THE SCIENTIFIC AND STATISTICAL COMMITTEE'S GROUNDFISH SUBCOMMITTEE REPORT ON TERMS OF REFERENCE FOR THE GROUNDFISH AND COASTAL PELAGIC SPECIES STOCK ASSESSMENT PROCESS FOR 2021 AND 2022

A stock assessment methodology review for length-based data-moderate stock assessments was held via webinar on May 12 - 14, 2020. The review focused on the application of Stock Synthesis incorporating catch and length data (SS-CL) as well as incorporating indices from research surveys (SS-CL-Index), which were both endorsed by the panel. The draft proposed language for the Terms of Reference (TOR) for length-based methods provided herein is based on the findings of the methodology review and suggestions from the panel. The language in track changes reflects feedback provided since the June Council meeting where the draft language was initially submitted for consideration in an SSC Groundfish Subcommittee Report. If the Council adopts the length-based assessment methods for use in management after final review of short-term requests by the Scientific and Statistical Committee (SSC) Groundfish Subcommittee and the full SSC, the proposed language should be incorporated into the TOR for 2021 stock assessments to facilitate application of the proposed methods. The following sections of the TOR for Stock Assessments would be amended to reflect the language provided to incorporate length-based assessment methods:

Page 9:

4. Data-moderate assessment: A continuum of "data-moderate assessment" methods exist between data-poor methods relying on catch data alone to fully integrated stock assessments incorporating all available data including up to date age data. Council-approved methods for data-moderate assessments are limited in that they do not use age data, even if available, and have simplified population dynamics (deterministic recruitment), which makes such assessments less complicated and enables more expeditious review. Two methods have been developed to conduct data-moderate assessment with historical catch data and one or more indices of abundance (or biomass) (e.g., survey data or fishery catch per unit effort [CPUE] indices) referred to as extended DB-SRA (XDB-SRA) using stand-alone programming and extended Simple Stock Synthesis (XSSS) using Stock Synthesis. In addition, length-based data-moderate stock assessment methods have been developed that incorporate only catch and length data in Stock Synthesis (SS-CL). Methods incorporating catches, lengths, and indices of relative abundance from fishery-independent surveys in Stock Synthesis have also been adopted by the Council for use in management (SS-CL+Index). Methods using length data do have the potential to estimate recruitment, thus adding more dynamics and complexity to the assessment and the review. Data-moderate assessments are reviewed by the relevant SSC Subcommittee (Groundfish or CPS) if an approved standard methodology is proposed to be used. They are reviewed by a STAR panel if a new or non-standard assessment methodology is proposed to be used.

Page 39:

9. Data-Moderate Assessments for Groundfish Species

Data-moderate assessments for groundfish species are a refinement over data-poor methods that result in category 3 assessments (described below) in that a data-moderate assessment includes length and/or abundance trend information in addition to the data informing a data-poor form of

the assessment (catch series plus prior information on productivity and status). Data-moderate assessments can result in category 2 (catch and length or catch and index) or category 1 designations (possible when catches, lengths, and index data are incorporated). One defining distinction between category 2 and category 3 assessments is that the length and/or abundance trend information is incorporated in a category 2 assessment enabling an estimate of stock status (Appendix F). While the SS-CL assessments have the potential to be category 1 assessments, simulation analyses indicate that if there are fewer than ten years of length data for a stock in question, there is substantial uncertainty in the results, potentially leading to designation as a category 3 assessment.

Two index-based data-moderate assessment methods have been endorsed since the 2013-14 assessment cycle, XDB-SRA and XSSS. In both cases, abundance trend information (e.g., survey or fishery CPUE indices) is included in the assessment. The length-based data-moderate assessment method using only catches and lengths is SS-CL, while SS-CL+Index uses catches, lengths and indices of abundance from fishery-independent surveys for which index derivation is well established. A flowchart describing the specific steps to take in conducting these assessments is provided in Appendix J. These specific applications provide assessments that are understood well enough to require only review by the SSC Groundfish Subcommittee. More complex data-moderate assessments within the SS framework incorporating fishery-dependent indices of abundance or use of age data are also possible and require review by a STAR panel review to address the added complexities of model fitting and index development.

The continuum of models should be accommodated to allow combinations of catch, lengths, ages and, indices to be applied to both new assessments and length-based extensions of existing benchmark assessments, though review processes may differ. A categorization of each assessment method is provided in Table 2, which defines the scope of each method in terms of data source and assessment type. The flow chart in Appendix J includes how to prepare catch estimates, length data, parameter estimation, model weighting, model convergence, and characterize uncertainty. The SS-CL+Index assessment method includes fishery-independent indices of abundance for which index development methods are well established. Implementation of assessments within the provided specifications allows for standardization and more streamlined review by the SSC Groundfish Subcommittee in May or June of odd years along with update assessments. The depth of potential reviews should increase with the estimation of more parameters. Review of one or more of length-based models that use an approved standard methodology can still be reviewed within a STAR Panel setting, though Groundfish Subcommittee review is generally sufficient. Intermediate methods using fishery-dependent indices or age data are subject to review at a STAR panel to allow further evaluation of model fitting and tensions between data sources (see Table 2 for categorization of assessment types). Assessments conducted with data-moderate methods may be category 2 or category 1 depending on whether only catch and length or catch and index data were used resulting in a category 2 assessment or if catch, length and index data were used, which can result in a category 1 assessment.

Intermediate assessments picking up where past full stock assessments left off, that exclude new age and/or index data, but maintain index or age data included in the previous assessment while adding length and catch data are subject to STAR panel review. These types of intermediate methods may be applied to stocks that were formerly assessed with a full benchmark assessment, but for which attainment has been low or conducting a full or update assessment is not a priority despite the assessment becoming dated. Addition of new age data to either type of length-based

Stock Synthesis assessment is discouraged to avoid confounding the nature of the assessment, reducing model tension arising from potential conflicts in age and length data, and increasing clarity in the related review process. Due to the complexities and potential data conflicts that can occur from fitting to age data in combination with indices and lengths, assessments that include current age data should be considered full benchmark assessments and reviewed in STAR panels.

Table 2. Model types, their data types and assessment type. All assume a known catch history.

Model	Lengths	Ages	Index	Assessment Type
DB-SRA/SSS	Ignore	Ignore	Ignore	Data-limited
XDB-SRA/XSS	Ignore	Ignore	Use	Data-moderate
SS-CL[1]	Use	Ignore	Ignore	Data-moderate
SS-CL+Index	Use	Ignore	Fishery-independent indices only (e.g., WCGBTS, H&L)	Data-moderate
SS (new config[2])-lite	Use	Ignore	Use	Likely Benchmark
SS (old config[3])	Use / update?	Use new / Ignore unread	Use / update?	Update
SS (new config[2])-heavy	Perhaps new data sources	Perhaps new data sources / Ignore unread	Perhaps new data sources	Benchmark

1: Flow chart for specifications related to fleets, life history parameters, selectivity etc.
2: New specifications for how the assessment is configured
3: Model specifications the same as the last assessment

Index-based Methods

The index-based method XSSS assumes that recruitment is related deterministically to the stock-recruitment relationship and allows index data to be used within a Bayesian framework. The Markov chain Monte Carlo (MCMC) or Sample Importance Resample (SIR) algorithm (perhaps implemented using Adaptive Importance Sampling) is used to quantify uncertainty for XSSS-based assessments. The XDB-SRA method is implemented within a Bayesian framework, with the priors for the parameters updated based on index data. The additional parameters in XDB-SRA compared with DB-SRA include the catchability coefficient (q), and the extent of observation variance additional to that inferred from sampling error (a). The priors for these parameters are a weakly informative log-normal and a uniform distribution, respectively.

Comparison of alternative methods (XDB-SRA and XSSS) is encouraged, but it is acceptable to present an assessment using a single modeling approach. The SSC Groundfish Subcommittee can make requests of the STATs for additional runs but should not impose an alternative method if STATs consider this is not appropriate for the stock concerned. If more than one model is presented, the SSC Groundfish Subcommittee should recommend adoption of a preferred model, if one can be identified, for use in management.

Length-based Methods

Applying SS-CL and SS-CL-Index is very similar to conducting a standard Stock Synthesis (SS) assessment since SS-CL is conducted in Stock Synthesis, and all equations for the model can be found in the SS documentation (Appendix A of Methot and Wetzel, 2013). Like SS, the data for these methods can include many fleets, sexes, etc. as desired, catches are a full time series and

assumed known and length compositions are assumed to be representative, with effective sample size treated in standard ways. Life history values (i.e., steepness, growth parameters (k, L_∞, t_0), natural mortality, fecundity, maturity) are initially pre-specified (some degree of this does happen in many standard SS models), but estimation of some values may be possible. Recruitment can be estimated, and standard bias correction procedures are applied. Selectivity can is also typically estimated and can be fixed. If multiple fleets have length composition data, data weighting approaches would follow standard procedures as outlined in Appendix B Section H. The starter and forecast files are specified as in traditional SS assessments. The performance and stability of SS-CL was better with smaller model dimensions (e.g., fewer fleets) and is sensitive to errors in the fixed values for L_∞ and the coefficient of variation (CV) of length-at-age, which, if fixed, should be explored in sensitivity analyses.

The limited scope of SS-CL and SS-CL+Index allow for more limited documentation requirements, described in Appendix E. For more complex intermediate models beyond the scope of these focused methods, reporting requirements should be developed in anassessment-specific TOR developed by the Chair of the STAR panel to provide flexibility to cover the range of possible applications, while still providing appropriate specificity and thorough evaluation. The reviews are expected to take between a half day and two days depending on the number, type, and novelty of the assessments. It may be beneficial to hold a half day preliminary review during a virtual meeting prior the Groundfish Subcommittee at which the review will be conducted. The number of SS-CL or SS-CL+Index assessments that can be conducted at a given STAR panel or the Groundfish Subcommittee of the SSC in combination with update assessments, depends on the complexity of the models, spatial areas, and novelty of the methods. Between two and four assessments in a review may be reasonable, and flexibility should be provided to the SSC in determining how many assessments should be reviewed and the process for each review.

The critical modeling steps for SS-CL and SS-CL+Index are included in the methodology step flow chart in Appendix J in addition to the following guidance. Jittering and alternative phasing should be used given the difficulties encountered by the analysts during this review finding the global minimum of the objective function. If there is dimorphic growth, then sex-specific information should be included, given increased uncertainty in simulation results with increasing variance in length at age, which is greater when sex data associated with lengths are not available or included in assessment of sexuallydimorphic species. This may be less of a concern if only males or females are predominantly sampled by the survey or caught in the fishery, but can be confounding if more equal sex ratios are observed in the catch or survey and sex data is unavailable for measured fish. Fleet consolidation is recommended if selectivity is similar among sectors or surveys to reduce model conflict and confounding affects. If a survey is included in an SS-CL+Index assessment, the length-composition from that survey should also be included, as well as length-composition from other fishery-independent or fishery-dependent data sources. Application of dome-shaped selectivity should be investigated when plausible, in addition to asymptotic selectivity. It is recommended that the model be run with asymptotic selectivity for at least one fleet if natural mortality is being estimated. Simplifying model structure and spatial areas will reduce complexity in the assessments and workload in both the analysis and review.

References:

Methot, R.D. and C.R. Wetzel. 2013. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* 142: 86-99.

Page 54:

Appendix E: Template for a Data-Moderate or Data-Poor Assessment Document

1. Title page and list of preparers – the names and affiliations of the stock assessment team (STAT).
2. Introduction: Scientific name, distribution, basic biology (growth, longevity, ecology), the basis for the choice of stock unit(s) (no more than 1-2 paragraphs).
3. Development of indices (used and rejected). Novel approaches should be fully documented.
4. Treatment of length composition data (weighting, addressing discards, etc.).
5. Survey of other data available for assessment: data available to inform indices of abundance, sample sizes by year and source of lengths, and ages (read and unread)--in case there is interest in conducting a full assessment in the future.
6. Selection of method: length-based (SS-CL), index-based (XSSS or XDB-SRA; authors are "encouraged" to do both) or hybrid method (SS-CL+Index).
7. Assessment model
    a. Specification of priors / production function (defaults are acceptable)
    b. Initial runs using catch-only methods (DB-SRA or SSS (or both))
    c. Diagnostics
        i. Evaluation of convergence
        ii. Residual plots
        iii. Posterior predictive intervals (if Bayesian)
        iv. Acceptable parameter estimates
        v. Time-trajectories of biomass, depletion, etc.
        vi. Sensitivity analyses using alternative catch streams, alternative priors for depletion, etc.
7. Estimates of OFL (median of the distribution), and
8. Estimates of stock status where applicable.

Page 59.

Appendix J: Flow chart for SS-CL

Take the following steps in completing an SS-CL assessment:

1. Prepare catch data

- Catch treated as known. Use total mortality (landings + dead discards).

2. Prepare length composition

- Determine length bins and frequency within bins across years. More than 10 years of data (with reasonable sample sizes) is recommended. Otherwise it is a category 3 assessment.
- This can be done for as many fleets as needed, but use the parsimony principle to define fleets, as model convergence may be more difficult with more fleets.
- Female, male and unknown data can be used.
- Determine effective sample sizes following standard protocol.
- Combine length data from landings and discards (or reasonable assumptions for the latter if no data) appropriately.

3. **Define life history parameters**
    - Natural Mortality: define using estimators (e.g., Hamel method (must include as a sensitivity at least, if an estimate of longevity/maximum age is available), Natural Mortality Tool (which includes the Hamel method)). Fix to central tendency (median value) and retain uncertainty for sensitivity analyses.
    - Growth parameters. Externally fit the von Bertalanffy growth function and use point estimates to fix in model. Choose a fixed valued for CV at length. Retain uncertainty for sensitivity analyses.
    - Steepness defined either through meta-analysis or expert opinion. Retain uncertainty for sensitivity analyses.
    - Recruitment variability also defined through meta-analysis or expert opinion. Retain uncertainty for sensitivity analyses.
    - Life history parameters will generally be pre-specified but consideration could be given to estimating these parameters (see Section 1.5.3).

4. **Parameter estimation**
    - Estimate R_0, recruitment deviations and selectivity parameters.
    - Life history parameters if likelihood profiles show information.
    - Selectivity can be logistic, dome-shaped, or whatever form in chosen in SS.
    - Bias correction to recruitment deviations can subsequently be applied.

5. **Model weighting**
    - Consider weighting the length compositions if multiple fleets.

6. **Model convergence**
    - Length-only models may take additional jittering to find convergence and avoid local likelihood minima.
    - Check model fits to length compositions.
    - Determine whether selectivity shapes and subsequent estimates make sense.
    - Review other parameters estimates for bounds and poor estimation (and whether they are reasonable).

7. **Characterize uncertainty**
    - Likelihood profile over, at minimum, M, L_2 (preferably parameterized as L_∞, though can also make the transformation for reporting) and k (retain correlation structure if possible), CV at length, and h.
    - Sensitivity analysis should be conducted, either based on likelihood profile information or identified model specification.
    - Ensemble modeling to quantify model specification error would be useful. This would need further discussion on how best to approach it.

PFMC
10/15/20