ACCEPTED PRACTICES GUIDELINES FOR GROUNDFISH STOCK ASSESSMENTS

The following guidelines are intended to supplement the Council's *Terms of Reference for Groundfish and Coastal Pelagic Species Stock Assessments* and provide groundfish STATs with default approaches they should use for dealing with certain stock assessment data and modeling issues. The STATs may diverge from the guidelines if they provide adequate justification for doing so. These guidelines are not intended to provide a comprehensive treatment of all potential issues, which are too numerous to list. Rather the guidelines focus on a limited number of issues that the SSC has so far considered. The purpose of having these guidelines is to lessen the time that might otherwise be spent during stock assessment reviews in discussions about how particular steps in the assessment process should have been conducted. The guidelines are subject to change as the SSC evaluates additional data sources and modeling approaches. STATs should consult with Council staff to obtain the most recent set of guidelines, which the SSC will finalize in April 2019 for use with 2019 stock assessments.

Biomass indices from bottom trawl surveys

The geostatistical delta-GLMM software (vector autoregressive spatial temporal model, VAST), developed by Dr. Jim Thorson (NWFSC) and the software wrapper for West Coast groundfish maintained by Dr. Kelli Johnson (NWFSC), is an acceptable tool for developing biomass indices from bottom trawl survey data, though exploration of other methods is encouraged. For survey data, the software includes a range of options that can either replicate previously recommended model configurations (e.g., delta-GLMM with vessel as a random effect) or use more advanced analytical methods, such as spatial autocorrelation. Analysts are strongly encouraged to compare model results with and without the spatial autocorrelation feature. If they use the geostatistical features they should provide appropriate diagnostic statistics. Assessment documents should include diagnostics supporting the selected model underlying each biomass index and should also include a comparison of the model-based biomass estimates with design-based estimates to gauge the uncertainty associated with the choice of methodology.

The following references offer guidance for using the VAST software, including recommended defaults and practices.

1. VAST wiki page (overview) – https://github.com/James-Thorson/VAST/wiki (and linked pages).
2. Guidance for VAST use – Thorson, J.T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat, and climate assessments. Fish. Res. 210: 143-161. https://www.sciencedirect.com/science/article/pii/S0165783618302820
3. The software wrapper used to describe the application of VAST to West Coast Groundfish Bottom Trawl data is available at https://github.com/nwfsc-assess/VASTWestCoast.

Biomass indices from fishery-dependent sources (e.g., logbooks)

The VAST software can also be used to standardize fishery CPUE data series for use as biomass indices. If the geostatistical option is used the approach provides an objective mechanism for imputing catch rates from regions with no fishing. STATs who apply the software to fishery-dependent data will need to provide the STAR Panels with substantive interpretation and diagnostics to demonstrate that the analysis appropriately considers issues such as changes in fishing power and truncation of large catches due to trip limits.

Spatial stock structure for groundfish species

STATs conducting assessments of groundfish species should explore regional differences in biology (or the underlying environmental conditions that influence biology) when defining stock structure in assessments. If there are separate regional models for a species the models should use consistent approaches for modeling productivity and for data weighting. STATs conducting assessments of nearshore groundfish species should explore state-specific or finer-scale stratifications for the assessment models to account for regional differences in exploitation and management history.

For STATs that explicitly include spatial structure within an assessment model, the SSC strongly recommends that STATs review both assessment documentation and STAR Panel reports of recent spatially explicit models (e.g., Canary Rockfish in 2015, Yelloweye Rockfish in 2017) to consider how they confront and evaluate the sensitivity of the model to particularly challenging aspects of spatially explicit models, such as parameterizing movement rates and the partitioning of new recruits across space. STATs should also consider the location of capture, not just the location of landings when considering either explicit or implicit spatial structure within assessment models. During model development, the STATs should consider the issues discussed in Berger et al. (2017) and Punt (2019).

Prior distributions for natural mortality (*M*)

Assessments for groundfish species should report the prior probability distribution for natural mortality (*M*) computed using the meta-analytical approach updated by Dr. Owen Hamel (NWFSC) based on maximum ages (Hamel, 2015; Then et al. 2015) and STATs should explore using the prior to inform the assessment models. This prior is defined as a lognormal distribution with median value (corresponding to the mean in log-space) = 5.40/maximum age and log-scale sigma = 0.438. Both parameters should include exactly three significant digits.

The maximum age values on which *M* priors are based should generally be from fish caught within the area of the assessment, not from Alaskan catches of the same species, for example. If the prior for *M* is used to provide a fixed value for *M*, the fixed value should be set equal to the median value of the prior (e.g., 5.40 / maximum age for the prior defined above).

Age- or gender-specific *M*

For assessment models with age-specific *M* the default modeling approach should be a step function rather than a linear ramp, which is a more complicated form of age-dependence. If the Lorenzen approach is used to model age-dependent *M* (Lorenzen, 1996) the assessment should also present a comparison run that uses constant *M* (i.e., no age-dependence).

STATs should exercise care when estimating sex-specific values for *M* because of the potential for confounding with sex-specific selectivity. In such cases, STATs should provide sensitivity analyses to explore consequences of potential confounding.

Weighting of compositional data

There are three accepted approaches for weighting age- and length-composition data: (1) the McAllister and Ianelli (1997) harmonic mean approach; the Francis (2011) approach; and the Thorson et al. (2016) Dirichlet multinomial likelihood approach. The first two methods have been used routinely in Council assessments, whereas the third method, which became available in Stock Synthesis in 2017, has yet to be used extensively. There is no clear consensus that one approach is superior in all circumstances. STATs are encouraged to provide a rationale for the method they select and are encouraged to conduct sensitivity runs with the other methods. STATs should explore correlations in residuals among bins and years to rationalize the weighting approach. Visual examination of bubble plots might provide evidence of substantial correlations between years and ages/lengths.

The calculation of the weighting coefficients for compositional data is done iteratively for the harmonic mean and Francis methods. Starting values are used and updated after each iteration. STATs should conduct several iteration steps to verify there is reasonable stability in the coefficients.

The starting values for the weighting coefficients for marginal compositional data (based on age or length) should be the number of bottom trawl survey tows or fishing trips contributing fish to the composition, or a formulaic combination of the two quantities (Stewart and Hamel 2014). The starting values for conditional age-at-length data should be the actual numbers of fish on which each composition is based.

Data Extractions

The STATs should record and report the versions of any databases they use and the dates of any database queries and data extractions so there can be verification that the most up-to-date data were used.

Landings Data

STATs should either (a) verify that the relevant unidentified fish category (e.g., URCK, UFLT) in PacFIN and RecFIN has no appreciable quantities of the species being assessed or (b) develop and apply an appropriate species proportion to the landings of unidentified fish to estimate corresponding landings of the species being assessed. Ideally, STATs will provide information regarding the data quality associated with species composition estimates in mixed species market categories.

STATs should consult with each of the state's data stewards, well in advance of the STAR, to verify that they have acquired the correct landings data series and that the series are complete. STATs should check with WDFW on the status of fish tickets included in PacFIN (or NORPAC for at-sea catches) for recent tribal landings and confirm there are complete tribal landings data.

The historical catch reconstruction developed for California currently does not account for fish landed into California that were caught off Oregon or farther north. STATs should establish if this portion of the historical fishery in California accounts for appreciable quantities of the species being assessed.

Discard Data

For discards in commercial fishing operations, the STATs should check in with the NWFSC Groundfish Observer program to obtain estimates of discards and summaries of any available biological information for discarded fish. Estimates of total commercial fishery discards and discard mortality are reported by the Observer Program in the Groundfish Expanded Mortality Multi-year (GEMM) annual report. The STATs should contact the state data stewards and RecFIN to obtain available data for discards by recreational fishers. Recreational discards should include both the "released dead" and "released alive" categories, and STATs should describe what they assumed for mortality rates for live-released fish. STATs should provide rationale for any assumed discard mortality rate.

The STATs should include an analysis to evaluate whether there is evidence of size-based discarding and determine if the assessment model should include size-based retention for either commercial or recreational catch.

Compositional Data

When combining compositional samples from different geographic strata, the composition proportions should be weighted by some appropriate measure of the numerical abundance in each stratum (catch in numbers for fisheries; numerical abundance for surveys). Catch weights would not be appropriate if the average weights of the fish vary appreciably among the regions. STATs should be mindful of the potential for size sorting of landings (and how sorting has changed over time) for species such as sablefish and petrale sole. Size sorting can influence the compositions reflected in commercial length and age data and has implications for how expansions are constructed.

A software package is available from the NWFSC to process biological sample data stored in PacFIN, in the Biological Data Samples (BDS), and to generate time series of compositional data that are formatted for use with Stock Synthesis. The STATs should use this software, or provide a rationale for why they do not. If a STAT uses other software, they should provide some comparison of the results of each approach, as well as a comprehensive rationale for why they have used alternative expansion approaches or data.

The composition data for the recreational fishery can be obtained from RecFIN. STATs should consult with the state data stewards regarding their use and expansion for contemporary and historical sampling programs. In addition to age- and length-composition data from landed catch, additional length-composition for discarded catch from onboard sampling in California (Type 3d data from the California Recreational Fishery Survey) can be informative of more recent recruitment patterns. Further evaluation of appropriate methods of treating discard lengths is an area for further exploration as both retention curves and including discards in a separate fleet have been employed in the past and may have unforeseen impacts on data weighting between fleets that may require examination in sensitivity analyses.

Recreational Catch-per-Unit-Effort Data

If a catch-per-unit-effort index is developed from a multi-species recreational data source that does not report fishing locations at a fine scale (e.g., the data were not collected by at-sea observers), the data should be screened using the Stephens and MacCall (2004) method to identify data records that were unlikely to include the species being assessed.

Modeling - Selectivity

Non-selected biomass, sometimes described as "Cryptic biomass", is a term used to describe the phenomenon whereby a model predicts biomass that is not directly observed in data from any fleet.

Non-selected biomass can only arise if all fleets have selectivity curves that are dome-shaped. An assessment that has all fleets with dome-shaped selectivity curves should as a sensitivity run include at least one fleet that has asymptotic selectivity and should provide a figure or estimate of the amount of non-selected biomass for cases where all selectivity curves are dome-shaped.

Modeling – Growth

For some species there may be length and age data available from special projects that fall outside normal port sampling programs (e.g., research samples from nearshore nursery areas). Such data may provide information that more completely informs the growth curve and can be used in an assessment model. Such data are not appropriate to use in modeling fishery selectivities (i.e., should be associated with their own fleet).

Check for stability in length-at-age.

Assessment models often assume that growth is time-invariant. A plot depicting observations of mean length-at-age by fleet over time would provide evidence to support or refute the assumption that growth has been constant.

Modeling - Fecundity

Rockfish stock assessments should consider the fecundity relationships from the meta-analysis in Dick et al. (2017), at the appropriate taxonomic scale, if better species-specific relationships are unavailable. If a size-dependent fecundity relationship is not used in the base model, the model should include a sensitivity run comparing spawning output proportional to mature female biomass versus increasing weight-specific fecundity.

Modeling – Diagnostics

In addition to the standard set of likelihood profiles identified in the Stock Assessment Terms of Reference (across the parameters $\ln(R_0)$[1], M and steepness), the STATs may wish to consider other diagnostics, such as those highlighted in Carvalho et al. (2017).

Modeling – Prior on Steepness – *Sebastes* Species

The SSC-approved steepness prior for rockfish species in 2019 has a mean value of 0.72 and standard deviation of 0.16. Both parameters are defined to exactly two significant digits. If the assessment model does not estimate steepness, the STAT should fix the steepness value at 0.72. This applies to all 2019 rockfish assessments, even for species that were included in the 2017 meta-analysis (i.e., no "Type-C" special case).

Modeling – Prior on Steepness – Other Species

If a prior for steepness is used to provide a fixed value for steepness, the fixed value should be set equal to the mean value of the prior.

Including extra variability parameters with an index.

STATs should be cautious to avoid adding variability to an index as a means of resolving model structure issues such as conflicts among data sources. Rather, variability should be added to account for sampling variance underestimating index uncertainty. STATs should provide *a priori*

[1] Parameter R_0 is the number of age-0 annual recruits in an unfished stock.

reasons for why the index variability input to the model has been underestimated (or underspecified).

Jittering to verify convergence.

In Stock Synthesis, the jitter fraction defines a uniform distribution in cumulative normal space +/- the jitter fraction from the initial value (in cumulative normal space). The normal distribution for each parameter, for this purpose, is defined such that the minimum bound is at 0.001, and the maximum at 0.999 of the cumulative distribution. If the jitter faction and original initial value are such that a portion of the uniform distribution goes beyond 0.0001 or 0.9999 of the cumulative normal, that portion beyond those bounds is reset at one-tenth of the way from the bound to the original initial value.

Therefore sigma = (max-min) / 6.18. For parameters that are on the log-scale, sigma may be the correct measure of variation for jitters, for real-space parameters, CV (= sigma/original initial value) may be a better measure.

If the original initial value is at or near the middle of the min-max range, then for each 0.1 of jitter, the range of jitters extends about 0.25 sigmas to either side of the original value, and the average absolute jitter is about half that. For values far from the middle of the min-max range, the resulting jitter is skewed in parameter space, and may hit the bound, invoking the resetting mentioned above.

To evaluate the jittering, the bounds, and the original initial values, a jitter_info table is available from r4ss, including sigma, CV and InitLocation columns (the latter referring to location within the cumulative normal – too close to 0 or 1 indicates a potential issue).

Strategies for phase sequencing.

In general, it is often best to evaluate parameters that scale the population (e.g., R_0, catchability, recruitment deviations, and initial abundance) in early phases before proceeding to phases that evaluate selectivity, growth, time blocks or time varying parameters. However, alternative phase sequences can have an impact on parameter estimation, likelihood minimization, and model convergence. STATs should consider alternative phase sequencing as a model diagnostic tool in addition to jittering.

Default assumptions for removals in projections and decision tables.

The default assumptions for the removals to include in projections are context dependent. In cases in which the fishery has been stable with low ACL attainment, considering scenarios in which future attainment is low is likely justified. The default should for removals to equal the projected ACLs as well as any removal scenarios (e.g., based on lower than 100% attainment) considered in the last assessment. A rationale should be provided if the removal scenarios differ from those in the last assessment. The STAT should work with the GMT / Council representatives at the STAR Panel to identify additional removals scenarios, which also should be justified if they differ from those in the last assessment.

References

Berger, A.M., Goethel, D.R., Lynch, P.D., Quinn II, T., Mormede, S., McKenzie, J., and Dunn, A. 2017. Space oddity: The mission for spatial integration. *Can. J. Fish. Aquat. Sci.* 74: 1698-1716.

Carvalho, F., Punt, A.E., Chang, Y.J., Maunder, M.N., and Piner, K.R. (2017). Can diagnostic tests help identify model misspecification in integrated stock assessments? *Fisheries Research* 192: 28-40.

Dick, E.J., Beyer, S., Mangel, M. and Ralston, S., 2017. A meta-analysis of fecundity in rockfishes (genus Sebastes). *Fisheries Research* 187: 73-85

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

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

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 54: 284–300.

Punt, A.E. 2019. Spatial stock assessment methods: A viewpoint on current issues and assumptions. *Fisheries Research* 213: 132-143.

Stephens, A. and MacCall, A. 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. *Fisheries Research* 70: 299-310.

Stewart, I.J., and O.S. Hamel. 2014. Bootstrapping of sample sizes for length- or age-composition data used in stock assessments. Canadian Journal of Fisheries and Aquatic Sciences 71:581-588.

Then, A. Y., Hoenig, J. M., Hall, N. G., and Hewitt, D. A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science* 72: 82-92.

Thorson, J.T., Johnson, K.F., Methot, R.D., and Taylor, I.G. 2017. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. *Fisheries Research* 192: 84-93.

Thorson, J.T., Fonner, R., Haltuch, M.A., Ono, K., and Winker, H. 2016. Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. *Can. J. Fish. Aquat. Sci.* 73.