its December 2012 minutes:

> *"The Team recommends that the Recruitment Working Group examine use of median recruitment (or other measure(s) of central tendency) as an alternative to mean recruitment for calculation of reference points."*

In 2013, SSC members Anne Hollowed and Farron Wallace, CPT members André Punt and Buck Stockhausen, and University of Washington graduate student Cody Szuwalski were added to the working group. With this expanded membership, work began on preparing the present (Phase III) report in February. One of the working group's main activities this year was to sponsor, along with CSIRO (Australia) and the US–ROK Joint Agreement Fisheries Panel, a "Workshop on Setting Biological Reference Points in a Changing Climate" (agenda shown in Appendix B). The other main activity was to complete this Phase III report in time for consideration by the Teams and SSC during their respective September and October meetings. The GPT comments from September and December 2012 are addressed in this report. In particular, topic B6 has been renamed, a new topic B7 has been added, and one of the alternatives under topic C1 has been revised. All alternatives for all topics now include at least a qualitative analysis, two appendices containing quantitative analyses of topics B1 and B7 have been added, and all "provisional recommendations" have been replaced by "recommendations." All recommendations made here are, of course, subject to review by the Teams and SSC. Note that some of the alternatives have been re-ordered to improve readability.

Topics and Alternatives

In the following, "SRR" stands for "stock-recruitment relationship."

A1: Current policy on identification of regime shifts

Alternative A1.1 (status quo):

For groundfish, the status quo approach is contained in a 1999 memorandum from James Balsiger (who was at that time AFSC Director) to the AFSC groundfish stock assessment authors, and consists of the following two sentences: *"Projections of future stock sizes and estimation of reference points should be based only on year classes spawned in 1977 or later, unless a compelling case can be made to begin the time series in some other year. The fact that earlier estimates are available does not in itself constitute a compelling case."*

For crab, the status quo approach is described in various parts of the policy listed in Appendix C. Briefly, this approach calls for identification of potential mechanisms to support regime shifts. Such identification should consider evidence of a change in magnitude and direction of life-history characteristics. Candidate life-history characteristics include natural mortality, growth, maturity, fecundity, recruitment, and recruits per unit of spawning. Candidate ecosystem characteristics include the "Overland method" of regime shift detection, change in production of benthic species in the Eastern Bering Sea, and consumption (from ecosystem model outputs). If stock-recruitment data are available, they are to be examined for evidence of multiple SRRs that are consistent with a proposed regime shift.

Because topic A1 is restricted to the status quo by definition, no other alternatives are presented for this topic. Also, because the status quo is a matter of fact, no recommendation is made for this topic.

A2: Possible improvements to current policy, including consideration of risk

Alternative A2.1: Do not consider effects of regime shifts.

Pro: 1) Extremely easy to implement. 2) Minimizes chance of a "false positive" regime shift identification. 3) If the regimes that occurred during the period spanned by the full time series of data constitute a random sample from the distribution of regimes that will occur in the long-term future, this method would give an unbiased estimate of future conditions over the long term.

Con: 1) Maximizes chance of a "false negative" regime shift (non)identification. 2) Given that regimes (almost by definition) persist for a period of at least several years, this method is likely to give a biased estimate of future conditions over the short term. 3) Because environmental regimes typically appear to persist over approximately decadal time scales and because most datasets for BSAI and GOA groundfish and crab typically extend back only a few decades, it is unlikely that the set of regimes that occurred during the period spanned by the data constitutes a random sample from the distribution of regimes that will occur in the long-term future; in which case this method is also likely to give a biased estimate of future conditions over the long term.

Alternative A2.2: Estimate breakpoints in the time series of recruits using an appropriate statistical test such as STARS, AIC_c, or likelihood ratio, and possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.

Pro: 1) Basing the analysis on the time series of recruits, without considering recruits per unit of spawning or a curvilinear SRR, is similar to existing practice for Tier 3 groundfish. 2) If the true SRR is of Beverton-Holt (or similar, asymptotic) form and spawning biomass has been sufficiently high throughout the time series (such that the recruitment predicted by the curve is almost independent of spawning biomass), this method will likely produce results similar to those that would be produced by the more complicated alternative of considering a fully parameterized SRR.

Con: 1) If spawning biomass has been sufficiently low for the most recent part of the time series, low recruitments from those recent years will be mistaken for a new regime even though the true SRR has not changed. 2) Because this method implicitly assumes that the true SRR is approximately horizontal across the observed range of spawning biomasses, productivity will be overestimated if the assumption is extrapolated all the way down to the origin.

Alternative A2.3: Estimate breakpoints in the time series of recruits per unit of spawning using an appropriate statistical test such as STARS, AIC_C, or likelihood ratio, and possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.

Pro: 1) Avoids the problems identified under "Con" for Alternative A2.2. 2) If spawning biomass has been severely depleted throughout the time series (such that spawning biomass is always close to zero), this method will likely produce results similar to those that would be produced by the more complicated alternative of considering a fully parameterized SRR.

Con: 1) If the true SRR is of Beverton-Holt (or similar, asymptotic) form and spawning biomass has been sufficiently high throughout the time series (such that the recruitment predicted by the curve is almost independent of spawning biomass) but spawning biomass has declined significantly during the most recent part of the time series, recent decreases in recruits per unit of spawning will be mistaken for a new regime even though the true SRR has not changed. 2) Because this method implicitly assumes that the true relationship between recruits and spawning is proportional across the observed range of spawning biomasses, productivity will be underestimated if the assumption is extrapolated far beyond the range of the data.

Alternative A2.4: Estimate breakpoints in the time series of an environmental time series such as the Pacific Decadal Oscillation (PDO) using an appropriate statistical test such as STARS, AIC_C, or likelihood ratio, and possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.

Pro: 1) The necessary data may be available even when recruitment data are not. 2) Breakpoints in environmental time series such as the PDO have already been well studied and shown to be significant predictors of many things. 3) This approach would eliminate the need to conduct a separate analysis for every stock.

Con: 1) If the productivity of a particular stock is not linked, directly or indirectly, to the environmental variable(s) used in the analysis, a "false positive" regime shift identification will result. 2) If the productivity of a stock changes only in response to some variable *not* used in the analysis, a "false negative" regime shift (non)identification will result.

Alternative A2.5: Estimate both parameters of a two-parameter SRR for every age- or length-structured stock assessment, with breakpoints estimated using an appropriate statistical test such as STARS, AIC_C, or likelihood ratio, and possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.

Pro: 1) Eliminates the need to use proxy reference points. 2) Does not imply functional forms for the SRR (e.g., horizontal or linear through the origin) that are almost certain to be implausible if extrapolated across the entire range of possible spawning biomasses.

Con: 1) Reliably estimating both parameters of a two-parameter SRR has proven to be very difficult for the vast majority of BSAI and GOA groundfish and crab stocks. 2) May result in imprecise estimates of breakpoints, or an inability to detect a breakpoint when one exists. 3) Different choices for the SRR may lead to different breakpoints.

*Alternative A2.6 (***recommendation; see also "Option" below***): Condition the productivity parameter of a two-parameter SRR on one or more F_{MSY} proxies specified or implied by the harvest control rules in the respective FMP, then estimate the scale parameter of the SRR for every age- or length-structured stock assessment, with breakpoints estimated using an appropriate statistical test such as STARS, AIC_C, or likelihood ratio, and possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.*

Pro: 1) Results in management recommendations that are consistent with existing F_{MSY} proxies. 2) Does not imply functional forms for the SRR (e.g., horizontal or linear through the origin) that are almost certain to be implausible if extrapolated across the entire range of possible spawning biomasses. 3) Eliminates the need to estimate the more difficult-to-estimate of the two SRR parameters, instead requiring estimation of only the scale parameter, which is analogous to the "average recruitment" currently estimated in all Tier 3 groundfish assessments. 4) This approach has been tested on 11 BSAI and GOA groundfish stocks using a very simple model, and the results appear to be reasonable wherever the assumptions are not violated too severely (6 of the 11 stocks were shown to have breakpoints that passed five statistical tests of significance, with the starting years of the current regimes for these 6 stocks ranging from 1968 to 1990). It has also been tested on 8 BSAI crab stocks and found to be reasonable for most of those stocks.

Con: 1) Requires use of F_{MSY} proxies. 2) Estimates of derived quantities such as B_{MSY} can be implausible if the F_{MSY} proxies are inconsistent with the data (however, this approach is intended only to estimate the *breakpoints*; estimates of other quantities obtained in the process of determining the breakpoints do not have to be used for management purposes).

*Option for any of the above except A2.1 (***recommendation; see also Alternative A2.6** *): Use a decision-theoretic approach to compute the optimal breakpoints, possibly employing additional constraints such as a minimum length for the current regime or a maximum permissible CV for parameter estimates.*

Pro: 1) Costs of mis-estimating a breakpoint are weighted appropriately.
Con: 1) Requires specification of a loss (cost) function. 2) More complicated than an approach that does not weight the costs of mis-estimating a breakpoint appropriately.

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

A simple but quantitative evaluation of the alternatives listed here is contained in Appendix D.

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).

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***): Defining* $A_{10\%}$ *as the first age with a survey selectivity of at least 10% and* X *as floor(1/(1–exp(–sqrt(*M*)))), for any species with a lifespan greater than* $A_{10\%}+1$ *years, exclude all year classes spawned within the last* $A_{10\%}+X$ *years.*

Pro: 1) Extremely easy to implement. 2) Always feasible, unless $A_{10\%}+X$ is higher than the largest age in the model.
Con: 2) No necessary relationship to precision of estimated year class strengths.

B2: Use of "conditioned" stock-recruitment parameters (e.g., $F_{MSY}=F_{35\%}$, $B_{MSY}=B_{35\%}$)

The following alternatives apply to Tier 3 stocks only. Appendix E contains some analyses relevant to this topic.

Alternative B2.1: Do not use conditioned stock-recruitment parameters.

Pro: 1) If successful, eliminates the need for proxy reference points, which may be inaccurate in any given instance.
Con: 1) Reliably estimating both parameters of a two-parameter SRR has proven to be very difficult for the vast majority of BSAI and GOA groundfish and crab stocks. 2) May result in imprecise (or implausible) estimates of F_{MSY} and B_{MSY}.

Alternative B2.2: Condition the SRR by forcing $F_{MSY}=F_{35\%}$ *and* $B_{MSY}=B_{35\%}$.

Pro: 1) Reduces dependency on difficult-to-estimate SRR parameters.
Con: 1) Requires use of B_{MSY} and F_{MSY} proxies, which may be inaccurate in any given instance. 2) Estimates of B_{MSY} and F_{MSY} may be inconsistent with the stock and recruitment data.

*Alternative B2.3 (***recommendation***): Condition the SRR by forcing* $F_{MSY}=F_{35\%}$*, but estimate* B_{MSY} *as a free parameter.*

Pro: 1) Reduces dependency on difficult-to-estimate SRR parameters. 2) For a two-parameter SRR, allows estimation of the easier-to-estimate parameter. 3) Imposes fewer constraints than Alternative B2.2.
Con: 1) Requires use of F_{MSY} proxies, which may be inaccurate in any given instance.

B3: Specification of priors, including hierarchical Bayes and other meta-analytic approaches

Alternative B3.1: Use non-constraining uniform priors only.

Pro: 1) Simple to implement.
Con: 1) May provide unrealistic estimates due lack of information in the data (e.g., "one-way trips," short time series, etc.).

Alternative B3.2: Use priors derived from hierarchical Bayes analysis of congeneric stocks.

Pro: 1) Already in use for some stocks (not necessarily in the BSAI or GOA).
Con: 1) May not reflect the true amount of prior uncertainty (e.g., if the sample of congeneric stocks is not representative of the stock in question).

*Alternative B3.3 (***recommendation***): Use priors that reflect the true amount of prior uncertainty.*

Pro: 1) Broadens the source of information about the potential SRR. 2) Likely to make estimation more tractable. 3) Likely to stabilize results and be more reliable. 4) If the prior distribution is to be used in a Bayesian sense, then this is the appropriate procedure. 5) Can incorporate Alternative B3.2 when appropriate.

Con: 1) Eliciting such prior distributions in a rigorous manner is difficult. 2) Absent formal elicitation, such priors are unlikely to be available.

B4: Alternatives for setting/estimating σ_R

Alternative B4.1: Set $\sigma_R = 0.6$.

Pro: 1) Extremely easy to implement. 2) Several precedents. 3) Seems to be a good estimate of a “typical” value for groundfish.

Con: 1) May fail to reflect true recruitment variability. 2) Does not allow an estimate of uncertainty in σ_R to be obtained.

Alternative B4.2: Estimate σ_R *iteratively.*

Pro: 1) Relatively easy to implement.

Con: 1) Will likely underestimate the true value, unless recruitment deviations are treated as random effects and integrated out of the likelihood. 2) Does not allow an estimate of uncertainty in σ_R to be obtained.

Alternative B4.3: Estimate σ_R *as a free parameter.*

Pro: 1) Relatively easy to implement. 2) Allows an estimate of uncertainty in σ_R to be obtained.

Con: 1) In some models, this approach is equivalent to Alternative B4.2, and so will have the problem as Alternative B4.2. 2) Even in cases where this approach gives a different result than Alternative B4.2, the resulting estimate is likely to be biased unless recruitment deviations are treated as random effects and integrated out of the likelihood.

*Alternative B4.4 (***recommendation***): Estimate* σ_R *according to the method presented at the 2012 recruitment workshop.*

This method consisted of the following three steps: 1) Estimate recruitment deviations when σ_R is set, provisionally, at a high (i.e., non-constraining value); label this vector **r**. 2) Estimate σ_R iteratively by matching the standard deviations of the estimated recruitment deviations; label this σ. 3) Obtain a final estimate of σ_R as sqrt(var(**r**) − σ·(stdev(**r**)−σ)). See Annex 2.1.1 of the 2012 BSAI Pacific cod assessment (p. 442-445).

Pro: 1) This algorithm gives the MLE for a linear random effects model with normal errors. 2) It may therefore be likely to be less biased than other methods that (when applied to a more general model) do not involve integrating random effects out of the likelihood. 3) Shown to be a good estimator for a variety of age-structured modeling scenarios (Hui-hua Lee et al., unpubl. manuscr., available from the senior author at the University of Hawaii’s Joint Institute for Marine and Atmospheric Research).

Con: 1) Requires more labor-intensive computations than other alternatives, particularly if multiple models are being developed. 2) Does not allow an estimate of uncertainty in σ_R to be obtained.

B5: Determining "reliability" of the F_{MSY} pdf

Alternative B5.1: Determine that the F_{MSY} *pdf is reliable if no parameter has an estimated standard deviation (obtained by inverting the Hessian matrix) greater than* X *or a CV greater than* Y *(values of* X *and* Y *to be determined).*

Pro: 1) This alternative provides measurable criteria by which reliability may be determined, and guards against acceptance of pdfs that imply a large degree of uncertainty.

Con: 1) This alternative reflects a fundamental misunderstanding of the role played by uncertainty in the Tier 1 control rules, which is to adjust the buffer between ABC and OFL by an amount appropriate to *whatever* degree of uncertainty exists; not to limit this adjustment to some pre-determined amount.

Alternative B5.2: Determine that the F_{MSY} *pdf is reliable if the Hessian matrix is positive definite.*

Pro: 1) This appears to be the *de facto* criterion for acceptance of many Tier 3 assessments, as goodness of fit or diagnostic statistics are often omitted from these assessments, so it should be good enough for Tier 1 assessments also.

Con: 1) This criterion, by itself, is far too weak to merit acceptance of the resulting F_{MSY} pdf.

*Alternative B5.3 (***recommendation***): Determine that the* F_{MSY} *pdf is reliable if: 1) the Hessian matrix is positive definite; 2) the average ratio of harmonic mean multinomial effective sample size to arithmetic mean multinomial input sample size exceeds unity for all size composition and age composition likelihood components; 3) the mean standardized log-scale residual for each survey abundance likelihood component is between –0.1 and 0.1; 4) the root-mean-squared standardized log-scale residual for each survey abundance likelihood component is between 0.9 and 1.1; 5) the assessment demonstrates that annual variability in selectivity at age (or length) and weight at age (or length) was considered during the process of model development, using either internal or external estimation of variability; and 6) the assessment demonstrates that sensitivity to alternative starting values for the parameters was examined before accepting the results from the model.*

Pro: 1) This alternative improves upon Alternative B5.2 by adding criteria for all of the major likelihood components typically included in age- or size-structured models of stocks managed by the NPFMC, guards against underestimation of uncertainty by requiring examination of time variation in selectivity at age/length and weight at age/length, and guards against acceptance of a model that has converged at a local minimum/maximum.

Con: 1) These criteria may result in downgrading one or more of the existing Tier 1 assessments, which may cause confusion. 2) Prescribing one-size-fits-all criteria such as these limits the ability of the SSC to consider contingencies that may arise in specific situations.

B6: Alternatives to estimation of SRR parameters

Alternative B6.1: Discontinue research into alternative assessment and management methods that are robust to lack of information about SRR parameters so as to free up more resources for research on estimation of SRR parameters.

Pro: 1) Estimation of SRR parameters is foundational to fishery science and management. 2) Much progress has been made in recent years, as evidenced by the recent upgrading of the BSAI yellowfin sole and northern rock sole assessments from Tier 3 to Tier 1. 3) Use of existing proxies provides a sufficiently reliable alternative for those cases where estimation of SRR parameters remains elusive.

Con: 1) Some recent publications (e.g., Haltuch et al., 2008, *Fish. Res*. 94:290-303; Conn et al., 2010, *CJFAS* 67:511-523; and Lee et al., 2012, *Fish. Res.* 125-126:254-261) suggest that the difficulty of accurately estimating SRR parameters is fundamentally difficult and may simply be impossible in many cases. 2) Existing "one-size-fits-all" proxies may be inaccurate in particular instances.

*Alternative B6.2 (***recommendation***): Continue trying to estimate SRR parameters whenever possible, but also continue research into alternative assessment and management methods that are robust to lack of information about these parameters.*

Pro: 1) Surplus production models can be used, either in place of age-structured models, or to condition the estimates of the SRR parameters in those models (e.g., by providing a prior distribution for the ratio of MSY to B_{MSY}). 2) The "survey/exploitation vector autoregressive" (SEVAR) model, which debuted this summer at the World Conference on Stock Assessment Methods, provides another alternative that is worth pursuing. 3) This alternative fits well with Management Strategy Evaluations, where alternative plausible SRR relationships (stationary and non-stationary) can be posed within the operating model and tested using simpler management measures, including assessment models and "management procedures" (in the sense that the term is used by the International Whaling Commission and elsewhere).

Con: 1) Surplus production models tend to underestimate uncertainty in estimates of parameters and may be badly biased (e.g, Punt and Szuwalski, 2013, *Fish. Res*. 134-136: 82-94). 2) Because the SEVAR model is very new (and so far unpublished), it remains to be seen whether this approach provides a reliable alternative to estimation of SRR parameters.

B7: Preferred measure of central tendency in recruitment

Appendix F contains an analysis relevant to this topic.

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 fishing effort and catch. 2) Compared to use of mean recruitment, use of the median results in higher average catch

Con: 1) Compared to use of mean recruitment, use of the median results in longer rebuilding times. 2) Compared to use of mean recruitment, use of the median results in lower average biomass (and thus CPUE). 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, 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 shorter rebuilding times. 2) Compared to use of median recruitment, use of the mean results in higher average biomass (and thus CPUE). 3) Estimates of Tier 3 reference points have been based on mean recruitment for well over a decade, with apparently general agreement that this is appropriate.

Con: 1) Compared to use of median recruitment, use of the mean results in greater variability in fishing effort and catch. 2) Compared to use of median recruitment, use of the mean results in lower average catch.

C1: Best practices for incorporating environmental forcing in stock assessments

Alternative C1.1: Do not incorporate environmental forcing in stock assessments.

Pro: 1) Easy to implement.
Con: 1) May miss important processes.

Alternative C1.2: Identify plausible environmental covariates of recruitment outside of the assessment model; then include them (adjusted for sign, as appropriate) as pseudo-surveys of recruitment in the assessment model.

Pro: 1) May include important processes that affect model results appropriately. 2) This is an established procedure for west coast groundfish.
Con: 1) May add complexity to model specifications. 2) Limited ability to know when to stop including covariates. 3) Some "plausible" covariates may actually be unimportant.

*Alternative C1.3 (***recommendation***): Identify plausible environmental covariates of recruitment outside of the assessment model; then include them as explanatory variables in the SRR, with parameters estimated inside the assessment model.*

Pro: 1) May include important processes that affect model results appropriately. 2) This approach is established in the literature (e.g., Wilderbuer et al., 2002, *Prog. Oceanogr.* 55:235-247), and is consistent with existing approaches for addressing environmental forcing in some BSAI groundfish assessments (e.g,. the relationship between temperature and catchability in the yellowfin sole assessment).
Con: 1) May add complexity to model specifications. 2) Limited ability to know when to stop including covariates. 3) Some "plausible" covariates may actually be unimportant.

C2: How knowledge of environmental forcing changes perceptions of reference points

Alternative C2.1: Use knowledge of environmental forcing to compare past, present, and projected stock sizes with past, present, and future values of environmentally forced reference points.

Pro: 1) Keeps BSAI and GOA groundfish and crab on the cutting edge of fishery science and management. 2) Avoids comparing apples and oranges in terms of stock status and reference points (i.e., for any given year, stock size would be compared to the reference point applicable to that year, as determined by the relevant past, present, or future values of the relevant environmental variables).
Con: 1) Extremely difficult to implement anytime in the near future. 2) Criteria used to make status determinations and to measure rebuilding will be moving targets, even for a fixed set of biological data. 3) Simulation studies have generally not found that this leads to better achievement of management goals (e.g., Punt et al., in press, *ICES J. Mar Sci.*).

*Alternative C2.2 (***recommendation***): Acknowledge that current knowledge of environmental forcing is insufficient to alter perceptions of reference points quantitatively.*

Pro: 1) Extremely easy to implement. 2) Probably an accurate description of the current state of knowledge for the vast majority (if not all) BSAI and GOA groundfish and crab stocks.
Con: 1) Does not advance the state of the art.

Appendix A: The April 2012 Workshop on Assessment/Management Issues Related to Recruitment

Agenda

	Wednesday, April 4	Speakers
0900	Welcome, purpose of workshop, introductions, appointment of rapporteurs	
A. Identification of regime shifts, either for an ecosystem or some subunit thereof		
1. Current policy on identification of regime shifts*		
0920	Estimating B_{MSY} for Tier 4 crab stocks and recruitment for Tier 3 crab stocks: Which years are representative?	B. Foy, D. Stram
0945	Jim Balsiger's memo of September 1999	Grant Thompson
0950	Discussion	
1010	- Break -	
2. Possible improvements to current policy, including consideration of risk*		
1020	A null hypothesis to explain regime-like transitions in ecosystem time series	Emanuele Di Lorenzo
1045	Considerations of biological factors affecting potential crab production regimes	L. Rugolo, J. Turnoc
1110	Identification and management of stocks with regime-based recruitment	Cody Szuwalski
1135	Risk-based selection of regime boundaries for a stock managed under a sloping, SPR-based control rule	Grant Thompson
1200	Discussion	
1220	- Lunch -	
B. Estimation of parameters (average recruitment, stock-recruitment relationships, σ_R)		
1. Establishing criteria for excluding individual within-regime year classes from estimates*		
1320	Criteria for excluding individual within-regime year classes from estimates: current practice for EBS pollock	Jim Ianelli
1345	Accounting for uncertainty in estimated recruitment when computing stock status reference points: an example from the 2010 BSAI blackspotted/rougheye rockfish assessment	Paul Spencer
1410	Choice of recruitment periods for OFL determination and its impacts on Bristol Bay red king crab	Jie Zheng
1435	Discussion	
1455	Break	
2. Use of "conditioned" stock-recruitment parameters (e.g., FMSY=F35%, BMSY=B35%)		
1505	Deriving steepness from F_{MSY} or F_{spr}	Steve Martell
1530	Discussion	
3. Specification of priors, including hierarchical Bayes and other meta-analytic approaches		
1550	Use of stock-recruit steepness priors based on meta-analysis in West Coast rockfish assessments	Martin Dorn
1615	Preliminary results for developing Bayesian priors for relative cohort strength of groundfishes off the U.S. West Coast using multi-species Stock Synthesis models	Jim Thorson
1640	Discussion	
1700	- Adjourn for the day -	

* Critical items for May 2012 Crab Plan Team meeting

Thursday April 5th		
B. Estimation of parameters, continued		
	4. Alternatives for setting/estimating σ_R	
0900	Problems associated with estimating recruitment and σ_R in a random effects model	G. Thompson
0925	Discussion	
	5. Determining "reliability" of the F_{MSY} pdf	
0945	Environmental factors affecting EBS pollock S-R relationships	Jim Ianelli
1010	Discussion	
1030	- Break -	
	6. Other issues involving the stock-recruitment relationship	
1040	Improving ecological validity and linkage among spawner recruitment, mortality, age structure, and harvesting models: An example from western rock lobster fishery neutrality harvesting model	Yuk W. Cheng
1105	Comprehensive analysis of the stock-recruitment relationship and reference points	Mark Maunder
1130	A new paradigm for stock-recruitment relationships: Viewing the stock-recruitment relationship as density dependent survival invalidates the Beverton-Holt and Ricker models	Mark Maunder
1155	Discussion	
1215	- Lunch -	
C. Forecasting environmental variability		
	1. Best practices for incorporating environmental forcing in stock assessments	
1315	Advice for estimating fishery management reference points given low frequency between-year environmental variability	Melissa Haltuch
1340	Multispecies modeling, including projections and effects of temperature variability and predators on mortality estimates	Kirstin Holsman
1405	Environmental forcing of recruitment in the Bering Sea and Gulf of Alaska and its use in stock assessments and stock projections	Franz Mueter
1430	Recruitment products and indices from FOCI and BASIS – new proposed products for the Plan Teams and SSC	Jeff Napp
1455	Discussion	
1515	- Break -	
	2. How knowledge of environmental forcing changes perceptions of reference points*	
1525	F_{msy} and B_{msy} proxies by regime	Jim Ianelli
1550	Discussion	
1610	Wrap-up	
1630	- Adjourn -	

* Critical items for May 2012 Crab Plan Team meeting

Modifications to the Agenda

1. Lou Rugulo and Jack Turnock's presentation under item A2 was withdrawn.
2. Unscheduled presentation by Andre Punt on use of surplus production models to estimate B_{MSY} in crab stocks was added in place of Rugulo and Turnock's presentation under A2.
3. Martin Dorn's presentation under item B3 was withdrawn.
4. Unscheduled presentation by Kerim Aydin on a multispecies model with an "emergent" stock-recruitment relationship was added under item C1.
5. Jim Ianelli's presentation under item C2 was withdrawn.

References

A'mar, T. Z., A. E. Punt, and M. W. Dorn. 2009. The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (Theragra chalcogramma) fishery. Canadian Journal of Fisheries and Aquatic Sciences 66:2222-2242.

Arregui, I., H. Arrizabalaga, D. S. Kirby, and J. M. Martin-Gonzalez. 2006. Stock–environment–recruitment models for North Atlantic albacore (*Thunnus alalunga*). Fish. Oceanogr. 15:5, 402–412.

Borja, A., Uriarte, A., Egana, J., Motos, L., and Valencia, V. 1998. Relationship between anchovy (Engraulis encrasicolus) recruitment and environment in the Bay of Biscay (1967-1996).

Brander, K. M. 2005. Cod recruitment is strongly affected by climate when stock biomass is low. ICES Journal of Marine Science, 62: 339-343.

Brooks, E. N., and J. E. Powers. 2007. Generalized compensation in stock-recruit functions: properties and implications for management. ICES Journal of Marine Science: Journal du Conseil 64(3):413-424.

Brooks, E. N., J. E. Powers, and E. Cortés. 2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES Journal of Marine Science: Journal du Conseil 67(1):165-175.

Chen, D. G.. 2001. Detecting environmental regimes in fish stock– recruitment relationships by fuzzy logic. Can. J. Fish. Aquat. Sci. 58: 2139–2148.

Chen, D. G., and D. M. Ware. 1999. A neural network model for forecasting fish stock recruitment. Can. J. Fish. Aquat. Sci. 56: 2385–2396.

Conn, P. B., E. H. Williams, and K. W. Shertzer. 2010. When can we reliably estimate the productivity of fish stocks? Canadian Journal of Fisheries and Aquatic Sciences 67(3):511-523.

Deriso, R.B., M. N. Maunder, and W. H. Pearson. 2008. Incorporating covariates into fisheries stock assessment models with application to Pacific herring of Prince William Sound, Alaska. Ecological Applications 18(5): 1270-1286.

Deriso, R.B., M. N. Maunder, and J. R. Skalski. 2007. Variance estimation in integrated assessment models and its importance for hypothesis testing. Can. J. Fish. Aquat. Sci. 64: 187-197.

Dorn, M. W. 1995. The effects of age composition and oceanographic conditions on the annual migration of Pacific whiting, *Merluccius productus*. CalCOFI Rep., Vol. 36, 1995.

Forrest, R. E., M. K. McAllister, M. W. Dorn, S. J. D. Martell, and R. D. Stanley. 2010. Hierarchical Bayesian estimation of recruitment parameters and reference points for Pacific rockfishes (Sebastes spp.) under alternative assumptions about the stockrecruit function. Canadian Journal of Fisheries and Aquatic Sciences 67(10):1611-1634.

Francis, R. I. C. C. 2006. Measuring the strength of environment–recruitment relationships: the importance of including predictor screening within cross-validations. ICES Journal of Marine Science: Journal du Conseil 63(4):594-599.

Haltuch, M. A., A. E. Punt, M. W. Dorn. 2009. Evaluating the estimation of fishery management reference points in a variable environment. Fisheries Research 100(42-56).

Haltuch, M. A., and A. E. Punt. 2011. The promises and pitfalls of including decadal-scale climate forcing of recruitment in groundfish stock assessment. Canadian Journal of Fisheries and Aquatic Sciences 68(5):912-926.

Haltuch, M. A., A. E. Punt, and M. W. Dorn. 2008. Evaluating alternative estimators of fishery management reference points. Fisheries Research 94(3):290-303.

Holt, C. A., and A. E. Punt. 2009. Incorporating climate information into rebuilding plans for overfished groundfish species of the U.S. west coast. Fisheries Research 100:57-67.

Humston, R., J. S. Ault, M. Lutcavage, and D. B. Olson. 2000. Schooling and migration of large pelagic fishes relative to environmental cues. Fish. Oceanogr. 9:2, 136-146.

Lee, H-H., Maunder, M.N., Piner, K.R., and Methot, R.D. (in press) Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fisheries Research. (http://www.sciencedirect.com/science/article/pii/S0165783612001099?v=s5)

Litzow, M. A. 2006. Climate regime shifts and community reorganization in the Gulf of Alaska: how do recent shifts compare with 1976/1977? ICES Journal of Marine Science, 63:1386-1396.

Maunder, M.N. 1998. Problems with using an environmental based recruitment index: examples from a New Zealand snapper (*Pagrus auratus*) assessment. In fishery stock assessment models, edited by F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.J. Schweigert, P.J. Sullivan, and C.I. Zhang, Alaska Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks, pp. 679-692.

Maunder, M.N. 2004. Population Viability Analysis, Based on Combining Integrated, Bayesian, and Hierarchical Analyses. Acta Oecologica 26: 85-94. Special issue for the Extinction Working Group of the National Center for Ecological Synthesis and Analysis.

Maunder, M.N. 2012. Evaluating the stock-recruitment relationship and management reference points: Application to summer flounder (Paralichthys dentatus) in the U.S. mid-Atlantic. Fisheries Research 125-126: 20-26. (http://www.sciencedirect.com/science/article/pii/S0165783612000811)

Maunder, M. N., and R. B. Deriso. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences 60:1204-1216.

Maunder, M. N. and R. B. Deriso. 2010. Dealing with missing covariate data in fishery stock assessment models. Fisheries Research 101: 80-86.

Maunder, M. N. and R. B. Deriso. 2011. A state–space multistage life cycle model to evaluate population impacts in the presence of density dependence: illustrated with application to delta smelt (*Hyposmesus transpacificus*). Can. J. Fish. Aquat. Sci. 68: 1285–1306.

Maunder, M. N., S. J. Harley, and J. Hampton. 2006. Including parameter uncertainty in forward projections of computationally intensive statistical population dynamic models. ICES Journal of Marine Science, 63:969-979.

Maunder, M.N. and Starr, P.J. 1998. Validating the Hauraki Gulf snapper pre-recruit trawl surveys and temperature recruitment relationship using catch at age analysis with auxiliary information. New Zealand Fisheries Assessment Research Document 98/15 23p.

Maunder, M. N., and G. M. Watters. 2003. A general framework for integrating environmental time series into stock assessment models: model description, simulation testing, and example. Fishery Bulletin 101: 89-99.

Methot, R. D., and I. G. Taylor. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68:1744-1760.

Myers, R.A. 1998. When do environment–recruitment correlations work? Rev. Fish Biol. Fish. 8(3): 285–305. doi:10.1023/ A:1008828730759.

Punt, A. E. 2011. The impact of climate change on the performance of rebuilding strategies for overfished groundfish species on the U.S. west coast. Fish and Fisheries 109(2-3):320-329.

Schirripa, M. J., C. P. Goodyear, and R. D. Methot. 2009. Testing different methods of incorporating climate data into the assessment of US West Coast sablefish. ICES Journal of Marine Science 66:1605-1613.

Sinclair, F., and W. R. Crawford. 2005. Incorporating an environmental stock–recruitment relationship in the assessment of Pacific cod (Gadus macrocephalus) Fish. Oceanogr. 14:2, 138–150.

Walters, C. J., R. Hilborn, and V. Christensen. 2008. Surplus production dynamics in declining and recovering fish populations. Canadian Journal of Fisheries and Aquatic Sciences 65(11):2536-2551.

Williams, E. H., and K. W. Shertzer. 2003. Implications of life-history invariants for biological reference points used in fishery management. Canadian Journal of Fisheries and Aquatic Sciences 60(6):710-720.

Zhu, J-F, Chen, Y., Dai, X.J., Harley, S.J., Hoyle, S.D., Maunder, M.N., Aires-da-Silva, A. 2012. Implications of uncertainty in the spawner-recruitment relationship for fisheries management: an illustration using bigeye tuna (Thunnus obesus) in the eastern Pacific Ocean. Fisheries Research 119–120:89–93. (http://www.sciencedirect.com/science/article/pii/S0165783611003869)

Participants

Teresa A'Mar	AFSC
Kerim Aydin	AFSC
Matt Baker	AFSC
Steve Barbeaux	AFSC
Bill Bechtol	UAF
Yuk Cheng	WDFW
Bob Clark	ADFG
Jane DiCosimo	NPFMC
Martin Dorn	AFSC
Bob Foy	AFSC
Scott Goodman	NRC/BSFRF
Melissa Haltuch	NWFSC
Toshida Hamazaki	ADFG
Dana Hanselman	AFSC
Jim Hastie	NWFSC
Alan Haynie	AFSC
Jon Heifetz	AFSC
Anne Hollowed	AFSC
Kirstin Holsman	AFSC/JISAO, UW
Pete Hulson	AFSC
Jim Ianelli	AFSC
Patricia Livingston	AFSC
Sandra Lowe	AFSC
Steve Martell	UBC
Franz Mueter	UAF
Jeffrey Napp	AFSC
Ivonne Ortiz	AFSC
Tom Pearson	AKRO
Andre Punt	UW
Terry Quinn	UAF
Lou Rugulo	AFSC
Kalei Shotwell	AFSC
Shareef Siddeek	ADF&G
Dave Somerton	AFSC
Paul Spencer	AFSC
Gary Stauffer	BSFRF
Diana Stram	NPFMC
Cody Szuwaski	UW
Ian Taylor	NWFSC
Grant Thompson	AFSC
James Thorson	UW
Jack Turnock	AFSC
Tadeyasu Uchiyama	UAF
Juan Valero	IPHC
Tom Wilderbuer	AFSC
Stephani Zador	AFSC
Jie Zheng	ADF&G

Appendix B: NMFS/CSIRO Workshop Agenda

Workshop on Setting Biological Reference Points in a Changing Climate
Sponsored by the NOAA Fisheries, CSIRO and the US – ROK Joint Agreement Fisheries Panel
Seattle, Washington, USA
Oceanographer Room Building 3

Agenda and Order of the Day
13-15 August 2013

Terms of Reference

a) Mechanisms underlying changes in production and analytical methods for detecting.
b) Identify challenges in separating the impacts of change to climate from fishing and natural variability.
c) Develop methods for quantifying uncertainty in projected changes.
d) Selecting harvest strategies that are robust to, or adjust appropriately in response to, climate change induced shifts in productivity.
e) Defining the precautionary approach under a changing climate.
f) Participants will contribute to papers addressing each TOR that will be submitted to a peer reviewed journal.

Agenda Day 1 Tuesday

09:00 Welcome, introductions.
09:15 Overview of goals and objectives of workshop

Session 1: Mechanisms underlying changes in production and analytical methods for detecting them.

09:30 **Nicholas Bond:** *What are climate models projecting for the Pacific Ocean?*
09:50 **Tim Essington:** *Using hierarchical approaches to explore synchrony in production dynamics and linkages to environmental drivers.*
10:10 **Teresa A'mar**: *Revisiting mechanisms underlying responses to climate variability in Gulf of Alaska walleye pollock*.
10:30 **Sukyung Kang:** *The impact of climate change on the distribution and abundance of mackerels in the Northwestern Pacific.*
11:10 **Sally Wayte:** A climate induced recruitment shift in Jackass Morwong.
11:30 Session 1: Discussion

Session 2: Challenges in separating the impacts of climate change from fishing and natural variability.

14:00 **Katyana Vert-pre:** *Frequency and intensity of productivity regime shifts in marine stocks (**by Web-ex**)*
14:20 **Cody Szuwalski:** *Examining common assumptions about recruitment using the RAM legacy stock assessment database.*
14:50 **Neil Klaer and <u>Tony Smith</u>**: *Criteria for accepting a productivity shift in fish stocks*.
15:10 **Malcolm Haddon**: *Species depleted before introducing a harvest strategy: changed productivity, continued overfishing, or just enforced ignorance?*

15:50 Session 2 Discussion

16:50 **Beth Fulton:** *Early insights into management challenges in SE Australia in 2050* ***(By Web-EX)***

17:10 Session ends

Wednesday Day 2

Session 3: Methods for quantifying uncertainty in projected changes

09:10 **Grant Thompson and Linsey Arnold:** *Risk-based selection of regime boundaries for a stock managed under a sloping, SPR-based control rule*.

09:30 **Jim Ianelli and Kirstin Holsman** *Developing simple multi-species trophic interaction models driven by climate (as mediated through complex models--i.e., FEAST) for testing management systems using single and multi-species assessment-control rules.*

10:10 Session 3 Discussion (includes time to share methods)

Session 4: Selecting harvest strategies that are robust to, or adjust appropriately in response to, climate change induced shifts in productivity

13:30 **Andre Punt:** *Fisheries management under climate and environmental uncertainty: Control rules and performance simulation*

13:50 **Lisa Pfeiffer:** *Why economics matters for understanding the effects of climate change on fisheries.*

14:10 **Andre Punt:** *An evaluation of stock-recruitment proxies for implementing the US sustainable fisheries act.*

14:50 Session 4 Discussion (includes time to share methods)

Thursday Day 3

Session 5: Defining the precautionary approach under a changing climate

09:10 **Lorenz Hauser:** *Time scales of adaptation in marine species with high larval mortality*

09:30 **Anne Hollowed:** *Is added precaution under one way trips a prudent course of action?*

10:10 Session 5: Discussion

13:30 Workshop synthesis and key findings

13:40 **Plenary Discussion synthesis action plan and recommendations**

14:00 Options for Global Partnerships and proposal writing

14:30 Closing statements

15:00 Workshop ends

Appendix C: Establishing Criteria in Estimating B_{MSY}

CPT (May 2011) with SSC revision (June 2011)

These criteria to select the time period to represent B_{MSY} or B_{MSYproxy} should be included in the analysis in each SAFE.

The time period should be representative of the stock fluctuating around B_{MSY}. The time period should be representative of the stock being fished at an average rate near F_{MSY}. For Tier 3 we are looking for an average recruitment and not an average biomass (B_{MSYproxy} formally only applies to Tier 4).

1. Provide an estimate of the production potential of the stock over the full time period of the assessment.
    a. Identify if the stock below a threshold for responding to increase production.
    b. For Tier-3 stocks, provide the time series of ln(R/S) and recruitment (R). For crab stocks, S is mature male biomass at the time of mating, and R is model estimate of recruitment.
    c. For Tier-4 stocks, provide a surplus production analysis using biomass and catch to evaluate the production potential over time. Give the formula for surplus production (units of MMB). Annual surplus production (ASP_t) is equivalent to the amount of yield that could have been taken in a given year that would have left the stock at equilibrium,

        $ASP_t = B_{t+1} - B_t + C_t$
        B_{t+1} = biomass in year t+1
        B_t = biomass in year t
        C_t = catch in year t

        Also, evaluate the time series of survey recruiting size class as a recruitment index. If it looks consistent look at time series of survey R/S.
    d. Identify potential mechanisms that should be considered to support production changes (i.e. Regime Shifts) based on a. and b. above. Consider evidence of a change in magnitude and direction of life-history characteristics that support a proposed change in production.

        Candidate life-history characteristics (empirical data) include:
        i. Natural Mortality (M)
        ii. Growth
        iii. Maturity (maturity schedule)
        iv. Fecundity
        v. Recruitment & recruits/spawner
        vi. Candidate ecosystem characteristics (empirical data) include:
            1. Overland method of Regime Shift detection
            2. Change in production of benthic spp. in EBS.
            3. Consumption (ecosystem model output).

2. Provide a plot of the history of the exploitation rate on MMB at the time of the fishery relative to F_{MSY} (Tier-3) or relative to the F_{MSY}=M proxy (Tier-4).

3. Provide a plot of the history of the exploitation rate on MMB at the time of the fishery relative to ln(R/S) (Tier-3) or relative to $\ln(R_{OBS}/MMB_{OBS})$ (Tier-4) where R_{OBS} is observed survey recruitment and MMB_{OBS} is observed survey MMB at the time of mating.
4. Examine the stock-recruitment relationship (SRR) for evidence of:
    a. Depensation in the SRR.
    b. Multiple SRRs consistent with a proposed regime shift paradigm.

The following methods were discussed by the CPT and SSC but considered not to be viable (see June 2011 SSC minutes). They are left in this version so that authors may comment on/ or consider their use.

5. For many crab stocks, historical rates of exploitation were higher or lower than current estimates of maximum rates fishing at F_{MSY}. The resultant B_{MSY} would be a biased (low or high) measure of reproductive potential since MMB at mating is tabulated after the extraction of the catch. If recruitment was maintained despite the difference, the extent of this bias is proportional to the magnitude of the catch above or below fishing at F_{MSY}. The recalculated B_{MSY} should be a better reference biomass estimate regardless of whether catches were larger or smaller than F_{MSY} catch.

6. For Tier-4 stocks, an alternative $B_{MSYproxy}$ can be estimated that adjusts for stock losses in excess of F_{MSY}. The analyst should estimate $B_{MSYproxy}$ based on the following approach:
    a. Using observed survey mature male biomass, estimate mature male biomass at the time of the fishery.
    b. Using the F_{MSY} proxy, estimate the catch using the biomass from (a).
    c. In years where exploitation rates exceeded those at F_{MSY}, replace the observed catch with that from (b) and recalculate MMB at mating.
    d. Produce a new time series of MMB at mating replacing those years where MMB was recalculated in (c).
    e. Recalculate $B_{MSYproxy}$ over the reference time period with the new time series of MMB at mating derived in (d).

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(∞) 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(∞) 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(∞) 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 E. Some examples of alternative ways to condition stock recruitment relationships for Tier 3 stocks

To illustrate how the implied stock recruitment relationship holds for proxy cases of F_{MSY} (e.g., $F_{35\%}$ for groundfish in Tier 3) some examples of “conditioned” values were constructed. This conditioning simply finds the value for the stock recruitment parameters (given selectivity, mean weight-at-age, maturity, and natural mortality) that satisfies the constraint that $F_{35\%} = F_{MSY}$. Here we refer to that process as “condition 1” An additional constraint considered satisfies the biomass proxy that $B_{35\%}$ is equal to B_{msy} and we refer to that as “condition 2”. For contrast, the stock recruitment parameters are also tuned to the “data” (output from the stock assessment model) as “unconditioned”. Ricker and Beverton Holt curves are applied to EBS pollock stock recruitment outputs.

Results show that for the Beverton Holt curve, the unconditioned fit resulted in a stock recruitment parameter value of 1.0 (constant recruitment) whereas when the curve was conditioned to satisfy the constraints (1 and 2) the slope of the stock recruitment curve is much shallower (Fig. E1). Note also that the scale changes between conditioning scenario 1 and 2 indicating that the assumption of $B_{35\%}$ being equal to B_{msy} results in a slight degradation of fit.

Results for the Ricker curve fits also shows considerable difference between the unconditioned fit and the conditioned curves (Fig. E2). The difference between the conditioning scenarios 1 and 2 are relatively minor. In summary, examining the implied stock recruitment by using F_{MSY} proxy can show how the results scale to the observations. Specifying a second level of conditioning (where $B_{35\%} = B_{msy}$) may be less defensible since $B_{35\%}$ is simply a function of historic recruitment levels (in this case averaged over the period from 1977 to the present).

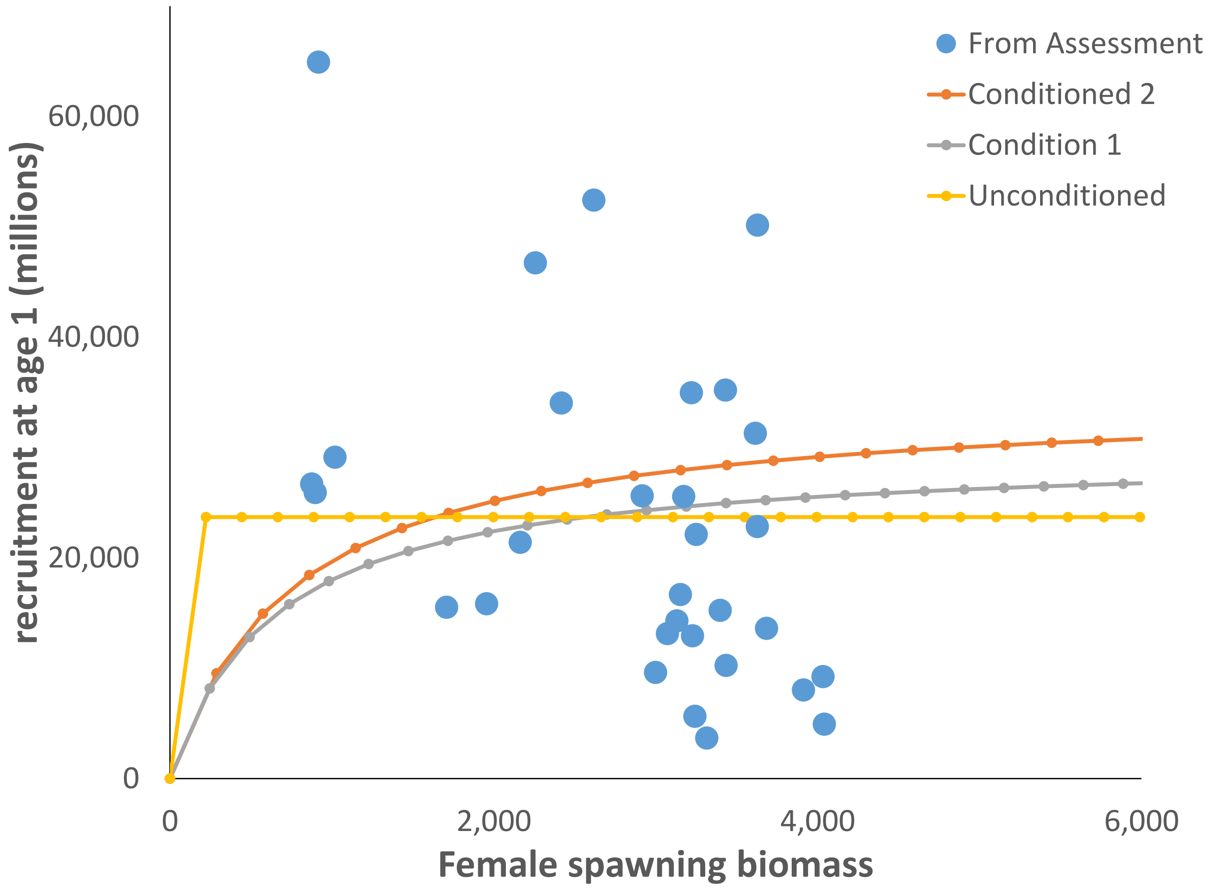

Figure E1. Fit of the Beverton Holt model for the EBS pollock data with different levels of model conditioning relative to proxy F_{MSY} and B_{MSY} values.

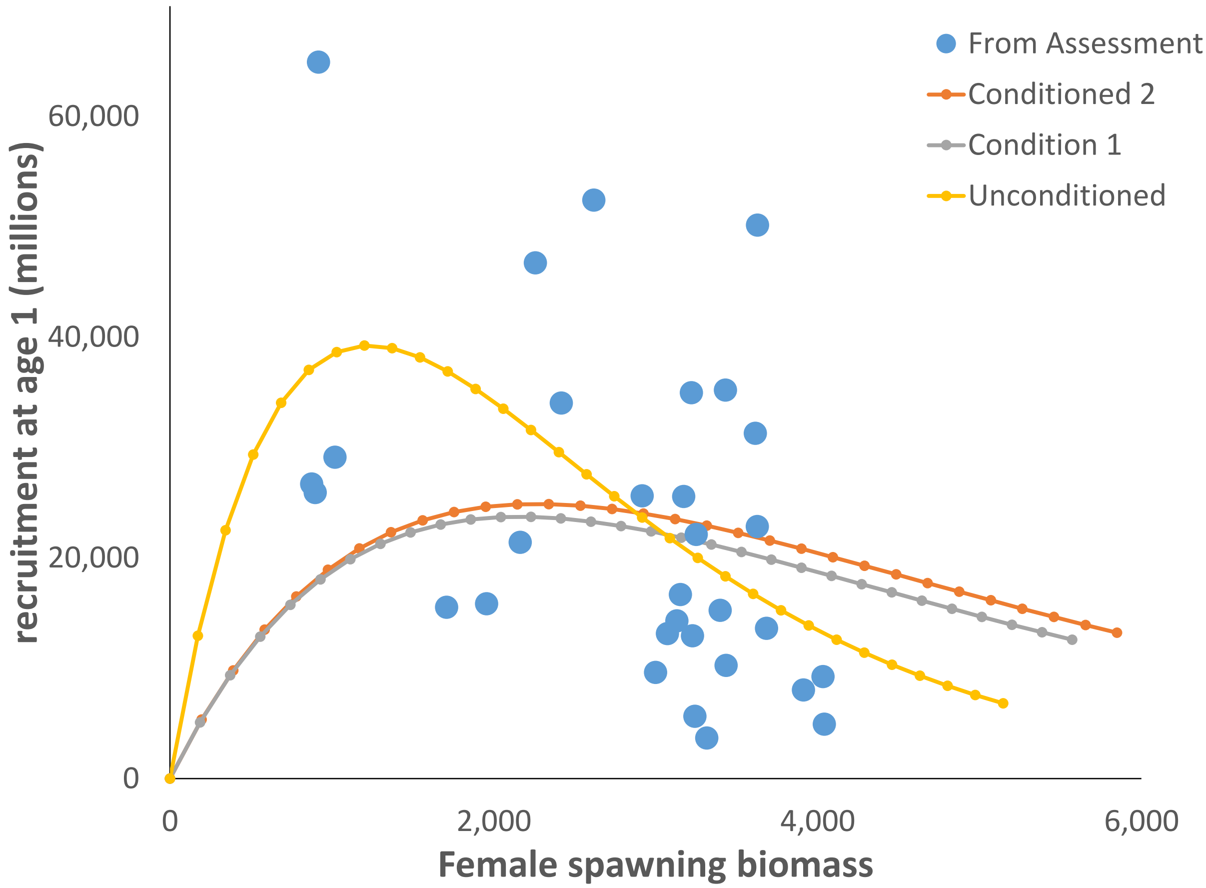

Figure E1. Fit of the Ricker model for the EBS pollock data with different levels of model conditioning relative to proxy F_{MSY} and B_{MSY} values.

Appendix F: A simple analysis of the B7 alternatives

To compare and contrast the performance of using mean and median 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. 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 E.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 were assumed to be known without error. Alternative values for the Tier 3 reference points were computed for each of the two estimators (mean recruitment and median recruitment). Catch was assumed to equal maxABC under the Tier 3 control rule in all years.

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. 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.

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 E.1
- Four statistics pertaining to rebuilding time (to the B_{MSY} proxy): upper and lower bounds of the 95% confidence interval, median rebuilding time, and mean rebuilding time; shown in Table E.2.

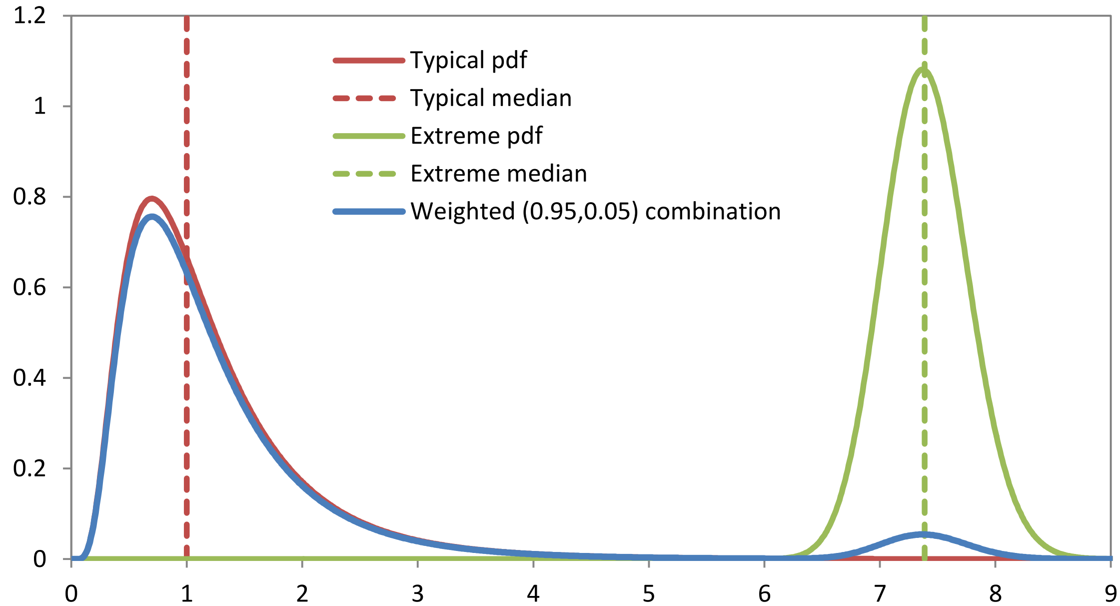

Figure E.1. Probability density function of recruitment used in the simulation.

Table E.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.

		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	0.58	0.13	1.03	0.16	0.68	0.22	1.04	0.23
Biomass	Median	0.56	0.12	1.00	0.18	0.65	0.21	1.00	0.25
Exploitation	Mean	0.55	0.14	0.95	0.07	0.65	0.19	0.93	0.09
Exploitation	Median	0.78	0.14	1.00	0.01	0.83	0.15	1.00	0.02
Catch	Mean	0.34	0.17	0.98	0.21	0.48	0.30	0.98	0.29
Catch	Median	0.45	0.18	1.00	0.18	0.56	0.27	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	0.82	0.34	1.05	0.33	0.89	0.42	1.06	0.40
Biomass	Median	0.77	0.33	1.00	0.35	0.84	0.42	1.01	0.42
Exploitation	Mean	0.73	0.21	0.90	0.12	0.76	0.21	0.88	0.13
Exploitation	Median	0.88	0.15	0.98	0.05	0.89	0.15	0.97	0.06
Catch	Mean	0.66	0.43	0.97	0.40	0.74	0.51	0.97	0.48
Catch	Median	0.71	0.39	0.99	0.36	0.79	0.47	0.99	0.44

Table E.2. Statistics related to rebuilding time under the two alternative estimators and a range of discrete annual mortality rates.

	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	49	69	32	44	22	31	17	24
Mean	21	27	13	16	8	11	7	8
Median	19	23	11	13	7	9	6	7
L95%CI	7	7	3	3	2	2	2	2