Agenda Item G.2
Attachment 1
June 2015

Pacific Mackerel

Stock Assessment Review (STAR) Panel Meeting Report

NOAA / Southwest Fisheries Science Center
La Jolla, California
April 27-29, 2015

STAR Panel Members:
André Punt (Chair), Scientific and Statistical Committee (SSC), Univ. of Washington
Will Satterthwaite, Scientific and Statistical Committee (SSC), NOAA Fisheries
Vivian Haist, Center for Independent Experts (CIE)
David Checkley, UC San Diego

CPSMT/CPSAS Advisors to STAR Panel
Diane Pleschner-Steele, CPSAS
Alan Sarich, CPSMT

Pacific Fishery Management Council (Council) Representatives:
Kerry Griffin, Council Staff

Pacific Sardine Stock Assessment Team:
Paul Crone, NOAA / SWFSC
Kevin Hill, NOAA / SWFSC

1) Overview

The Pacific Mackerel Stock Assessment Review (STAR) Panel (Panel) met at the Southwest Fisheries Science Center, La Jolla, CA from April 27-29, 2015 to review a draft assessment by the Stock Assessment Team (STAT) for Pacific Mackerel. Introductions were made (see list of attendees, Appendix 1), and the agenda was adopted. A draft assessment document and background materials were provided to the Panel in advance of the meeting on a PFMC FTP site.

Paul Crone and Kevin Hill presented the assessment methodology and the results from a draft assessment utilizing the Stock Synthesis Assessment Tool, Version 3.24s to the Panel. David Demer presented aspects of the methodology and results for the acoustic-trawl method (ATM) surveys for Pacific mackerel. The assessment report included several model runs. The major changes between the base model (D) in the draft assessment and the model selected during the last full assessment reviewed in 2011 (model XA) were (a) the California Recreational Fishery Survey (CRFS) index was excluded from the assessment, (b) the recreational and commercial catches were combined into a single fleet, and (c) the ATM index and its associated length-composition were introduced into the assessment. Model D also differed from model XA in terms of data weighting, the assumed standard deviation of the log-deviations about the stock-recruitment relationship, allowance for a bias ramp for the recruitment deviations, whether the coefficient of variation (CV) for the length-at-the-oldest age is estimated, and whether the initial fishing mortality is estimated.

The review and subsequent explorations of the assessment through sensitivity analyses were motivated primarily by the desire to find a model that fit the data adequately, made realistic assumptions regarding selectivity and was consistent with prior expectations for the catchability coefficient for the ATM survey (i.e. whether the estimates of biomass were realistic given auxiliary information). The Panel and STAT agreed on some changes to model D (denoted by models H1b and H3 depending on whether the ATM indices and length-composition data were (H1b) or were not (H3) included in the assessment). These changes were (a) the indices and length-compositions for the ATM surveys were split into spring and summer series, with selectivity "mirrored" for the two surveys but with different catchability coefficients for spring and summer (model H1b), (b) selectivity for the ATM survey was assumed to be a logistic function of age (model H1b), (c) selectivity for the CPFV index was a 2-parameter (logistic) rather than 3-parameter (double-normal) asymptotic function (models H1b and H3), and (d) whether the ATM indices and length-composition data should be included in the assessment (models H1b vs H3). The STAT provided the Panel with models in which the fishery selectivity was asymptotic and dome-shaped. However, the model-estimated ATM survey catchability was unrealistically high (~1) when fishery selectivity was asymptotic, and unrealistically low (often <0.01) when fishery selectivity was dome-shaped.

Ultimately, the STAT decided to propose model H3 which did not include the ATM survey data, as an interim model.

The fit to the ATM survey index and length-composition data was poor for all model configurations. The poor fit to the survey indices is perhaps related to the variable proportion of the stock biomass in the area surveyed (the issue of spatial coverage is also a problem for the CPFV index, but large inter-annual variation is not apparent for the CPFV index). The poor fits to the survey length-composition data may be related to the very small number of mackerel encountered in trawls and measured (as low as three on one year).

In the time available, the STAT and Panel could not find a model configuration that fit the data, and led to plausible ATM catchability and hence biomass estimates. The question of how to estimate the 'scale'[1] of the biomass could not be resolved during the Panel meeting, with some model configurations fitting the data well but leading to implausibly large estimates of biomass.

The Panel highlights the research recommendations, in particular those related to improving the coverage of the ATM surveys. Only through changes to the survey design to cover more of the range of Pacific mackerel is it likely that the uncertainty associated with the assessment can be reduced.

The stock assessment exhibits a very marked retrospective pattern (see Request 18 and Fig. 1). In particular, the estimates of 1+ biomass are consistently updated downwards with the addition of new information. For example, the estimate of 1+ biomass for 2009 from an assessment with a terminal year of 2009 is 485,135 metric tons (mt), but the estimate of 2009 1+ biomass based on the model proposed by the STAT (model H3) is only 66,763mt. The retrospective bias should be accounted for when the extent of scientific uncertainty is quantified by the SSC.

The Panel thanked the STAT for their hard work and willingness to respond to Panel requests, and the staff at the SWFSC La Jolla laboratory for their usual exceptional support and provisioning during the STAR meeting.

2) Requests to the STAT

<u>Day 1 requests made to the STAT during the meeting – Monday, April 27</u>

Request 1: Start from model D. Drop the 1985-89 CPFV length-composition data from all further analyses. Use a logistic function to model asymptotic selectivity in all further analyses. This is new model H1.

Rationale: The 1985-89 "length" data appear to be conversions from weight and not trustworthy. Using the logistic function for asymptotic selectivity should help address convergence concerns.

Response: Model outputs are robust to dropping the 1985-89 CPFV length data. The STAT felt that the 2-parameter logistic formulation of selectivity was too inflexible (never yielding a positive definite Hessian matrix) when applied to the fishery although it worked well when applied to the CPFV index (and also the ATM survey in the models where ATM selectivity was asymptotic). A double normal selectivity pattern, with three parameters fixed to force an asymptotic form, worked better for the fishery, and was used in later models. The logistic function was used for the CPFV index (and ATM survey in later models) in all cases when their selectivity was asymptotic.

Request 1a: As with request 1, but also treat the spring and summer ATM surveys as two separate surveys with two separate *q*'s (but shared selectivity). This is model H1a. If this model appears to converge and produces "reasonable" scale/*q*/etc., use model H1a as the basis for all further modifications requested today. Otherwise, use model H1 as the basis for all further analyses requested today.

Rationale: These are two different surveys that differ in timing and survey areas, and in some years their biomass estimates differ by an amount too large to plausibly represent a true change in total population biomass. It would not be unexpected that survey *q* would vary seasonally due to the changes in the spatial distribution of Pacific mackerel.

[1] The term 'scale' is used here to refer to absolute size of the biomass.

Response: Model H1a was somewhat unstable[2] (would not converge to the same likelihood/solution from multiple starting values). This model has dome-shaped selectivity for the ATM surveys (now split into spring vs. summer), and asymptotic selectivity for the fishery and CPFV index. It yielded ATM *q*'s of 2.4 for the spring surveys and 1.5 for the summer surveys. Fits to CPFV lengths do not appear visibly better than those for Model H1a, nor do the fits to CPFV index. The fits to ATM length-composition data are not good, and the model does not appear to track the ATM index. The STAT chose to run an additional model, H1b, which made all selectivities asymptotic (i.e. the change relative to model H1a is that the ATM survey has asymptotic selectivity). Model H1b seems more stable, and estimates ATM *q*'s of 1.3 for spring and 0.9 for summer. Model H1a has 58 parameters and a negative log-likelihood of 1188.9, while H1b has 54 parameters and a negative log-likelihood of 1198.4.

Request 2: Run a variant of model H1a with dome-shaped selectivity for the fishery and asymptotic (logistic) selectivity for the surveys. This is model H2.
Rationale: Similar to previous model E in the draft assessment report.
Response: The STAT presented two "solutions" obtained, which were very different in scale. Model H2-HI had a negative log-likelihood of 1154 and model H2-LO had a negative log-likelihood of 1191, suggesting that model H2-LO corresponded to a local minimum, but model H2-HI may have fully converged. Model H2-HI provides a much better fit to the data than models H1a and H1b. However, its scale was deemed implausible (billions of tons) and estimated *q*'s for the ATM surveys were implausibly low at 0.004 and 0.006 for spring and summer, respectively. The Panel discussed the performance of model H2-HI relative to model H1b. Model H1b did not appear to have issues with parameter estimates hitting boundaries, while H2-HI hit lower bounds on ATM survey *q*. Model H2-HI has larger estimates of R_0 and R_1 than model H1b. Model H1b has lower negative log-likelihoods for most likelihood components, but model H1b has a lower contribution by recruitment to the overall objective function. Based on the likelihood tables presented, model H2-HI seems to fit the CPFV index much better than model H1a. Visually, model H2-HI seems match the CPFV index data in the 1990s and track the increase in the CPFV index from 2005-2010 better than model H1b. According to the likelihoods, model H2-HI also fits the CPFV length-composition data and commercial ages better, but this was not readily apparent in the figures showing the fits. Differences in recruitment deviations were not readily apparent visually. Fishery selectivity is higher for age 1 for model H2-HI and then drops off at age 6, whereas CPFV selectivity on age 1 is estimated to be higher in model H2-HI. Taken together, these results indicate that the available data provide very little information on scale.

Request 2a: Run a variant of model H1a, but use a simpler parameterization for dome-shaped selectivity, e.g. selectivity pattern 22 in SS 3.24s. This is model H2a.
Rationale: This may address convergence concerns.
Response: Selectivity pattern 22 is for length-based but not age-based selectivity. Therefore this run was not completed.

Request 3: Run a variant of model H1a without the ATM index and length-composition data, setting the phases for the selectivity and catchability parameters of the ATM survey to a negative number (so that SS does not attempt to estimate them). This is model H3.

[2] Henceforth, the term "stable" is used to refer to a model which converges to the same solution from different starting values.

Rationale: To better understand the conflicting information coming from the two surveys.
Response: Model H3 yields slightly higher estimates of biomass than model H1a with generally similar fits, suggesting that the ATM data have little influence on the assessment outcomes. (This request could equivalently have been based off model H1b, since the only difference between Models H1a and H1b was the form of the ATM selectivity curve.)

Request 4: Run a variant of model H1a in which the post-2009 commercial fishery age-composition data is dropped. This is model H4.
Rationale: To see if this allows the model is able to capture the apparent recruitment pulses/cohorts present in the ATM length-composition data.
Response: Model H4 did not improve the fit to the ATM survey length-composition data. Recruitment deviations estimated for this model do not indicate a 2011 recruitment (age-1) pulse, and relative to model H1a this model suggests more negative recruitment deviations in recent years. Model H4 does not provide a marked improvement in fit relative to model H1a, suggesting that something else is responsible for the conflict with the ATM length-composition data.

Request 5: Run a variant of model H1a without the CPFV index and length-composition data, setting the phases for the selectivity and catchability parameters for the CPFV survey to a negative number. This is model H5.
Rationale: To better understand the conflicting information from the two surveys and see if removing the CPFV data allows the model to capture the apparent recruitment pulses/cohorts observed in the ATM length-composition data.
Response: Model H5 estimates implausibly high fishing mortality (*F*) rates for recent years (F>4 for the terminal year). In combination with the results from Request 3, this indicates that the ATM surveys provide little information on biomass as the model is currently configured. Inclusion of the CPFV index in the assessment model is required to avoid implausible results.

Request 6: Run a variant of model H5, except with weights (lambdas) on the ATM length-composition data are increased 10-fold. This is model H6.
Rationale: To better understand the conflicting information coming from the two surveys and see if greater emphasis on the ATM length-composition data allows the model to capture the apparent recruitment pulses/cohorts observed in these data.
Response: The STAT inadvertently did this analysis based on model H1a rather than model H5. The fit to the ATM length-composition data is not visibly improved, although there is a positive recruitment deviation estimated for 2010. The performance of model H5 was so poor that the Panel no longer needed to see model H6 as originally requested.

Request 7: Run variants of model H1a, fixing R_0 to low and high values (i.e., those values of R_0 corresponding to approximately two negative log-likelihood units higher than the negative log-likelihood for the base model) and showing the corresponding biomass trajectories.
Rationale: To quantify uncertainty in scale.
Response: The biomass trajectories for the two runs were very similar, suggesting that some other parameters changed when R_0 was varied. The Panel noted a likelihood profile on current biomass instead of on R_0 would be more useful for this assessment given the life history of Pacific mackerel (i.e., R_1 can easily swamp the importance of R_0 in determining scale).

Request 8: Plot the catch sex ratio data to explore whether a sex-aggregated model is appropriate.
Rationale: To investigate whether the current sex-aggregated model is appropriate.
Response: The STAT plotted the sex ratios of all commercial biological data (Fig. 2). For ages above age 1, the sex ratio is very close to 50-50. There was a high incidence of unknown sex animals at younger ages, likely reflecting slower maturing females. The Panel concluded that a sex-aggregated model was appropriate for the assessment of Pacific mackerel.

Request 9: Redraw Figure 21 of the draft assessment document (average length frequencies of ATM, commercial catch, and CPFV catch) using data only from common years (i.e., 2005-14).
Rationale: The current Figure 21 is based on data from different years for the ATM vs. the two fisheries, confounding comparisons.
Response: The long right-side limb was reduced for the CPFV and commercial fishery data making them more similar to the ATM survey, but the CPFV fishery still tended to catch (or at least retain for sampling) larger fish on average than the commercial fishery (Fig. 3).

Request 10: Run a variant of model H1a with higher σ_R (i.e., 1.2).
Rationale: To see if this model configuration can provide a better fit to the apparently strong 1981 year-class (as seen in the fishery age-composition data). This is model H8.
Response: The STAT ran two versions of model H8, one with σ_R=1.0 and one with σ_R=1.2 (compared to σ_R=0.75 in model H1a). There was relatively little change in estimated biomass.

<u>**Day 2 requests made to the STAT during the meeting – Tuesday, April 28**</u>
Request 11: Plot the length-distribution of combined recreational catch (all modes of recreational fishing) compared to commercial fishery and ATM length-composition data for the same (common) years as in Request 9 (i.e., 2005-14).
Rationale: Model stability is improved if asymptotic selectivity is assumed for the combined (commercial-dominated) fishery, but this would not be consistent with strong evidence that the recreational fishery caught larger fish than the commercial fishery (implying dome-shaped selectivity of the commercial fishery). The distinction from Request 9 is that in Request 9 only CPFV data are used for the recreational lengths whereas Request 11 uses all modes of the recreational fishery.
Response: There is a difference in length-distribution between the commercial and recreational fisheries, with the recreational fishery catching slightly larger fish (Fig. 3). The CPFV length-distribution is shifted even further to the right than both the commercial and the recreational length distributions.

Request 12: Run a variant of model H1b, but with length-based asymptotic selectivity for the fishery and both surveys (3-parameter double-normal formulation for the fishery).
Rationale: To determine whether a move to length-based selectivity is technically feasible.
Response: The fit was poorer by 10 log-likelihood units in total, mostly due to a poorer fit to the age-composition data.

Request 13: Run a variant of model H1a, but with length-based and dome-shaped fishery selectivity (3-parameter double-normal formulation). This is similar to Request 2, except that selectivity is length-based rather than age-based.
Rationale: To test whether a model with estimated dome-shaped fishery selectivity can produce a plausible biomass estimate if selectivity is length- rather than age-based.

Response: A model formulated as requested was developed and presented by the STAT, who noted that they did not have time to assure that the model had fully converged. The model as presented to the Panel resulted in an approximate 3- or 4-fold increase in scale (biomass) and reduced estimated *q* for the ATM surveys to about 0.3 and 0.4 for spring and summer, respectively. The negative log-likelihood of the presented fit was approximately 1178.

Request 14: As for request 13, except down-weight the fishery age-composition data and fit to the fishery length-composition data.
Rationale: Attempts to estimate dome-shaped selectivity within the assessment consistently yielded numerically unstable models and often produced implausible estimates of biomass and/or ATM *q*, but nevertheless the responses to Requests 9 and 11 argue strongly for dome-shaped fishery selectivity, as does the much lower negative log-likelihood of model H2-HI (over 44 log-likelihood units) compared to any model formulations with asymptotic fishery selectivity.
Response: This model fit the fishery length-composition data better than that pf model H1b. However, the fit to the time-aggregated length-composition data for the CPFV fleet was still mis-specified and ATM catchability was very low (<0.1). The Panel concluded that estimation of dome-shaped selectivity within the assessment model will not lead to plausible estimates of ATM survey catchability.

Request 15: Run a variant of model H1b with dome-shaped fishery selectivity.
Rationale: This is similar to Request 2, except that selectivity for the ATM survey is asymptotic (as in model H1b) rather than dome-shaped (as in model H1a) and reflects the STAT's expressed preference for model H1b over model H1a.
Response: This is model H1b_fishdome, which has 57 parameters and a negative log-likelihood of 1155.0, but with ATM *q*'s <0.01 for both seasons.

Request 16: Run a variant of model H1b that excludes the ATM data.
Rationale: Based on Request 5, it appears that the ATM survey (as currently configured) provides little information on biomass, so Requests 16 and 17 together were intended to explore whether model performance (convergence and plausibility of biomass estimates) was improved by excluding the ATM survey data, particularly in the case of dome-shaped fishery selectivity (see Request 17).
Response: This is model H3 (since models H1a and H1b are identical once the ATM survey data are dropped) except that the phase for ATM *q* should have been set negative as well. However, correcting this showed that the model had converged properly even without the phases set negative, and the model yielded a negative log-likelihood of 1077.7.

Request 17: Run a variant of model H1b that excludes the ATM data and also uses dome-shaped selectivity for the fishery.
Rationale: Requests 16 and 17 together were intended to explore whether model performance (convergence and plausibility of biomass estimates) was improved by excluding the ATM data, particularly in the case of dome-shaped fishery selectivity.
Response: This is model H3_fishdome. This model has 53 parameters and leads to a negative log-likelihood of 1032.3. The difference in objective function value between models H3 and H3_fishdome is very similar to the difference in objective function value between models H1b and H1b_fishdome, confirming that the ATM survey data contribute little information to the assessment.

Request 18: Present diagnostics (such as retrospective analysis, likelihood profiles, and convergence tests) for the STAT's preferred model formulation.
Rationale: If the STAT puts forward a preferred model, it should be accompanied by the standard diagnostics.
Response: This was done based for model H3. Model H3 shows a strong retrospective pattern, similar to model D of the draft assessment and the adopted model in the 2011 assessment (Fig. 1). The total objective function profile on R_0 was consistent with the model having converged to a global optimum (Fig. 4). The profile for US commercial age data was very different from the others, suggesting much lower R_0, while the other data sources preferred higher R_0. The model converged to the same solution repeatedly in the face of jittering, different starting points for R_0, and phasing. Fig. 5 shows the value for the objective function and several of the likelihood components of the objective function as a function of current 1+ biomass. This plot suggests that model H3 converged and that the age, length, and index data provide comparable information related to current biomass.

3) Technical Merits and/or Deficiencies of the Assessment

The base model proposed by the STAT (model H3) incorporates the following specifications:

- Time period from 1983-2014.
- Sexes combined and maximum modelled age of 12 years.
- Natural mortality fixed at M=0.5yr^{-1}, and constant over age and time.
- Maturity pre-specified.
- Commercial and recreational fisheries combined.
- The mean growth curve estimated within the model with CVs of length-at-age set to 0.1.
- Fishery selectivity age-based and modelled using a 3-parameter (asymptotic-like) double-normal curve.
- CPFV selectivity age-based, and assumed to be governed by a logistic curve. Fitted to weighted fishery age-compositions, fishery mean length-at-age time-series, and CPFV index and length-compositions.
- Virgin recruitment, initial recruitment offset and steepness estimated; σ_R set to 0.75.
- Recruitment deviations estimated (with bias ramp) for 1978 to 2013.
- Initial fishing mortality rate set to 0.
- Fishery age-compositions weighted by monthly catches, with effective sample sizes equal to the numbers sampled divided by 25.
- CPFV length-compositions, with effective sample sizes equal to the numbers sampled divided by 25.

The issues with the Pacific mackerel stock assessment relate to the information content of the available data, and the STAT did a comprehensive and thorough job in investigating alternative model configurations that might provide a credible assessment. One of the overriding objectives for the assessment was to develop a parsimonious model, limiting model parameters to those necessary to model the key data and dynamics. The decisions made to this end were all appropriate and supported by the available information. A single sex model is consistent with the sex-specific length-at-age estimates (Figure 3 of draft assessment document) and fishery sex ratios (Request 8). Modeling a single fishery is appropriate given age-composition data are available from only the California commercial fishery, and the recreational fishery is relatively small (<5% of total catch, US and Mexico combined) with only length-composition data. The basis for compiling composition data was updated from the previous assessment by weighting compositions by monthly landings.

The fit to the ATM survey index and length-composition data was poor for all model configurations that included these data. The poor fit to the survey indices is perhaps related to the variable proportion of the stock biomass in the area surveyed (this is also a problem for the CPFV index, but large inter-annual variation is not apparent for the CPFV index). The poor fits to the survey length-composition data may be related to the very small number of mackerel encountered in trawls and measured (as low as three in one year). However, the seemingly clear progression of cohorts through the length-composition data from the ATM surveys suggests that the samples sizes are at least adequate to monitor general trends in size over time. The fit to the CPFV length-composition data for all model configurations also suggests model mis-specification (Fig. 6).

The STAT and Panel did not agree upon a model configuration for the Pacific mackerel assessment. Rather, the STAT recommended model H3 as the best available interim model and the Panel could not identify a model that was adequately defensible, in terms of fitting the data and leading to a plausible value for the catchability coefficients for the ATM survey indices.

Although models H1b and H2 differed only in their parameterization of fishery selectivity, asymptotic in model H1b and dome-shaped in model H2, they resulted in very different estimates of recent stock biomass as indicated by their estimates of the ATM survey *q*'s (1.2 and 0.8 for model H1b and 0.004 and 0.006 for model H2-HI). Both sets of *q* estimates were considered implausible by the STAT and Panel, with those from model H1b too high for a stock with considerable distribution outside the ATM survey area and those from model H2 below any credible lower bound. The implied ATM survey *q* for the best fit of model H3_fishdome would also be implausibly low.

The asymptotic fishery selectivity assumption of model H3 is inconsistent with the length-composition data; average length frequencies (2005-14) indicate smaller proportions of large Pacific mackerel in the commercial fishery than in the CPFV fishery or the ATM survey (Fig. 3). Model H3 did, however, result in stable estimation, while Model H3_fishdome, which includes dome-shaped fishery selectivity, was unstable and "converged" to different local minima following relatively minor changes to the initial parameter values. Model H3_fishdome tended to either estimate effectively asymptotic fishery selectivity, with biomass estimates and objective function values similar to those from model H3, or to estimate dome-shaped fishery selectivity with implausibly high biomass and much lower objective function values (i.e. reduction of around 40 log-likelihood units in objective function between models H3_fishdome and H3). Clearly, the domed fishery selectivity allows a much better fit to the data but does not result in credible biomass estimates. The much better fit of models with a dome-shaped selectivity pattern is a feature of the current assessment. The 2011 assessment review examined this issue and found little evidence for dome-shaped selectivity for the fishery (~5 log-likelihood units better fit, at the cost of three extra parameters) and model configurations with dome-shaped selectivity were, as was the case for this assessment, unstable.

During the review meeting considerable time was spent attempting to find models with dome-shaped fishery selectivity that were stable and had realistic biomass (ATM *q*) estimates. The efforts were ultimately futile although Request 13 did lead to a model configuration with ATM *q*'s which the Panel considered were plausible, but this result may have been a local minimum.

The inability to estimate 'scale' is troubling but characteristic for stocks such as Pacific sardine and Pacific mackerel that have few age-classes and whose dynamics can be modeled as being driven primarily by changes in recruitment. This issue was resolved for Pacific sardine by assuming that the ATM surveys provided estimates of absolute abundance (acoustic q=1), but this assumption is considered implausible for Pacific mackerel.

Commercial fishery composition data (age and length) are only available for the California commercial fishery. Although commercial landings of Pacific mackerel off Washington and Oregon are generally small relative to those from California, it is believed that these fisheries tend to capture larger fish that are more northerly distributed. As such, composition data from these fisheries could allow modeling a separate northern fishery with asymptotic selectivity (while retaining dome-shaped fishery selectivity in California) that could lead to more stable model estimation. This is speculative, but without having data from those fisheries not possible to test.

In principle, an ATM survey is the ideal way to index the abundance of Pacific mackerel and other coastal pelagic species (CPS). ATM survey data were included in models H1b and H1b_fishdome. However, this survey, as currently implemented, is unlikely to provide a consistent index of the abundance of Pacific mackerel because the proportion of the stock in the survey area likely varies among years (primarily because of its distribution south of the US-Mexican border). Survey biomass estimates suggest that this (and/or some other aspect of the survey such as species composition estimates) is a considerable problem because the 95% confidence intervals from consecutive surveys do not overlap in some consecutive years in particular seasons, i.e., the extremely high estimated biomass from the spring 2010 survey being most illustrative of this point. Although it is unclear whether this survey in its current configuration can contribute useful information to the Pacific mackerel stock assessment, the Panel felt that it did provide a basis for rejecting some model runs on the basis of implausible survey *q* estimates.

4) Areas of Disagreement

There were no major areas of disagreement within the Panel. However, while the Panel and STAT agreed on some aspects of modeling, they did not agree upon a model configuration for the Pacific mackerel assessment. Rather, the STAT recommended model H3 as the best available interim model and the Panel could not identify a model that was adequately defensible in terms of fitting the data, modeling fishery selectivity using a dome-shaped function, and leading to plausible values for the ATM survey catchability coefficients.

5) Unresolved Problems and Major Uncertainties

Selection of a base model and the scaling problem

The STAT proposed model H3. This model is 'stable' as shown in jittering analyses and is generally able to fit the data. However, it ignores the ATM survey data owing to concerns about the representativeness of the data. As noted, if the ATM survey data are considered reliable, the catchability coefficient implied by model H3 would be close to 1.

There was some discussion about including priors for the ATM survey *q*'s in the assessment. This could potentially result in plausible *q* estimates and stabilize model variants with dome-shaped fishery selectivity. In discussing this option, there was general agreement that while the ATM *q*'s were unlikely to be less than 0.1 or greater than 0.9, there was no objective basis for fully defining an appropriate prior. The current Pacific mackerel Harvest Control Rule (HCR) has a *distribution* parameter of 0.7 which implies that 70% of the stock is available in US territorial waters. However, the *distribution* value was determined in 1987 its provenance is unclear and proportion of the stock in US waters changes over time, so this parameter provides little basis for developing a prior. Also, given the large decrease in objective function value when switching from asymptotic to dome-shaped fishery selectivity (~40 log-likelihood units), it is unlikely that anything less than a very strong *q* prior would affect results without additional model changes (i.e. data reweighting). Priors are included as penalty functions in Stock Synthesis assessments. Ideally, if priors for the ATM *q*'s were available, perhaps by year, the assessment should be moved to a fully Bayesian framework.

Indices of abundance

The draft assessment for Pacific mackerel included two indices of abundance: (a) a catch-rate index based on data collected from the CPFV fleet and (b) a survey index of abundance from the ATM surveys. Neither of these indices is ideal for indexing the abundance of Pacific mackerel.

The CPFV index

The concerns with the CPFV index include the fact that the CPFV fleet does not target Pacific mackerel and the accuracy of the catch reporting may be questionable. Moreover, the targeting practices of the CPFV fleet will have changed in response to management actions such as the implementation of the rockfish closures and the reduction in allowable harvests of rockfishes, but the standardization procedure does not account for this potential change in targeting. This fleet is also localized so large-scale environmentally-driven changes in the distribution of Pacific mackerel may be mis-interpreted as changes in abundance. The Panel identified alternative ways to analyse the CPFV data which may address some of these issues (see Section 8, Research Recommendations). However, it is unclear whether a re-analysis of the data will remove the model mis-specification evident in the fits to the CPFV index.

The ATM index

In principle, an ATM survey is the ideal way to index the abundance of coastal pelagic species such as Pacific mackerel. A Methodology Panel held in 2011 (PFMC, 2011a) concluded that acoustic-trawl surveys, as well as the associated methods of data collection and analysis, are adequate for the provision of advice on the abundance of Pacific sardine, jack mackerel, and Pacific mackerel, subject to caveats, in particular related to the survey areas and distributions of the stocks at the times of the surveys. It concluded that the estimates of abundance for Pacific mackerel are very uncertain as measures of absolute abundance because a sizable, and variable, fraction of the stock is outside of the survey area. The PFMC SSC, in its review of the 2011 full assessment (PFMC, 2011b), noted that these surveys should "be expanded to encompass Mexican waters, and ideally Canadian waters as well, to be useful for Pacific mackerel assessment". The current Panel agrees with the conclusions of the SSC and the Methodology Panel and strongly recommends that efforts be made to ensure that future surveys cover a larger area, particularly in latitude, to reduce the effects of uncertainty regarding the proportion of the population in the surveyed area. The Panel also notes the importance of implementing previous research recommendations related to the ATM survey to improve the confidence in the ATM survey estimates, including the investigation of potential species selectivity effects by comparing the ratios of catch rates and acoustically-estimated densities in areas where single species dominate. (PFMC, 2011a). The within- and between-year variation in estimates of Pacific mackerel abundance from the ATM surveys is unexpected and makes use of these data for assessment purposes more difficult.

The model is unable to fit the length-compositions for the ATM surveys, which appear to indicate stronger 2006 and 2011 cohorts (at age 1). This suggests that there is a conflict between the data from the fishery and the ATM surveys. This is explored in Requests 3-6. However, the ATM length-compositions are based on few animals, even though a recent effort has been made to increase sample sizes. Moreover, assumed effective sample sizes for the ATM length-compositions were based on the number of positive hauls with Pacific mackerel and are hence smaller than the effective sample sizes for the fishery age-compositions.

In principle, even if they cannot currently provide estimates of absolute abundance for the entire stock, the ATM surveys might be able to provide minimum estimates of absolute abundance. However, the use of the ATM data in this way relies *inter alia* on whether the target strength for Pacific mackerel is estimated in an unbiased manner, and this is uncertain. Moreover, there is no straightforward way to include minimum estimates of biomass into an assessment implemented using Stock Synthesis.

6) Issues raised by the CPSMT and CPSAS representatives during the meeting

a) CPSMT issues

The CPSMT representative appreciates and commends the STAT for their hard work and dedicated efforts during a difficult process. Pacific mackerel are currently a data-rich but information-poor species, and that fact is continuing to cause difficulties. Recommendations are made in this report for more data gathering specific to Pacific mackerel, and should be given prompt attention. Pacific sardines will be closed to non-tribal directed commercial harvest until at least July 2016, and mackerel are presumed to take a place of increased importance for the present and near future. Meanwhile, management of the mackerel resource will be based on a multitude of uncertainties.

The ATM surveys presently constitute the best potential avenue for gathering adequate data to inform these stock assessments and begin to reduce the most glaring uncertainties. Methodology should be developed to be more specific, with data collection aimed toward Pacific mackerel and other CPS assemblage species while the survey’s physical range should also be expanded. The last methodology review was conducted in 2011 and the CPSMT reiterates its support for a new methodology review. The STAT’s recommendation of conducting a MSE on mackerel and CPS management would require a large commitment of time and workload, but the results could be advantageous towards future mackerel and other CPS stock assessments and management.

b) CPSAS issues

The CPSAS representative extends grateful thanks to the STAT for their hard work, and shares the frustration of both the STAT and the Panel regarding the uncertainties inherent in this Pacific mackerel stock assessment review, which in many ways mirror the problems experienced at the 2011 Pacific mackerel STAR panel meeting.

One apparent problem seems to be that the model and data are not consistent: the model is attempting to estimate the entire Pacific mackerel biomass, but the two independent indices of abundance initially considered measure only a part of it. For example, CPFV surveys record only incidental mackerel landings in California, yet mackerel landings are increasing in the Pacific Northwest, and to date biological composition data are unavailable to account for Pacific Northwest catches. Moreover, party boat logs and dockside surveys also may be underreporting the occurrence of Pacific mackerel. The catchability coefficient for the ATM surveys, that were designed for Pacific sardine, was estimated to be close to 1 in model H1b even though the ATM surveys cover only a portion of the Pacific mackerel range, excluding Mexico, and an unquantified but likely substantial portion of the stock. Including ATM surveys in the model scaled biomass downward, but generally provided little information on biomass. The final model excluded the ATM data until survey methodology can be improved.

One recommendation that the CPSAS representative believes is critical, in light of recognized 'spikes' in Pacific mackerel abundance in favorable conditions, is for the Council to have flexibility to adjust management measures as needed between scheduled stock assessment reviews. It is important to point out that with the closure of the sardine fishery, and the probable decline in squid abundance due to the current El Nino cycle, effort will increase on Pacific mackerel during 2015, both in California and the Pacific Northwest.

Appendix 2 provides additional information from the CPSAS representative, including additional recommendations for research and monitoring priorities.

7) Other issues

Catch projections

The STAT requested the Panel provide advice on how to conduct projections to inform the setting of harvest specifications. The SSC will recommend Overfishing Limits (OFLs) and the value for σ (the extent of scientific uncertainty). The Council will select the probability of overfishing, P* and hence the Acceptable Biological Catches (ABCs), and the Annual Catch Limits (ACLs). OFLs, ABCs and ACLs will be selected for the 2015-16 and 2016-17 fishing years at the June 2015 Council meeting. It will be necessary to conduct a projection to compute the OFL (and hence Harvest Guideline) for the 2016-17 fishing year. This will involve selecting a basis for forecasting the 2015 recruitment and the 2015-16 catch. Generating the 2015 recruitment off the stock-recruitment relationship seems appropriate given the estimates of the recruitment deviations for recent years are at or above average (Fig 7). The catch for 2015-16 could be set to the Harvest Guideline for 2015-16 in the absence of a way to predict the fraction of the Harvest Guideline that will be taken.

Review process

The Panel discussed how best to identify the primary issues that would be considered during the review, noting that this issue is not documented in the Terms of Reference for Stocks Assessments. Some Panel members noted they used a 'hierarchical' approach for reviewing assessment documents:

- Stage 1. Evaluate whether there are major 'red flags' such as unrealistically high exploitation rates, unrealistically high or low survey catchability coefficients, or data sets that are available but were not used in the assessment.
- Stage 2. Evaluate whether there are obvious signs for model mis-specification and whether the range of models explored is sufficiently broad, including consideration of the appropriateness of pre-specified values for parameters such as natural mortality, *M*.
- Stage 3. Refinement of weighting factors.

The meeting also discussed the value of reviewers providing written comments before a Panel meeting to help prepare the STAT for the review process. It was noted that this might streamline the process, but agreed that the STAT would not be required to provide written responses to any written comments as that would defeat the aim of making the process more efficient.

The STAT stated that future reviews for addressing a new and improved assessment model for Pacific mackerel would benefit from a two-phase meeting approach: 1) the first meeting would be held with members of the CPS subcommittee of the SSC before the next formally scheduled full assessment review to critique/discuss a revised model, e.g., during 1-day meeting held in concert with a previously scheduled SSC meeting—potentially, summer/fall 2016; and 2) a second, more typical STAR meeting would then be conducted that fully meets the CPS terms of reference for purposes of providing management advice for the coming fishing year(s). Although the current assessment schedule for this species stipulates that the next review meeting should take place in spring 2017 (catch-based projection only), the STAT stated that they felt the best deliverable would entail using a summer/fall 2016 meeting with the CPS subcommittee for guidance concerning the type of assessment that should go forward for review in spring of 2017, e.g., update or full assessment, rather than a simple catch-based projection.

The meeting did not draw conclusions regarding the review process but highlights this issue for consideration by the SSC.

8) Research Recommendations

The STAT commented on how the recommendations from the 2011 full assessment were addressed in the draft assessment document.

General

1. Develop a way to automatically profile over current biomass. It is relatively easy to profile over parameters such as R_0 (e.g. Fig. 4). However, CPS management is based on the estimate of current biomass so that quantity rather than R_0 should be the focus for likelihood profiles (e.g. Fig. 6) and sensitivity analyses.

High Priority

1. Improve collaboration with fishery researchers from Mexico. As noted in previous assessment reviews, a large fraction of the catch is taken off Mexico, and efforts should be made to obtain length, age and related biological data from the Mexican fisheries. Inclusion of the ATM surveys in the assessment has increased the need for Mexican data from comparable surveys because such information could be used to develop an index that is close to being a measure of the absolute abundance of the transboundary stock of Pacific mackerel.
2. Continue to refine the indices of abundance: The Panel considers an ATM survey to be the ideal way to index the abundance of CPS such as Pacific mackerel. The following should be addressed to better realize the potential of the ATM survey for Pacific mackerel:
    a. PFMC (2011) conducted a review of the ATM surveys. Some of the recommendations of that review have been implemented (Zwolinski and Demer, 2014). However, most of the recommendations have yet to be addressed. Given the results of the ATM surveys are likely to be used in several assessments, there may be value in conducting a second PFMC Methodology Review for these surveys. The review would follow up on the recommendations from the 2011 PFMC and any other reviews of the ATM surveys.
    b. Efforts should be made to ensure that future surveys cover a larger area, particularly in latitude, to reduce the effects of uncertainty regarding the proportion of the population in the surveyed area.
    c. The sample sizes for the ATM survey length-compositions can be very small. Further identify and implement way to increase the number of fish caught during the trawling associated with the ATM surveys.
    d. Refine the target strength estimates for Pacific mackerel.
    e. Develop an informative prior for the relative proportion of the population in the survey area when the spring and summer surveys are conducted.
3. Continue to refine the CPFV index of abundance. The CPFV index is used in the assessment of Pacific mackerel and could be included in other assessments. This index is based on fitting a fixed-effects model to catch rates by year, quarter and spatial region. This index can be improved by:
    a. Developing a single database that includes the raw trip-level data.
    b. Conducting analyses in which the trip is the unit of analysis and trip-within-vessel is treated as a random effect and the factors associated with blocks within region are explicitly modelled.
    c. Conducting analyses in which an attempt is made to include catch-rates of other classes of target species as covariates.

Note that it is unclear whether a re-analysis of the data will remove the model mis-specification evident in the fits to the CPFV index.

4. Increase support for current port sampling and laboratory analysis programs for CPS, particularly in the Pacific Northwest.
5. Biological (e.g. length, age, sex) data on mackerel caught in the Pacific Northwest should be collected. These data could further assist in understanding whether and to what extent selectivity for the commercial fishery is dome-shaped. The aging of Pacific sardine in the Pacific Northwest should be co-ordinated with researchers conducting ageing in California.
6. Standard data processing procedures should be developed for CPS, similar to those developed for groundfish species, and a 'data document' developed that provides, in considerable detail, how the basic data sources (e.g., catches, CPFV indices, etc.) are constructed. Much of this information has been published in the past, but a single (and 'living') document describing the basic data will assist assessment authors and future review panels.
7. Investigate the spatial distribution, especially the range, of the Pacific mackerel population over time and whether this changes with population size and/or environmental conditions. In particular, an environmentally-based index of spatial distribution might prove useful for developing priors for ATM catchability for use in future assessments.

Medium Priority

1. Revisit the basis for the current estimate of *M* and hence longevity; explore the use of historical tagging data to estimate *M*.
2. Examine whether parameters such as growth rate and asymptotic size have changed over time.
3. Ageing error should be revisited. As noted during the 2011 STAR Panel report, few otoliths have currently been read multiple times, so additional readings need to be made. An age validation study should be conducted for Pacific mackerel. Such a study should compare age readings based on whole and sectioned otoliths and consider a marginal increment analysis and other validation methods.
4. Conduct a study to update the information used to determine maturity-at-length (and maturity-at-age).
5. Compare catch rate trends of CPFV observer data and CPFV logbook data for the years 1985-89. This work may help validate trends in the logbook data.

Low Priority

1. Explore the feasibility of modeling nonlanded mortalities of sublegal-sized fish in the Mexican fishery

References

PFMC. 2011a. Acoustic-Trawl Survey Method for Coastal Pelagic Species Report of Methodology Review Panel Meeting. Agenda Item C.3a Attachment 1, April 2011.

PFMC. 2011b. Scientific and Statistical Committee Report on Coastal Pelagic Species (CPS) Survey Methodology Review. Agenda Item C.3.b. Supplemental SSC Report, April. 2011.

Zwolinski, J. and D. Demer. 2014. Progress related to the recommendations from ATM survey review. Appendix 3 of the Pacific Sardine STAR Panel Meeting Report, Agenda Item H.1a Attachment 3, April 2014.

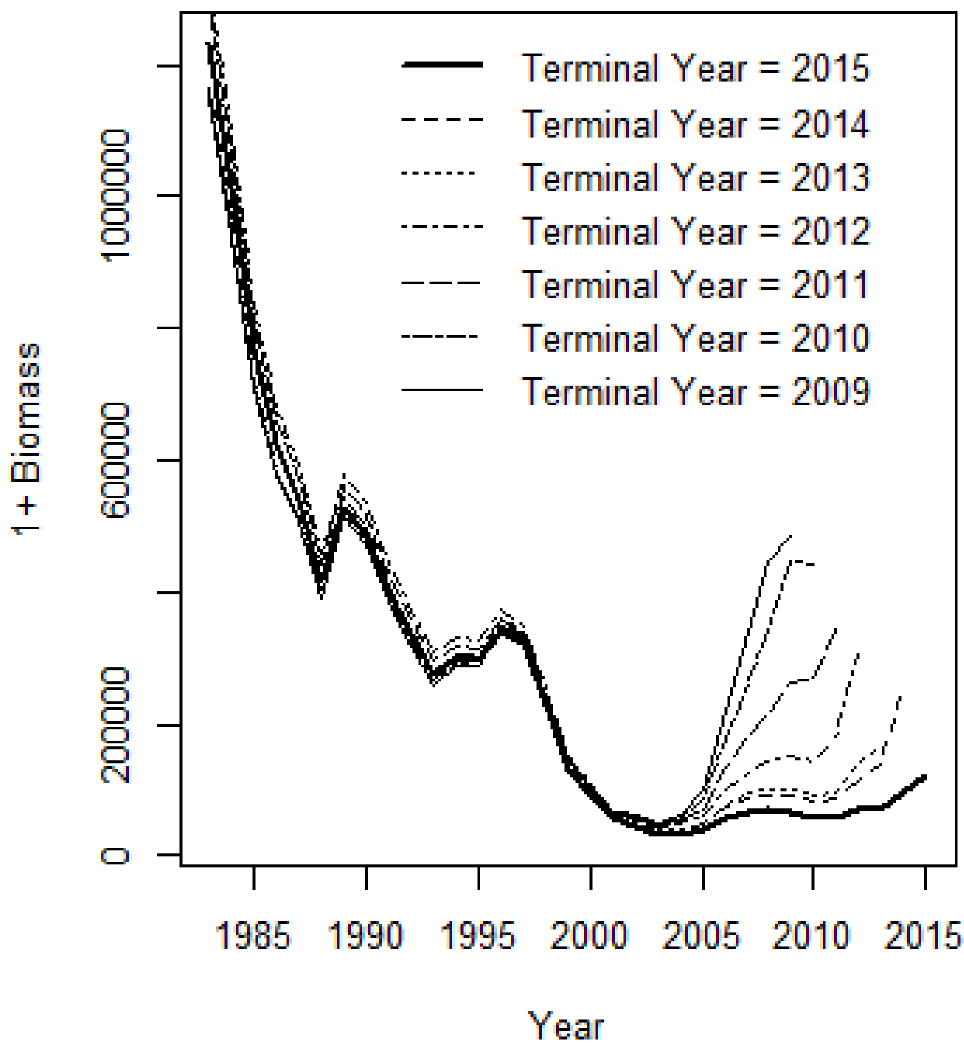

Figure 1. Retrospective pattern for model H3.

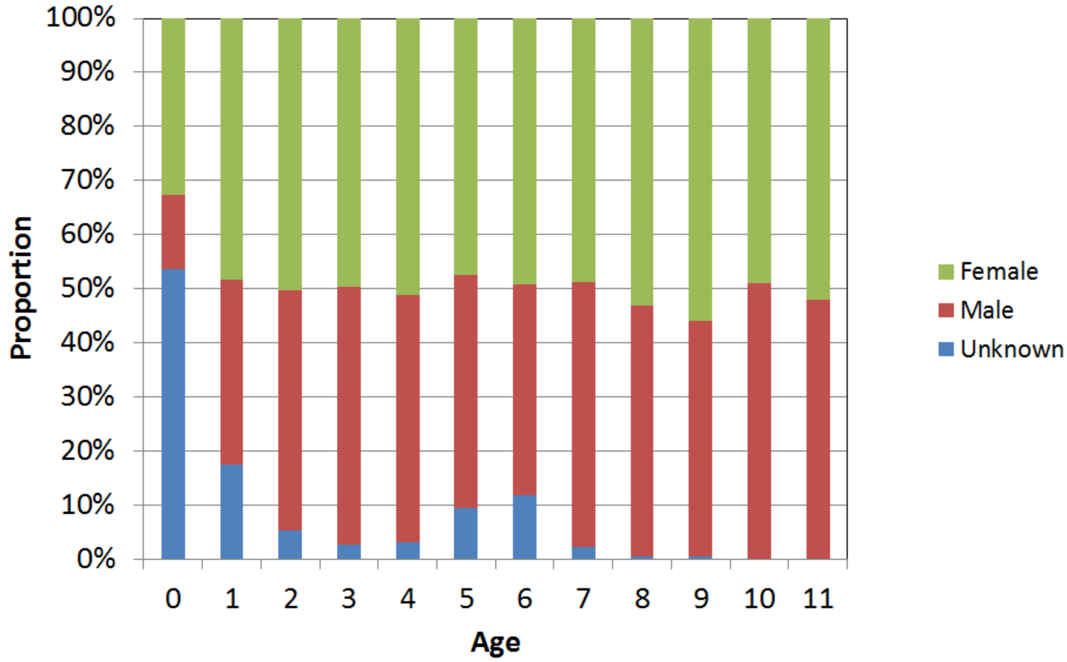

Figure 2. Sex ratios by age for the fishery catches.

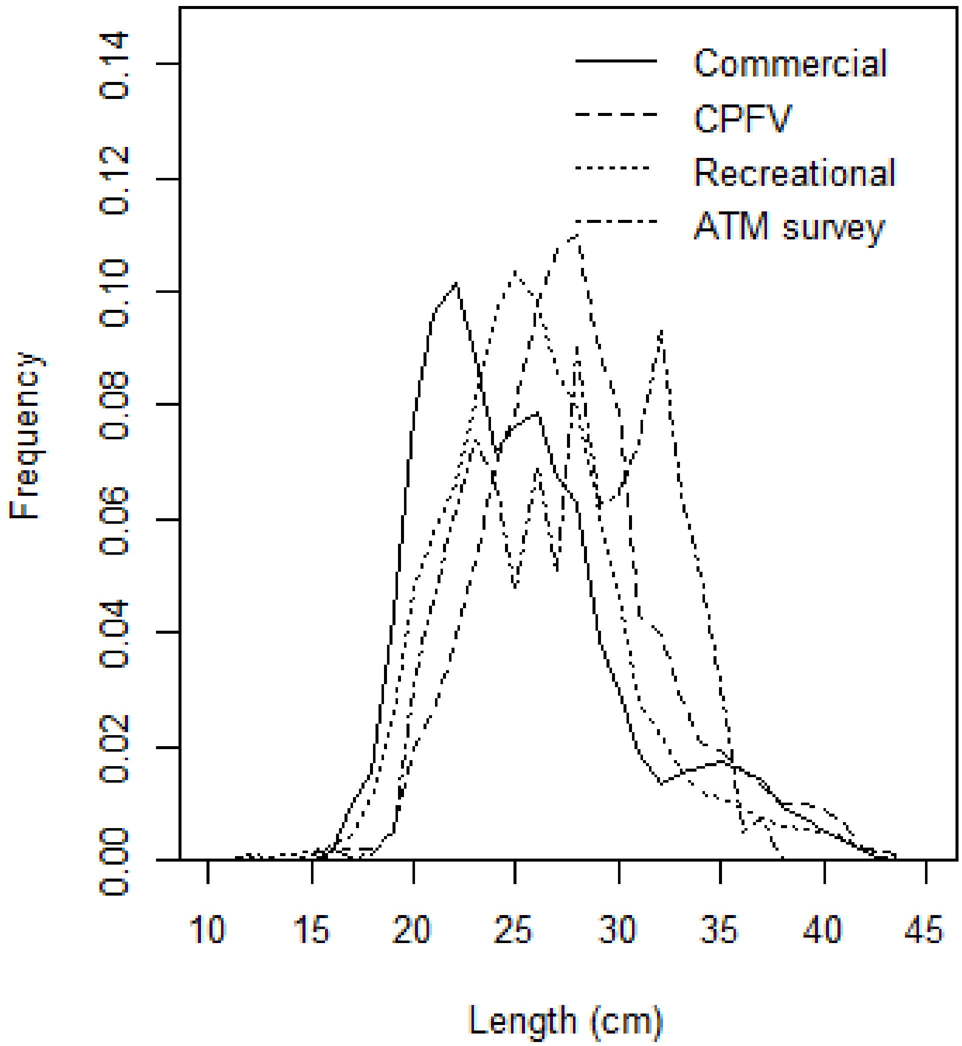

Figure 3. Commercial, recreational, CPFV and ATM survey length-composition data aggregated over samples collected during 2005-14.

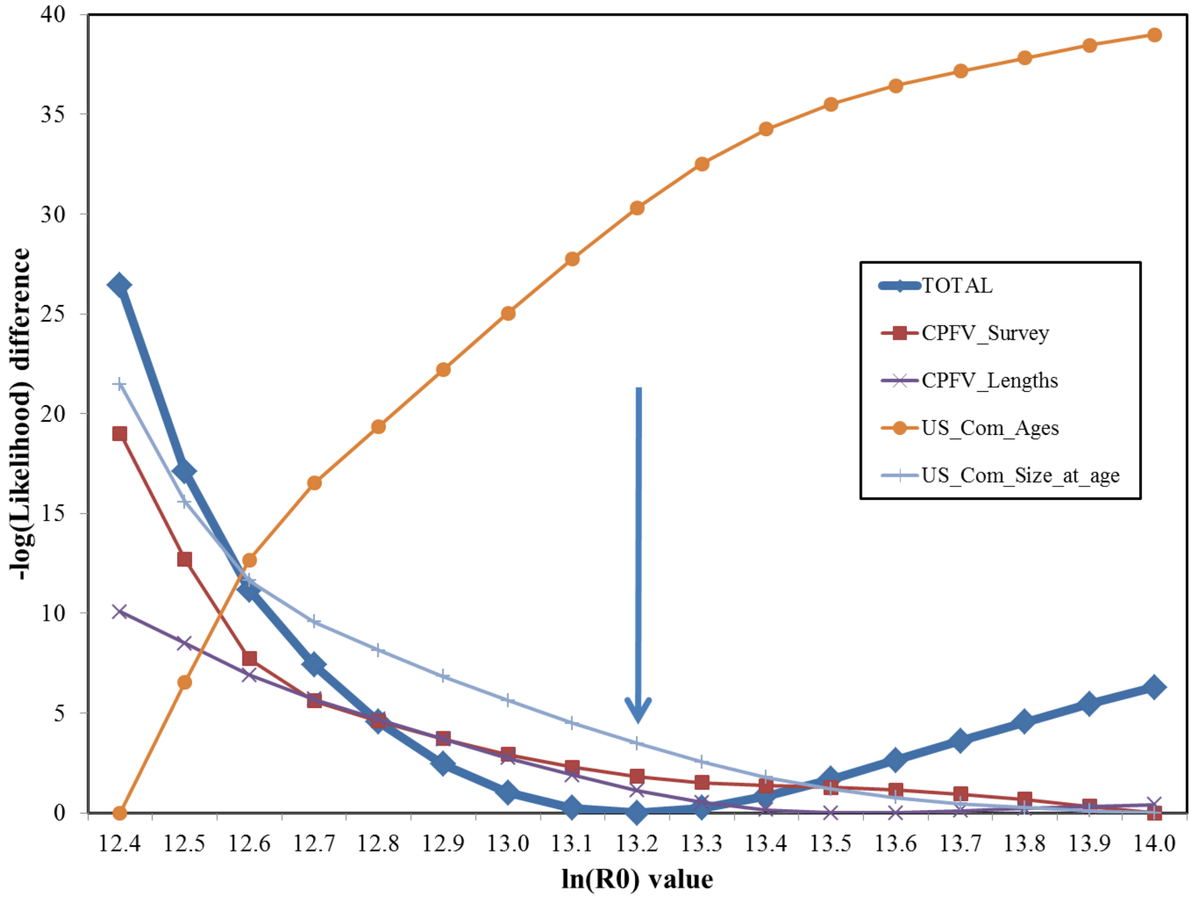

Figure 4. Likelihood profile (total negative log-likelihood on y-axis vs R_0 on x-axis) for model H3.

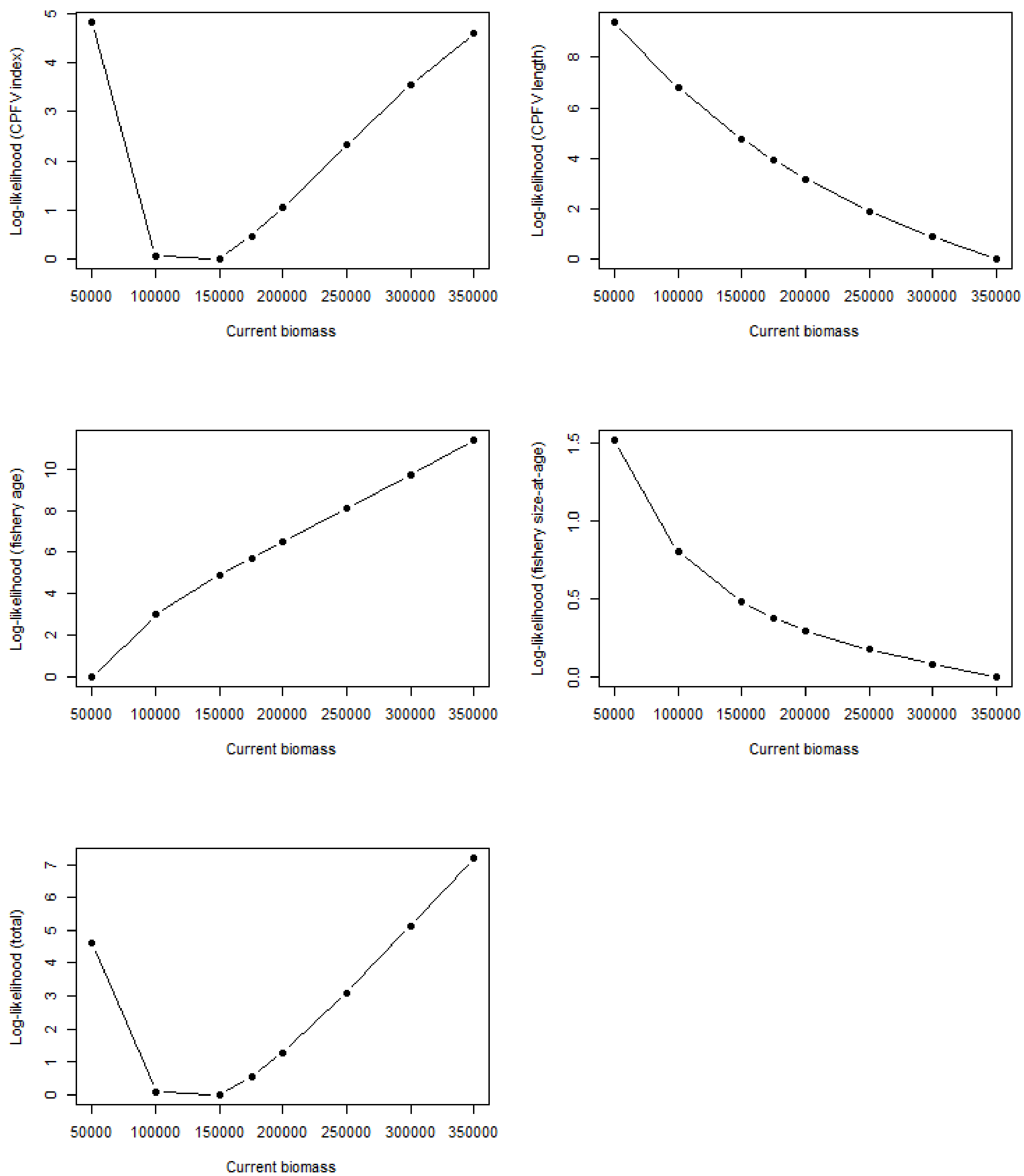

Figure 5. Negative log-likelihood (total and by data component) against current biomass.

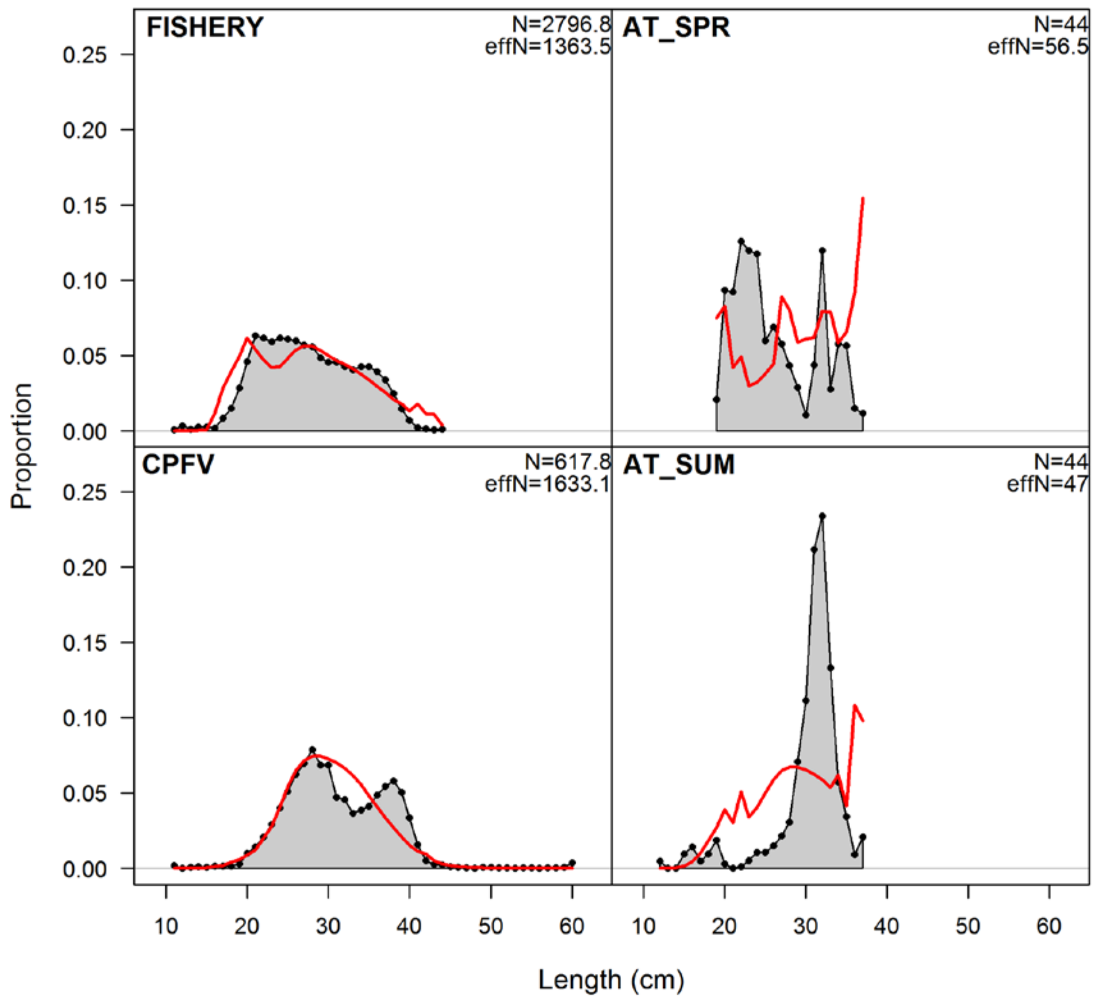

Figure 6. Fits of model H1b to the length-compsosition data aggregated over years

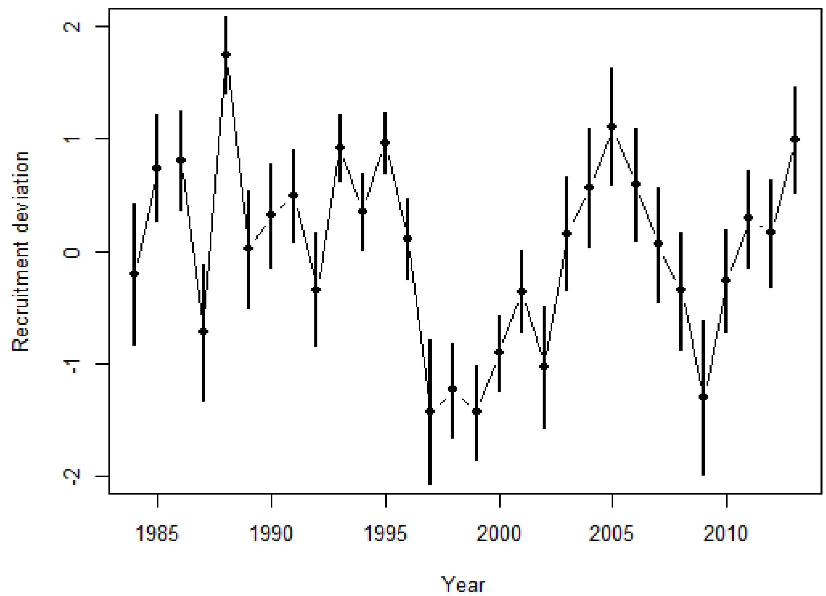

Figure 7. The estimates of the log-deviations about the stock-recruitment relationship from model H3.

Appendix 1
2015 Pacific Mackerel STAR Panel Meeting Attendees

STAR Panel Members
André Punt (Chair), SSC, University of Washington
Will Satterthwaite, SSC, SWFSC
Vivian Haist, Council for Independent Experts (CIE)
David Checkley, Scripps Institution of Oceanography (SIO)

CPSMT/CPSAS Advisors to STAR Panel
Diane Pleschner-Steele, CPSAS
Alan Sarich, CPSMT

STAT Members
Kevin Hill, SWFSC
Paul Crone, SWFSC

Other Attendees
David Demer, SWFSC
Kirk Lynn, CPSMT, CDFW
Dale Sweetnam, SWFSC
Josh Lindsay, CPSMT, NMFS WCR
Kevin Piner, SWFSC
Tony Temp-Lily, SWFSC
Jeannette Miller, CDFW
Emmanis Dorval, CPSMT, SWFSC
Hui-Hua Lee, SWFSC
Gilly Lyons, Pew Trust

CDFW – California Department of Fish and Wildlife
CPSAS - Coastal Pelagic Species Advisory Subpanel
CIE – Council on Independent Experts
CPSMT - Coastal Pelagic Species Management Team
NMFS - National Marine Fisheries Service
SSC - Scientific and Statistical Committee (of the Pacific Fishery Management Council)
SWFSC - Southwest Fisheries Science Center (National Oceanic and Atmospheric Administration)
WCR – West Coast Region of the National Marine Fisheries Service

Appendix 2

Additional concerns raised by the CPSAS representative during the meeting

The CPSAS representative extends grateful thanks to the STAT for their hard work, and shares the frustration of both the STAT and the Panel regarding the uncertainties inherent in this Pacific mackerel stock assessment review, which in many ways appears to mirror the problems experienced at the 2011 Pacific mackerel STAR panel meeting.

One problem seems to be that the model and data are not consistent: the model is attempting to estimate the entire Pacific mackerel biomass but the two independent indices of abundance initially considered in the model measure only a part of it. For example, CPFV surveys record only incidental mackerel landings in California, yet mackerel landings are increasing in the Pacific Northwest, and to date biological composition data are unavailable to account for Pacific Northwest catches. Further, as noted in the 2011 STAR panel report, the CPSAS representative voiced concern with "…inconsistent reporting of Pacific mackerel encounters [in CPFV logs], whether the fish are caught and retained for consumption or for bait, or caught and released. Recreational anglers surveyed also reported that dockside surveys did not always ask about Pacific mackerel that were used for bait or retained. Therefore the party boat logs and dockside surveys may be underreporting the occurrence of Pacific mackerel." This problem continues.

The catchability coefficient for the ATM surveys, that were designed for Pacific sardine, was estimated to be close to 1 in model H1b even though the ATM surveys cover only a portion of the Pacific mackerel range, excluding Mexico, and an unquantified but likely substantial portion of the stock. Including ATM surveys in the model scaled biomass downward, but provided little information on biomass. The final model excluded ATM data until survey area and methodology can be improved.

Recommendations include:

- Data collection programs need to be substantially expanded:
    - Recreational catch data collection programs should emphasize the need to report all incidental Pacific mackerel catches, whether retained, returned or used for bait.
    - Biological (age, size, weight, length) composition data in the Pacific Northwest should be collected.
- AT survey methodology should be modified to increase the spatial boundaries of the survey grid, ideally into Mexico either independently or cooperatively, as well as to add side-looking sonar acoustics to capture fish in the upper water column (a hope that may be resolved at least in part with the deployment of the Ruben Lasker in future CPS surveys).
    - Trawl sample size in ATM surveys also should be substantially increased.
- Efforts should be continued to encourage collaborative Tri-national research and data exchanges, and to collaborate with the fishing industry toward improving the knowledge of Pacific mackerel.

The CPSAS representative notes that similar concerns are also expressed in some form in the Research and Data Needs section of the Pacific Mackerel Stock Assessment Report.

One overarching recommendation that the CPSAS representative believes is critical, in light of the dynamics of the Pacific mackerel fishery, is for the Council to have flexibility to adjust management measures as needed between scheduled stock assessment reviews. The rationale for this that the current schedule calls for management to be adjusted on two-year cycles, with a full assessment only once in four years, and a catch-only projection/update after two years. Thus, the management measures set in this stock assessment will determine the harvest guidelines for at least the next two years, and possibly longer. Pacific mackerel are acknowledged to have rapid 'spikes' in abundance in favorable conditions. The CPSAS representative expresses concern with the apparent dichotomy between current scientific survey data and observations on the grounds. Model runs reviewed, especially those including the ATM index, reflected a declining trend in abundance and scaled the population lower, while observations on the grounds both in California and the Pacific Northwest indicate an increase. Recent catches, of course, could not be considered in the 2014 surveys. Although 2014 catches declined in California, the fleet was largely focused on market squid, caught close to port, rather than running offshore, spending time and fuel for a lower valued product. Both small and larger size Pacific mackerel have been observed and landed recently in southern California, with fish sizes ranging from 100 to 500 gm. Pacific mackerel abundance also has increased in the Pacific Northwest and catches also increased in 2014. The current El Niño and "warm blob" phenomenon may have pushed Pacific mackerel north. Anchovy are also increasing in abundance in both California and the Northwest, and veteran fishermen report anecdotally that Pacific mackerel typically follow anchovy. The fleet anticipates an increase in Pacific mackerel abundance and catches.

The industry representative understands the purpose of the STAR panel is to attempt to find a <u>risk-neutral</u>, best estimate of biomass, and consideration of management is outside the terms of reference. However, it is important to point out that with the closure of the sardine fishery, and the probable decline in squid abundance due to the current El Nino cycle, will increase effort on Pacific mackerel in 2015, both in California and the Pacific Northwest. Flexibility to adjust the Pacific mackerel harvest guideline upward if survey(s) or catches in intermediate years warrant, perhaps as an "emergency" action, would enable the fleet and markets to continue operating in the absence of other CPS.