ACCEPTED PRACTICES GUIDELINES FOR COASTAL PELAGIC SPECIES STOCK ASSESSMENTS IN 2024

The guidelines in this document are intended to supplement the Council's *Terms of Reference for Coastal Pelagic Species Stock Assessment Review Process.* The guidelines provide coastal pelagic species (CPS) stock assessment teams (STATs) with default approaches for stock assessment data and modeling issues. The STATs may diverge from the guidelines prior to stock assessment review (STAR) panel or other review meetings, if adequate justification for doing so is provided. These guidelines are not intended to provide a comprehensive treatment of all potential issues. Rather the guidelines focus on issues that the Scientific and Statistical Committee (SSC) has so far considered. The purpose of these guidelines is to provide advice about how particular steps in the assessment process should be conducted.

Data Considerations

Removals Data

The STAT should obtain landings (and discard) data from all sources, including commercial, recreational and charter boat fisheries in Mexico, the USA and Canada, with USA catches taken from the Pacific Fisheries Information Network (PacFIN) and the Recreational Fisheries Information Network (RecFIN). Discards in non-directed fisheries that take CPS (including those stored in the At-Sea Hake Observer Program database) should be included in the catch history.

It is common for catches to be unavailable for specific regions and/or time periods. In these cases, the STAT should use the most recent observations (specific to region and time period) to fill in missing values.

Compositional Data

When combining compositional samples from different geographic regions into a single fleet, the fishery-dependent composition proportions should be weighted by catch weights. Fishery age-composition data are weighted by the total monthly landings (L_m). For example, port samplers off California generally biologically sample 25 individual fish per landed haul, and the input sample sizes for assessment purposes are usually the number of total sampled fish by time period divided by 25 The steps below are used to develop the weighted age-composition data (assuming the full 25 fish sample was obtained):

- Enumerate the number of individual fish (n) sampled in each month (m), age (a), and calendar year (y): $n_{m,a,y}$
- Sum the total biological sample weights (B) by month and calendar year ($B_{m,y}$), and calculate the mean weight ($\underline{w}$) of sampled fish by month, age and calendar year ($\underline{w}_{m,a,y}$).
- Calculate proportions in the biological samples (A) by month, age and calendar year:

$$A_{m,a,y} = \frac{\underline{w}_{m,a,y} * n_{m,a,y}}{B_{m,y}}$$

- Calculate the total landings (L) by month, age and calendar year:

$$L_{m,a,y} = A_{m,a,y} * L_{m,y}$$

- Calculate number of fish (F) by month, age and calendar year:

$$F_{m,a,y} = L_{m,a,y} / \underline{w}_{m,a,y}$$

- Sum by age and model time period (T), which spans month 1 (m_1) of year 1 (y_1) to month 2 (m_2) of year 2 (y_2). The months and calendar years corresponding to T vary by species. Generally, the model year aligns with the fishing year, which starts in the summer of calendar year y and ends in the summer of calendar year $y+1$. For example, for Pacific Sardine, the data for model year 2005 are summed by semester: S1, which spans July-December of calendar year 2005 and S2, which spans January-June of calendar year 2006. For Northern Anchovy, data for model year 2005 are also summed by semester: S1, which spans June-December of calendar year 2005, and S2, which spans January-May of calendar year 2006. Pacific mackerel has an annual time step, and model year 2005 sums data from July of 2005 to June of 2006.

$$F_{a,T} = \sum_{z=m1,y1}^{m2,y2} F_{a,z}$$

- The final proportion P by age and time period is normalized across ages 0 to the plus group age ($maxA$).

$$P_{a,T} = F_{a,T} / \sum_{z=0}^{maxA} F_{z,T}$$

Age-length keys are an acceptable practice for generating fishery-independent age-composition data (i.e., acoustic-trawl survey age compositions). Estimates of abundance-at-length are converted to abundance-at-age with survey-specific age-length keys for summer survey data. Depending on the species, these data may come from spring and summer surveys or from commercial samples (Crone et al., 2019), although the goal is to generate survey–specific age-length keys, when possible. Ideally, age data should come from samples collected close in space and time to the collection of length data.

In the case of low biological sample sizes, the age-length key can be generated from multiple surveys at the discretion of the STAT. Pooling age and length data across years should be limited to adjacent years, when possible. The age-length keys are constructed using ordinal generalized additive regression models. A generalized additive model with an ordinal categorical distribution fits an ordered logistic regression model in which the linear predictor provides the expected value of a latent variable following sequentially ordered logistic distributions. Fishery-independent age compositions are weighted (i.e., input sample size in Stock Synthesis) by the number of positive clusters in each survey.

Constructing Indices of Abundance

The California Current Ecosystem Survey (CCES), also known as the Summer CPS survey, is scheduled to take place annually and applies an acoustic-trawl method (AT) to typically cover waters ranging from at least Cape Flattery, WA to the US-Mexico border (Renfree et al., 2023). Coverage in Canada, to the northwest end of Vancouver Island, and Mexico, to Punta Eugenia, Baja CA, varies among years. The SSC has deemed that the AT survey is appropriate to estimate the biomasses of the management unit finfish CPS in the California Current: Pacific Sardine, Pacific Mackerel, Jack Mackerel, and Northern Anchovy. The sampling domain is likely to encompass the entire distribution of the northern stocks of Northern Anchovy and Pacific Sardine, as well as a variable portion of the southern stock of Pacific Sardine, the central stock of Northern Anchovy, Pacific herring, and both mackerels.

Nearshore Biomass
The AT survey has had three methods of extrapolating or observing nearshore biomass: model extrapolation, unmanned surface vehicles, and fishing vessel acoustic-trawl methods (Stierhoff et al., 2020). Aerial surveys can also estimate nearshore biomass. For anchovy and sardine assessments, estimates from one of the nearshore methods discussed below are added to biomass estimates associated with the core survey area. The four accepted approaches for estimating inshore biomass are:

1. Fishing vessel acoustic-trawl methods involve equipping vessels with acoustic echosounders and conducting nominally one purse seine set per transect, generally during daylight hours. Sets are conducted at night in the case of abundant coastal pelagic species or an unsuccessful daytime set. Weights and lengths are recorded and otoliths collected for up to 50 randomly selected specimens of target species, prioritizing Pacific Sardine and Northern Anchovy. This survey protocol and the subsequent biomass calculation most closely matches the methods used in the core grid of the AT survey.
2. With model extrapolation, the easternmost portions of transects are extrapolated to the 5-m isobath in the unsampled nearshore areas. Thus, the length and species compositions associated with the end of the transects are extrapolated to the 5-m isobath.
3. Unmanned surface vehicles (USVs) generally cover portions of the coast rather than the entire coast. The ability to collect USV observations has depended on the number of USVs available for use and on local wind conditions. The USVs collect acoustic data but do not collect associated biological samples. As a result, the nearest trawl compositions in space are assumed to be representative of the nearshore acoustic observations when calculating species-specific biomass values.
4. The California Department of Fish and Wildlife has conducted an aerial survey off the coast of central and southern California. There are age-composition data associated with the aerial survey.

Modeling

Prior Distributions for Natural Mortality (M)
Assessments for CPS species should report the prior probability distribution for natural mortality (M) computed using the updated meta-analytical approach (Hamel and Cope, 2022) based on maximum ages (Hamel, 2015; Then et al., 2015) at a minimum (other approaches can also be considered e.g., age-specific M or another method when maximum age is not reliably estimated), and STATs should explore using the Hamel-Cope prior to inform the assessment models. This prior is defined as a lognormal distribution with median value (corresponding to the mean in log-space) of 5.40/maximum age and log-scale sigma of 0.31. The M parameter 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. If a 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 Sex-specific M
The Lorenzen approach (Lorenzen, 1996, 2022; Methot and Wetzel, 2013) should be the default modeling approach for assessment models with age-specific *M*. If the Lorenzen approach is used to model age-dependent *M* the assessment should also present a comparison run that uses constant *M* (i.e., no age-dependence).

Currently, CPS do not display strong sexual dimorphism. If data support modeling sexes separately, 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: the McAllister and Ianelli (1997) harmonic mean approach; the Francis (2011) approach; and the Thorson et al. (2017) 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 been used less frequently or has yet to be used extensively. There is no clear consensus that one approach is superior in all circumstances. The Francis method has become the most used method and provides a basis for comparison to the other methods in evaluating the preferred method for the stock in question. 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 may need to conduct multiple iteration steps (usually two to three) for the McAllister-Ianelli and Francis methods 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 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.

Weight-At-Age
Fishery empirical weight-at-age values are calculated by nominally averaging weights for each age. If age-0 values are missing, the average age-0 value across available years serves as the substitute. Missing values between ages are linearly interpolated by cohort. If values are missing above a certain age (for example ages 6, 7, and 8), the last observation is used to fill the data gap (age 5 value).

Two of the outputs of the AT surveys are abundance-at-length and biomass-at-length (Zwolinski et al., 2019). Calculations of abundance-at-age, biomass-at-age, and weight-at-age rely on the constructions of age-length keys. An age-length key (ALK) is a model that describes the probability of a fish of a known length belonging to an age-class (Stari et al., 2010). ALKs are used often to calculate abundance and catch-at-age from fisheries-dependent and -independent sources (e.g., Kimura, 1977; Clark, 1981; Hoenig and Heisey, 1987; Robotham et al., 2008). Their use is common when only a subsample of all the fish sampled for lengths are aged. The use of an ALK relies on the assumption that the conditional distribution of ages given length in the subsample is representative of that in the population (Kimura, 1977; Westrheim and Ricker, 1978). The sampling scheme to build an ALK requires a sufficient number of individuals to estimate the conditional age-distribution over a set of fixed length intervals.

When the number of individuals sampled for age is large, an empirical age-length key can be built by computing the proportion of individuals of all ages across all discrete length classes (Ailloud and Hoenig, 2019). However, when sample size is small and there is ageing error, empirical age-length keys might be dominated by error (Stari et al., 2010). In these cases, creating a smooth ALK relying on some sound underlying process is preferable (e.g., Martin and Cook, 1990; Berg and Kristensen, 2012).

There are numerous analytical approaches to build smooth or model-based ALK (e.g., references above; Stari et al., 2010 and references therein). An approved method is to assume for ages a (in years) such as, for example $a \in \{0,1,\ldots,6+\}$, the probability distribution conditioned on length l, $P_a(l) = \{p_0(l), p_1(l), \ldots, p_{6+}(l)\}$, follows an ordered categorical distribution. $P_a(l)$ could be modeled using the *gam* function in the *mgcv* package (Wood et al., 2016), with distribution *ocat.* Below is a brief explanation of the model fitting in R.

For a data set with a variable *age.ordinal* – coded by natural numbers from 1 to 7, corresponding to ages 0, 1, 2, … 6+ years, and *standard.length* – coded as a continuous variable in mm, the *gam* model can be fitted by:

```
R = 7 # number of age categories
model <- gam(age.ordinal ~ s(standard.length) , data = data , family= ocat(R= R)) # the ordinal model as
    smooth function of length
```

and the resulting ALK can be created by:

```
prob.matrix <- predict(model, newdata = data.frame(standard.length = seq(20,200, by =10)), type =
    "response")
```

which results in a 19 by 7 matrix in which each row is the estimated vector of probabilities $P_a(l)$ of a fish of length l (in cm) with $l \in \{2,3,\ldots,20\}$ belonging to an age group a, with $a \in \{0,1,\ldots,6+\}$. Considering a vector of abundances at length $N_l = n_2, n_3, \ldots, n_{20}$, the elements of vector of abundances-at-age N_a are calculated by $n_a = \sum_{l=2}^{20} P_a(l) n_l$. Similarly, the elements of biomass at age B_a are given by $b_a = \sum_{l=2}^{20} P_a(l) n_l w_l$, where w_l is the average weight of an animal in the l-th length class derived from a length-to-weight relationship. Finally, mean weight-at-age is obtained by dividing B_a by N_a.

Specification of, and Priors for, Survey q

The survey catchability (q) can be specified for specific model periods. Example approaches are:

- The ratio of the US biomass to the biomass of US and Mexican waters (for summer AT surveys).
- The ratio of biomass estimates between the spring and summer AT surveys (for spring AT surveys).

- The ratio of offshore biomass to inshore plus offshore biomass (e.g., with inshore biomass defined by the aerial survey) for surveys missing inshore estimates of biomass.

The STAT should consider all available data and explore alternative q calculations in developing base models. Future configurations may or may not be identical to these examples depending on future data sets. The STAT should attempt to fully document the uncertainty associated with survey q (e.g., by using bootstrapping).

Diagnostics
In addition to the standard set of likelihood profiles identified in the CPS Stock Assessment Terms of Reference (across the parameters $\ln(R_0)$[1], M, and steepness), the STATs should consider other diagnostics, such as those highlighted in Carvalho et al. (2017) as well as a likelihood profile for terminal year biomass.

Including Extra Variability Parameters with an Index
STATs should be cautious to avoid adding extra variability to an index as a means of resolving model structure issues such as conflicts among data sources. When adding additional variance to indices, one should look for possible over-inflation of the added variance due to conflicts with other data (e.g., biological compositions). In those instances, it may be more appropriate to determine what data sources contain the most representative population signal and justify the need to add more variance to index values, and conduct sensitivity analyses to assumptions about which data sets and types are most representative.

Jittering to Verify Convergence
Jittering involves changing the initial values of parameters and re-applying the estimation method to evaluate whether convergence to a local minimum in the negative log-likelihood may have occurred (i.e., different jitter runs should ideally converge to the same or nearly identical parameter estimates). 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 fraction 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.

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

To evaluate the jittering, the bounds, and the original initial values, a jitter info table is available from r4ss (an R package), 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 estimate 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. Alternative phase sequences can have an impact on parameter estimation, negative log-likelihood minimization, and model convergence. STATs should consider alternative phase sequencing as a model diagnostic tool in addition to jittering.

Forecast Configuration

The STAT should input future catches as fishing mortality rates rather than catch biomass. Recruitment predictions should be based on the stock-recruit relationship, and selectivity curves should represent an average of recent estimates. The impact of assuming alternative selectivities on forecast biomass should be evaluated.

Applying Harvest Control Rules

Future Removals

The catch data for the last year of the assessment is usually incomplete and OFLs, ABCs and HGs are needed for the year after the assessment and often additional future years. Unobserved or projected catches are typically set to the mean of annual catches estimated over the period since the last full assessment.

Determining the Value of Sigma

A potential alternative to the category-specific default sigma is one based on the assessment's internal estimate of uncertainty. Ideally, uncertainty should be calculated for the OFL rather than biomass or spawning output alone. For CPS, the OFL uncertainty calculation should be for the OFL in the year the assessment was conducted and would first be adopted for management. In this case, the comparison should be against the updated category-specific defaults of 0.5 and 1.0 for categories 1 and 2, respectively. If uncertainty is calculated for biomass (ideally summary or age-1+ biomass as opposed to total biomass in the case of CPS), the defaults should be 0.36 and 0.72, respectively. If the biomass uncertainty exceeds the category-specific biomass default sigma, the OFL sigma should be adjusted up from the OFL sigma default by the same proportion. For example, a biomass sigma of 0.48 (33% higher than 0.36) should lead to an OFL sigma of 0.67 (33% higher than 0.50). When evaluating uncertainty based on arithmetic scale CVs for comparison against log-scale sigmas, note that sigma=$\sqrt{\log(CV^2+1)}$, where log is the natural (base e) logarithm.

Additions Identified for Future Consideration

- Given the linkage between the input sample size and the Dirichlet Multinomial data-weighting approach, future research should be conducted to provide improved guidance on developing input sample size for weighting compositional data (particularly for the Dirichlet Approach).
- Explore the use of Markov chain Monte Carlo (MCMC) runs for assessments to explore uncertainty in a probabilistic fashion, akin to what is currently being provided in the Pacific

Whiting stock assessment report. The time it takes to run an MCMC may be time prohibitive for benchmark assessments given the compressed time frame between getting final data and document deadlines as well as issues with running alternative model configurations during a review. Application to update assessments may be more reasonable given the few changes and less consideration in need of evaluation.

- Investigate a truncated age structure. These are short lived species and most of the biomass is in ages 0 – 3. Assess stock assessment outcomes if all ages >= 4 are lumped in 1 age category.

References

Ailloud, L.E., and Hoenig, J.M. 2019. A general theory of age-length keys: combining the forward and inverse keys to estimate age composition from incomplete data. ICES Journal of Marine Science 76: 1515-1523.

Berg, C.W., and Kristensen, K. 2012. Spatial age-length key modelling using continuation ratio logits. Fisheries Research 129: 119-126.

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.

Clark. W.G. 1981. Restricted least-squares estimates of age composition from length composition. Canadian Journal of Fisheries and Aquatic Sciences 38: 297-307.

Crone, P.R., Hill, K.T., Zwolinski, J.P, and Kinney, M.J. 2019. Pacific mackerel (*Scomber japonicus*) stock assessment for U.S. management in the 2019-20 and 2020-21 fishing years. Pacific Fishery Management Council. 112 p.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 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 Journal of Marine Science 72: 62-69.

Hamel, O.S., and Cope, J.M. 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. Fisheries Research https://doi.org/10.1016/j.fishres.2022.106477.

Hoenig, J.M., and Heisey, D.M. 1987. Use of a log-linear model with the em algorithm to correct estimates of stock composition and to convert length to age. Transactions of the American Fisheries Society 116: 232-243.

Kimura, D.K. 1977. Statistical assessment of age-length key. Journal of the Fisheries Research Board of Canada 34: 317-324.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology 49: 627–647.

Lorenzen, K. 2022. Size- and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. Fisheries Research 255: 106454.

Martin, I., and Cook, R.M. 1990. Combined analysis of length and age-at-length data. ICES Journal of Marine Sciences 46:178-186.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences 54: 284–300.

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

Renfree, J.S., Beittel, A, Bowlin, Noelle M., Erisman, B.E., James, K, Mau, S.A., Murfin, D.W., Sessions, T S., Stierhoff, K.L., Vasquez, L., Watson, W., Zwolinski, J.P., and Demer, D.A. 2023. Report on the Summer 2022 California Current Ecosystem Survey (CCES) (2207RL), 27 June to 30 September 2022, conducted aboard NOAA ship Reuben Lasker, fishing vessels Lisa Marie and Long Beach Carnage, and uncrewed surface vehicles. NOAA technical memorandum NMFS-SWFSC; 678.

Robotham, H., Young, Z.I., and Saavedra-Nievas, J.C. 2008. Jackknife method for estimating the variance of the age composition using two-phase sampling with an application to commercial catches of swordfish (*Xiphias gladius*). Fisheries Research 93: 135-139.

Stari, T., Preedy, K.F., McKenzie, E., Gurney, W.S.C., Heath, M.R., Kunzlik, P.A., and Speirs, D.C. 2010. Smooth age length keys: Observations and implications for data collection on North Sea haddock. Fisheries Research 105: 2-12.

Stierhoff, K.L., Zwolinski, J.P., and Demer, D.A. 2020. Distribution, biomass, and demography of coastal pelagic fishes in the California Current Ecosystem during summer 2019 based on acoustic-trawl sampling. NOAA Technical Memorandum NMFS-SWFSC-626.

Stewart, I.J., and Hamel, O.S. 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.

Westrheim, S.J., and Ricker, W.E., 1978. Bias in using an age-length key to estimate age-frequency distributions. Journal of the Fisheries Research Board of Canada 35:184-189.

Wood, S.N., Pya, N., and Safken, B. 2016. Smoothing parameter and model selection for general smooth models. Journal of the American Statistical Association 111 :1548-1563.

Zwolinski, J.P., Stierhoff, K.L., and Demer, D.A. 2019. Distribution, biomass, and demography of coastal pelagic fishes during summer 2017, estimated from acoustic-trawl sampling. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-610.