Pacific Sardine
STAR Panel Meeting Report

NOAA / Southwest Fisheries Science Center
La Jolla, California
March 3-5, 2014

STAR Panel Members:
André Punt (Chair), Scientific and Statistical Committee (SSC), University of Washington
Meisha Key, SSC, California Department of Fish and Wildlife
José De Oliveira, Center for Independent Experts (CIE)
John Simmonds, Center for Independent Experts (CIE)

Pacific Fishery Management Council (Council) Representatives:
Kerry Griffin, Council Staff
Diane Pleschner-Steele, CPSAS Advisor to STAR Panel
Chelsea Protasio, CPSMT Advisor to STAR Panel

Pacific Sardine Stock Assessment Team:
Kevin Hill, NOAA / SWFSC
Paul Crone, NOAA / SWFSC
Dave Demer, NOAA / SWFSC
Juan Zwolinski, NOAA / SWFSC
Emmanis Dorval, NOAA / SWFSC
Beverly Macewicz, NOAA / SWFSC

1) Overview

The Pacific Sardine Stock Assessment and Review (STAR) Panel (Panel) met at the Southwest Fisheries Science Center, La Jolla, CA from March 3-5, 2014 to review a draft assessment by the Stock Assessment Team (STAT) for Pacific Sardine. 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 SWFSC 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 (SS model) to the Panel. The assessment report included many model runs. However, two "blended" models (G and H) were the focus for Panel discussion. Model G included the following features: (a) the data were updated through 2013, (b) the catches for the MexCal fleet were derived from the environmental-based method, (c) the weight-length and maturity-at-length relationships were updated, (d) the data for the aerial survey were omitted from the assessment, (e) the acoustic-trawl (ATM) survey was split into spring and summer surveys (with separate catchability and selectivity parameters), with catchability parameters (*q*s) no longer fixed, (f) no additional data weighting for survey abundance data beyond input coefficients of variation (CVs) (i.e., lambda=1), (g) no additional data weighting for the length composition data for fisheries/surveys beyond the input effective sample sizes (lambda=1), (h) weighting for the conditional age-at-length data in addition to the input effective sample sizes (lambda=0.5), (i) the value for σ_R was rounded and fixed to 0.75, and (j) recruitment was related to spawning stock size according to a Beverton-Holt stock-recruitment relationship with pre-specified steepness (set to 0.8). Model H differed from Model G by assuming age- rather than length-specific selectivity patterns, by fitting to age-composition data rather than length-composition and conditional age-at-length data, and by fixing the parameters of the growth curve. Model H included no additional data weighting to the abundance or composition data.

David Demer presented the environmental-based method for identifying the catches from the northern subpopulation (NSP). This method led to excluding some of the data (catches and associated composition data) for the Ensenada and San Pedro fisheries from the assessment as those catches were predicted to have come from the southern subpopulation. The Panel welcomed this new approach, noting that past Panels had recommended that developing and applying a method for a more appropriate splitting of catches between the northern and southern subpopulations was a high research priority. The Panel noted that adopting this new catch series meant that there would be no assessment for the population (southern) part of which is subject to being caught during the fall off southern California. The CPS representative commented during the Panel that a pragmatic way to address issues of stock structure might be to conduct an assessment based on catches from US waters only, since the proportions of the southern and northern stocks landed at San Pedro and Ensenada respectively were approximately equal.

David Demer and Emmanis Dorval presented aspects of the methodology and results for the ATM and Daily Egg Production Method (DEPM) respectively. No representative of the Northwest Aerial Survey was available to present the results from the 2013 aerial survey, but Tom Jagielo provided a summary of the results by email on February 27, just prior to the review (Appendix 2). Chris Francis (NIWA, retired) provided a presentation regarding data weighting and the use of conditional age-at-length data in assessments. The Panel noted, and was particularly appreciative of, the efforts made by the STAT to respond to the recommendations from past panels and the SSC. The draft assessment report did not include a summary of progress relative to the recommendations from the ATM survey methdology panel that was held in 2011. Appendix 3 was produced by the end of the Panel meeting, which summarizes this progress. This document was not reviewed by the Panel, but is included in this report for completeness.

The review and subsequent explorations of the assessment through sensitivity analyses were motivated primarily by the need to determine the weightings assigned to the compositional data, particularly given the obvious sensitivity of the results of the assessment to how the conditional age-at-length data are weighted. The Panel also explored various configurations for how selectivity and catchability are parameterized for the ATM survey.

The STAR 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) Day 1 requests made to the STAT during the meeting – Monday, March 3rd

[Note: Request numbers do not necessarily correspond with the model numbers given in Table 1.]

A. *Request*: Compare the yearly length-composition data for the Ensenada fishery that are included in the MexCal data set for the NSP scenario with the corresponding southern California length compositions. Also, compare the yearly length-composition data for the Oregon-Washington catches with those for the British Columbia fishery.
Rationale: There are no age-length data for the Ensenada fishery or for the British Columbia fishery available for use in the assessment at this time, but model H implicitly assumes that the length frequencies for the Ensenada fishery are the same as those for the southern California fishery and that the length-frequencies for the British Columbia fishery are the same as those for the Oregon-Washington fishery.
Response: This request was not required because the Panel focused on model G (length-based) that was presented as the potential base case model and not model H (age-based). Model H was not a focus for the Panel review because it was not as fully tested as model G, and because the construction of the catch age-composition data ignored the length data for Mexico and British Columbia. However, this request has been put forward as a research recommendation.

B. *Request*: Compute age-compositions for the ATM survey by multiplying the survey length-frequencies by the associated age-length keys. Compare the mean age-at-length time-series north and south of 40°10' from the ATM survey.
Rationale: The age data for the ATM survey presented in the draft report were unweighted.

Response: This request was not required because the Panel focused on model G (length-based) that was presented as potential base case model and not model H (age-based). However, this request has been put forward as a research recommendation.

C. ***Request***: Construct catch time series using a one month shorter and longer monthly duration for when the San Pedro and Ensenada fisheries are catching southern subpopulation fish.
Rationale: To evaluate the sensitivity of the catches to the cutoff (50%) that is used to assign catches to the NSP.
Response: Figure 1 shows that the results are likely to be somewhat sensitive to the cut-off chosen to define catches from the northern subpopulation. A research recommendation was raised to examine this issue further.

D. ***Request***: Overlay the habitat map with the spring survey results for the 2013 ATM survey.
Rationale: The survey did not go north of San Francisco. The Panel was interested to know whether the areas north of San Francisco would have been expected to have been suitable habitat for Pacific sardine.
Response: The plots showed no evidence of substantial suitable habitat north of San Francisco in the two weeks around the time the survey was conducted, which suggests that the survey should have provided an adequate sample of the population.

E. ***Request***: Provide additional information regarding the apparent discrepancy between the biomass estimates from the ATM survey in the Washington / Oregon area and the landings in this area, based on the information from 2012.
Rationale: The Panel wished to have more information on this apparent discrepancy.
Response: Juan Zwolinski noted that the ATM survey sampled the region between 44° 47.2’N and 48°18’N and from the 50m to the 1500m depth isobaths from 07/31/2012 to 08/10/2012. The resulting point estimate of sardine biomass was 13,333 mt. The sampling variance was high, resulting in a 95% confidence interval of [3,918, 27,559] mt. During the same time period, the commercial fishery off Oregon and Washington caught 9,747 mt. The ATM surveyed the area to the north, including northern Washington and western Vancouver Island, B.C. There, the sardine biomass was estimated at 18,675 mt, with a 95% confidence interval of [2,661, 54,017] mt. It was likely that by 08/10/2012, 32,008 mt of sardine, with 95% confidence interval [12,439, 68,945] mt, would have been available for the Oregon and Washington fisheries, assuming that all the sardine observed off western Vancouver Island migrated from the south.

F. ***Request***: With model G (from initial draft), reweight the fishery and survey length-composition and conditional age-at-length data by applying the Francis (2011) weighting method (Equation TA1.8). The weighting factors should be implemented as changes to the lambdas in the SS model.
Rationale: The compositional data may not be appropriately weighted.
Response: The upper panel of Table 2 lists the factors to weight the input sample sizes (which are lower than the actual number of fish sized and aged), for each length-composition and conditional age-at-length data component that needs to be weighted. The

response to this request (and requests L, M, and N) was based on model ‘K’ in which the conditional age-at-length data are not downweighted by 0.5 (see Table 1 for the specifications for the models investigated during the Panel requests). The Francis method suggested that the length-compositions needed to be downweighted substantially. In contrast, this method also suggested that the conditional age-at-length data for the MexCal fleets and the ATM survey need to be upweighted. Implementing these weighting factors (model F) led to a markedly lower biomass trajectory and substantially changed selectivity patterns for the two MexCal fisheries. The results from this request led to requests L, M, N and O.

G. *Request*: With model G (from initial draft), include the NWSS aerial survey data. Summarize the results in terms of residual patterns and the information given in Table 8 of the draft document.
Rationale: The Panel wished to understand whether the aerial survey data would be influential if they were included in the assessment.
Response: The biomass trajectory was lower than for model G when the NWSS aerial survey was included in the assessment, but otherwise the results were not substantially different. The Panel did not see evidence to disagree with the STAT’s recommendation to leave this survey out of the assessment.

H. *Request*: With model G (from initial draft), examine scenarios in which catchability is the same for the spring and summer ATM surveys. Consider values for ATM survey catchability from 0.7 to 1.1 in steps of 0.2. Summarize the results in terms of residual patterns and the information given in Table 8.
Rationale: The Panel noted that the ATM survey scientists expressed the view that the spring and summer surveys were directly comparable and wished to understand whether this view is supported by the data included in the assessment.
Response: There is no evidence to support having separate q’s for the spring and summer ATM surveys in terms of the change to the value of the objective function. The single q is closer to that from the spring surveys, which is expected given the relative number of ATM survey data points for spring (6) and summer (3). The spring survey selectivity pattern switches to being less knife-edged for the higher qs, but the change for this and the biomass trajectory did not occur in a systematic way as the ATM survey catchability was changed from 0.7 to 1.1. This request led to an additional request (P).

I. *Request*: With model G (from initial draft), replace the Beverton-Holt stock-recruitment relationship with the Ricker form of this relationship. Estimate steepness rather than assuming it equals 0.8.
Rationale: Several past assessments were based on the Ricker form of the stock-recruitment relationship, with steepness estimated. The Panel wished to explore the sensitivity to this change from prior assessments.
Response: The scale of biomass is slightly lower with the Ricker stock-recruitment relationship, with no difference in likelihoods between the two model runs. Steepness was estimated at 2.05.

J. *Request*: With model G (from initial draft), set $M = 0.5\text{yr}^{-1}$.

Rationale: The analysis of Zwolinski and Demer (2013) suggests that *M* is higher ($0.52yr^{-1}$) than the model G assumption of $0.4yr^{-1}$.
Response: As expected, the scale of the biomass was higher, and the ATM survey *q*'s were lower (spring=0.58, summer=0.63). The change in likelihood was 3 units with the higher *M*, but given the concerns with the weights assigned to the length and conditional age-at-length data, this is not considered to be a substantial change.

Day 2 requests made to the STAT during the meeting – Tuesday, March 4th

K: ***Request***: Conduct an assessment where all the weighting factors (lambdas) are set to 1 and compare the results for this model to those for model G (from the initial draft assessment).
Rationale: The selection of the factors to weight the length-composition and conditional age-at-length data was based on this model.
Response: The STAT provided model K which showed increasing the weights on the conditional age-at-length data from 0.5 to 1 substantially lowered the biomass trajectory.

L. ***Request***: Based on model K, apply the Francis method to estimate weighting factors for the length-composition and conditional age-at-length data, pooling the two MexCal fleets, pooling the spring and summer ATM survey data and analyzing the PacNW separately.
Rationale: Some of the weighting factors are based on very few compositions and consequently the weighting factors are uncertain (Table 2, upper).
Response: This was model L. The weighting factors for the pooled fleets are as expected, but the confidence intervals, particularly for the ATM survey, are narrower (Table 2, lower). The Panel considered it appropriate to pool across fleets when computing the weights for the length-composition and conditional age-at-length data.

M. ***Request***: Based on model K, change only the weights assigned to the length-composition data using the weighting factors from Request F.
Rationale: The Panel wished to understand whether the length-frequency or conditional age-at-length data were most influential.
Response: This was model M. The biomass estimates for the early years were sensitive to changing the weights assigned to the length-frequency data. However, the trend in abundance over recent years was unchanged, and the biomass scale was largely unchanged. The Panel concluded that how the conditional age-at-length data are weighted was the major cause of the change in results observed for request F.

N. ***Request***: Based on model K, change only the weighting factors assigned to the conditional age-at-length data using the weighting factors from Request F.
Rationale: The Panel wished to understand whether the length-frequency or conditional age-at-length data were most influential.
Response: The biomass trajectory for model N was markedly lower (and survey *q* markedly higher) when the conditional age-at-length data were changed.

O. ***Request***: Same as for request N, except that the weighting factor for the conditional age-at-length data sets for the PacNW fishery is assumed to equal 1.

Rationale: The weighting factor for the conditional age-at-length data for the PacNW fleet was less than one, in contrast to the weighting factors for the MexCal fleets and the ATM survey.
Response: The results for model O were essentially identical to those for request N.

P. ***Request***: Same as for model G, except that catchability and selectivity for the spring and summer ATM surveys are assumed to be the same.
Rationale: The Panel wished to understand whether there is support for separating the two surveys.
Response: The fits to the survey length-frequency data for model P were not as good as for model G, even after accounting for there being three fewer parameters. The biomass trajectory was lower than for model G, and the ATM survey catchability was 2.38, a value considered implausible. The single ATM survey selectivity was less knife-edged and to the right of those for the spring and summer ATM survey selectivities from model G, which was unexpected. The model appeared to increase the selection at smaller lengths to account for the summer survey which had appreciable catches at these lengths. The consequence was to then reduce selection at the greater lengths that were previously fully selected when the surveys were fitted with separate selection patterns.

Q. ***Request***: Same as for model P, except that the weight assigned to ATM survey length-frequency data was increased from 1 to 20.
Rationale: The Panel wished to understand whether it is possible to fit the length-frequency data for the ATM survey, at least in principle.
Response: The fits to the ATM length-frequency data for model Q were better, but the model was still unable to adequately mimic all of the length-frequencies.

R. ***Request***: Conduct models R, S, T, W and U.
Rationale: The Panel wished to understand the trade-offs in results among various treatments of ATM survey catchability and selectivity. Some of these models ignore the ATM survey conditional age-at-length data because these data were not computed accounting for the sampling scheme for the survey.
Response: Figure 2 summarizes the biomass trajectories from these models. Models R and S, in which selectivity for the spring and summer ATM surveys was assumed to be the same, led to higher estimates of biomass compared to model G, whereas model T which estimated separate selectivity patterns for the spring and summer ATM surveys, led to lower estimates of biomass; in contrast model W, which is the same as model T but estimates separate catchabilities for the ATM surveys, led to higher estimates of biomass than even model S. Model U in which the conditional age-at-length data for the MexCal and PacNW fisheries were markedly downweighted led to much lower biomass estimates and unrealistically high estimates of survey catchability.

S. ***Request***: Repeat request Q, but omit the ATM survey length-frequency data for spring 2012.
Rationale: This length-frequency was considered unreliable by the ATM survey team.
Response: This model (V) was not able to adequately fit the remaining ATM survey length-frequencies.

T. Request: Conduct analyses for a range of values for the extent which the conditional age-at-length data are downweighted. The analyses should be conducted for model specifications G-2, W-2, W-3, and T-2 (See Table 1).
Rationale: The Panel wished to understand the impact of different weighting factors on the results of the model.
Response: The outputs for models based on configuration W-3 all led to values for the ATM survey catchability coefficients which were considered unrealistically low (~0.25). The biomass trajectories for recent years were more robust for the models based on configuration T-2, but there was considerable sensitivity of biomass estimates for the early years (Figure 3). The biomass trajectories for recent years fell into two groups (one group based on weighting factors on the conditional age-at-length data of 0.1, 0.2 and 0.4; another group based on weighting factors of 0.3, and 0.5 and larger). The biomass trajectories were more stable for model runs based on configuration W-2 than configuration W-3. The weighting factor is 0.035 for configuration W-2 if it is chosen so that the average ATM (spring and summer) survey catchability is 1. Alternatively, this weighting factor is ~0.7 if the analysis is based on configuration G-2. Downweighting is more severe for model configuration W-2 because this model configuration ignores the ATM conditional age-at-length data which tends to support lower biomass estimates. However, the STAT noted that choosing a weighting factor to achieve a given average ATM survey catchability coefficient may not be a robust way to provide management advice. The Panel concurred with this view.

Day 3 requests made to the STAT during the meeting – Wednesday, March 5th

At this point in the meeting, the STAT and Panel agreed to proceed with models which are variants of configuration T-2, i.e. the weighting factors for the length-frequency data are set to 1, catchability is set to 1 for both the spring and summer ATM surveys, separate selectivity patterns are estimated for the spring and summer ATM surveys, and the ATM survey conditional age-at-length data are ignored. The STAT and Panel agreed to focus on two models: T-2_0.2 and T-2_0.7. The difference between these two models is the weight assigned to the fishery conditional age-at-length data. These choices for weighting factors were selected because they are representative of the two groups in Figure 3.

U. Request: Apply models T-2_0.2 and T-2_0.7 when the length-frequencies for the 2011 and 2012 spring ATM surveys are ignored.
Rationale: It was speculated that some of the model sensitivity was due to attempts to fit these two length-frequencies (the fits to these length-frequencies are always poor).
Response: The results when the weighting factor for the conditional age-at-length data was set to 0.7 were similar to those when the weighting factor was set to 0.2 (Figure 4), suggesting that at least one reason for the two groups of results in Figure 3 are conflicts when fitting to the length-frequencies for the 2011 and 2012 spring ATM surveys.

V. Request: Apply models T-2_0.2 and T-2_0.7 when the data for the last four years are ignored.

Rationale: The Panel wished to understand whether a retrospective analysis might help to distinguish between these two models.
Response: The results from both models changed markedly when the data for last four years were ignored (Figure 5).

The STAT and Panel agreed that model T-2_0.2 would be the base model given the relative lack of sensitivity to omitting data (see request U).

3) Technical Merits and/or Deficiencies of the Assessment
Recruitment estimation and environmental variables
The estimate of the most recent recruitment (age 1 in 2013) is uncertain and estimated to be close to the expected value from the stock-recruitment function (Figure 6). Deviations of sardine recruitment from a fitted stock-recruitment model of either Ricker or Beverton-Holt form are observed to be correlated in time, such that there appear to be periods of 'high' recruitment and separate periods of 'low' recruitment. Investigations of the potential for environmental factors to be informative have been conducted by Zwolinski and Demer (in press). They showed that the variability in sardine recruitment in the California Current during the last three decades mimics aspects of the environment in the North Pacific indicated by the Pacific Decadal Oscillation (PDO) index. Research indicated that the average number of recruits per biomass during "warm" periods was more than threefold higher than during "cold" periods. In addition to the environmental conditions experienced by sardine larvae, variability in sardine recruitment is also partially explained by both the environmental conditions several months before the spawning season and the adult's condition factor prior to spawning.

Management of the stock uses information on the biomass of age 1+ sardine when applying the Overfishing Level and Acceptable Biological Catch control rules. Recruitment in the last few years has been lower than expected from the stock-recruitment relationship used in the assessment model. Improved estimation (or prediction) of age-1 recruitment for the most recent year would improve management of the stock given that the assessment model currently leads to a rather imprecise estimate of this quantity (Figure 6). There are a number of potential approaches to do this.

1. A prediction model based on recent recruitment and observed autocorrelation could be used to provide more likely estimates of recruits in the final year without assigning any specific underlying reason for the recruitment.
2. A recruitment prediction index such as that proposed by Zwolinski and Demer (in press), could be used outside the assessment model to replace the assessed value with an alternative value based on a weighted mean of the assessed and index-derived values. One method of determining appropriate weights is given by Shepherd (1997).
3. Inclusion of informative environmental indices in stock-recruitment estimation within the assessment model.

When investigating environmental drivers to explain recruitment, a number of issues need to be considered:

1. The spawning biomass and recruitment pairs estimated in an assessment are subject to uncertainty, and this needs to be accounted for when estimating the prediction intervals for any potential index.
2. Development of environmental indices (for recruitment) through regression analysis needs to be undertaken with care. There are often many explanatory environmental variables. The approach is often to examine many variables to establish the most significant explanatory set. However, to understand the significance of the conclusions, it is important to recognise that exclusion of unsuitable variables is effectively setting the coefficient for the relationship to zero. This needs to be accounted for correctly in tests for overall significance by, for example, removing one degree of freedom for every variable (or variable at lag) rejected. This can be done easily for variables formally tested, but may be more difficult to include when variables are rejected at an early stage based on simple graphical investigation. Currently there are 20 stock-recruitment pairs for Pacific sardine; rejection of 18 potential variables (and or lags) while a relationship is being developed should result in a perception of no significant fit. Failure to consider this can lead to an over-optimistic conclusion of the utility of explanatory functions; see for example Gröger *et al*. (2010) who examined many potential indices and a wide variety of lags, and concluded they had found significant drivers for recruitment.

DEPM Survey

The analysis of the egg survey has some minor issues, mostly to do with the raising of density to survey area. The survey design is intended to sample the region of higher density, because, ideally, the survey obtains lower values around the periphery. A high density stratum is then drawn around a group of observations that contain the higher values, by creating a 'simple' (relatively smooth) boundary using the location of the points. The main idea behind this approach is that the survey objective is to map a peak density in space. There is therefore an assumption that the survey will have higher values towards the centre of the area and lower values around the edges. This is then analysed using a two stratum analysis approach that has two minor issues:

1. the current method for placing the boundary between the high and low density areas by placing the boundary on the observation locations means the higher density area is smaller than the region represented by those observations, and conversely the low density area is a little larger, resulting in a small underestimate. The method should be changed so that the correct area allocation is used for each point in each of the two strata. The effect is likely small on the index value used in the assessment because the current procedure is applied for all years.
2. The post stratification and CV calculations may not be correctly calculating the CV used to weight the survey index values in the assessment. The use of post stratification may result in underestimation due to the separation into strata based on the observed values. The use of a simple variance based on the within-stratum observations in the two strata may result in overestimation given there is expected to be some spatial trend within each stratum. A method that accounts for transect-based sampling, and correlated observations, and reflects the presence of a spawning aggregation would be an improvement.

Construction of conditional age-at-length for the ATM survey

Currently fish aged during the ATM survey are combined into an unweighted age-length key, and subsequently used to construct the conditional age-at-length data for each complete ATM survey. This treatment is not considered to be optimal given the possibility for age- and size-specific distribution of sardine. The use of separate age-length keys for the MexCal and PacNW fleets suggests that there may be differences in age-length keys from these regions. The implication of the current method for the ATM surveys is that this is not occurring. The alternatives are to develop separate age-length keys for the different regions covered by the ATM survey, or to use appropriate biomass-based weighting for each part of the survey area.

Sensitivity of biomass estimates

During its deliberations (see Section 2 of this report) the Panel found, as have several previous Panels, that the trend in abundance for Pacific sardine is generally well-determined by the available data. However, the absolute scale of the population is not well-determined by the data and seemingly small changes to the specifications of the assessment (e.g. the relative weighting of the composition data) can lead to marked changes to the scale of the population. The sensitivity to scale is most obvious in the early years of the assessment period, for which the only index data are the (relatively uninformative) DEPM and Total Egg Production (TEP) estimates. The 2011 assessment addressed this "stability" issue by fixing the q for one of the surveys. The 2011 Panel noted that this is not an ideal approach, and it recommended that this assessment include the development of informative priors for the q parameters for the DEPM, aerial and ATM surveys. However, it also noted that development of informative priors is a non-trivial task and should involve people in addition to the STAT, in particular the survey teams. The last assessment imposed the assumption q=1 for the ATM survey because (a) there are more estimates of abundance for this series than for the aerial survey, (b) the ATM survey is more synoptic (in terms of area coverage) than the aerial survey, (c) the estimates are generally more precise than those for the aerial survey, and (d) the assumption q=1 for the DEPM survey leads to unrealistic values of q for the aerial and ATM surveys (>1.8).

The current assessment team and Panel examined sensitivity to weighting factors (lambdas), and the ATM survey q and selectivity options, and concluded the following:

1. *Sensitivity to the weighting of the ATM conditional age-at-length data*: Estimates of biomass were particularly sensitive to this factor (see models G, K, F, L, N), and the time series were not appropriately assembled (see "Construction of conditional age-at-length for the ATM survey" above). Due to both of these considerations, the ATM conditional age-at-length data were excluded from the final model.
2. *Sensitivity to the weighting of the ATM length-composition data*: When compared to weighting by haul (model K), model results for recent years were insensitive to alternative weighting of the ATM length-composition data, including the use of Francis weights (model M) and arbitrary up-weighting (by a factor of 20; models Q and V).

3. *Sensitivity to weighting of the fishery conditional age-at-length*: A range of weighting factors less than 1 were explored (see models G-2, W-2, W-3, T-2). The sensitivity observed depended on whether the ATM *q* was estimated or fixed (*q*=1). Model outputs were more stable when *q* was fixed (model T-2).
4. *Sensitivity to weighting of the fishery length-composition data*: Two options were investigated: weighting by haul and using the Francis data weighting method. When *q* is estimated (W-3), the use of Francis weights resulted in unrealistically low estimates of *q* (0.2-0.3). For haul-based weights (G-2, W-2), estimates of *q* included the value of 1 over the range of weights considered.
5. *Sensitivity to estimation of ATM q*: Three options were explored: (a) separate estimated *q*s for the spring and summer surveys, (b) a single estimated *q* for both surveys, and (c) a fixed *q*=1 for both surveys. The sensitivity to how the fishery conditional age-at-length data are weighted was considerably reduced for recent years when fixing *q*=1 (e.g. compare models W-2 and T-2). Given the rather arbitrary conditional age-at-length weights being applied for Model G, and that the sensitivity to these could be considerably reduced by fixing *q*=1, it was decided to choose this option in the final model, thereby reducing the sensitivity of the model results to weighting. Generally similar reasoning was used in past assessment reviews (e.g., PFMC, 2011).
6. *Sensitivity to selectivity options for ATM survey*: Two options were explored: (a) a single selectivity pattern for both ATM surveys (spring and summer) or (b) separate selectivity patterns for each survey. When estimated separately, selectivity for the spring survey was nearly knife-edged at around 16cm, and in comparison, that for the summer survey shifted to higher lengths (e.g., model G). When estimated as a single selection pattern, the result was a much shallower curve, starting in a similar place to that estimated for the spring survey and extending to even greater lengths than that estimated for the summer survey (e.g. model P). This probably results from a requirement to include fish between 15 and 18cm in the spring survey, while giving reduced selection at around 20cm for the summer survey and thereby implying a reduction in selectivity for a range of lengths greater than 22cm that were fully selected with separate selection patterns.

The final base model incorporates the following specifications:

- catches for the MexCal fleet computed using the environmentally-based method;
- two seasons (semesters, Jul-Dec=S1 and Jan-Jun=S2) for each assessment year from 1993 to 2013;
- sexes were combined;
- two fisheries (MexCal and PacNW fleets), with an annual selectivity pattern for the PacNW fleet and seasonal selectivity patterns (S1 and S2) for the MexCal fleet;
    - MexCal fleet:
        - dome-shaped length-based selectivity with two periods of time blocking (1993-1998, 1999-2013);
    - PacNW fleet:
        - asymptotic length-based selectivity for a single time period;
    - length compositions with effective sample sizes calculated by dividing the number of fish sampled by 25 (externally) and lambda weighting=1 (internally);

    - conditional age-at-length compositions with effective sample sizes calculated by dividing the number of fish sampled by 25 (externally) and lambda weighting=0.2 (internally);
- Beverton-Holt stock-recruitment relationship "steepness" was fixed (0.8);
- *M* was fixed (0.4 yr^{-1});
- recruitment deviations estimated from 1987-2012;
- virgin (R_0), and initial recruitment offset (R_1) were estimated, and σ_R fixed (0.75);
- initial *F*s set to 0 for all fleets (non-equilibrium model following the initial age composition method in SS);
- DEPM and TEP indices of spawning biomass with *q* estimated for both surveys;
- ATM survey biomass 2006-2013, partitioned into two (spring and summer) surveys, with *q*=1 for each survey;
    - length compositions with effective sample sizes set to 1 per haul (externally) and lambda weighting=1 (internally);
    - asymptotic length-based selectivity for spring and summer surveys;
    - conditional age-at-length data from the ATM surveys excluded;
- NWSS aerial survey index of abundance (biomass) and associated length compositions excluded.

The Panel agrees that the final base model represents the best available science regarding the status of the northern subpopulation of Pacific sardine. The Panel wishes to highlight that the level of variation in terminal biomass evident from the retrospective pattern (on the order of 100,000s of tons from one year to the next; Figure 7 of this report) is not unexpected and has been seen in previous assessments (e.g., PFMC, 2011). Changes in terminal age-1+ biomass estimates used for management of this magnitude may occur when the 2015 assessment update takes place.

On the final day of the review, the STAT provided the Panel with a model in which ATM survey catchability was assumed to be 1 or estimated, separate selectivity patterns were estimated for the spring and summer ATM surveys, the weighting factors for all the length-frequency data were set to 1, and for the conditional age-at-length data were set to 1 for the fishery data and to zero for the ATM surveys, and there were three time blocks for selectivity for the PacNW fishery. There was insufficient time to fully evaluate these models, but the Panel agreed that it would be a valuable model configuration to consider for a future full assessment. That is, model configurations that include time-varying selectivity for suspect fishery/survey composition data that potentially influence absolute abundance estimation is an alternative to downweighting data sources as was largely conducted during this review.

Figure 8 shows time-trajectories of biomass based on applying the final base model (T-2_0.2) in which the catch series is constructed by assuming that all catches in the MexCal fleet are from the northern subpopulation. This model could be used to form the basis for management advice if the model using the environmentally-based catch series cannot be used for management purposes.

4) Areas of Disagreement

There were no major areas of disagreement between the STAT and Panel, nor among members of the Panel.

5) Unresolved Problems and Major Uncertainties

1. The ongoing uncertainties, in particular regarding absolute biomass, are likely to persist until the information content of the data increases substantially, and perhaps not even then.
2. The Panel wishes to highlight that the level of variation in terminal biomass evident from the retrospective analysis (on the order of 100,000s of tons from one year to the next; Figure 7 of this report) is not unexpected, and changes in terminal age 1+ biomass estimates of this extent may occur when the 2015 assessment update takes place.
3. The indices of abundance do not exhibit consistent trends even after allowing for the differences in their respective selectivities, and remain in conflict even when the age and length data are greatly down-weighted.
4. The data set is able to estimate general trends in abundance fairly robustly, but the likelihood is flat over a wide range of current biomass levels, which means that relatively small changes to the data sets or assumptions can lead to marked changes in current abundance.

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

a) CPSMT issues

The CPSMT representative commends the STAT for their efforts accomplished prior to and during the meeting. The CPSMT representative notes that the Panel thoroughly reviewed the stock assessment and the survey data informing the stock assessment. The CPSMT representative appreciates the STAT's effort in addressing data weighting, specifically related to the conditional age-at-lengths from the ATM survey and fisheries. The CPSMT representative agrees with the Panel's attempt to dampen the sensitivity of weighting the data.

The Panel recognized the scaling in the model is not defined given the available data and has been a recurring concern for Pacific sardine and mackerel assessments. Given this instability often seen in the model, the CPSMT representative urges careful consideration when establishing sardine harvest management measures. Ultimately, it is only through further data collection and refinement of data collected that these uncertainties may be resolved. An increase in trawl sampling during the ATM survey could help to increase the amount of size/age data in the model and to potentially reduce conflict between the survey and fishery data.

b) CPSAS issues

The CPSAS representative commends the Panel and STAT for their significant body of work throughout the 2014 sardine STAR panel. Unfortunately, the 2014 sardine assessment encountered the same basic difficulty with scaling issues observed in the 2011 assessment. The SS model is very sensitive to weighting of the input length and conditional age-at-length data from the ATM surveys. Most of the work at the meeting was spent making further analyses to resolve the source of these problems, which included very high variability in the biomass estimates for the first half of the time series. It became apparent from sensitivity runs that data weighting matters. The STAT and Panel attempted to find a solution that made results less sensitive by down-weighting certain conditional age-at-length data.

The sardine assessment model was improved by a more realistic separation of the landings from the northern and southern stocks (excluding the landings of southern stock sardine from Ensenada and Southern California). This reduces the biomass estimates and largely resolves problems associated with the distribution parameter in the harvest guideline.

The final base model ultimately fixed catchability (Q) at 1 for the ATM surveys, as in prior years, attempting to achieve model stability. The CPSAS has voiced concern in the past that acoustic surveys as currently deployed have been unable to measure the full biomass, particularly in the Pacific Northwest. The point is that fishermen observed and caught significantly more fish in the area than the point estimate of the ATM cruise – which measured only one spot in time but contributed to a low overall sardine biomass estimate.

The CPSAS also voices concern that stock assessments seem to be gravitating toward one independent index based on ATM surveys. We encourage a continuation of multiple surveys as each survey type has similar constraints. We acknowledge and applaud the acquisition of the RV Reuben Lasker and its capability to survey with forward and side-scanning sonar. We can support the ATM with the use of sonar to augment acoustic search of water columns that the downsounder does not effectively measure (i.e. the top 10 meters of the water column).

On behalf of the CPSAS and industry at large, the CPSAS representative also expresses disappointment that the aerial survey has been dropped from consideration in this and presumably future stock assessments. Ultimately, industry wants to see a sustainable resource (to the degree that environmental conditions will allow) that is in no danger of being overfished. Current sardine stock assessments and harvest policy are very precautionary. We sincerely hope that going forward we can develop a truly collaborative research program for the CPS complex.

Appendix 4 elaborates on the above concerns and provides recommendations for future stock assessments.

7) Research Recommendations

High priority

A. The assessment would benefit not only from data from Mexico and Canada, but also from joint assessment activities, which would include assessment team members from both countries during assessment development.
B. Modify Stock Synthesis so that the standard errors of the logarithms of age-1+ biomass can be reported. These biomasses are used when computing the Overfishing Level, the Acceptable Biological catch, and the Harvest Level, but

the CV used when applying the ABC control rule is currently that associated with spawning biomass and not age-1+ biomass.

C. Explore models that consider a much longer time-period (e.g. 1931 onwards) to determine whether it is possible to model the entire period and determine whether this leads to a more informative assessment as well as provide a broader context for evaluating changes in productivity.
D. Investigate sensitivity of the assessment to the threshold used in the environmental-based method (currently 50% favourable habitat) to further delineate the southern and northern subpopulations of Pacific sardine. The exploration of sensitivity in the present assessment was limited given time available, but indicated potential sensitivity to this cut-off.
E. Compute age-composition data for the ATM survey by multiplying weighted length-frequencies by appropriately constructed age-length keys (i.e. taking account of where the samples were taken).
F. Investigate alternative approaches for dealing with highly uncertain estimates of recruitment that have an impact on the most recent estimate of age-1+ biomass that is important for management. Possible approaches are outlined in Section 3 of this report.
G. Validation of the environmentally-based stock splitting method should be carried out if management is to be based on separating the northern and southern subpopulations using the habitat model. It may be possible to develop simple discriminant factors to differentiate the two sub-populations by comparing metrics from areas where mixing does not occur. Once statistically significant discriminant metrics (e.g. morphometric, otolith morphology, otolith micro-structure, and possibly using more recent developments in genetic methods) have been chosen, these should be applied to samples from areas where mixing may be occurring or where habitat is close to the environmentally-based boundary. This can be used to help set either a threshold or to allocate proportions if mixing is occurring.
H. Continue to investigate the merits/drawbacks of model configurations that include age compositions (e.g., model H) rather than length-composition and conditional age-at-length data, given some evidence for time- and spatially-varying growth.

Medium priority

I. Continue to explore possible additional fishery-independent data sources. However, inclusion of a substantial new data source would likely require review, which would not be easily accomplished during a standard STAR Panel meeting and would likely need to be reviewed during a Council-sponsored Methodology Review.
J. The reasons for the discrepancy between the observed and expected proportions of old fish in the length and age compositions should be explored further. Possible factors to consider in this investigation include ageing error / ageing bias and the way dome-shaped selectivity has been modelled.
K. The Panel continues to support expansion of coast-wide sampling of adult fish for use when estimating parameters in the DEPM method (and when computing

biomass from the ATM surveys). It also encourages sampling in waters off Mexico and Canada.

L. Consider spatial models for Pacific sardine that can be used to explore the implications of regional recruitment patterns and region-specific biological parameters. These models could be used to identify critical biological data gaps as well as better represent the latitudinal variation in size-at-age.

M. Consider a model that explicitly models the sex-structure of the population and the catch. An analysis of length-at-age samples did not indicate sexual dimorphism for this stock (see Figure 4a in Hill et al. 2014), so all models presented were combined-sex configurations. Nevertheless, it was felt that a sex-specific model was needed minimally as a sensitivity test to investigate the possibility that accounting for sex will have an impact on stock-assessment results for this resource.

N. Consider a model that has separate fleets for Mexico, California, Oregon-Washington and Canada.

O. Compare annual length-composition data for the Ensenada fishery that are included in the MexCal data sets for the NSP scenario with the corresponding southern California length compositions. Also, compare the annual length-composition data for the Oregon-Washington catches with those from the British Columbia fishery. This is particularly important if a future age data/age-based selectivity model scenario is further developed and presented for review.

P. Further explore methods to reduce between-reader ageing bias. In particular, consider comparisons among laboratories and assess whether the age-reading protocol can be improved to reduce among-ager variation.

Q. Change the method for allocating area in the DEPM method so that the appropriate area allocation for each point is included in the relevant stratum. Also, apply a method that better accounts for transect-based sampling and correlated observations that reflects the presence of a spawning aggregation.

R. Consider future research on natural mortality. Note that changes to the assumed value for natural mortality may lead to a need for further changes to harvest control rules.

Low Priority

S. Develop a relationship between egg production and fish age that accounts for the duration of spawning, batch fecundity, etc. by age. Using this information in the assessment would require that the stock-recruitment relationship in SS be modified appropriately.

Finally, the Panel notes that value of the *Small Pelagic Ageing Research Cooperative*, which should improve consistency in age-reading methods generally, and in particular for Pacific sardine. Lack of consistency in age estimates was the reason for not using age data for British Columbia.

8) References

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* 68: 1124–1138.

Gröger, J.P., Kruse, G.H. and N. Rohlf. 2010 Slave to the rhythm: how large-scale climate cycles trigger herring (Clupea harengus) regeneration in the North Sea. *ICES Journal of Marine Science* 67: 454–465.

Hill, K.T., Crone, P.R. et al. 2014. Assessment of the Pacific sardine resource in 2014 for U.S.A. management in 2014-15. Preliminary draft for Star Panel Review, 3-5 March 2014: 137pp.

PFMC. 2011. Pacific Sardine STAR Panel Meeting Report. Pacific Fishery Management Council. Agenda Item F.2.b. Attachment 5. 24pp.

Shepherd, J.G. 1997. Prediction of year-class strength by calibration regression analysis of multiple recruit index series. *ICES Journal of Marine Science* 54: 741–752.

Zwolinski, J.P. and D.A. Demer. 2013. Measurements of natural mortality for Pacific sardine (Sardinops sagax). *ICES Journal of Marine Science* 70: 1408–1415.

Zwolinski, J.P. and D.A. Demer. In press. Environmental and parental control of Pacific sardine (Sardinops sagax) recruitment. *ICES Journal of Marine Science* 00: 00-00. doi:10.1093/icesjms/fst173.

9) Tables and Figures

Table 1. Summary of the models requested of the STAT during the review. "F" indicates that the weights assigned to the composition type were based on the Francis (2011) TA1.8 method, "F-pool" indicates that factor to weight the composition concerned pooled information across fleets / seasons, "split" under the "ATM Q" and "ATM sel" (selectivity) columns indicates that separate parameters were estimated for the spring / summer surveys, "equal" under the "ATM Q" and "ATM sel" columns indicates that the parameters concerned were assumed to be the same for the spring / summer surveys, "1" indicates that survey catchability was assumed to be 1. "profile" in the last three lines implies that the STAT were requested to profile over the weighting factor concerned.

	Lambda: Length composition			Lambda: Conditional age-at-length			ATM	ATM
	MexCal (1+2)	PacNW	ATM	MexCal (1+2)	PacNW	ATM	Q	Sel
G	1	1	1	0.5	0.5	0.5	split	split
K	1	1	1	1	1	1	split	split
F	F	F	F	F	F	F	split	split
L	F-pool	F	F-pool	F-pool	F	F-pool	split	split
M	F	F	F	1	1	1	split	split
N	1	1	1	F	F	F	split	split
O	1	1	1	F	1	F	split	split
P	1	1	1	0.5	0.5	0.5	equal	equal
Q	1	1	20	0.5	0.5	0.5	equal	equal
R	1	1	1	0.5	0.5	0	equal	equal
S	1	1	1	0.5	0.5	0	1	equal
T	1	1	1	0.5	0.5	0	1	split
U	1	1	1	0.01	0.01	0.5	split	split
V	1	1	20, excl spr12	0.5	0.5	0.5	equal	equal
W	1	1	1	0.5	0.5	0	split	split
G-2	1	1	1	profile	profile	profile	split	split
W-2	1	1	1	profile	profile	0	split	split
W-3	F-pool	F	F-pool	profile	profile	0	split	split
T-2	1	1	1	profile	profile	0	1	split

Table 2. Weighting factors and 95% confidence intervals. Results are shown when the Francis (2011) method TA1.8 is applied separately by fleet,and when it is applied to data pooled over fleets or surveys.

	Weighting factors	
Fishery/Survey	**Length**	**Conditional age-at-length**
Single data source		
MexCal_S1	**0.17** (0.11-0.43)	**1.79** (1.43-2.33)
MexCal_S2	**0.15** (0.10-0.31)	**1.69** (1.40-2.11)
PacNW	**0.11** (0.08-0.22)	**0.39** (0.30-0.54)
Aerial	NA	NA
ATM_Spr	**0.15** (0.09-1.13)	**2.11** (1.52-3.49)
ATM_Sum	**0.04** (0.03-Inf)	**1.61** (1.0-3.64)
Pooled data source		
MexCal_S1-S2	**0.17** (0.12-0.28)	**1.66** (1.40-1.98)
PacNW	**0.11** (0.08-0.22)	**0.39** (0.30-0.53)
ATM_Spr-Sum	**0.09** (0.06-0.42)	**1.87** (1.37-2.85)

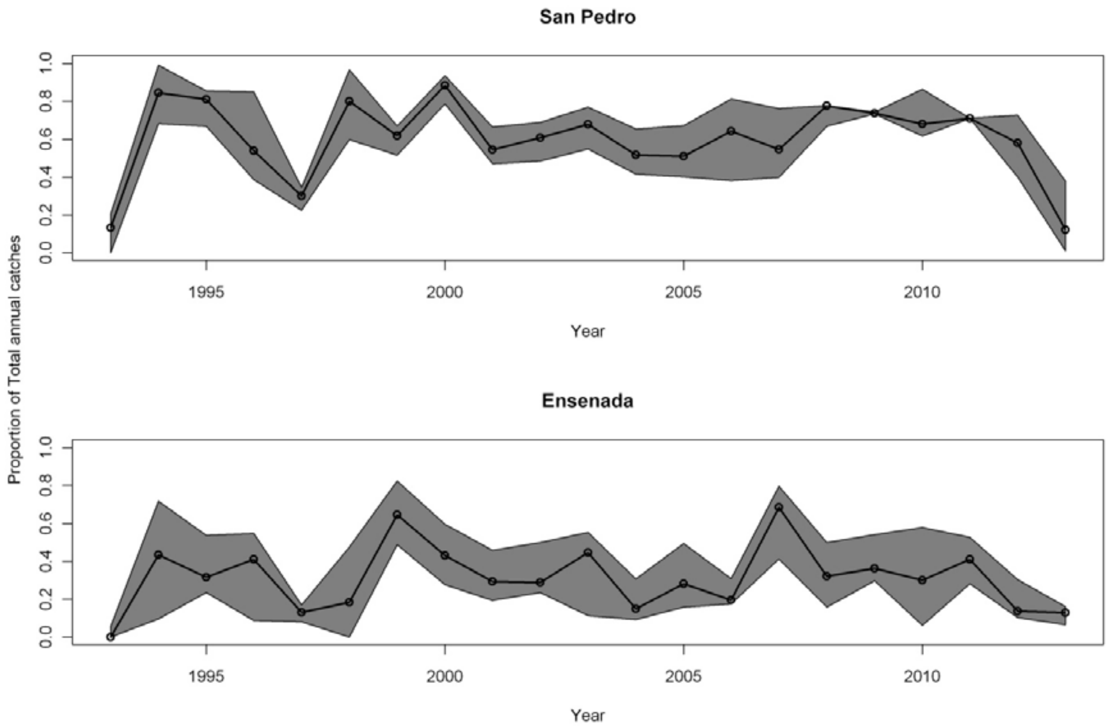

Figure 1. Sensitivity of the proportion of the total catch off San Pedro and Ensenada that is estimated to be from the northern subpopulation to basing the apportionment method on one additional and one fewer month.

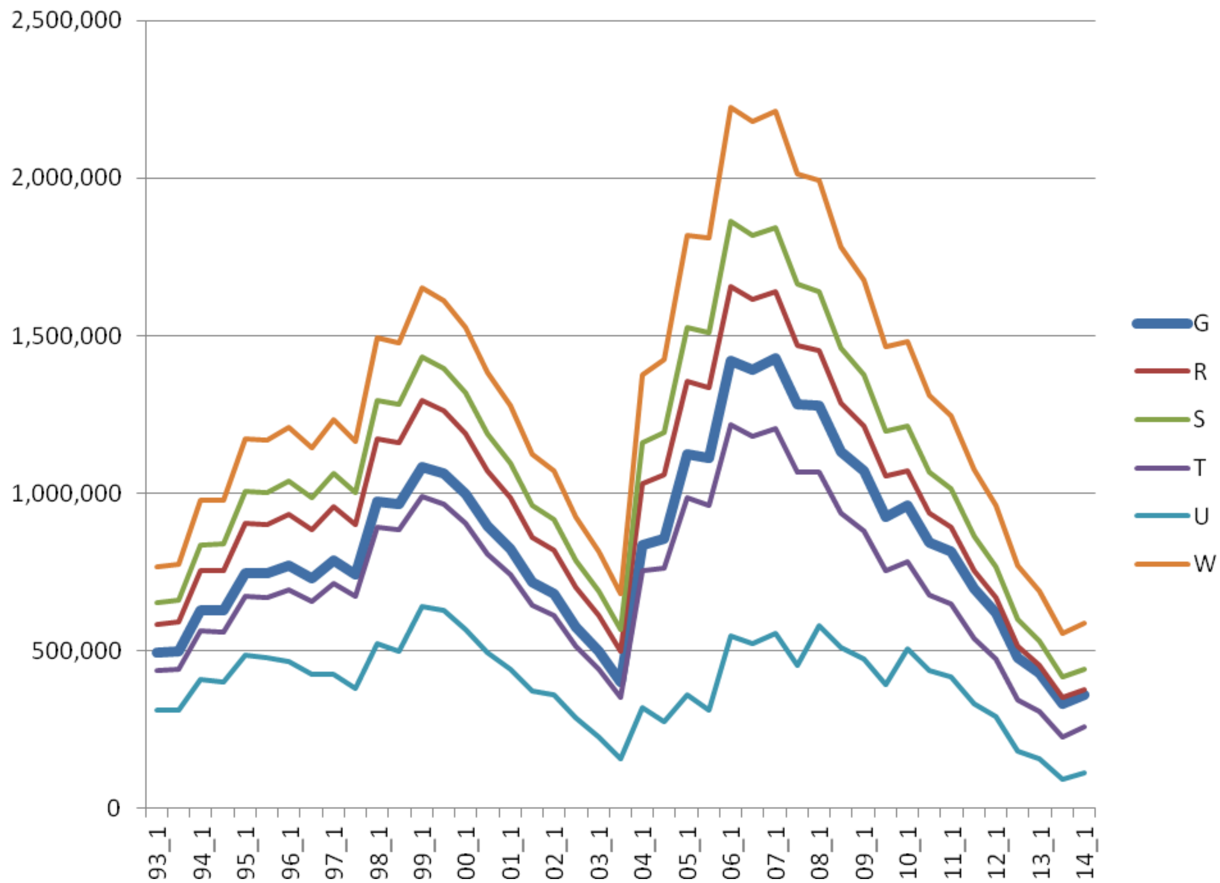

Figure 2. Sensitivity of the results of model G to varying the treatment of the ATM survey selectivity and catchability.

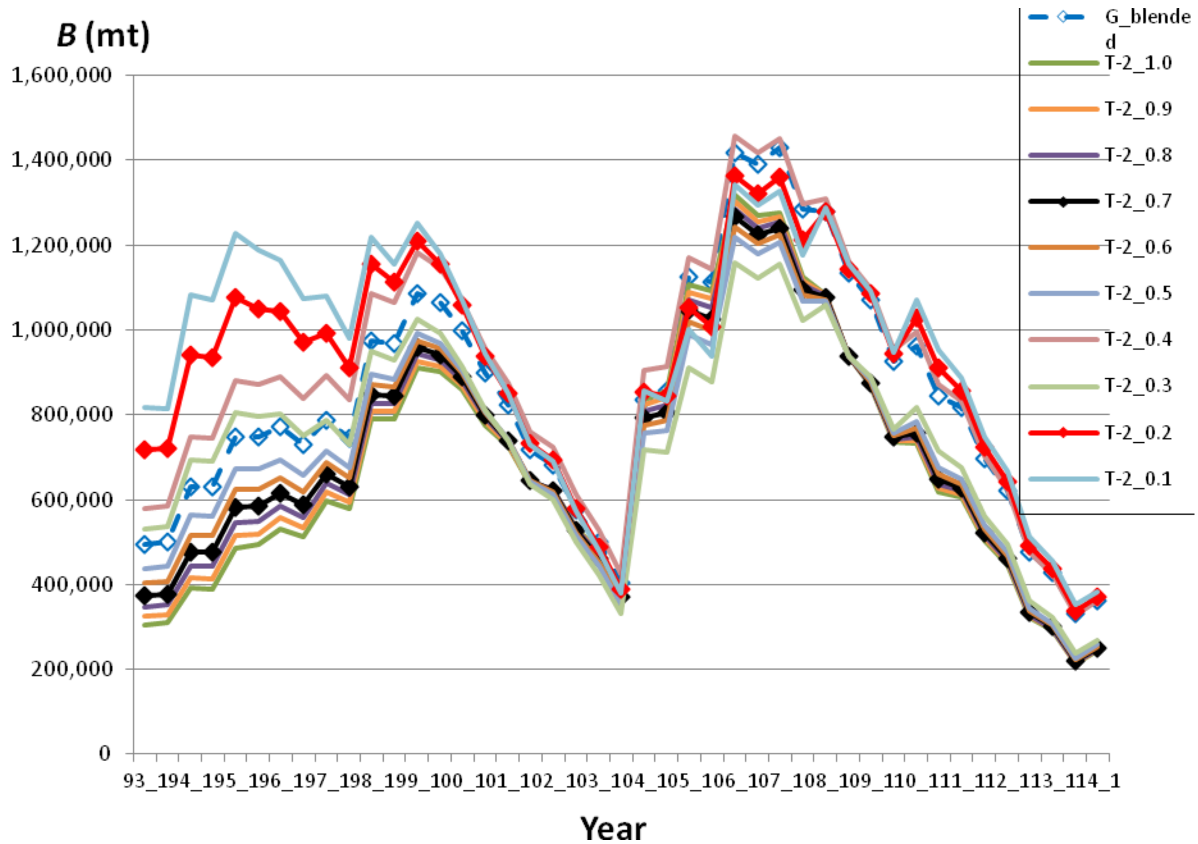

Figure 3. Biomass trajectories for variants of model configuration T-2 constructed by changing the weighting factor for the conditional age-at-length data for the MexCal and PacNW fisheries.

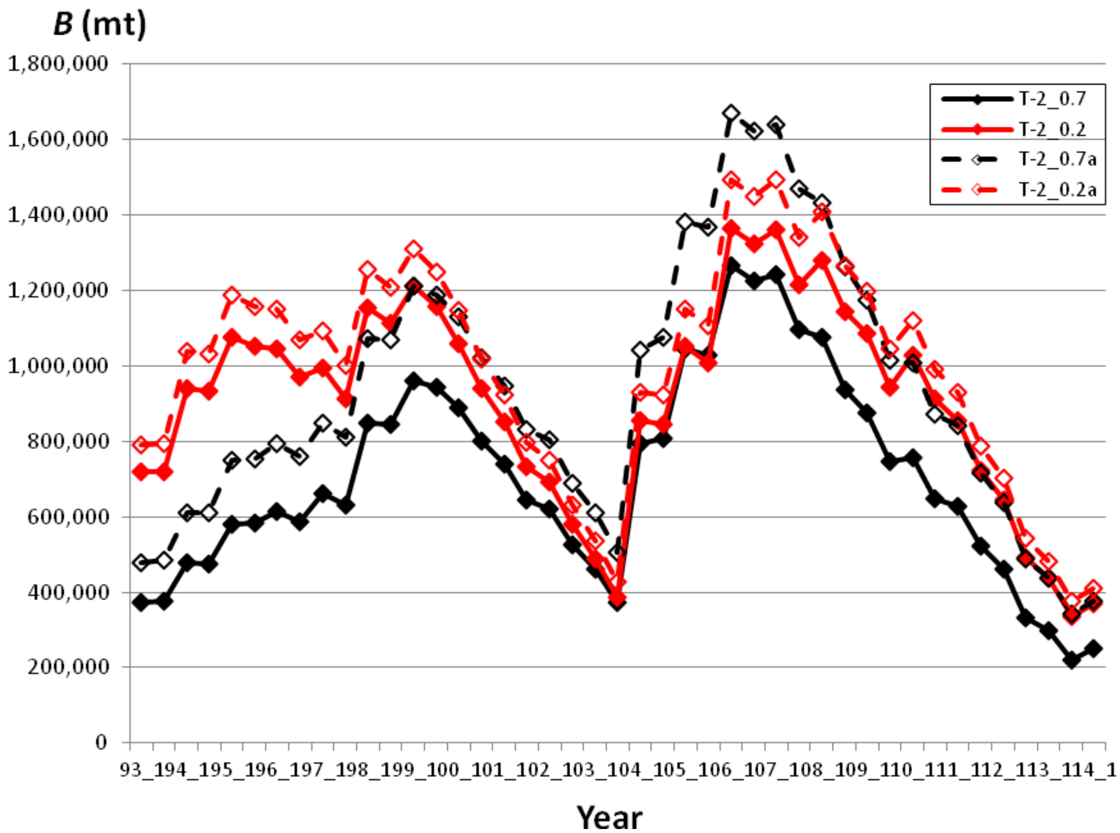

Figure 4. Biomass trajectories for models T-2_0.2 and T-2_0.7 and variants thereof that ignore the length-frequencies for the 2011 and 2012 spring ATM surveys.

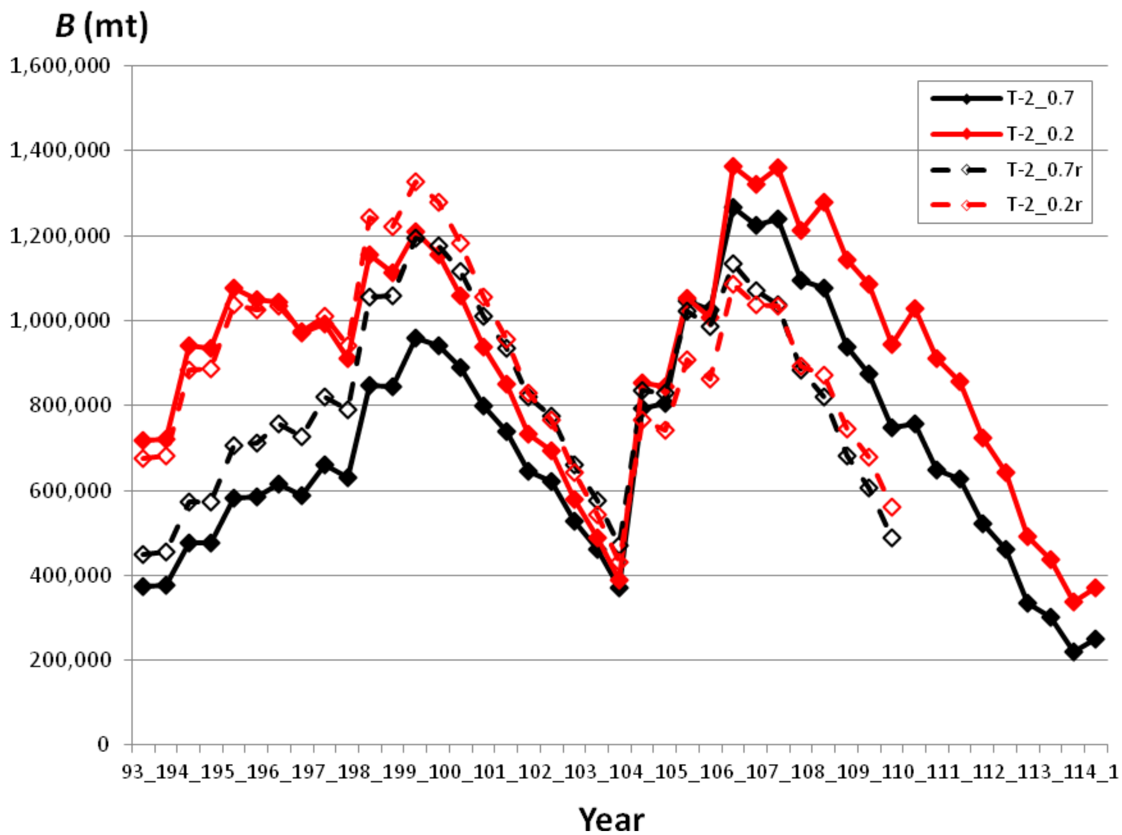

Figure 5. Biomass trajectories for models T-2_0.2 and T-2_0.7 and variants thereof that ignore data for the last four years.

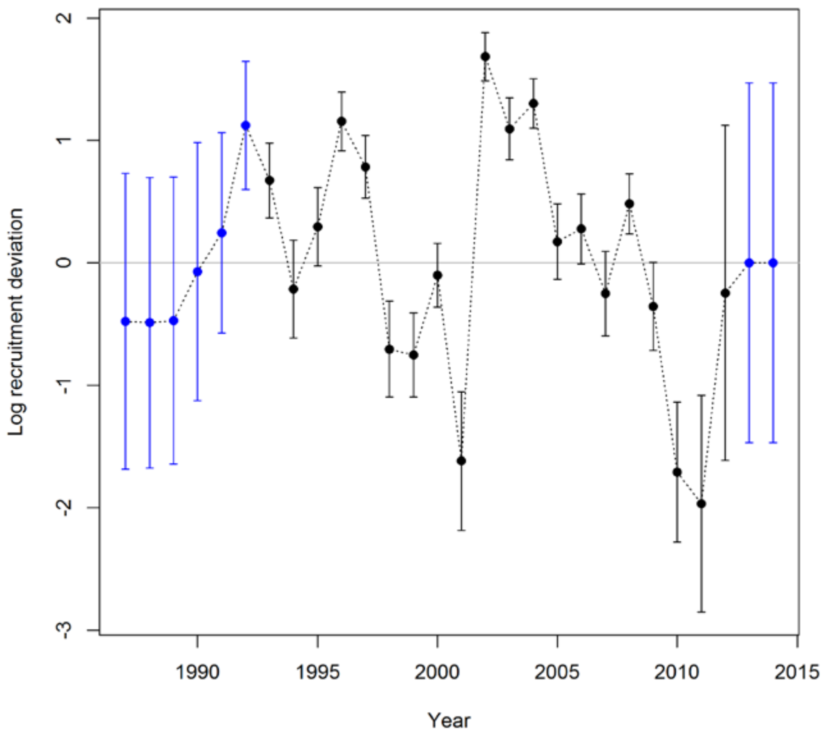

Figure 6. Estimates of the recruitment deviations for model G with their asymptotic standard errors.

Figure 7. Results of a retrospective analysis based on the final base model T-2_0.2.

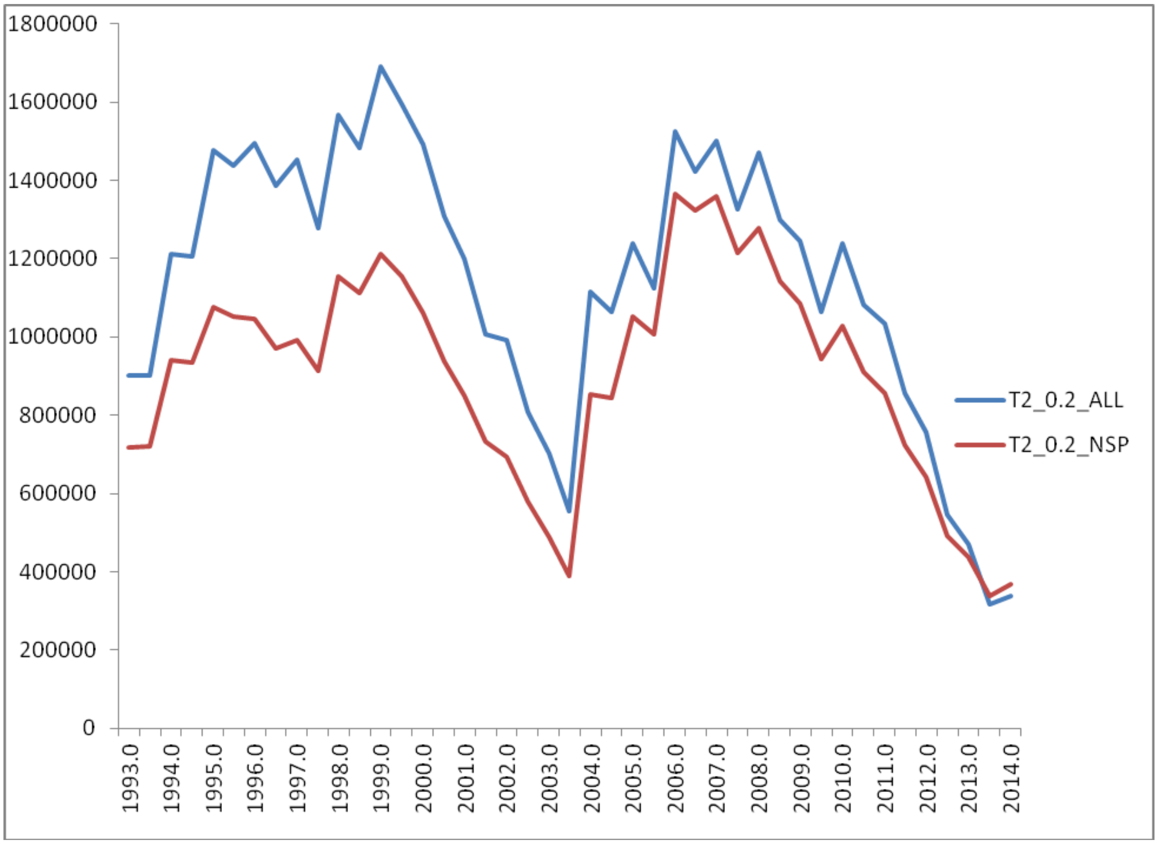

Figure 8. Comparison of the biomass trajectory for model T-2_0.2 when it is applied to the NSP only (differentiating catches) and the total catch time series.

Appendix 1
2014 Pacific Sardine STAR Panel Meeting Attendees

STAR Panel Members
André Punt (Chair), SSC, University of Washington
Meisha Key, SSC, CDFW
José De Oliveira, CIE Reviewer, CEFAS
John Simmonds, CIE Reviewer, ICES

Pacific Fishery Management Council Represenatives
Diane Pleschner-Steele, CPSAS Advisor to STAR Panel
Chelsea Protasio, CPSMT Advisor to STAR Panel
Kerry Griffin, PFMC

STAT Members
Kevin Hill, SWFSC
Paul Crone, SWFSC
Dave Demer, NOAA / SWFSC
Juan Zwolinski, NOAA / SWFSC
Emmanis Dorval, NOAA / SWFSC
Beverly Macewicz, NOAA / SWFSC

Other Attendees
Jenny McDaniel, SWFSC
Kirk Lynn, CDFG
Dale Sweetnam, SWFSC
Erin Reed, SWFSC
Ed Weber, SWFSC
Josh Lindsay, NMFS WCR
Russ Vetter, SWFSC
Al Carter, Ocean Companies
Richard Carroll, Jessie's Ilwaco Fish Company
Elizabeth Helmers, CDFW
Nancy Lo, SWFSC
Sam McClatchie, SWFSC
Richard Parrish, NMFS Emeritus
Yukong Gu, SWFSC
Jeff Laake, AFSC
Kevin Piner, SWFSC
William Watson, SWFSC
Elaine Acuňa, SWFSC
Anna Holder, CDFW
Joel Van Nord, CWPA
Noelle Bowlin, SWFSC
Mike Okoniewski, Pacific Seafood
Cisco Werner, SWFSC

Sarah Shoffler, SWFSC
Kristen Koch, SWFSC
Chris Francis, NIWA
Emily Gardner, SWFSC
Alex Da Silva, IATTC
Steven Teo, SWFSC
George Cutter, SWFSC
Mark Maunder, IATTC

AFSC – Alaska Fisheries Science Center
CDFW – California Department of Fish and Wildlife
CEFAS - Centre for Environment, Fisheries & Aquaculture Science
CPSAS - Coastal Pelagic Species Advisory Subpanel
CIE – Council on Independent Experts
CPSMT - Coastal Pelagic Species Management Team
CWPA – California Wetfish Producers Association
IATTC – Inter-American Tropical Tuna Commission
ICES – International Council for the Exploration of the Sea
NIWA - National Institute of Water and Atmospheric Research
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

Appendix 2
Email from Tom Jagielo regarding the 2013 aerial survey

Hi Kevin,

I just completed crunching the numbers for the 2013 aerial sardine survey. We are now in the process of preparing a survey report with all the details about the 2013 sampling season, but I wanted to forward the "bottom line" to you in advance of finishing that.

The survey occurred on 8-12-2013 and 8-13-2013 and covered a latitudinal distance of about 48 miles, ranging from The Columbia River to the area offshore of Garibaldi, OR. A total of 21 transects were used for the analysis.

Biomass = 160,763
CV = 0.3488

As noted previously, no new point sets were conducted in 2013. Thus, the biomass estimate was derived using the same point set data as last year (n=123 collected from 2008-2012).

Also noted previously, no bio-data were collected in 2013. Thus, I have no new length composition data for you. In previous years, we saw very good agreement between length comps from the fishery and the point sets sampled. In general, both operate in the same area using the same gear. This suggests that fishery length comps could serve as a proxy for estimating selectivity for the survey, depending on what you may have from the fishery in 2013.

Please do not hesitate to call me with any questions.

Thank you for your consideration,

Tom

Appendix 3
Progress related to the recommendations from ATM survey review
Juan Zwolinski and David Demer

1. Immediate (prior to the next stock assessments)

a. Analyses be conducted using auxiliary information (e.g. trends in density along transects, information from ichythoplankton surveys south of the survey area, and catch information) to provide best estimates for the biomass outside of the survey area, as well as the range of possible biomass levels.

Response: During spring surveys (i.e., April and early May), the northern stock of Pacific sardine resides ~30-70 m deep and spawn offshore of central and southern California. During summer surveys, (i.e., June through August), the same stock resides shallower and closer to the shore off central California, Oregon, Washington, and Vancouver Island. The sardine biomass estimates from the spring and summer ATM surveys during 2008 (Demer et al., 2012), 2012 (Zwolinski et al. in Hill et al. 2012), and 2013 (Zwolinski et al. in Hill et al. 2013) were not statistically different, indicating that any biomass outside of the survey areas are small compared to the stock biomass and the survey precision.

b. The CVs for the estimates need to be modified to fully account for the uncertainty of the trawl data.

Response: In the case that the trawl information was used to characterize independently the length and species composition of each transect (i.e., by having at least one transect per trawl), bootstrapping of the transect means would provide an unbiased of the sampling CV (Demer et al., 2012). Since 2011, efforts were made to obtain a larger number of trawls in order to get closer to the full independence of the transects.

2. Short-term

a. Investigate potential species selectivity effects by comparing the ratios of catch rates and acoustically-estimated densities in areas where single species dominate.

Response: There are strong limitations on the use of the surface, night-time trawls as quantitative measurements of fish density that preclude us to compare them to the measurements of daytime, depth-integrated fish densities from acoustics. The three main ones are: 1) There is strong vertical variability on the opening of the net by trawling at the surface, especially under bad weather; 2) It is difficult to determine with accuracy the horizontal dimension of the net to be used in the calculation of the swept area. Some studies suggest that the herding of fish begins at the doors, which have a distance much larger than that of the horizontal dimension of the net; 3) For the data already collected, there is no way to determine if all the fish that were vertically integrated by the echosounder are contained in the depth interval spanning the surface and the foot rope.

b. Compare total CPS backscatter along transects to trawl catch rates using statistical techniques.

Response: Positive trawls were associated with acoustic samples with significantly higher than average backscatter (Zwolinski et al., 2012).

c. Conduct sensitivity tests in which stations are pooled and allocated to acoustic values over a larger area.

   Response: The trawl catches from each night are pooled. Species and size composition data from these "trawl clusters" are associated to the most proximate acoustic samples (see Appendices A and B in Hill et al., 2012).

d. Consult experts in trawl design to evaluate the current trawl design in relation to the survey objectives.

   Response: Trawl experts have been consulted.

e. Develop methods that categorize the acoustic record and thus support automatic species identification and continue to work on definition and precision of the VMR process

   Response: Due to the overlap in size of the various schooling CPS, acoustic classification of species is inherently difficult when the number of samples within a school is small (for example, when using a large interval between pings when recording acoustic data over 750 m depth while conducting at a survey 10 kts). The first approach to ameliorate the quality of the data was the development the EK60 Adaptive Logging software (EAL). This software allows the reduction of the interval between acoustic pings when the bottom is shallower than 750 meters, effectively increasing the sampling intensity of schools observed over the continental shelf and slope.

   The VMR is part of a larger algorithm aiming to identify and eliminate the backscatter of non-CPS targets from echograms. The algorithm is tested on a survey basis to ensure that the retained backscatter of the echoes identified as CPS is at least 95% of the original backscatter.

f. Evaluate the potential use of the echosounder in a non-vertical position.

   Response: Multibeam observations have been made of CPS schools since the initial ATM survey in 2006. These data have been used to evaluate potential avoidance of CPS to the survey ship (see report of the PFMC/CIE review of the ATM). The new FSV Reuben Lasker is equipped with Simrad EK60, ME70, MS70, and SX90 echosounders/sonars, which will facilitate improved characterizations of fish behaviours and abundances.

g. Check the filtering algorithm every year to ensure that it is still suitable under changing conditions.

   Response: The filtering results are checked on a subset of fish schools during every survey to ensure that at least 95% of the acoustic backscatter of CPS schools is retained in the filtered echograms.

h. Study trends in frequency response over depth strata in schools.

   Response: We observed that the CPS echoes of tightly schooling fish in areas with positive trawls for anchovy, mackerels, and sardine had very little depth contrasts due to their association with the mixed layer. There, there were no obvious patterns of variability in the frequency response of the schools.

i. Compare results from the 18-kHz and other transducers to examine possible avoidance reactions.

   Response: The recommendation is unclear.

j. Continue to consider the advantages and disadvantages of conducting ATM surveys at different times of the year.

Response: This was addressed in the January 2014 CIE review of the summer sardine-hake survey (SaKe).

k. Evaluate the potential to give age-based abundance or biomass estimates for sardine and consider their utility in the SS3 assessment, given the lack of contrast in length-at-age at older ages and the ability to directly estimate total mortality from the survey result.

   Response: Age-based abundances can be estimated from the ATM using age-to-length keys derived from sardine collected on the survey themselves, or from a composite age-to-length key from the fisheries.

   The ATM survey showed the persistence of dominant cohorts over time, allowing the estimation of total and natural mortality (Zwolinski and Demer, 2013).

l. Conduct standard (ICES) vessel noise measurements for all vessels.

   Response: Vessel noise measurements are made for all NOAA FSVs. Noise measurements have not been made for RV Ocean Starr, formerly RV David Starr Jordan.

3. Long-term

a. Evaluate if different trawling practices or gears, or both would be beneficial.

b. Use the current variance estimation procedure to investigate the trade-offs in terms of variance of different time allocations between acoustic transect and trawl data collection.

c. Use a trawl/vessel configuration that can support directed trawl sampling.

d. Conduct repeated trawl sampling experiments to obtain a better understanding of small-scale variability.

   Response: The current sampling technique involves three trawls per night with inter-trawl distance of less than 10-nmi.

e. Test the efficiency and selectivity of the trawl by comparing samples from same area taken with the survey trawl and purse seine.

f. Apply state-of-the-art acoustic and optic technology to investigate fish behavior and escapement at various critical positions of the trawl.

   Response: Cameras attached to the trawl in front of the cod end have been developed and used extensively in the spring and summer 2013 surveys to observe and quantify fish behaviour and MMED performance.

g. Conduct validation tows on various kinds of backscatter to assure that the filtering algorithm is performing as intended to apportion backscatter to CPS.

h. Make efforts to obtain TS measurements for *in situ* CPS in the California Current Ecosystem.

i. Focus on utilizing more advanced instrumentation and resource-demanding research for studying vessel impacts.

   Response: The state-of-the-art instrumentation aboard the FSV Reuben Lasker (EK60s, ME70, MS70, SX90) should facilitate studies of fish behaviour that could potentially impact the estimations of abundances.

References

Demer, D. A., and Zwolinski, J. P. 2013. Corroboration and refinement of a method for differentiating landings from two stocks of Pacific sardine (*Sardinops sagax*) in the California Current ICES Journal of Marine Science.

Demer, D. A., Zwolinski, J. P., Byers, K., Cutter Jr., G. R., Renfree, J. S., Sessions, S. T., and Macewicz, B. J. 2012. Seasonal migration of Pacific sardine (*Sardinops sagax*) in the California Current ecosystem: prediction and empirical confirmation. Fishery Bulletin, 110: 52-70.

Hill, K., Crone, P. R., Lo, N. C. H., Demer, D. A., Zwolinski, J. P., and Macewicz, B. J. 2012. Assessment of the Pacific sardine resource in 2012 for U.S. Management in 2013. Pacific Sardine Assessment Update Report, Agenda Item G.3.b. Supplemental Attachment 1: 51 pp.

Zwolinski, J. P., and Demer, D. A. 2013. Measurements of natural mortality of Pacific sardine (*Sardinops sagax*). ICES Journal of Marine Science.

Zwolinski, J. P., Demer, D. A., Byers, K. A., Cutter, G. R., Renfree, J. S., Sessions, S. T., and Macewicz, B. J. 2012. Distributions and abundances of Pacific sardine (*Sardinops sagax*) and other pelagic fishes in the California Current ecosystem during spring 2006, 2008, and 2010, estimated from acoustic-trawl surveys. Fishery Bulletin, 110: 110-122.

Zwolinski, J. P., Emmett, R. L., and Demer, D. A. 2011. Predicting habitat to optimize sampling of Pacific sardine (*Sardinops sagax*). Ices Journal of Marine Science, 68: 867-879.

Appendix 4
Full CPSAS representative comments
Diane Pleschner-Steele

The CPSAS representative commends the Panel and STAT for their significant body of work throughout the 2014 sardine STAR panel. Unfortunately, the 2014 sardine assessment encountered the same basic difficulty with scaling issues observed in the 2011 assessment. The SS model is very sensitive to weighting of the input length and conditional age at length (CAAL) compositions from the ATM surveys. Most of the work at the meeting was spent making further analyses to resolve the source of these problems; which included very high variability in the biomass estimates for the first half of the time series. It became apparent from sensitivity runs that data weighting matters. The STAT and Panel attempted to find a solution that made results less sensitive by down-weighting certain conditional age-at-length data.

The sardine assessment model was improved by a more realistic separation of the landings from the northern and southern stocks (excluding the landings of southern stock sardine from Ensenada or Southern California). This reduces the biomass estimate and largely resolves problems associated with the distribution parameter in the harvest guideline.

The final base model ultimately fixed catchability (Q) at 1 for the ATM surveys, as in prior years, attempting to achieve model stability. The CPSAS has voiced concern in the past that acoustic surveys as currently deployed have been unable to measure the full biomass, particularly in the Pacific Northwest. For example, in 2012 the ATM survey went through waters from Newport to the Canadian border in 11 days and estimated the total biomass for that area at 13,000mt. We understand that the CV for that survey leg was estimated at 0.63. In the same 11 days the fishery landed 9,747mt. Previous to the arrival of the NOAA vessel the harvest in that area was 35,531mt. After the NOAA vessel left those waters the harvest was 32,781mt for the remainder of the season. The point is that fishermen observed and caught significantly more fish in the area than the point estimate of the ATM cruise – which measured only one spot in time but contributed to a low overall sardine biomass estimate. In contrast, the NWSS-sponsored aerial survey for that summer (which was later down-weighted due to too few point sets) estimated more than 900,000mt in the PNW. The inconsistency in the two data points remains unresolved.

On behalf of the CPSAS and industry at large, the CPSAS representative also expresses disappointment that the aerial survey has been dropped from consideration in this and presumably future stock assessments. It should be noted that the rationale for eliminating the aerial survey, "*vulnerability of this survey method to prevailing ocean conditions potentially affecting q over short and long time frames (water clarity, sea state, water column stratification, and associated changes in vertical distribution,…*" *co*uld be applied to other fishery independent indices as well. Moreover, the aerial survey assumption that daylight-photographed schools represented sardines was questioned by comparing species composition from night-time ATM trawls. The CPSAS notes that schooling patterns day vs. night differ and should not be compared.

The CPSAS also voices concern that stock assessments seem to be gravitating toward one independent index based on ATM surveys. We encourage a continuation of multiple surveys, recognizing that each survey type has issues with varying ocean conditions and assumptions. Although the CPSAS and industry express serious reservations about use of only one index for sardines developed solely around the ATM survey, we acknowledge and applaud the acquisition of the RV Reuben Lasker and its capability to survey with forward and side-scanning sonar. We can support the ATM with the use of sonar to augment acoustic search of water columns that the downsounder does not effectively measure (i.e. the top 10 meters of the water column). Further, sonar can offer clues to school behavior. As stated by a sitting Council member who has had many years of experience fishing for sardines: First choice: sonar: second choice spotter plane: third choice downsounder.

Ultimately, industry wants to see a sustainable resource (to the degree that environmental conditions will allow) that is in no danger of being overfished. Current sardine stock assessments and harvest policy are very precautionary. We sincerely hope that going forward we can develop a truly collaborative research program for the CPS complex.

Recommendations:

- Continue to involve industry in collaborative research.
- Recognize that the 2014 assessment is "déjà vu all over again" and most of the unresolved problems and major uncertainties listed in the 2011 STAR panel report still exist
- Also, many of the research recommendations in 2011 also are applicable in 2014, i.e.
    - Explore models which consider a much longer time-period (e.g. 1931 onwards) to determine whether it is possible to model the entire period
    - Consider model configurations which use age-composition rather than length-composition and conditional age-at-length data given evidence for time- and spatially-varying growth.