The Scientific and Statistical Committee's Groundfish Subcommittee Report on the Review of Assessment Methodologies Proposed for Use in 2017 Groundfish Assessments

January 25
National Marine Fisheries Service
Northwest Fisheries Science Center
The Auditorium
2725 Montlake Boulevard E
Seattle, WA 98112

January 26
Seattle Yacht Club
Heritage Room
1807 E Hamlin St.
Seattle, WA 98112

25 – 26 January 2017

Introduction

The Groundfish Subcommittee of the Scientific and Statistical Committee (SSC) met to review new methods proposed for use in groundfish stock assessments prepared for the Pacific Fishery Management Council. The meeting was held at the Northwest Fisheries Science Center facility in Montlake on 25th January 2017 and at the neighboring Seattle Yacht Club on the 26th. The meeting began with a welcome by the Subcommittee chair, Dr. David Sampson (Oregon State University), followed by a round of self-introductions from the attendees, a brief review of the agenda and assignment of reporting duties. Appended to this report are a list of reviewers and attendees, as well as the meeting agenda.

Use of the Dirichlet Multinomial Likelihood for Compositional Data

Dr. Jim Thorson (Northwest Fisheries Science Center, NWFSC) gave a presentation on weighting compositional data using the Dirichlet multinomial, which is an option in the new 3.30 version of Stock Synthesis. Thorson et al. (In Press) provides details about the Dirichlet multinomial approach, which in essence treats the effective sample size for a multinomial compositional data series as an estimable parameter and accomplishes the tuning of compositional weights (relative adjustments to effective sample sizes) without the need for iterative reweighting as done in recent assessments using the methods of McAllister and Ianelli (1997) or Francis (2011). The Dirichlet multinomial appears to have a very similar effect on assessment results as the McAllister-Ianelli (MI) reweighting approach but has the advantage of being fully objective and repeatable. Thus using the Dirichlet approach would automatically take care of reweighting during auxiliary analyses (e.g., bridging runs) that typically do not include reweighting the compositional data.

The Dirichlet approach, like the MI approach but unlike the Francis method, does not account for correlations among the compositional data. Dr. Thorson noted that the Francis method only addresses correlated residuals at one level (among bins in a given year) but there are also correlations among years in a given bin, between sexes, and among fleets. He suggested that these correlations should be accounted for in the structure of the model itself rather than in the weighting of the compositional data. He also proposed that the iterative McAllister-Ianelli,

iterative Francis method, and the Likelihood-based Dirichlet-multinomial option could all be used in the upcoming round of assessments, with the selection of the method being left to the STAT.

Dr. Thorson noted that the Dirichlet approach, as currently implemented in SS 3.30, has the property that the lower bound for the effective sample size is one, which could have the unintended effect of upweighting small samples. However, the code could potentially be changed to allow for a consistent proportional change to the effective sample sizes, even if the resulting effective sample size would be less than one.

There was general consensus by review panel members that the Dirichlet multinomial approach seems to be at least as good as the MI method (mainly due the automation of the tuning and thus reproducibility). When setting initial composition sample sizes the STATs should use the number of tows (for survey data) or trips (for fishery data), or a set of initial values based on a composite of the numbers of fish sampled and the number of tows or trips.

The Groundfish Subcommittee recommends that STATs by default use the Francis method (TA1.8) for weighting age-, and length-, and conditional age-at-length compositional data. Assessment documents should include sensitivity runs that use (a) the MI harmonic mean weighting approach as well as (b) the Dirichlet multinomial likelihood approach, as a mechanism to gauge the uncertainty associated with the choice of methodology.

C. Application of the Generalized Linear Mixed Model with Spatial Autocorrelation to Survey Data

Dr. Jim Thorson (NWFSC) gave a presentation on the application of generalized linear mixed model (GLMM) with spatial autocorrelation to survey data. The geostatistical GLMM approach is similar to what has been used during several assessment cycles for the analysis of bottom trawl survey data series. Predicted fish biomass density is derived as the product of a "delta" portion for the probability of a non-zero catch and a second portion for the magnitude of the non-zero catches. Further, the geostatistical GLMM framework can accommodate spatial autocorrelation. Additional information about the approach and the software package it is implemented in are available from www.fishstats.org. Dr. Thorson indicated that he plans to phase out the single-species version of the software in favor of a multispecies version called VAST (vector autoregressive spatial temporal model, where the single species version remains as a special case of the multispecies version).

Dr. Thorson's software package makes use of special purpose software that implements the integrated nested Laplace approximation (INLA) approach, which he described as the ADMB for spatial correlation models. INLA uses a simplification for representing the variance of multivariate normal distributions to reduce complexity and speed up computer processing. A triangulated mesh is used to represent the relationships between points in space. There was discussion regarding why triangulation would be a better approach than a gridding approach, and on what basis a user would choose the number of "knots" to use for the triangulated mesh.

Dr. Thorson raised the problem of "the garden of forked paths" in reference to how an analyst could predetermine the outcome of an analysis by their selection of how many knots they include in an analysis. He cautioned that the SSC should be suspicious of analyses based on a small number of knots. He suggested that a solution would be for the SSC to pre-specify the minimum

number of knots used in any analysis that develops indices for assessments. Having too many knots is costly in terms of the amount of computer time needed to complete an analysis.

There was discussion about using depth (or some other auxiliary data) to influence the placement of the knots or as a covariate and whether aspects of the sampling design can be ignored in geostatistical models (because they have littler influence on the results). Dr. Thorson identified the problem of preferential sampling, as likely occurs in fishery-dependent data, versus design-based sampling. Whether or not depth should be included in a geostatistical analysis of survey data remains a topic of research. The geostatistical GLMMs conducted in the last assessment cycle for the canary rockfish and darkblotched assessments did not directly include depth information (but depth was used to define stratum boundaries).

There was discussion about including vessel as a random effect in the analysis of survey data given that the same vessels have been contracted for the survey for many years. It was noted that although the vessels have not changed much there have been changes in skippers and the vessel gear such as winches. Dr. Cadigan recommended that assessment analysts and reviewers confirm there are no temporal trends or other unusual patterns in effects that should be random, such as the vessel.

With regard to model diagnostics Dr. Thorson provided examples of residual plots by year and over space that are now produced by his software package. There was discussion of how residual plots or other diagnostics should be interpreted as indicating possible problems in the model. Large spatial areas having large numbers of large positive or negative residuals, or temporal patterns in clumps of residuals might be interpreted as indications that the model is missing some key spatial feature. Dr. Thorson suggested Q/Q plots are not always a very informative diagnostic but whether or not the Hessian has converged is a key diagnostic.

Dr. Thorson presented results from a simulation study that showed that the geostatistical GLMM approach can capture upward, downward and stable trends in biomass. It was noted that geometric anisotropy is important on our coast, but the effect varies depending on whether a species is southerly or northerly distributed. There was discussion about whether analyses of survey data should use an optional feature in the software that corrects for bias associated with nonlinear functions that have random effects. The bias-correction feature is memory-intensive. Tests of the bias-correction have shown it to produce results that are consistent with even more computationally intensive MCMC sampling.

The Groundfish Subcommittee recommends that the geostatistical GLMM software developed and maintained by Dr. Jim Thorson should be considered as the first choice for developing biomass indices from bottom trawl survey data, though exploration of other methods is encouraged. Dr. Thorson will provide a document describing recommended defaults and practices for using the software.

D. Application of the Generalized Linear Mixed Model with Spatial Autocorrelation to Fishery CPUE Data

Dr. Jim Thorson gave a presentation on the application of the geostatistical GLMM approach to fishery CPUE data, based on the VAST software package that he has developed. As in the application to fishery independent (survey) data, the model combines a component for the probability of a zero density with a component for spatio-temporal variability in density to

predict total density across space and time. In the example applications presented, years were treated as being independent (not autocorrelated). Temporal autocorrelation can be included in the models but this feature is not recommended for use at present. One particularly novel aspect of the package is that it can accommodate multispecies applications, in which model estimates of covariance among species provide information about the expectation of a given species in a given observation. In the example provided, Atlantic cod and haddock were shown to have positive covariation with respect to habitat preferences (spatial variation) and in their annual response to environmental signals (spatio-temporal variation), such that the catch rates of one species helped inform the expected catch rates for the other species. Dr. Thorson suggested that the approach might also be applied in the future for standardizing compositional data, with cohorts being treated as "species".

Dr. Thorson's suggested that the software package could be applied to the analysis of fisheries dependent data, particularly for the exploration of multispecies data sets that include detailed spatial information. The assessments planned for California Scorpionfish and blue/deacon rockfish might benefit from an exploration with the software of the CDFW onboard CPFV observer data; the assessment for yellowtail rockfish might benefit from an exploration of the bycatch CPUE data from the at-sea whiting fishery. There was discussion of the challenges and merits associated with the analysis of CPUE data from commercial logbooks, such as accounting for the effects of regulations, vessels, skippers and gear. After some debate there was general agreement that the VAST software would likely not replace the Stephens-MacCall approach that has been used in several Council assessments to filter dockside intercept data and identify trips that could plausibly have caught the species of interest. The issue with applying the VAST software to the dockside intercept data is that these are trip-level data that do not have spatial information associated with them.

A number of additional features of the VAST software relevant to the analysis of fishery dependent data were also discussed, such as the decomposition of the covariance in catchability into portions thought to be controllable (vessel behavior) and those that are not controllable (e.g., gradual changes in the fleet). Many of these features in the software and the statistical approach it implements have not been fully developed but are areas of active exploration. One point of discussion was that if an analyst wants to apply the VAST software to fisheries dependent data, the analyst should first evaluate the spatial and temporal extent of the data to understand if there are large gaps in coverage. Most of the features in the VAST software and approach are fairly consistent with methods developed and applied in the past (e.g., the underlying geostatistical GLMM has been used for index development in the past). However, some aspects are novel and untested (e.g., modeling the targeting behavior of fishers) and will require building a foundation of models of growing complexity in the review process. There are some non-trivial impediments in using fishery dependent data to develop indices (e.g., accounting for changes in fishing power).

The Groundfish Subcommittee does not recommend at the present time use of the VAST software for developing abundance or biomass indices based on fishery logbook data series or dockside interviews. However, the software could be a useful tool for analysis and standardization of data series based on information from at-sea observers.

E. Revised Set of Priors for Natural Mortality

Dr. Owen Hamel (NWFSC) gave a presentation of a revised prior for natural mortality (*M*). The new prior is a revision to a prior developed and described in Hamel (2015), based on an approach and information from Then et al. (2015). The presentation considered five published meta-analyses that examined relationships between *M* and life history correlates such as maximum age, the von Bertalanffy growth coefficient and asymptotic size, water temperature and a gonadosomatic index.

The approach recommended by Dr. Hamel is a simplification of the model for *M* as a non-linear function of maximum age derived by Then et al. (2015). The Hamel prior, which takes a different approach than Then et al. to account for variability and provides a much better fit to the underlying data, is based on the relationship $M = 5.4 / A_{max}$, where A_{max} is the maximum age.

There was discussion of how the original studies (on which the meta-analyses were based) derived their estimates of natural mortality and the life history correlates. For example, one would expect that observations of maximum age would depend on how many fish were sampled and would be influenced by the properties of the sampling gear and ageing error. It was pointed out that estimates of *M* based on growth parameters could be biased from distortion of the growth curve due to size-based selection (e.g., smaller fish for a given length are under-represented). There was also discussion of whether the maximum ages used to derive *M* from the prior should be restricted to samples from the region of the assessment and whether derivations of *M* should consider differences in maximum age by gender.

The Groundfish Subcommittee considers it important that different assessments use consistent approaches for deriving values of *M* because this parameter is difficult to estimate but often has high influence on assessment results. The Subcommittee recommends that (a) groundfish assessments during the 2017 assessment cycle report the prior probability distribution for natural mortality based maximum ages as updated by Dr. Owen Hamel and (b) that STATs explore using the prior to inform the assessment models. Further, the maximum age values on which *M* priors are based should be from fish caught within the area of the assessment (e.g., not from Alaskan catches of the same species). If a prior for *M* is used to provide a fixed value for *M*, the fixed value should be set equal to the mean value of the prior.

F. Revised Prior for Steepness

Dr. Jim Thorson (NWFSC) gave a presentation on a revised prior for rockfish steepness for use in this year's groundfish assessment cycle. This is the sixth iteration of specifying a prior for steepness based on likelihood profiles from previous rockfish assessments. Providing recommendations to assessment authors on a prior for steepness is helpful to ensure (a) that assessments use a consistent approach for specifying this important parameter and (b) that the approach is based on a synthesis of available information. Seven of the 12 likelihood profiles used in the revised meta-analysis were from new or revised assessments. The re-run of the meta-analysis was straightforward and encountered no technical issues. The resulting mean of the prior was 0.72, which is decline from the mean steepness of 0.77 of the distribution in 2015, but still higher than when the meta-analysis was first done for the 2007 assessment cycle. The standard deviation of the prior decreased very slightly from 0.16 to 0.15. The Groundfish Subcommittee endorses the use of the updated prior for steepness in this year's rockfish stock assessments.

There are several possible ways that the prior for steepness could be used in the 2017 stock assessments. The Groundfish Subcommittee recommends the following procedures be followed:

1. If the stock assessment is not in the set of stocks used to estimate the steepness prior and you chose not to estimate steepness, then fix steepness at the mean of the predictive distribution.
2. If the stock assessment is not in the set of stocks used to estimate the steepness prior and you chose to estimate steepness, then use the mean and standard deviation of the predictive distribution as the mean and standard deviation of your prior on steepness.
3. If the stock assessment is in the set of stocks used to estimate the steepness prior and you chose not to estimate steepness, then fix steepness at the mean of the predictive distribution.
4. If the stock assessment is in the set of stocks used to estimate the steepness prior and you chose to estimate steepness, use a "Type-C" value that is recalculated while excluding that stock (for 2017 assessments, this will apply to Pacific Ocean Perch and Yellowtail). This ensures that the prior distribution does not "double count" data for that stock when estimating steepness (Minte-Vera et al. 2005). For Type-C priors, assessment authors should contact James Thorson with at least one month lead prior to when the value is needed.

The Groundfish Subcommittee also requests that Dr. Thorson report values of steepness at each quantile from the predictive distribution from 0.025 to 0.975 (separated by 0.05) as well as the value of steepness at the following set of quantiles of the predictive distribution {0.05, 0.25, 0.75, and 0.95}. Quantiles may be used by STAT teams to define values of steepness used in decision tables, or they could be used to integrate over the distribution of steepness to more fully characterize assessment uncertainty. To gauge the utility of the approach the Groundfish Subcommittee is interested in reviewing several examples where this approach has been explored, but the Subcommittee does not request that it be done for all assessments in the upcoming cycle.

The Groundfish Subcommittee notes that concerns have been raised previously (and at the methodology review) about the reliability of the meta-analysis approach used to develop the steepness prior. One concern, that ignoring autocorrelation in recruitment within stocks is causing bias in steepness estimates, was addressed by a paper discussed by Dr. Thorson at the productivity workshop. His work indicated that biases in steepness caused by autocorrelation tend to be relatively minor; this finding helps to alleviate this particular concern.

Other concerns include potential correlation in recruitment across stocks due to similar environmental forcing, and potential bias in the estimation of steepness and the steepness profile. The Groundfish Subcommittee encourages research to address these issues, some of which is underway. In particular, the Groundfish Subcommittee recommends continuation of the research being done at the SWFSC to evaluate the conditions under which stock assessment models produce reliable estimates for steepness and the steepness profile.

G. New Features in the Revised Stock Synthesis Software

Drs. Chantel Wetzel and Teresa A'mar (NWFSC) presented information on the new version (3.30) of the Stock Synthesis (SS) software that is now available. Documentation and the software can be found at the secure website https://vlab.ncep.noaa.gov/home. Individuals who

are not part of NOAA need to request access to the website by sending their email address, name, phone number, and organization to nmfs.stock.synthesis@noaa.gov.

The new SS version has many features that will make it easier to develop input files, such as providing more flexibility in how fleets are specified and list-oriented inputs that use an end-of-data indicator rather than requiring specification of the number of data rows. New modeling features of version 3.30 include the ability to specify one to many recruitment events within a year, a catch multiplier that can use time-blocks or any other time-varying approach to scale catches (e.g., to explore the effects of uncertainty in historical catch series), an option to use the Shepherd stock-recruit function, time-varying stock-recruitment parameters that could be used to model regime shifts, an option for dome-shaped retention functions (as an alternative to the logistic form), the ability to include autocorrelation in parameters that have deviations, and an option to use the Dirichlet multinomial for representing sampling variability in compositional data. The SS 3.30 User Manual, which is available from the website, describes all of the changes and new features.

A translation routine is available that will convert files constructed for use with SS version 3.24 to formats that will run with version 3.30. Some aspects of the automatic conversion may require the analyst to make additional changes.

The new version of Stock Synthesis has undergone extensive testing. Model comparisons between SS 3.30 and SS 3.24 were successfully produced for sixteen assessments. Some features of these models required manual conversion (e.g. time-varying selectivity bounds for retention, time-varying *Q* setup). The testing conducted so far has found no more than 5% differences in results for depletion, SB_0, and SB_{final}. Outside of the assessment for Pacific hake, the largest differences have been less than 2%. Differences of this small magnitude are not totally unexpected. The two versions of SS do not use identical internal calculations. One important change is in how the two versions do age-length-key calculations. Another difference is in the transformation of bounds of time-varying parameters.

Drs. Wetzel and A'Mar recommended the following approach be used for bridging an existing SS assessment model from version 3.24 to version 3.30. The file format conversions should be done before any new data are added, then step from the old SS version to the new SS version, and then add new data. Finally, add new features, parameters, or options, step by step in any order.

There is no requirement to show perfect agreement between results when converting to the new version. This is especially the case for very old assessments (e.g., yellowtail rockfish, blue rockfish, and California scorpionfish). However, it is important for analysts who are bridging to the new version to try and understand the source(s) of the differences.

During discussion there was a question about whether the new version has fixed a problem identified during the last round of assessment that the minimum effective sample size for composition data could not be reduced below one, but it was unclear whether this had been taken care of yet. There was also a question about whether there was an option available in the new version for the Ricker stock-recruitment relationship with a power parameter.

H. Review of the Draft "Accepted Practices Guidelines for Groundfish Stock Assessments

During the final hours of the meeting the group reviewed the draft Accepted Practices Guidelines document and made some changes to the text. The proposed changes are flagged in a revised version of the Guidelines that has been included in the Briefing Book for the March 2017 Council meeting.

References

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

Hamel, O.W. 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.

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. In Press. Model-based estimates of effective sample size in stock assessment models using the Dirichlet-multinomial distribution. Fisheries Research.

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.

Reviewers Present:

Dr. David Sampson, Oregon State University, SSC (Chair)
Dr. Aaron Berger, National Marine Fisheries Service (NMFS) / Northwest Fisheries Science Center (NWFSC)
Dr. John Budrick, California Department of Fish and Wildlife (CDFW)
Dr. Noel Cadigan, Memorial University, St. Johns, Newfoundland, Canada
Dr. Martin Dorn, NMFS / Alaska Fisheries Science Center
Dr. John Field, (NMFS) / Southwest Fisheries Science Center (SWFSC)
Dr. Owen Hamel, NMFS / Northwest Fisheries Science Center (NWFSC)

Attendees at the Meeting:

Dr. Teresa A'mar, NMFS / NWFSC
Mr. John DeVore, Pacific Fishery Management Council
Ms. Jin Gao, University of Washington
Dr. Melissa Haltuch, NMFS / NWFSC
Dr. Jim Hastie, NMFS / NWFSC
Ms. Stacey Miller, NMFS / NWFSC
Ms. Merrill Rudd, University of Washington / NMFS / NWFSC
Dr. Ian Taylor, NMFS / NWFSC
Dr. James Thorson, NMFS / NWFSC
Mr. John Wallace, NMFS / NWFSC
Dr. Chantel Wetzel, NMFS / NWFSC
Dr. Haikun Xu, NMFS / NWFSC

Attendees on WebEx:

Dr. E.J. Dick, NMFS / SWFSC
Ms. Jessi Doerpinghaus, Washington Department of Fish and Wildlife (WDFW)
Dr. Xi He, NMFS / SWFSC
Ms. Melissa Mandrup, CDFW
Dr. Melissa Monk, NMFS / SWFSC
Mr. Corey Niles, WDFW
Dr. Andi Stephens, NMFS / NWFSC
Dr. Theresa Tsou, WDFW

PROPOSED AGENDA
Groundfish Assessment Methods Review
Pacific Fishery Management Council
Scientific and Statistical Committee's
Groundfish Subcommittee

January 25	January 26
National Marine Fisheries Service Northwest Fisheries Science Center The Auditorium 2725 Montlake Boulevard E Seattle, WA 98112	Seattle Yacht Club Heritage Room 1807 E Hamlin St. Seattle, WA 98112

January 25-26, 2017

This meeting is open to the public and public comments will be accepted at the discretion of the meeting chair. Agenda times are approximate and are subject to change.

WEDNESDAY, JANUARY 25, 8:30 AM

A. Call to Order

1. Call to Order and Introductions — David Sampson
2. Approve Agenda and Rapporteur Assignments
   (*8:30 a.m., 0.5 hours*)

B. Use of the Dirichlet Multinomial Likelihood for Compositional Data — Jim Thorson
(*9 a.m., 1.5 hours*)

C. Application of the Generalized Linear Mixed Model with Spatial Autocorrelation to Survey Data — Jim Thorson
(*10:30 a.m., 1.5 hours)*

LUNCH (12 P.M. - 1 P.M.)

D. Application of the Generalized Linear Mixed Model with Spatial Autocorrelation to Fishery CPUE Data — Jim Thorson and Melissa Monk
(*1 p.m., 2 hours)*

E. Revised Set of Priors for Natural Mortality — Owen Hamel
(3 *p.m., 2 hours*)

THURSDAY, JANUARY 26, 8:30 A.M.

F. Revised Prior for Steepness — Jim Thorson
(*8:30 a.m., 1.5 hours*)

G. New Features in the Revised Stock Synthesis Software Chantel Wetzel and Teresa Amar (*10 a.m., 2 hours*)

H. Review of the Draft "Accepted Practices Guidelines for Groundfish Stock Assessments" (*1 p.m., 2 hours*)