Pacific Sardine
STAR Panel Meeting Report

NOAA / Southwest Fisheries Science Center
La Jolla, California
February 21-23, 2017

STAR Panel Members:
André Punt (Chair), Scientific and Statistical Committee (SSC), Univ. of Washington
Will Satterthwaite, SSC, Southwest Fisheries Science Center
Evelyn Brown, SSC, Lummi Natural Resources, LIBC
Jon Vølstad, Center for Independent Experts (CIE)
Gary Melvin, Center for Independent Experts (CIE)

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

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

1) Overview

The Pacific Sardine Stock Assessment and Review (STAR) Panel (Panel) met at the Southwest Fisheries Science Center (SWFSC), La Jolla, CA from February 21-23, 2017 to review a draft assessment by the Stock Assessment Team (STAT) for the northern subpopulation of 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 Council FTP site.

Drs. Paul Crone, Kevin Hill, and Juan Zwolinski presented the assessment methodology. Paul Crone first outlined the assessment philosophy, which focused on selecting an approach that made most use of the data source considered by the STAT to be most objective, i.e. the Acoustic Trawl Method (ATM) survey. The STAT provided results for two assessment approaches: (a) use of the summer 2016 ATM survey estimate and associated age-composition projected to 1 July 2017, and (b) a model-based assessment that provides an estimate of age 1+ biomass on 1 July 2017.

Juan Zwolinski described the survey-based method for estimating age 1+ biomass on 1 July 2017, which involved estimating numbers-at-age on 1 July 2016 from the summer 2016 ATM survey from numbers-at-length using an age-length key that pooled data over multiple summer surveys, and projecting these numbers forward accounting for natural mortality and growth, and adding the estimated recruitment for 2016. The recruitment for 2016 was based on the stock-recruitment relationship estimated by model ALT, and the spawning stock biomass for 2016 was estimated by back-projecting the summer 2016 numbers-at-age to 1 January 2016.

Kevin Hill and Paul Crone described the data on which the model-based assessment was based, as well the results from a draft assessment utilizing the Stock Synthesis Assessment Tool, Version 3.24aa. Model ALT differed from the model on which the 2016 update assessment was based by starting the assessment in 2005 rather than 1993, excluding the Daily Egg Production Method (DEPM) and Total Egg Production (TEP) indices, estimating rather than pre-specifying stock-recruitment steepness, pre-specifying weight-at-age rather than estimating it within the assessment, assuming that selectivity for the ATM survey is zero for age 0 and uniform for age 1 and older, estimating survey catchability (Q), assuming that selectivity is age- rather than length-based, modelling ages 0-10+yr rather than ages 0-15+yr, assuming natural mortality (*M*) is 0.6yr^{-1} rather than 0.4yr^{-1} for all age classes and fitting the catch and ATM survey age-composition data (rather than the associated length-composition data). Unlike the 2016 and earlier assessments, model ALT included additional live bait landings, which generally reflected a minor contribution to the total landings in California. However, model ALT did not include biological composition data from the live bait catches, given this fishery sector had not been regularly sampled in the past, with samples being available for only the most recent year of the time period modelled in the assessment.

The review and subsequent explorations of the assessment through sensitivity analyses were motivated primarily by the need for the survey-based method to provide an estimate of age 1+ biomass and its CV, to better understand the rationale for the changes made to the model on which the last full assessment was based that led to model ALT, and to identify the best approach for providing an estimate of age 1+ biomass on 1 July 2017. The Panel had several comments and concerns regarding the ATM survey methodology and ways in which estimates of close-to-absolute abundance can be obtained. However, this was not a review of the ATM survey, since a second Council-sponsored ATM methodology review is planned for early 2018. Therefore, comments regarding the ATM survey and how estimates of abundance from that survey are constructed are reflected primarily in the Research Recommendations section of the report.

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) <u>Day 1 requests made to the STAT during the meeting – Tuesday, February 21</u>

Request 1: Provide documentation on the procedures used to calculate the survey age-composition data, including how age-length and age-biomass keys are constructed.

Rationale: These calculations are critical to projecting biomass after accounting for natural mortality, somatic growth, and recruitment; but the draft assessment document did not describe these calculations in sufficient detail for them to be reproduced. In addition, the age-compositions for the ATM survey in model ALT were computed using the method.

Response: Dr. Zwolinski presented written documentation and figures. The function "multinom" from the R package "nnet" fits a multinomial log-linear model using neural networks. The response is a discrete probability distribution (see Fig. 1). It is simpler to use than the alternative (sequential logistic models), and it provides a smoother transition between classes than an empirical age-at-length key. The age and lengths used for constructing the age-length key were from surveys from 2004 to the present. Due to the assumption of a July first date and its effect on ageing, the STAT built a season-specific age-length key using data pooled across time, separately for spring/summer.

The Panel agreed that aggregation across years is not appropriate if some length classes represent multiple ages, which is the case for Pacific sardine. Moreover, substantial spatial and temporal variation occurs in size-at-age, and merging the data from several years creates bias in annual estimates of age compositions of varying magnitude and direction.

Request 2: Provide full specification, including equations, of the calculations used to 1) project from the ATM survey biomass estimate to the estimated age-1+ biomass on July 1 of the following fishing year, and 2) calculate the uncertainty associated with that biomass estimate.

Rationale: The projection calculations need to be reproducible. Management advice (Overfishing Level OFL, Acceptable Biological Catch ABC, and Harvest Guideline HG) for Pacific sardine requires an estimate of age 1+ biomass (OFL, ABC, HG) and its uncertainty (ABC) on July 1, 2017.

Response: For 1), Dr. Zwolinksi walked the Panel through a spreadsheet that made these calculations and the Panel agreed that the calculations were sensible, conditional on the age-weight key. For 2), assuming independence of age 1 and age 2+ biomass, the total variance was calculated by summing the respective variances. This calculation is negatively biased because it ignores uncertainty in age-composition and weight-at-age. It was noted that the resultant coefficient of variation (CV) for age 1+biomass is lower than the CV for either component (age 1 versus age 2+) due to their assumed independence.

Request 3: Plot cohort-specific rather than year-specific growth curves (weight-at-age) for the ATM survey and overlay raw data/information on sample sizes. Make it clear which values are estimated versus inferred. Do this for the fisheries data as well.

Rationale: Cohort-specific curves are easier to interpret as growth trajectories than year-specific curves. It is important to understand how much data drives these estimates, and to understand the consequences of applying the same age-length key for all years with survey data to calculate the weight-at-age and age-composition for the ATM survey.

Response: Dr. Hill presented tables including sample sizes and estimated means for each cohort-season-age combination. The tables were formatted to highlight entries that were inferred versus estimated. Dr. Hill calculated means whenever 3 or more samples were available. However, these means were sometimes overwritten based on the assumption that animals did not shrink. The ATM data showed substantial variation in weight-at-age across years (Fig. 2), and possibly increasing size-at-age in recent years. The MexCal catch data appeared less variable overall, and it was noted that fishery sample sizes were generally larger than the ATM sample sizes. The smoothing was not applied to the PNW catch.

The Panel noted that the adopted method ended up discarding data for cohorts with unusually large mean sizes for (for example) age-0 fish by not allowing "shrinkage", whereas it may have been the age-0 means that were anomalous rather than the means calculated for older ages. The Panel also noted that in many cases, the sample sizes were very small. The weight-at-age key used within the survey-based projection did not exclude "shrinkage". Using the weight-at-age key in model ALT produced an imperceptible difference in model-estimated age 1+ biomass.

Request 4: Verify that model ALT was run with ATM survey selectivity set equal to 0 for age-0 fish. Contact Dr. Rick Methot to better understand how selectivity is being modeled under the chosen selectivity option in SS.

Rationale: The model outputs appear to indicate that the model predicts non-zero catches of age-0 fish despite the intent to specify selectivity to be zero on age-0 fish. This may have significant unintended consequences for the likelihood calculations.

Response: This question was not fully resolved. It appears that Stock Synthesis predicts some catch of nominal "age 0" even given selectivity of zero on true age-0 fish because aging error leads to the expectation that some age-1 fish will be caught and mis-categorized as age 0. Further, model runs revealed that the model was unable to converge if aging error was set to zero or made very small, but reductions in the specified aging error led to the expected reduction in the predicted age-0 catch. It was noted that surveys likely include a mix of age-1 fish mis-categorized as age-0, as well as fish that are truly age 0.

Dr. Methot also noted that Stock Synthesis had not been as thoroughly debugged for semester-based models as for strictly annual models.

See also Requests 5, 8, and 9.

Request 5: Re-run model ALT with age 0 fish removed from the input file for the ATM survey.

Rationale: Similar to Request 4, the model likelihood should not be influenced by data on age-0 fish if it is assumed selectivity on age-0 fish is zero, but the model appears to be generating non-zero predictions and comparing these against the input data.

Response: The model still predicted catch of age-0 fish in this scenario. This is consistent with the explanation suggested for this pattern under Request 4.

Request 6: Report the CV of the estimate of terminal biomass based on changes in how the compositional data are weighted.

Rationale: The weighting of composition data appeared to have little effect on the point estimate of biomass, but it is important to understand implications of alternative weighting schemes for uncertainty as well.

Response: Data weighting increased the CV by 2-3%. The base model had a CV of approximately 36%, Francis-weighting led to a CV of approximately 38%, and harmonic mean weighting led to a CV of about 39%.

Request 7: Show more outputs from T_2017 and T_2017_No_New_AT_Comp.
Rationale: These outputs would help the Panel evaluate the reasons for proposing a move away from a strict update of the previously accepted model structure, i.e. identify problems with a strict update that the new model structure addresses.
Response: Selectivity curves for the spring and summer ATM surveys were noticeably different depending on whether the two most recent survey length-compositions were included in the assessment or not (Fig. 3). These models appeared to yield acceptable fits to abundance indices, but the fits to observed length-compositions were poor. It appears that the model estimates very low selectivity on small fish for the summer survey (since selectivity does not vary across years, and very few small fish are encountered most years) such that when small fish are encountered, they are expanded to a very large number. During Panel discussion, it was noted that this unexpected behavior should not happen if selectivity were forced to be the same for the spring and summer surveys.

Day 2 requests made to the STAT during the meeting – Wednesday, February 22
Request 8: Develop a model in which selectivity for age-0 animals in the survey is time-varying.
Rationale: The availability of age-0 animals to the survey seems to be highly variable among years, but influential on the results. A selectivity function in which age-0 selectivity varies among years should "discount" the influence of occasional catches of age-0 animals.
Response: A model was presented that assumed essentially full selection on age-1+ animals, and time-varying age-0 selectivity. The model estimated nearly zero selectivity on age-0 fish in all years except 2015, when estimated selectivity on age-0 fish was nearly 1.0 (atypically large pulse of small/young fish observed in summer 2015). Fits to composition data were similar to those for model ALT, except that the spike of age-0 fish in 2015 was captured better. The estimate of age 1+biomass on 1 July, 2017 for this model was 77,845 t.

Request 9: Run a variant of model ALT in which the age-composition data are assigned to a new fleet (6) that has logistic selectivity (estimated separately for the spring and summer periods).
Rationale: Selectivity for the ATM survey is assumed to be uniform on animals aged 1 and older so age-composition data are not required for this survey. The selectivity pattern for the trawl component of the survey is not uniform on age-1+ animals (some age-0 animals are caught) and it may be possible to represent this using a logistic selectivity function.
Response: This model performed generally similar to a logistic formulation applied to the ATM survey for both age-composition and as an abundance index, but it misses the summer 2016 ATM survey estimate of biomass from above whereas the logistic fits that estimate closely. However, the logistic model had a negative log-likelihood of approximately 311, compared to 305 for this variant, and 333 for model ALT. Thus, both a model with logistic ATM selectivity and a model that assumed 1+ selectivity for ATM survey estimates and logistic selectivity for the associated age-composition data fit the data somewhat better than model ALT.

Request 10: Conduct a retrospective evaluation of how well alternative assessment methods can predict the biomass from the summer ATM surveys. For each year Y for which there is a summer ATM survey estimate for year Y and year Y+1, report predictions of year Y+1 biomass based on

(a) the estimate of biomass from the results of the ATM survey during summer of year Y, (b) the estimate of biomass based on applying the projection method to the results from the ATM survey in summer of year Y, and (c) model ALT based on data through year Y.
Rationale: The Panel wished to understand which method was able to predict the ATM survey estimate of biomass most accurately.
Response: The STAT provided results for the three selected approaches as well estimates of age 1+ biomass obtained by projecting the actual assessments used for 2012, 2013, 2014 and 2015 forward ("Past assessment" in Fig. 4) and estimates of age 1+ biomass obtained by projecting the model used for 2014, 2015 and 2016 management advice ("2014 formulation"). Model ALT generally came closest to predicting the survey biomass estimate the following year, doing so by a substantial margin for 2014. "Past assessment" was usually the worst. Model ALT had the lowest residual variance. Relative errors were a CV of 1.07 for Model ALT, 1.26 for the 2014 model formulation, 1.50 for the last survey without projection, 1.62 for the values adopted in management specifications, and 1.70 for projections from the previous ATM survey (see Appendix 2 for the specifications for the method).

Day 3 requests made to the STAT during the meeting – Thursday, February 23
Request 11: Develop a method for estimating recruitment solely from ATM data, explain how these recruitment estimates could be used to project forward from an ATM biomass estimate, and then add results for that method to the retrospective comparison described in Request 10.
Rationale: During discussion of Request 10, it was clear that much of the concern regarding the currently proposed method of projecting from the survey was its dependence on model ALT for stock-recruitment estimates for conducting the projection, resulting in its dependence on the same assumptions the STAT was hoping to avoid by moving away from an integrated assessment. It was pointed out that it could be possible to develop estimates of age 1 biomass on 1 July, 2017 strictly from the ATM data.
Response: The STAT modified the survey projection method so that projected biomass of 1-year-olds was the average over the most recent five years (see Appendix 2 for details). As desired, this approach was not tied to the model ALT. However, the residual standard deviation for this approach ("Survey projection 2"), while better than "Survey projection", was still worse than Model ALT and the 2014 model formulation (1.45) (Fig. 4).

3) Technical Merits and/or Deficiencies of the Assessment

Alternative assessment approaches
The Panel considered four ways to estimate age 1+ biomass on 1 July 2017: (a) use the estimate of biomass from the summer 2016 ATM survey, (b) project the estimate of biomass from the summer 2016 ATM survey to 1 July 2017 using the 'survey projection' model (or an alternative approach), (c) model ALT, and (d) the model on which the 2014-16 assessments were based. The Panel had concerns with, and comments on, all of these methods:

- Assuming that the 1 July 2017 biomass equals the estimate of biomass from the summer 2016 ATM survey ignores mortality (from natural causes and from fishing), growth and recruitment from July 2016 to July 2017. However, this method is simple to implement because it does not rely on a model, nor does it rely on estimates of age composition for which sample sizes are low.
- Projecting the biomass from the 2016 ATM survey to 1 July 2017 accounts for mortality, growth and recruitment from July 2016 to July 2017. However, the approach used to

convert from length composition to age composition is incorrect, and the method used to derive the CV of age 2+ biomass does not allow for uncertainty in population age composition, projected weight-at-age and maturity-at-age. In addition, the method relies heavily on model ALT because approximately half of the age 1+ biomass on 1 July 2017 consists of age-1 animals, i.e. the estimate of this biomass is based to a substantial extent on the stock-recruitment function from model ALT. Finally, the value for M of 0.6yr^{-1} has no clear justification. The version of the projection model provided initially to the Panel did not account for catches so it could not be applied were the targeted sardine fishery to be re-opened, and does not account for the limited catches during 2016.

- Model ALT has several of the problems associated with the 'survey projection' model, i.e. the age-composition data are based on a year-invariant age-length key, and the basis for M=0.6yr^{-1} lacks strong empirical justification (and indeed likelihood profiles indicate some support for lower M than the value adopted for model ALT). In addition, the model presented to the Panel predicted age-0 catch in the ATM survey even though it is assumed that age-0 animals are not selected during the ATM survey. It appears that the model predictions of age-0 animals in the ATM survey are actually model-predicted numbers of age-1 animals that are predicted to be mis-read as age-0 animals. However, examination of the ATM survey length-frequencies suggests that that some age-0 animals (or animals that were spawning earlier in the year) are encountered during the surveys (Fig. 5). Model ALT estimates Q to be 1.1, which is unlikely given some sardine are not available to the survey owing to being inshore of the survey area.
- The model on which the 2014-16 assessments were based was approved for management by the 2014 STAR Panel. However, that assessment had some undesirable features, including extreme sensitivity to the occurrence of small (<~15cm fish) in the ATM surveys, poor fits to the length-composition and survey data, and sensitivity to the initial values for the parameters (i.e. local minima). These sensitivities and the resultant high uncertainty about population scale were noted in previous reviews.

The Panel explored alternatives to the current selectivity formulation to better understand why model ALT was predicting age-0 catch when selectivity for age-0 fish was set to zero. It was noted that the results are generally robust to assuming that selectivity is a logistic function of length (but that implies that some age-1+ animals are not available to the ATM survey), allowing for time-varying age-0 selectivity, and estimating a separate selectivity pattern for ATM survey age-composition data.

The Panel noted that the 'survey projection' model and model ALT both rely on the samples from the ATM surveys to compute weight-at-age and survey age-composition data. The samples sizes for age from each survey are very small (16 – 1,051), which means that estimates of, for example, weight-at-age are highly uncertain. The procedure of ensuring that weight-at-age for a cohort does not decline over time seems intuitively correct. However, if the estimated mean weight of young fish in a cohort is anomalously high or low due to sampling errors (owing to small samples), it can impact the weight-at-age of that cohort for all subsequent ages.

Model ALT estimated steepness rather than fixing it equal to 0.8. The results were not sensitive to fixing versus estimating steepness, but the estimate of 0.36 was low.

Selection of an assessment approach

The Panel considered the merits of the various approaches. It concluded that:

- The approach on which 2014-16 management was based exhibited undesirable assessment diagnostics, and produced extremely high estimates of recruitment when large numbers of small fish were observed in the ATM survey length-frequencies. The approach also performed poorly in retrospective analysis (Fig. 4)[1]. The Panel and STAT agreed that this approach should not be used for 2017 management.
- The survey projection method (and the modified version, "Survey projection 2") seems a viable and defensible way to estimate age 1+ biomass using the ATM survey results, especially if the method could be modified to not use the results from model ALT. However, as currently formulated, this method performs no better than assuming that the age 1+ biomass in July 2017 equals the survey estimate of biomass for summer 2016 (Fig. 4). Thus, while viable, this approach requires further development and review prior to adoption.
- Estimating the biomass on 1 July of year Y+1 based on the ATM survey estimate for year Y is simple, but the Panel was concerned that this method ignored catches during year Y and may lead to additional risk. Thus, the basic approach is viable, but needs additional testing prior to adoption.

Given the current management approach that requires an estimate of age-1 biomass at the start of July, the Panel and STAT agreed that model ALT was the best approach at present for conducting an assessment for the northern subpopulation of Pacific sardine, notwithstanding the concerns listed above. The results from the assessment are robust to changes to how selectivity is modelled, the value for steepness and data weighting, but there were several concerns with this model that could not be resolved during the Panel meeting. Assuming uniform selectivity leads to lower estimates of current 1+ biomass, but this assumption reflects the expectation that all fish in the survey area are vulnerable to detection during an acoustic survey.

The final model (model ALT) 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 2005 to 2016;
- sexes were combined; ages 0-10+.
- 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: age-based selectivity (one parameter per age)
  - PacNW fleet: asymptotic age-based selectivity;
  - age-compositions with effective sample sizes calculated by dividing the number of fish sampled by 25 (externally) and lambda weighting=1 (internally);
- Beverton-Holt stock-recruitment relationship with "steepness" estimated;
- *M* was fixed (0.6 yr^{-1});
- recruitment deviations estimated from 2005-2015;
- virgin recruitment estimated, and σ_R fixed at 0.75;
- initial *F*s estimated for the MexCal S1 fleet and assumed to be 0 for the other fleets;

[1] Care needs to be taken interpreting Fig. 4 given the low number of years involved and the fact the observed 1+ biomass is subject to considerable sampling error.

- ATM survey biomass 2006-2013, partitioned into two (spring and summer) surveys, with Q estimated;
  - age-compositions with effective sample sizes set to 1 per cluster (externally);
  - selectivity is assumed to be uniform (fully-selected) above age 1 and zero for age 0.

The estimate of age 1+ biomass on 1 July 2017 from model ALT is 86,586t (CV 0.363). Model ALT indicates that age 1+ biomass has rebuilt close to that in 2014, owing to a substantial increase in biomass based on the indices from the survey (Fig. 6). The estimate of age 1+ biomass is less than the estimate of age 1+ biomass on 1 July 2016 from the 2016 stock assessment (106,137t). This is a consequence of the change in assessment methodology, in particular that selectivity for the ATM survey is assumed to be uniform for fish aged 1 and older (assuming that selectivity is logistic in model ALT increases the estimate of 1+ biomass from 86,586t to 153,020t).

Future directions

The STAT strongly supports that management advice for Pacific sardine be based on the estimates of biomass from the ATM survey rather than a projection model or an integrated assessment. The Panel notes the following ways in which management could be based on the ATM survey results.

- Change the start-date of the fishery so that the time between conducting the survey and implementation of harvest regulations is minimized.
- Use Management Strategy Evaluation to evaluate the risk to the stock of basing management actions on an estimate of biomass that could be a year old at the start of the fishing season (if the fishery start date is unchanged). Review of an updated MSE would likely not require a Methodology Panel, but could instead be conducted by the SSC.

The Panel notes that there may be benefits to attempting to use both the spring and summer ATM surveys as the basis for an ATM survey-only approach and that moving to an assessment approach that relies on the most recent ATM survey (or two) may be compromised by reductions in ship time and/or problems conducting the survey. It agrees with the STAT that there is value in continuing to collect biological data and to update model ALT even if management moves to an ATM survey-only approach.

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

The core issues for stock assessments continue to be related to the temporal and spatial scale of the surveys and insufficient sample sizes of age-length for sardine in the ATM survey. The ability of a single boat following fixed transects along the entire sardine NSP region over a single period to sufficiently observe and sample a highly mobile schooling fish that exhibits high variability in recruitment, migratory patterns and timing, school structure, and depth distribution remains a core challenge. The relatively small sample size of sardine for biological analysis remains a concern related to acoustic expansions, population model estimates, and projection forecasts that depend on age composition and size-at-age information. A solution may require more resources than SWFSC has at its disposal so that will require Council action; resolution of this issue is outside of the ability of the Panel to address.

The Panel identified concerns with all of the proposed assessment approaches as highlighted in Section 3 of this report. In relation to model ALT, the Panel was unable to fully resolve the issue of observations of age-0 animals in the ATM survey age compositions, and how to compute age-composition and weight-at-age for the ATM survey.

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

a) CPSMT issues

The CPSMT (MT) representative appreciates the substantial efforts by the STAT and the constructive Panel discussion, and offers the following comments.

The STAT proposed the ATM survey as the preferred approach over an integrated model for estimating sardine biomass. However, because the ATM survey at this time does not better estimate biomass projected to the start of the 2017-18 fishing year, the integrated model (Model ALT) was ultimately recommended. The MT representative agrees this was a reasonable approach to meet management requirements for a July 1, 2017 biomass estimate, but nevertheless also supports further consideration for shifting to the ATM survey to estimate biomass. The MT representative notes that issues of spatial and temporal coverage, and sample size remain for the survey. This has implications for the model ALT as well.

The review noted problems associated with some very small sample sizes produced by the trawl component of the ATM survey. Given that fish captured in trawls informs the species composition of the acoustic signals, as well as providing biological data, additional effort is required to refine and improve trawling operations. Additionally, more of the fish (particularly during the summer survey) that are collected need to be processed for ageing. The MT representative notes small sample size was flagged as a concern in the last full update conducted in 2014 and strongly supports the Panel recommendation that the SWFSC conduct analyses to estimate optimal sample size and to refine the survey methodology.

The lack of nearshore coverage by the ATM survey persists. Research needs to be conducted to explore possible approaches for surveying this area. Collaborative projects with industry should be encouraged to leverage their expertise. Further, emphasis should be placed on ensuring that the survey has sufficient sea-days to effectively cover the entire west coast irrespective of whether the ATM survey is used within a model or if the ATM survey is to be considered the preferred approach to inform the biomass estimate for management. The current plan to reduce the number of sea-days from 80 in 2016 to 50 in 2017 is concerning. The 50-day summer survey planned for 2017 does not include the area south of Monterey. If distance between transects were increased, the survey could possibly be extended to Point Conception, which would still not include the Southern California Bight. Fewer days at sea and the corresponding likely decrease in number of trawls also reduces the data upon which to base species composition and to produce biological data.

An MSE to evaluate the effects of using the ATM biomass estimate to inform the following year's harvest control rules is proposed as a high research priority (G). If the MSE were to find the one-year lag does create unacceptable outcomes one approach would be to develop an improved projection model. Another proposed fix would be to move the fishing year start date. While possible, the MT representative would like to highlight that the start date was adjusted beginning in 2014 to afford the STAT more time between the conclusion of field seasons and the deadline for STAR review of stock assessments. More significantly, shifting the start date can raise management issues because embedded in it is the period-based catch allocation scheme. Selecting an existing allocation period start date (January 1, July 1 and September 1) is perhaps more straightforward and would not necessarily require substantial analysis. Selecting any other starting point would likely necessitate an analysis of impacts and therefore more time to implement (i.e. two to three Council meetings). How to best accomplish aligning a shift to using only an ATM survey-derived biomass estimate with a change to the fishing year will require additional deliberation.

b) CPSAS issues

The CPSAS representative commends the Panel and STAT for their extensive and thoughtful body of work throughout the 2017 sardine STAR panel. Unfortunately, the 2017 sardine assessment again encountered the same difficulties observed in previous STAR panels. Most of the unresolved problems and major uncertainties listed in the 2011 and 2014 STAR panel reports still exist.

Earlier panels pointed out significant scaling issues. The 2017 assessment also encountered issues with ageing, notably an age-length key that was deemed incorrect. One persistent problem is the very small sample size for biological composition data obtained during ATM surveys and other sampling; another is the high variability in length-at-age observed in sardine year-to-year. As pointed out during the meeting, an age/length key averaged over seasons is not valid; it ignores differential cohort strengths. This presents a major problem in model projections, and adds another layer of uncertainty considering the current time lag between field surveys and the development of either ATM survey-based or model-based management advice for the fishery.

Assigning July 1 as the standardized birth date for sardine also presents problems, particularly in light of recent year ocean conditions that have precipitated sardine spawning earlier in the year, too early to be observed in April DEPM surveys, and producing age-0 fish assumed too small to be captured in ATM surveys. Yet an abundance of small fish exists! In fact, the 2015 summer ATM survey did encounter a spike of very small fish. A record number of pelagic juvenile sardines (and anchovies) also was found in the 2015 juvenile rockfish cruise. However, the length-composition data for the small fish were omitted from the assessment model in 2015 because the biomass estimate produced was “unrealistic.”

Ironically, none of the approaches considered at this STAR panel meeting found adequate evidence of recruitment in 2016 to boost the stock assessment “number” in 2017. In fact, the projected biomass estimate for 2017 is lower than 2016 at a time that sardines are increasing in abundance, apparently coast-wide, but certainly in California. The current report attributed this to a change in assessment methodology.

Fishermen from the Pacific Northwest and California who attended the STAR panel meeting reported that they have observed an abundance of 3-6 inch fish for the past couple of years, particularly in live bait catches. California fishermen delivered samples of these fish to the SWFSC and California Department of Fish and Wildlife (CDFW). But while the 2016 draft stock assessment did include a small number of live bait catches (now the only active non-treaty fishery for sardine on the West Coast), the corresponding biological-composition data were not aged and hence included in the assessment.

In the opinion of the fishermen, an opinion shared by this CPSAS representative, none of the four approaches considered during the panel meeting accurately reflect the biomass of sardine now in the ocean. The Panel also voiced concerns with all the methods presented; those concerns are reflected in the body of this report under **Technical Merits and/or Deficiencies of the assessment.**

The CPSAS representative highlights major concerns, including:

- The STAT now recommends the ATM survey as the most objective survey method. However, ATM surveys at present do not capture fish in the upper water column, nor a large biomass of young fish (sizes 3 inches and up) that fishermen have observed in nearshore waters since late 2014; this biomass is largely inside ATM survey tracks. But the ATM survey is assigned a catchability quotient (Q) of 1 nonetheless, meaning it "sees" all the fish. The Q for Model ALT, which is based largely on ATM survey data, is estimated at 1.1, which the STAR Panel report calls into question, given for example the unquantified volume of fish in nearshore waters.
- The summer 2016 ATM survey reported a fourfold increase in age 1+ biomass, but the biomass estimate produced is substantially lower than the estimate used for management in 2016. The STAR panel found fault with the methodology used to project the 2016 biomass to 2017. So do we – but using the 2016 ATM biomass estimate without adjusting for recruitment ignores reality.
- In addition, the proposal to simply use the biomass estimate from the summer ATM survey directly, to avoid uncertainty in model assumptions, could bypass surveying a substantial portion of the biomass if/when cruises are shortened, or disrupted. For example, the 2017 summer survey schedule is only 50 days, down from 80 days in 2016. This means the survey may not extend much below San Francisco, which will miss a substantial portion of California's historical fishing grounds.
- Also, a proposal to change the fishing season start date to more closely follow the survey, thus avoiding the need to project recruitment, is not as simple as it sounds. The current seasonal structure is tied to an allocation framework that would require serious discussion and analysis before any change could be implemented.
- At the end of the day, the STAR panel cautiously recommended proceeding with Model ALT, as the "least-worst" way to produce the age 1+ biomass estimate and CV required for management in 2017. The CPSAS hopes the SSC and Council will acknowledge all the caveats, and recognize that this is a "stop-gap" approach until the ATM methodology review can be accomplished in 2018, along with further review and improvement of Model ALT input and assumptions and potential review of other assessment indices.
- The CPSAS representative again voices concern that stock assessments appear to be gravitating toward one independent index measuring one point in time, based on ATM surveys. We strongly encourage a continuation of multiple surveys as each survey type has strengths and weaknesses. Other fishery-independent research, i.e. the juvenile rockfish survey, was informative in 2016 and should be approved to provide information for future sardine stock assessments, as this could serve as another indicator of recruitment.
- Clearly the small sample size and inadequate biological composition data are causing serious problems in assessing the sardine (and anchovy) resource. Industry has offered to help collect data, and we hope this offer will be acted upon in a way that such information can be incorporated into future stock assessments.

- As we have noted in the past, 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.

Other recommendations:

- Please work collaboratively with industry to resolve persistent data deficiencies, including assessing the nearshore, upper water column, and the need for substantial increase in sample size and biological composition data for sardine (and other CPS), particularly ageing.
- Recognize that the 2017 assessment is "déjà vu all over again" and most of the unresolved problems and major uncertainties listed in the 2011 and 2014 STAR panel reports still exist.
- Prior panel, SSC, CPSMT and CPSAS reports have recommended a methods review of the ATM survey ASAP as a high priority research and data need. We continue to emphasize this need, and further recommend that such review also encompass review of Model ALT and other potential data collection options, including the juvenile rockfish survey, CDFW/CWPA aerial survey and any other promising data collection prospects available by the time of the scheduled ATM review in January 2018.
- We also support the STAT high-priority recommendation to address: "*technical issues related to echosounder deployment and associated signal interpretation (e.g., uncertainty surrounding species-specific target strength [TS], sonar bias related to backscatter uncertainty, and areas of the upper water column that potentially are not capable of being surveyed*)."

Dr. Zwolinski noted that target strength is currently based on "similar" fish, not Coastal Pelagic Species (CPS) found in the California Current. The STAT and Panel recognized that incorrect target strength could result in both over or under-estimation of biomass

Finally, the CPSAS representative points out that improving survey and assessment methodology to accurately reflect abundance of sardine (and other CPS) is absolutely essential: the future of the industry hangs in the balance.

7) Research Recommendations

High priority

A. Conduct an analysis of effect of fish sample size on the uncertainty in the ATM biomass estimates and model outputs. Use this information to re-evaluate and revise the sampling strategy for size and age data that includes target sample sizes for strata
B. The clusters (the Primary Sampling Units, PSUs) with age-length data should be grouped into spatial strata (post-strata, or collapsed post-strata used in ATM biomass estimators). The variance in estimates of age-length compositions can then be estimated by bootstrapping of PSUs, where age-length keys are constructed for each bootstrap replicate. The sub-sample size of fish within clusters that are measured for lengths should be increased, and length-stratified age-sampling should be implemented. This approach would likely increase coverage of age samples per length class and reduce data gaps.

C. The survey projection method should be developed further. Specifically, the survey age-composition should be based on annual age-length keys, and the uncertainty associated with population age-composition, weight-at-age and maturity-at-age needs to be quantified and included in the calculation of CVs. A bootstrapping procedure could be used to quantify the uncertainty associated with population age-composition and projected weight-at-age. Uncertainty in weight-at-age could also be evaluated using a retrospective analysis in which the difference between observed and predicted weight-at-age for past years was calculated. Ultimately, improved estimates of weight-at-age and measures of precision of such estimates could be obtained by fitting a model to the empirical data on weight-at-age.
D. The methods for estimating 1 July age 1+ biomass based on the results of the ATM survey during the previous year currently use only the results of the summer survey. Improved precision is likely if the results from the spring and summer surveys were combined. This may become more important if the number of days for surveying is reduced in future. Consideration should be given to fish born after 1 July.
E. 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.
F. Modify Stock Synthesis so that the standard errors of the logarithms of age-1+ biomass can be reported. These biomasses are used when computing OFLs, ABCs and HGs, but the CV used when applying the ABC control rule is currently that associated with spawning biomass and not age-1+ biomass.
G. The approach of basing OFLs, ABCs and HGs for a year on the biomass estimate from the ATM survey for the previous year should be examined using MSE so the anticipated effects of larger CVs and a possible time-lag between when the survey was conducted and when catch limits are implemented on risk, catch and catch variation statistics can be quantified.
H. 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.
I. The assessment would benefit from the availability of estimates of 1+ biomass that include quantification of the biomass inshore of the survey area and in the upper water column.
J. It is unclear how the habitat model is applied to determine survey design. Is this an *ad hoc* decision or is there a formal procedure? The next Panel should be provided with comprehensive documentation on how the habitat model is applied.
K. 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.
L. Explore the potential of collaborative efforts to increase sample sizes and/or gather data relevant to quantifying effects of ship avoidance, problems sampling near-surface schools, and currently unsampled nearshore areas.
M. Reduce aging error and bias by coordinating and standardizing aging techniques and performing an aging exchange (double blind reading) to validate aging and estimate error. Standardization might include establishing a standard “birth month” and criteria for establishing the presence of an outer annuli. If this has already been established, identify labs, years, or sample lots where there is deviation from the criteria. The outcome of comparative studies should be provided with every assessment.

Medium priority

N. Continue to explore possible additional fishery-independent data sources such as the SWFSC juvenile rockfish survey and the CDFW/CWPA cooperative efforts (additional sampling and aerial surveys). 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.
O. 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; this should include an analysis of age-structure on the mean distribution of sardine in terms of inshore-offshore (especially if industry partner-derived data were available).
P. Consider a model that has separate fleets for Mexico, California, Oregon-Washington and Canada.
Q. Compare annual length-composition data for the Ensenada fishery that are included in the MexCal data sets for the northern sub-population 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.

Low priority

R. Consider a model that explicitly models the sex-structure of the population and the catch.
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.
T. 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.

Recommendations that should be addressed during the 2018 review of the ATM survey

A. In relation to the habitat model
    a. 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.
    b. Further validate the environmentally-based stock splitting method. The habitat model used to develop the survey plan and assign catches to subpopulation seems to adequately predict the spawning/egg distribution in the CalCOFI core DEPM region, but eggs were observed where they were not expected in northern California, Oregon and Washington during one of the two years when the survey extended north. 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.

    c. Consider including environmental covariates in model-based approaches that would account quantitatively for environmental effects on distribution and biomass. The expertise from a survey of fishermen could be extremely useful in identifying covariates that impact the distribution of clusters.

B. The SWFSC plans to examine ship avoidance using aerial drone sampling; there is an ongoing significant effort by Institute of Marine Research in Norway to understand the same issue using sonar, and the SWFSC acoustics team should communicate and coordinate with those researchers.

C. The effect of population size affecting the number and spacing of school clusters likely affects the probability of acoustic detection in a non-linear way; this could create a negatively biased estimate at low population levels and potentially a non-detection threshold below which the stock size cannot be reliably assessed. A simulation exercise should be conducted using the current, decreased and increased survey effort over a range of simulated population distribution scenarios to explore this.

D. The consequences of the time delay and difference in diurnal period of the acoustic surveys versus trawling need to be understood; validation or additional research is critical to ensure that the fish caught in the trawls from the night time scattering layer share the same species, age and size structure as the fish ensonified in the daytime clusters.

E. The ATM survey design and estimation methods need to be more precisely specified. A document must be provided to the ATM review (and future assessment STAR Panels) that:
    - delineates the survey area (sampling frame);
    - specifies the spatial stratification (if any) and transect spacing within strata planned in advance (true stratification);
    - specifies the rule for stopping a transect (offshore boundary);
    - specifies the rules for conducting trawls to determine species composition;
    - specifies the rule for adaptive sampling (including the stopping rule); and
    - specifies rules for post-stratification, and in particular how density observations are taken into account in post-stratification. Alternative post-stratification without taking into account density should be considered.

References

Venables,W.N. and D.B. Ripley, B.D., 2002. Modern Applied Statistics with S, 4th ed. Springer-Verlag, New York.

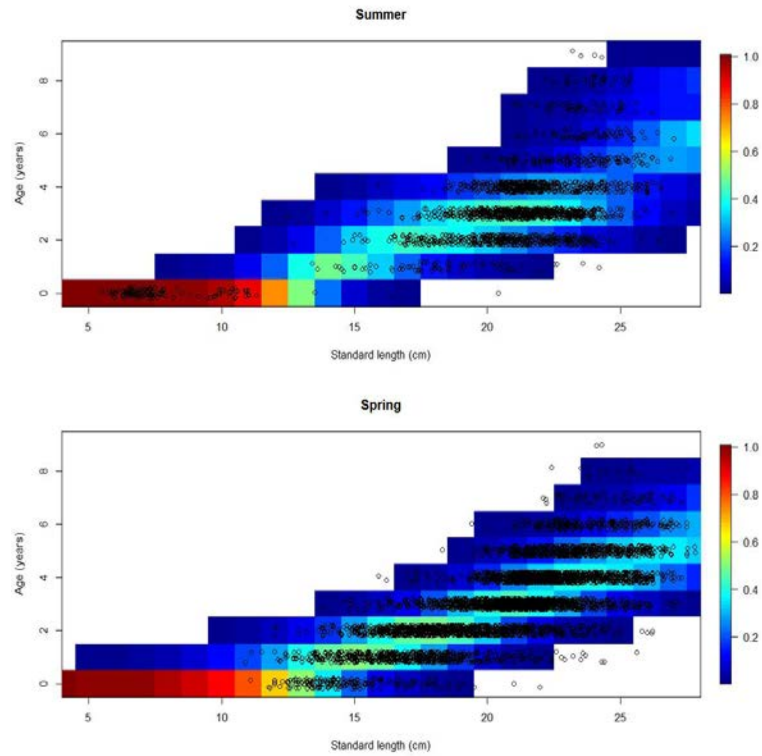

Fig. 1. Age-length key constructed using age and length information from sardine collected during Spring (upper panel) and Summer (lower panel) ATM surveys from 2004 to the present. The colored surface in the background is the multinomial surface $P(x = i|length)$ for $i \in \{0,1,\ldots,8,9+\}$ fit using the *multinom* function available in the *nnet* package for R (Venables and Ripley, 2002). The points in the foreground represent the pairs of data used to fit the model.

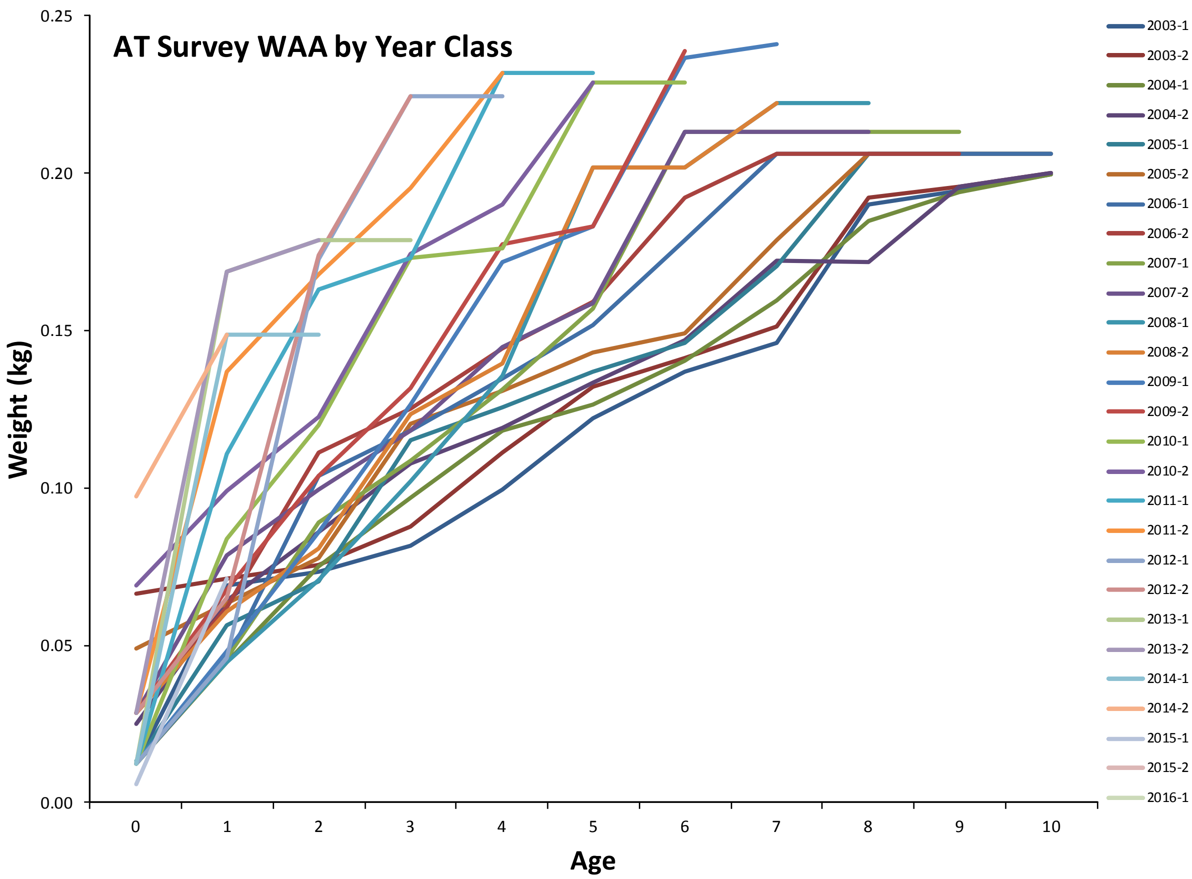

Fig. 2. Weight-at-age by cohort for the ATM survey.

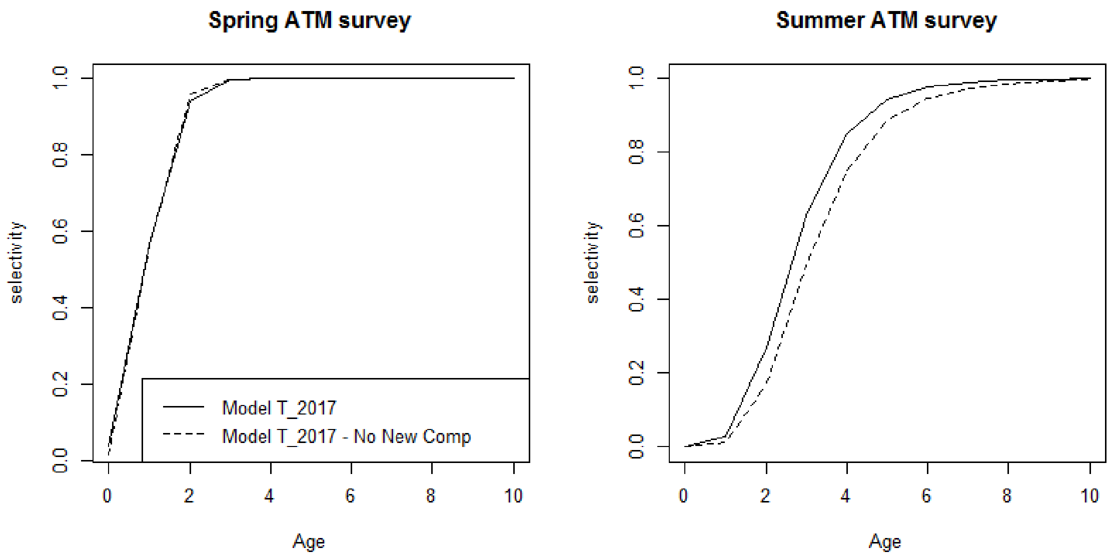

Fig. 3. ATM survey selectivity for the spring and summer surveys from Model T2017 and a variant of that model in which the last two ATM length-compostions are dropped from the model.

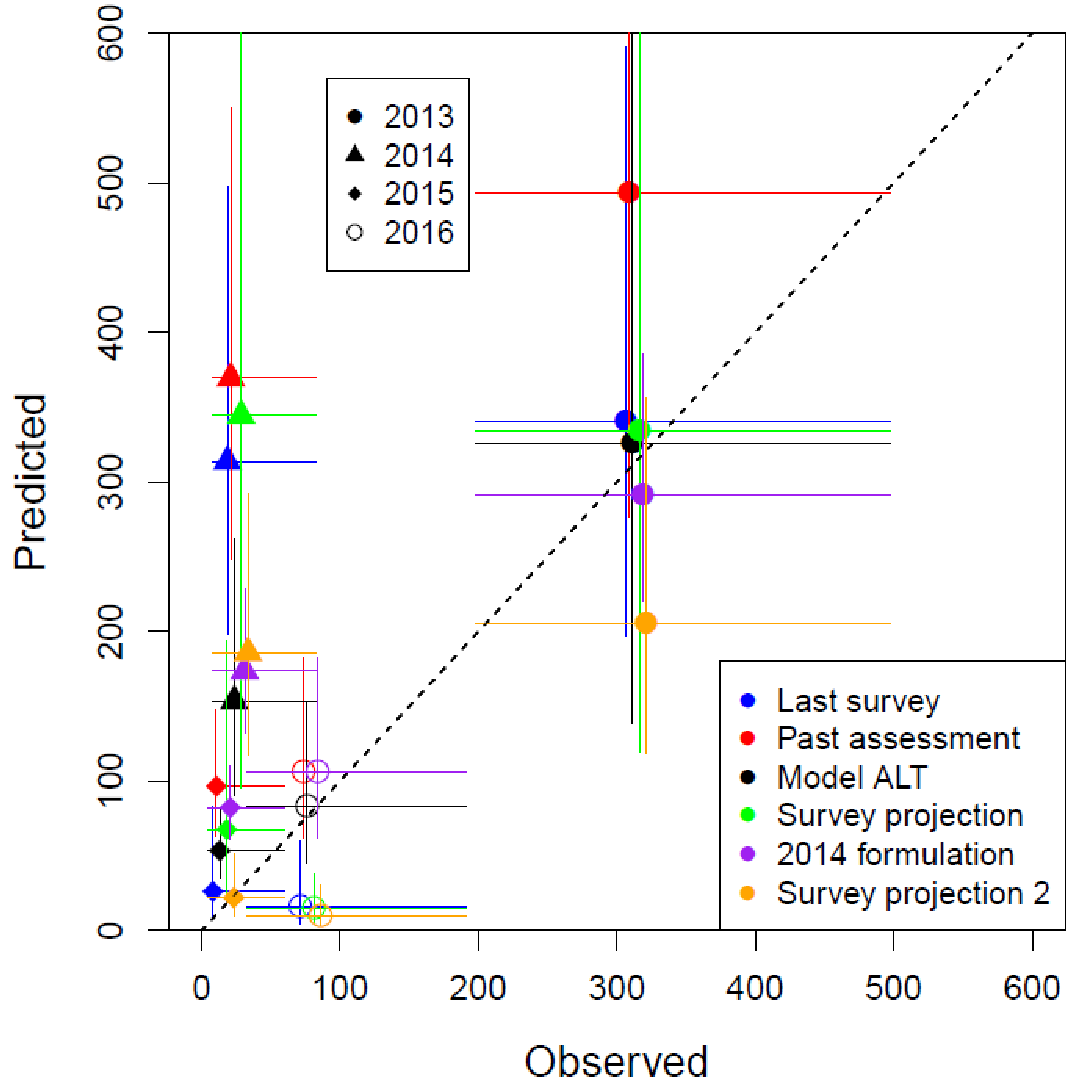

Fig. 4. Observed (x-axis values, ATM survey biomass estimates) and model-predicted (y-axis values) biomass on 1 July of each of 2013, 2014, 2015 and 2016. The observed values are the summer ATM survey estimates. The lines indicate 90% confidence intervals under the assumption of log-normal error. The x-axis values are jittered for ease of presentation.

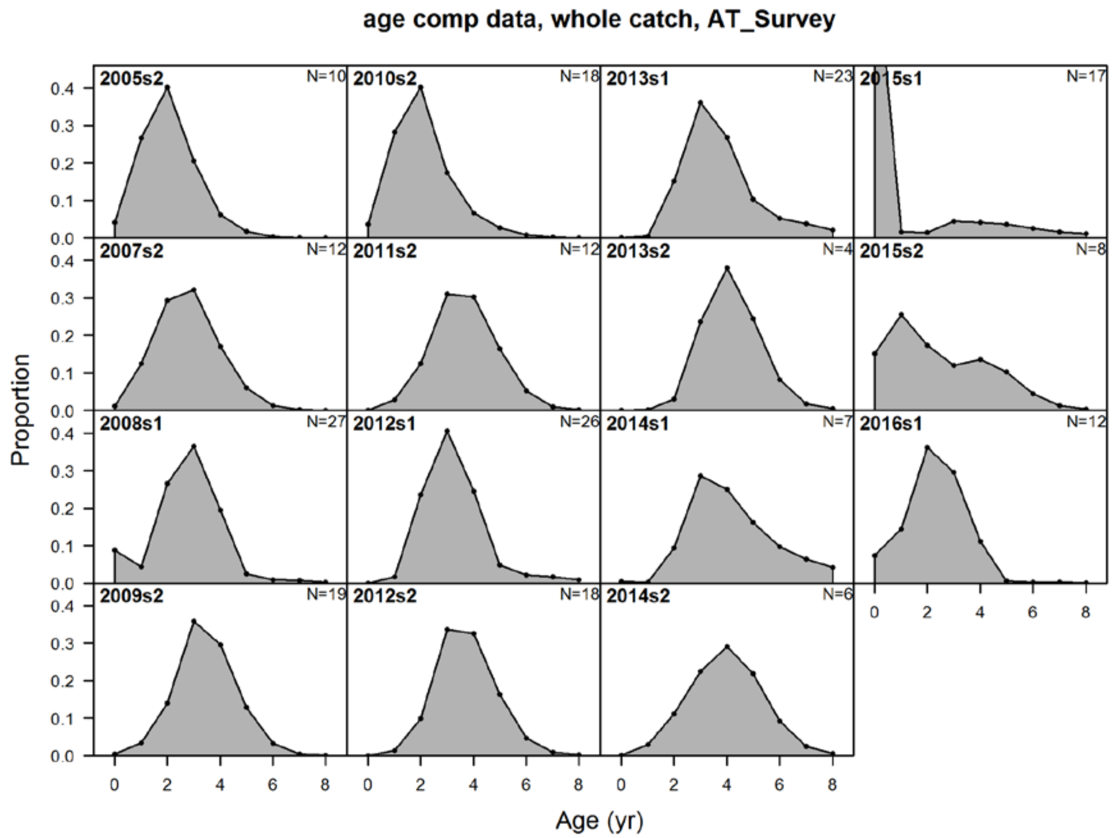

Fig. 5. The ATM survey age-compostion data.

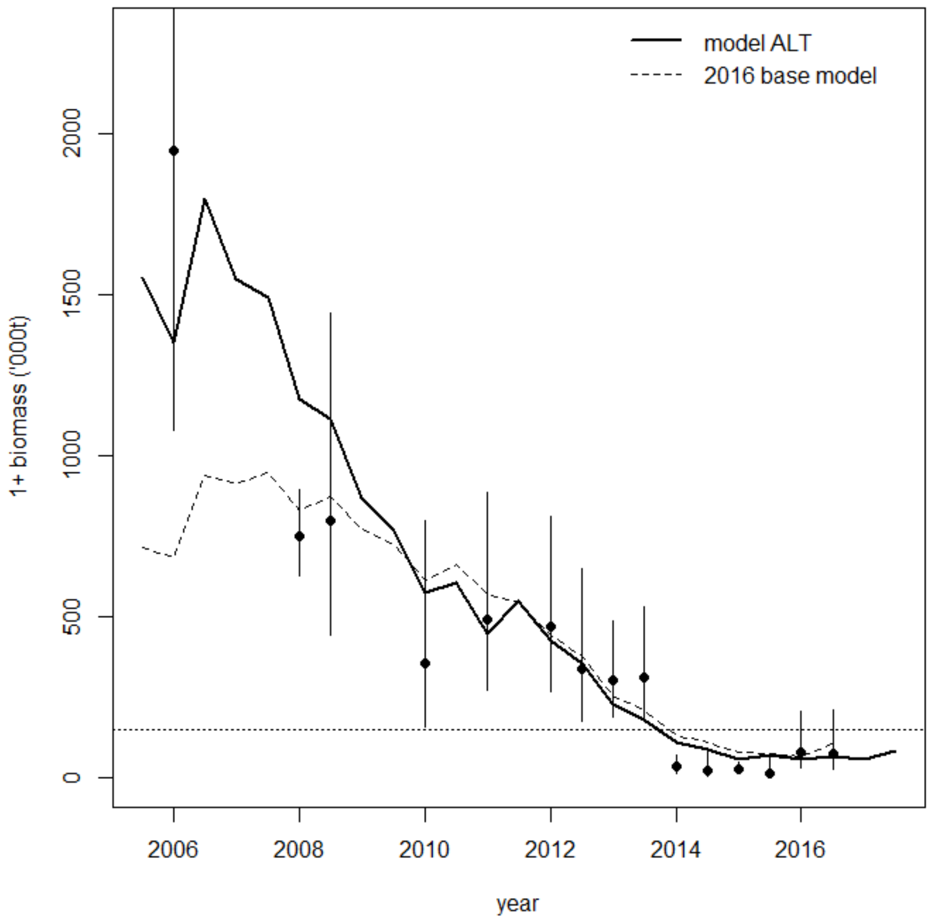

Fig. 6. Time-trajectories of 1+ biomass from model ALT and the 2016 base model. The ATM survey estimates of biomass and their 95% confidence intervals are indicates by the dots and the vertical bars, respectively.

Appendix 1
2017 Pacific Sardine STAR Panel Meeting Attendees

STAR Panel Members:
André Punt (Chair), Scientific and Statistical Committee (SSC), Univ. of Washington
Will Satterthwaite, SSC, Southwest Fisheries Science Center
Evelyn Brown, SSC, Lummi Natural Resources, LIBC
Jon Vølstad, Center for Independent Experts (CIE)
Gary Melvin, Center for Independent Experts (CIE)

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

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

Other Attendees
Dale Sweetnam, SWFSC
Alan Sarich, CPSMT/Quinault Indian Nation
Emmanis Dorval, SWFSC
Chelsea Protasio, CPSMT/CDFW
Kirk Lynn, CPSMT/CDFW
Ed Weber, SWFSC
Josh Lindsay, NMFS WCR
Erin Kincaid, Oceana
Al Carter, Ocean Gold
Jason Dunn, Everingham Bros Bait
Nick Jurlin, F/V Eileen
Neil Guglielmo, F/V Trionfo
Andrew Richards, Commercial
Hui-Hua Lee, SWFSC
Bev Macewicz, SWFSC
Chenying Gao, Student
Steven Teo, SWFSC
Kevin T.R. Piner, SWFSC
Andy Blair, Commercial
Jamie Ashley, F/V Provider
John Budrick, CDFW
Steve Crooke, CPSAS
Gilly Lyons, Pew Trusts

<u>Acronyms</u>
CDFW – California Department of Fish and Wildlife
CPSAS - Coastal Pelagic Species Advisory Subpanel
CIE – Council on Independent Experts
CPSMT - Coastal Pelagic Species Management Team
CWPA – California Wetfish Producers Association
SSC - Scientific and Statistical Committee
SWFSC - Southwest Fisheries Science Center (National Oceanic and Atmospheric Administration)
WCR – West Coast Region

Appendix 2
Projection of summer AT biomass 1 year into the future (Juan Zwolinski)

Given a vector of abundance-at-age from a summer survey during year *t* $\hat{\mathbf{a}}_{\mathrm{t}} = [\hat{a}_{0t}, a_{1t}, \ldots, a_{9+t}]$, with ages 0 through 9 and above, and where $\hat{a}_{0t}$ is the expected abundance of age-0 sardine estimated in one of the two possible ways described below, the abundance of sardine age 1 and older (zge-1+) at year *t+1* can be estimated by $\hat{\mathbf{a}}_{t+1} = \hat{\mathbf{a}}_{\mathrm{t}} \times e^{-(M+F)}$, where M and F are natural and fishing instantaneous mortality coefficients relative to one year, respectively. The corresponding biomass is obtained by the pointwise product $\hat{\mathbf{a}}_{\mathrm{t+1}} \times \mathbf{w}_{\mathrm{t}}$, where the empirical mean weight-at-age $\mathbf{w}_t = [w_{1t}, \ldots, w_{9+t}]$ is estimated from the survey during year *t*. If fishing mortality is expressed in catch, then $\hat{\boldsymbol{a}}_{t+1}$ can be approximated by $\hat{\boldsymbol{a}}_{t+1} = (\,\hat{\mathbf{a}}_t \times e^{-(M/2)} - \mathbf{c}_t) \times e^{-(M/2)}$, where $\hat{\boldsymbol{c}}_t = [c_{0t}, c_{1t}, \ldots, c_{9+t}]$ is the expected catch in numbers per age class.

Estimating a_{0t}

Summer AT surveys are not reliable estimators of the abundance of age-0 sardine at time *t* ($a_{0\mathrm{t}}$). Therefore, any projection of biomass from a survey at year *t* to year *t+1* requires a_{0t} to be estimated. Assuming that no fishing occurs for age-0 sardine, the expected age-0 abundance $\hat{a}_0$ can be estimated as the mean of the implied age-0 abundances calculated from *n* surveys such that:

$$E[a_0] = \hat{a}_0 = \frac{1}{\mathrm{n}} \sum_n a_1 \times e^M \,.$$

Alternatively, a_{0t} can be estimated using the stock-recruitment relationship from the most recent assessment. In order to do so, the abundance $\boldsymbol{a}_t = [a_{1t}, \ldots, a_{9+t}]$ from the summer survey has to be regressed 6 months and converted into spawning stock biomass (SSB) at *t-0.5*. Using empirical mean weight-at-age in winter $\mathbf{w}_{\mathrm{t-0.5}} = [w_{0t-0.5}, \ldots, w_{8+t}]$, and the vector of proportions of mature fish per age class $\mathbf{s}_{\mathrm{t-0.5}} = [s_{0t-0.5}, \ldots, s_{8+t}]$, $\mathrm{SSB}_{t-0.5}$ is obtained by the sum of the pointwise-product $\mathbf{a}_{\mathrm{t-0.5}} \times \mathbf{w}_{\mathrm{t-0.5}} \times \mathbf{s}_{\mathrm{t-0.5}}$, where $\mathbf{a}_{\mathrm{t-0.5}}$ can be calculated by $\hat{\mathbf{a}}_{\mathrm{t-0.5}} = \hat{\mathbf{a}}_{\mathrm{t}} \times e^{(M+F)/2}$ in case *F* is reasonably known. If fishing is expressed in catch, then $\hat{\mathbf{a}}_{\mathrm{t-0.5}} = (\,\hat{\mathbf{a}}_{\mathrm{t}} \times e^{(M/4)} + \mathbf{c}_{\mathbf{t-0.5}}) \times e^{(M/4)}$. There, $\mathbf{c}_{\mathbf{t-0.5}}$ is the vector of catch-at-age that occurred in the 6 months prior to the survey.