Pacific Sardine

STAR Panel Meeting Report

NOAA / Southwest Fisheries Science Center
La Jolla, California
October 4-7, 2011

STAR Panel Members:
André Punt (Chair), Scientific and Statistical Committee (SSC), Univ. of Washington
Ray Conser, SSC, Southwest Fisheries Science Center (SWFSC)
Larry Jacobson, External Reviewer, Northeast Fisheries Science Center
Chris Francis, Center for Independent Experts (CIE)

Pacific Fishery Management Council (Council) Representatives:
Lorna Wargo, Coastal Pelagic Species Management Team (CPSMT)
Mike Okoniewski, Coastal Pelagic Species Advisory Subpanel (CPSAS)
Kerry Griffin, Council Staff

Pacific Sardine Stock Assessment Team:
Kevin Hill, NOAA / SWFSC
Paul Crone, NOAA / SWFSC
Nancy Lo, NOAA / SWFSC
Beverly Macewicz, NOAA / SWFSC
Emmanis Dorval, NOAA / SWFSC
Jennifer McDaniel, NOAA / SWFSC
Yuhong Gu, NOAA / NWFSC

Acoustic- Trawl Survey Team
David Demer, NMFS, SWFSC
Juan Zwolinski, NMFS, SWFSC

Aerial Survey Team
Tom Jagielo, Tom Jagielo Consulting

1) Overview

The Pacific Sardine Stock Assessment and Review (STAR) Panel (Panel) met at the Southwest Fisheries Science Center, La Jolla, CA Laboratory from October 4-7, 2011 to review a draft assessment by the Stock Assessment Team (STAT) for Pacific Sardine. Introductions were made (see list of attendees, Appendix 1), the agenda was adopted, and Kerry Griffin reviewed the Terms of Reference (TOR) for CPS assessments with respect to how the Panel would be conducted. A draft assessment document and background materials were provided to the Panel in advance of the meeting on a SWFSC FTP site. The Chair, André Punt, noted that the assessment report included analyses related to estimating F_{MSY}, but that reviewing this analysis was beyond the scope of the TOR for the Panel.

Kevin Hill presented the assessment methodology and the results from a draft assessment utilizing the Stock Synthesis Assessment Tool, Version 3.21d (SS3) to the Panel. The model on which the draft assessment was based differed from that on which the 2009 assessment was based in several respects. The draft assessment included: (a) two rather than four fleets, (b) a later start-date for the assessment (1993 rather than 1981), (c) fewer time-blocks for selectivity, (d) no time-blocking for growth, (e) inclusion of the indices of abundance from the acoustic-trawl surveys, (f) revised age-reading error matrices, and (g) the aerial (and acoustic-trawl) surveys were assumed to be relative rather than absolute indices of abundance. The draft assessment benefited from a number of improvements to the abundance data and an improved understanding of the precision of the age data for sardine. The assessment was also based on other updated data streams, in particular additional age and length data for the Ensenada fishery.

David Demer, Nancy Lo, and Tom Jagielo respectively presented aspects of the methodology and results for the acoustic-trawl, Daily Egg Production Method (DEPM), and aerial surveys. The Panel agreed that the current approach of calculating spawning fraction for DEPM estimates should be continued and no futher work related to a Bayesian analysis of spawning fraction was required. 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 review and subsequent explorations of the assessment through sensitivity analyses were motivated primarily by the reasons for the changes from the last assessment, the poor residual patterns for some of the fits, understanding the best way to weight the various data sources, the considerable sensitivity of the estimate of current 1+ biomass to what would seem to be minor changes to the specifications of the assessment (see, for example request U below), and the assumptions related to catchability for the aerial and acoustic-trawl surveys. The Panel supported the effort by the STAT to simplify the assessment; with the aim of finding a more stable assessment (likelihood profiles presented to the Panel indicated that even though the assessment includes many data points, these are largely uninformative regarding current 1+ biomass).

The Panel noted that the approach to computing effective Ns in Appendix 2 differs from that used in most assessments of west coast coastal pelagic and groundfish species. This approach accounts for correlations among residuals within years, unlike the conventional method of McAllister & Ianelli (1997), which is used in SS3 to calculate 'output' effective sample sizes. These correlations are often substantial (those shown in Figure 2 of Appendix 2 are typical). The SSC should consider whether the approach of Appendix 2 should be used regularly when conducting stock assessments for Council-managed stocks.

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 exceptional support and provisioning during the STAR meeting.

2) Discussion and Requests Made to the STAT during the Meeting

Tuesday AM

A. Tabulate and plot the annual mean size-at-age in the catch by fishery (Mexico, California and Oregon/Washington) for semester 1; and superimpose the growth curve estimated in the model and, if possible, growth curves from the literature. ***Rationale***: To determine if there is evidence in the data for differences in growth by fishery and over time (mean size-at-age by fishery is not reported in the assessment document). These diagnostics may also provide some insight into possible model misspecification, and allow an evaluation of whether the estimated growth curve is biologically realistic. ***Response***: Mean size-at-age (averaged over years) was plotted for the various regions along the west coast. Mean size-at-age increased with latitude but decreased over time within region. The reduction in mean size-at-age over time was most apparent in the Pacific Northwest (PacNW) region, but most of the change occurred before 1991 (the assessment modeling begins in 1993).

B. Smooth the ageing error standard deviation (SD) relationship for California ages in 2007 (Figure 8 of the assessment report). ***Rationale***: Ageing error data are very noisy for fish older than 3.5 yr. The ageing error SD for age 4.5 is clearly an artifact. ***Response***: The spike in SD at age 4.5 was eliminated and linear extrapolation was used for all older ages. This change led to no changes in the 1+ biomass and became part of the base case for all subsequent model runs.

C. Conduct a run that does not use the ageing error matrix, or downweights the ageing error to near zero. ***Rationale***: To determine whether ageing error has an important effect on key assessment results. ***Response***: This change smoothed the recruitment estimates, but did not cause an appreciable change in the time-series of 1+ biomass.

D. Add the recommendations from the September 2010 SSC CPS Subcommittee review and the November 2010 SSC report to the recommendation list from the 2009 STAR Panel (see 2010 assessment document, p 135+). ***Rationale***: This will complete the assessment review history of requests and actions taken. ***Response***: This request could not be completed before the end of the Panel meeting and was added to the list of changes that need to be made to the final document.

Tuesday PM

E. Progressively estimate fewer recruitment deviations (2007-11) at the end of the time series. Carry out retrospective analyses (2007-11) to ascertain if estimating fewer recruitment deviations improves the retrospective pattern. Determine the appropriate number of recruitment deviations to estimate using this analysis. Keep the number of recruitment deviations not estimated constant. ***Rationale***: There are few data near the end of the time series to inform estimation of annual recruitment. ***Response***: Changing the number of year classes forced to fall on the S/R curve near the end of the time series led to fairly large changes in 1+ biomass, especially near the end of the time series. The retrospective pattern seen in the base case generally persisted.

F. Check the estimate of biomass from the acoustic-trawl survey for summer 2008 and the CVs of these biomass estimates for all years. ***Rationale***: Values in Table 5 of the assessment document appear to differ from those shown in the acoustic-trawl survey presentation. ***Response***: The values were corrected. This change led to no difference in the estimates of 1+ biomass and the revised estimate of abundance became part of the base case for all subsequent runs.

G. Conduct a sensitivity run which replaces the CV for the spring 2008 acoustic-trawl survey with the average CV from the other acoustic-trawl surveys. ***Rationale***: The CV for the spring 2008 acoustic-trawl survey (9.2%) appears to be too small given the CVs for the other acoustic-trawl surveys and the sampling issues experienced during the 2008 survey. ***Response***: The CV was changed to the average value (CV=33%). This change led to no appreciable difference to the 1+ biomass.

H. Examine the effect on the biomass estimates from the aerial survey of using complete point sets observed from altitudes less than 4000 feet when fitting the density vs. school area relationship. ***Rationale***: A considerable amount of potentially useful data are currently not being used in biomass estimation because of the operating constraint that requires the 4000 foot altitude. ***Response***: The biomass estimate increased less than 10% and the CV decreased slightly. There was no appreciable change to the fitted curve to the density vs school size data.

I. Modify Table 7 (p.43) of the aerial survey report to include the sum of the biomass for each column, and do a paired t-test on the effect of different readers. ***Rationale***: The Panel wanted to get a better understanding of the possible effects from the two independent readers. ***Response***: While the paired t-test showed a difference at the $\alpha=0.05$ level of significance, the biomass estimates from the two readers were quite similar. There appears to be no practical difference between the two readers.

J. Compute the autocorrelation function among positive transects from the 2011 aerial survey. ***Rationale***: Strong autocorrelation will violate the assumption of independence among transects on which method used to calculate the CV for the 2011 aerial survey is based. ***Response***: The correlation was 0.25 at lag 1; similar or smaller correlations were found for lags greater than 1. The transects appear to be sufficiently independent for application of the chosen method of variance estimation.

K. Compute the mean length of fish in each school from the point sets from the 2009, 2010 and 2011 aerial surveys, and plot by latitude. ***Rationale***: To examine whether the size data from the point sets are representative of the sardine population in the Pacific Northwest; in particular, to determine whether the shift (to the right) in length compositions over 2009-11 (Figure 11 of the aerial survey assessment report) are an artifact of the latitude at which the point sets were made. ***Response***: There are clearly year effects in mean length-at-age from the point sets, and some trend with latitude, but not enough to explain the misfitting of the length compositions in the assessment.

L. Plot catch weight vs. school area for the 2011 point sets and add a fitted line. ***Rationale***: This relationship may be an alternative to the density vs. school area relationship. ***Response***: The plot of catch weight vs. school area showed large variance and confirmed that density vs. school area is more likely to produce a useful predictive relationship.

M. Create a likelihood profile for q for the acoustic-trawl survey (q = 0.25 - 1.75). Tabulate the likelihood components for each discrete value of q used in the profile. ***Rationale***: To determine the key likelihood components over a range of biomass scalings. ***Response***: The total likelihood was flat across all values of acoustic-trawl q (less than 2 units difference over the entire range). The likelihood components for the indices of abundance and the age compositions favoured q at the high end of the range profiled (other than the PacNW age-at-length data), but the length compositions favored q at the low end of the range. However, the overall difference in likelihood units was small (~ 5 units) for all individual components over the full range of q ($0.25 \rightarrow 1.75$).

N. Conduct a run with initial F set to zero and continue to estimate the recruitment deviations starting in 1987. ***Rationale***: The initial F estimate in the base case model is not credible (F=4 yr^{-1}), and the estimated recruitment deviations are not significantly different from zero. Setting F=0 may result in better recruitment deviation estimates as a means of initializing the model, i.e. creating numbers-at-age at the start of 1993. ***Response***: This run led to a trend in 1+ biomass that was nearly identical to that for the base case, but overall 1+ biomass was approximately 50% greater than for the base case. The recruitment trend was also similar, but recruitment was ~30% larger than for the base case. Some of the later early deviations became significantly different from zero and R_0 increased approximately 35% compared to the base case. Early recruitment deviations were negative rather the zero as for the base case, indicating lower than average recruitment during late 1980s. The q estimates were more reasonable (all less than 1.0). The Panel and STAT agreed that this run (which also reflects the modifications from Requests B and F, above) was more plausible than the base case in the assessment document, and should serve as the base case for all subsequent runs.

O. Conduct a run with one vector of recruitment deviations, i.e. do not model early and main recruitment deviations separately. ***Rationale***: It was not clear to the Panel why the early and main recruitment deviations need to be modeled separately. ***Response***: This run was not carried out due to lack of time and the low priority given to it by the Panel.

P. Plot the sex ratio by length for each fishery. ***Rationale***: The model is not sex-specific. This plot will help to assess whether the data support a single-sex model. ***Response***: The sex ratios were plotted by length bin and region. The proportion of males decreases appreciably above the 21 cm size bin in all regions. It was also noted that the sex ratio data by weight from the DEPM surveys also showed that the percentage of females in the spawning population is consistently greater than 50%. Future modeling may wish to consider sex explicitly (see research recommendations, below).

Q. Do a profile over S/R variability (σ_R) using the base case in the assessment document. Show the 1+ biomass trend for each σ_R. ***Rationale***: σ_R from the base case (σ_R=0.622) may be smaller than is typical for a small pelagics. ***Response***: As σ_R increases from σ_R=0.622, the 2011 1+ biomass increases considerably through σ_R=1.0, but 1+ biomass decreases markedly when σ_R>1.

R. Do a sensitivity run dropping the TEP index. ***Rationale***: The DEPM time series is now much longer that when the TEP index was first introduced. It may not be necessary to continue to use the TEP index which ignores variation among years in biological parameters. ***Response***: Removing the TEP index had little effect on the time series of 1+ biomasses.

Based on the requests, above, the Panel and STAT considered the run from Request N to be the candidate base case subject to the additional requests, below.

Wednesday

S. Create a separate Canadian fishery with selectivity mirrored to the USA portion of the PacNW fishery. Present length and conditional age-at-length residuals by fishery. If possible, keep the annual effective sample sizes the same as in the base case model. ***Rationale***: While this change should not affect model fitting and results greatly, it will provide additional diagnostics for understanding the poor fits to the length compositions from the PacNW fishery and to assess whether it is justified to pool data for Oregon, Washington and Canada. ***Response***: The residual pattern for the Canadian fishery is quite different than that for the USA PacNW fishery (the former has many more positive residuals at the larger sizes). The next stock assessment should consider establishing a separate Canadian fishery.

T. Create a separate Mexican fishery with selectivity mirrored to the USA portion of the MexCal fishery. Present length and conditional age-at-length residuals by fishery. If possible, keep the annual effective sample sizes the same as in the base case model. ***Rationale***: While this change should not affect model fitting and results greatly, it will assist the Panel examine whether it is justified to pool data across Mexico and California. ***Response***: The residual pattern for the Mexican fishery is somewhat different than that for the USA portion of the MexCal fishery (the former has more positive residuals at the larger sizes, particularly during semester 2). The next stock assessment should consider re-establishing a separate Mexican fishery.

U. Drop the 2008-10 conditional age-at-length data for the PacNW fishery. ***Rationale***: The age readings from these years appear to be quite different from all other years. ***Response***: The trend in 1+ biomass is similar to the base case (run N), but the average biomass is much reduced - current 1+ biomass is ~20% less that for run N.

V. Reduce the multipliers for the effective sample sizes for the length composition data using the Francis vector (Appendix 2 of this report) and reduce the multipliers for the effective sample sizes for the conditional age-length data by 90%. ***Rationale***: Considerable among-length / -age correlation is evident in both the length composition and conditional age-at-length residuals, but the method used to infer effective sample sizes in SS3 assumes independence among residuals. Hence, the presence of strong correlation, combined with the method used in SS3 to compute downweighting factors, effectively over-weights the age and length data. ***Response***: The trend in 1+ biomass differed from that for the base case (run N) and all other runs examined to date. The average 1+ biomass was lower than for run N, but closer to that run than to the average biomass from run U. The fit to the indices were similar to those seen in all earlier runs.

W. Apply a model that fits predominately to age-based data. Use the age composition data rather than the combination of length and conditional age-at-length data, whenever available; do not use length data whenever acceptable age data are available; fix growth using the base case (run N) parameter estimates; continue using length-based selectivity for the fisheries (as in the base case); and use the effective sample sizes and lambda multipliers for the length data from the base case for the age data. ***Rationale***: The sardine assessment is unusual in that a large proportion of the sampled fish are aged. The additional information from length compositions may be marginal, and the model has difficulty fitting the length compositions. This should be considered an exploratory model, i.e. not one that is likely to be used as a base case for this year's assessment. ***Response***: Selectivity at length did not differ greatly from for the base case run (some selectivity curves were steeper at small sizes, but had similar points of inflection). The recruitment deviations for recent years differed markedly from those for run N (all were highly positive). Fits to indices of abundance were generally similar; as were fits to the age compositions. The trend in 1+ biomass differed from that for run N (two roughly equally high peaks) and the average 1+ biomass was slightly lower than for run N. The next stock assessment should consider an approach similar to the one explored here.

Thursday

X. Conduct six additional model runs based on the current base-case model (run N):
   1. fix DEPM survey q=0.5 and retain length and conditional age-at-length composition weighting as in run N;
   2. fix DEPM survey q=0.5 and weight the length and conditional age-at-length composition data as in run V;
   3. fix aerial survey q=1 and retain length and conditional age-at-length composition weighting as in run N;
   4. fix aerial survey q=1 and weight the length and conditional age-at-length composition data as in run V;
   5. fix acoustic-trawl survey q=1 and retain length and conditional age-at-length composition weighting as in run N;
   6. fix acoustic-trawl survey q=1 and weight the length and conditional age-at-length composition data as in run V.

Rationale: The results of these runs are needed to address two issues: (i) the scale of biomass in the assessment is not well determined; fixing *q*=1, one survey at a time, should better inform the scale issue; and (ii) the length and conditional age-at-length data appear to be over-weighted relative to the indices of abundance (see Request V, above), but the full impact of alternative weighting needs to be more fully examined. ***Response***: The estimate of 2011 1+ biomass (used in the PFMC control rule) was greater in run N than in any of runs X.1 through X.6. The trend in 1+ biomass was similar in runs X.1, X.3 and X.5 to that for run N, but those for runs X.2, X.4, X.6 (when the age and length data were further down-weighted relative to the indices) differed from that for run N. The fits to the indices of abundance were similar across all runs. Biomass scaling differed most from run N for runs X.1, X.2, and X.6. The realized S/R variability was noticeable smaller for run X.6 (σ_R=0.39). The estimated *q*'s for the aerial and acoustic-trawl surveys were most plausible for runs X.3 through X.6 (i.e., except when the DEPM indices were assumed to be absolute).

Y. Use run X.5 (above) as the reference run (i.e. a candidate for a new base case) and conduct six additional runs:
   1. drop the conditional age-at-length data from the PacNW fishery for 2008-10 (analogous to run V);
   2. constrain only the last recruitment such that it falls on the S/R curve;
   3. constrain the last three recruitments such that they fall on the S/R curve;
   4. fix σ_R = 0.4;
   5. fix σ_R = 0.8; and
   6. fix σ_R = 1.0.

   Rationale: Run N has been the candidate base case, but it exhibited some instabilities – particularly in biomass scale (see Requests E, Q, and U, above). The *q* for the acoustic-trawl survey was fixed (*q*=1) in run X.5 in an effort to provide more stability. This set of runs was designed to examine the stability of run X.5 relative to the stability of run N. ***Response***: Run Y.1 showed the largest effect on biomass scaling (relative to run X.5), but the amount of change in biomass scaling was much less than was seen for the comparable sensitivity run based on run N (*cf.* Request U). The biomass scaling effect was not greatly different for Run Y.2 than that for the comparable runs based on the base case in the assessment document (*cf.* Request E). However, runs Y.5 and Y.6 did show improved stability in biomass scale relative to the comparable sensitivity runs based on run N (*cf.* Request Q). The biomass series for runs Y.3 and Y.4 differed from that for run X.5, but SS3 failed to converge for these runs so the Panel could not draw conclusions regarding stability.

Z. Consider run X.5 to be the new base case and make a final set of sensitivity runs:
   1. jitter to the 10% level; for each jitter, present total likelihood, *q* for all surveys, terminal year 1+ biomass and exploitation rate;
   2. create a likelihood profile on *M* [0.25-0.75yr^{-1}; step size 0.125yr^{-1}]; for each *M*, present total likelihood, *q* for all surveys, terminal year 1+ biomass and exploitation rate;

3. create a likelihood profile on the *q* for the acoustic-trawl survey [0.25-2.00; step size 0.25]; for each *q*, present total likelihood, *q* for all surveys, terminal year 1+ biomass and exploitation rate;
4. conduct a retrospective analysis over the last 5 years (2007-11); for each terminal year, present time-series of 1+ biomass and recruitment;
5. conduct a prospective analysis over the first 5 years (1993-97); for each initial year, present time series of 1+ biomass and recruitment.

Rationale: Additional runs are needed for the candidate base case (run X.5) to check for local minima; to identify the major axis of uncertainty and to quantify same; and to check for retrospective and prospective patterns. ***Response***:

1. **Run Z.1** (test for local minima). The full jitter was not completed, but will be included in the final assessment document. A few runs with R_0 changed converged to the same minimum as run X.5.
2. **Run Z.2** (*M* profile) showed that the total likelihood and the conditional age-at-length likelihood tend to strongly favor higher natural mortality rates than assumed in the base case; the length compositions favored a somewhat higher *M*. Increasing *M* reduces 2011 1+ biomass and increases the exploitation rate. The *M* profile is quite similar to the corresponding profile from the 2010 assessment.
3. **Run Z.3** (*q* profile) indicated that the length compositions do not inform the choice of acoustic-trawl *q*, but the conditional age-at-length data do have some influence. Overall, however, the likelihood surface is quite flat (even after fixing the acoustic-trawl *q*) – the profile showed a difference of only 2 units over the entire range of *q* (0.25 - 1.75). As expected, terminal year biomass and *F* were greatly affected by *q*.
4. **Run X.4** (retrospective analysis) showed an appreciable retrospective variability (up to 400,000 t changes among years in terminal biomass), but no systematic effect (i.e. the pattern is mixed - some high some low).
5. **Run X.5** (prospective analysis) showed modest changes in early year biomass estimates (and no systematic pattern), but virtually no change in 2011 biomass.

3) Technical Merits and/or Deficiencies of the Assessment

During its deliberations (see Section 2 of this report) the Panel identified a number of issues which should be explored for the assessment of Pacific sardine (see Section 6) including (a) further downweighting of the age and length data; (b) use of age-compositions rather than the combination of length-compositions and conditional age-at-length data, given within-year growth and among-region variation in growth; (c) additional fleets; and (d) inclusion of spatial- and sex-structure. Several analyses were conducted by the STAT to examine whether such changes warrant consideration in future. However, the STAT stated that major changes to the structure of the assessment should not be made without full and careful analyses of model structure and weights. The Panel agreed with the STAT that making these types of changes was not feasible in the time available and therefore focused on model configurations with two fleets and no spatial- or sex-structure. Some of these suggested changes may lead to more complicated models that cannot be supported by available, largely uninformative, data, and which may exhibit the types of undesirable behaviours seen in previous assessments. These changes should therefore only be implemented if there are clear benefits to the assessment and management of the stock.

Although trends in 1+ biomass do not change much given changes to the specifications to the assessment (although not necessarily to marked changes in data weighting), absolute biomass is poorly determined. The STAT and Panel therefore agreed that an appropriate way to increase stability in the assessment was to fix the *q* for one of the surveys. This is not an ideal approach, and the Panel recommends that the next full assessment include the development of informative priors for the *q* parameters for the DEPM, aerial and acoustic-trawl surveys. Development of informative priors is a non-trivial task and should involve people in addition to the STAT, in particular the surveys teams; therefore this task should start before the analytical work on the assessment itself, perhaps in the form of a workshop. The STAT and Panel agreed to impose the assumption q=1 for the acoustic-trawl survey because (a) there are more estimates of abundance for this series than for the aerial survey, (b) the acoustic-trawl survey is more synoptic 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 acoustic-trawl surveys (>1.8). While the SSC recommended that strong evidence is needed to assume q=1 for any survey, the STAT and Panel agree that in this instance it is best available science to make this assumption. The use of q=1 for this assessment is, however, *not* an endorsement of this assumption for future assessments. Rather it is preference of the STAT and Panel to use informative *q* priors in future. However, this is not feasible at present.

The STAT and Panel strongly agreed that it would be better in principle to downweight the age and length data using an approach such as that of Appendix 2 of this report. However, runs with the downweighted data led to lower than expected values for the root mean square error of the recruitment deviations (0.391 for the acoustic-trawl q=1 run), and to a growth curve which did not match the size-at-age data well. Further work on models with downweighted age and length data should form part of the next full assessment, but there was insufficient time during the Panel to find a model configuration which downweighted the data and did not exhibit poor behaviour in other respects.

The final base model incorporates the following specifications:

- two seasons (Jul-Dec and Jan-Jun) (assessment years 1993 to 2011);
- sex is ignored;
- two fleets (MexCal, PacNW), with an annual selectivity pattern for the PacNW fleet, and seasonal selectivity patterns for the MexCal fleet;
- length-based, double-normal selectivity with time-blocking (1993-1998, 1999-2011) for the MexCal fleet; asymptotic length-selectivity for the PacNW fleet;
- Ricker stock-recruitment relationship with estimated "steepness";
- $M = 0.4$ yr^{-1}; $\sigma_R = 0.622$ (tuned value);
- initial recruitment estimated; recruitment residuals estimated for 1987-2009;
- length-frequency and conditional age-at-length data for all fisheries;
- virgin (R_0) and initial recruitment offset (R_1) were estimated;
- initial *F*s set to 0 for all fleets;

- DEPM and TEP measures of spawning biomass; *q* estimated;
- aerial survey biomass, 2009-2011, *q* estimated, domed selectivity; and
- acoustic-trawl survey biomass, 2006-2011, *q*=1, asymptotic selectivity.

The Panel agrees that the final base model represents the best available science regarding the status of the northern subpopulation of Pacific sardine.

It is difficult to fully characterize uncertainty in the assessment. However, estimates of 1+ biomass from sensitivity analyses about run N, including runs with *q*=1 for each survey (Figure 1 of this report), are a crude depiction of the underlying uncertainties.

An important uncertainty not addressed elsewhere stems from the differences in biomass scale and trend indicated by the acoustic, DEPM and aerial surveys (see Figure 15 in the assessment report). In trying to fit all of the surveys, the final base case model estimates an average trend that does not match the trends in any of the individual surveys. In particular, the final model does not match or explain the relatively substantial and consistent decline in the acoustic-trawl survey during 2007-2011. In future assessments, it would be advisable to examine models that may better fit the trend in each of the individual surveys.

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.
2. 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 2 of this report) is not unexpected, and changes in terminal 1+ biomass estimates of this extent may occur when the 2012 assessment update occur.
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 set or assumptions can lead to marked changes in current abundance. The current assessment has somewhat reduced the influence of this lack of information by fixing survey catchability. Ultimately, it is only through further data collection (or the development of informative priors for survey catchability) that these uncertainties may be overcome.
5. The STAT evaluated a large number of model configurations to identify a more stable model that fits the data better. However, the residual patterns for the composition data and indices remain unsatisfactory. Furthermore, attempts to split the data by fleet to reduce some of these patterns led to unrealistic results (e.g. Fs > $2yr^{-1}$ in recent years for the MexCal fishery). The Panel identified the need to consider models with sex-

and spatial-structure, but there was insufficient time to develop, test, and evaluate such models during the Panel meeting.

6. Further downweighting the age and length data is warranted given the analyses in Appendix 2 of this report. However, time is needed to find a model configuration that does not lead to undesirable diagnostics (such as a low value for the root mean square error for the recruitment deviations, or a poor fit to the size-at-age data, as found in initial models examined during the meeting).
7. The period covered by the current assessment starts in 1993 (rather than in 1981 as in past assessments). This change was necessary because of a variety of factors, including lack of precise abundance estimates for the years 1981-92, lack of age and length data for the Ensenada fishery (only three years of data), and the fact that the age and length data for southern California were collected from an incidental fishery for sardine for much of this period. In addition, the growth data for these years is inconsistent with the later growth data and was one reason for the previous assessment invoking the assumption of time-varying growth. While the Panel supports the change in start year, dropping the early data means that it is no longer possible to assess the state of the stock prior to 1993, which adds to uncertainty about the dynamics of this population and current biomass levels.
8. The scarcity of old and large sardines in the data relative to model estimates is a fundamental tension in the assessment that may be due to assumptions about, for example, growth, selectivity, natural mortality, and data weighting.

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

a) CPSMT issues

The CPSMT representative commends the Panel and STAT for the significant amount of work accomplished prior to and during the meeting, and for a conducting a well-run review. The CPSMT representative notes that poor fitting of age data from fisheries in the Pacific Northwest by the model was identified as potentially an age reading issue and encourages efforts to evaluate whether or not this is the case, or if there is another reason. The upcoming ageing workshop in December 2011 offers an excellent opportunity to pursue future exchanges of otoliths for comparison among readers in the various laboratories. Previous recommendations have called for new indices to be incorporated into the sardine stock assessment. The CPSMT representative is encouraged to see the acoustic-trawl survey and aerial survey as recent additions, and notes that another survey (Canadian trawl survey) may be under consideration as well. The CPSMT representative suggests that in addition to considering new surveys in the next assessment, that a comparable effort to further refine and improve all data sources should be made to ensure these data are as informative as possible.

The Panel's consensus is that the model is very sensitive to relatively minor changes in parameters and data, and thus the biomass estimate is subject to significant variations of several hundred thousand metric tons. Given this uncertainty inherit in the model, the CPSMT representative suggests careful consideration of this fact when establishing sardine harvest management measures.

b) CPSAS issues

The CPSAS representative commends the Panel and STAT for integrating a new acoustic-trawl survey into the SS3 model. Previous Panels, the CPS Advisory Bodies, and the SSC have remarked that additional work was needed in the areas of surveys to enrich the data sources that are use when fitting the model.

Industry wants to see a sustainable resource that is not in danger of being overfished. Overfishing makes a poor platform for economic investment. That said, the CPSAS representative does not believe there is any immediate danger that overfishing is taking place at present. Anecdotal reports from Ensenada to the Queen Charlottes suggest that the sardine biomass is larger at this point in the expansion cycle than at any time since the last expansion. Boats in Westport Washington and Monterey California were often able to do "daily doubles" when there was sufficient processing capacity during the brief fishing periods this summer. Canadian vessels now report a "solid wall" of fish in October the entire length of West Vancouver Island.

The CPSAS representative does not have concerns about the model work, but it is very complex. The model demands data to function rationally. Slight tweaks to data and assumptions can lead to huge swings in outputs, particularly for the original base model. The model cannot operate effectively without robust data. The acoustic-trawl survey is a welcome tool, but when strictly coupled with the habitat model, migration theory, and certain assumptions on vessel avoidance we believe that this survey capacity is not fully utilized. The 2011 Sardine Workshop recommended utilization of the acoustic-trawl survey with application of a powerful sonar during the height of the summer feeding season, when the sardines are in peak abundance simultaneously in the Northwest and Canada. These stocks should be surveyed in Canada to the northern end of their range.

It is now known that the Canadian swept-trawl survey CV reported previously was an over-estimate. A recommendation of the 2009 STAR Panel was to consider possible use of the Canadian data in the stock assessment. One reason for not doing so in the current assessment was the high CV. The CPSAS representative recommends that this important data source be utilized as soon as feasible, and believes that there well may be, an older, and as large a biomass in Canada at peak season as inhabits the Northwest at the same time. None of this information is presently available for the modeling platform. To advance use of the Canadian survey data will require a methodology review for the swept trawl survey. This should be undertaken in 2012.

The CPSAS representative would like to thank the STAT, the SWFSC, the survey teams, and the Panel, along with the public for their hard work, dedication, and time.

7) Research Recommendations (not in priority order)

A. Continue to explore possible additional fishery-independent data sources. As noted by previous Panels, there would be value in attempting to include the data from the mid-water trawl surveys off the west coast of Vancouver Island (see Appendix 3 of this report for an overview) in the assessment. However, inclusion of a substantial new data source would likely require review which would not be easily accomplished during a standard STAR Panel meeting so would likely need to be reviewed during a

Council-sponsored Methodology Panel. Similarly, the information provided on presence of sardine in the SWFSC juvenile rockfish survey should be explored further for possible inclusion in the future assessment.

B. 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 acoustic-trawl surveys). It also encourages sampling in Mexican and Canadian waters (aerial and acoustic-trawl surveys).

C. Temperature at catch could provide insight into stock structure and the appropriate catch stream to use for assessments, because the southern subpopulation is thought to prefer warmer water. Conduct sensitivity tests to alternative assumptions regarding the fraction of the MexCal catch that comes from the northern subpopulation

D. The assessment would benefit not only from data from Mexico and Canada, but also from joint assessment, which includes assessment team members from these countries.

E. Conduct additional studies on stock structure - otolith and microchemistry studies are useful tools for this purpose.

F. The relationship between environmental correlates and abundance should be examined. In particular, the relationship between environmental covariates and overall recruitment levels as well as recruitment deviations should be explored further.

G. Consider spatial models for Pacific sardine, which 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.

H. Explore models which 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 and to provide a broader context for evaluating changes in productivity.

I. Modify Stock Synthesis so that the standard errors of the logarithms of 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 1+ biomass.

J. In relation to the aerial survey: (a) provide the otoliths collected from the point sets to the SWFSC for possible ageing, (b) explore different functional forms for the mean relationship between school density and area (e.g. splines) as well as the variation about the mean curve (e.g. gamma), and (c) consider possible covariates (e.g. average fish size) in the relationship between catch weight and area.

K. Modify the r4ss package to include a plot of correlations among the residuals for the length and data data, as well as the fit of the model to the mean length or age in each composition (see Appendix 2 of this report).

L. Consider a model which explicitly models the sex-structure of the population and the catch.

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

N. Develop a relationship between egg production and age which accounts for the duration of spawning, batch fecundity, etc. by age.
O. 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.
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. The reasons for the discrepancy between the observed and expected proportions of old animals 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 modeled.
R. Any future management strategy evaluation work to compare control rules should focus on alternatives which are as robust as possible to uncertainty regarding absolute abundance.
S. Profiles on key parameters should be included in future draft assessment to facilitate initial review.

Suggestions for modifications to the assessment report

A. Add a section on 'data sources considered but not used.'
B. Add a description of the derivation of the acoustic-trawl estimates in an appendix to the assessment report.
C. Add text to the report to explain why selectivity blocking was changed. Discuss whether the resulting selectivity patterns are consistent with auxiliary information on the behaviour of sardine and the fishery.
D. Add an update to Table 5a from the previous aerial survey report to the current report, and add the intended and achieved distribution of point sets by weight.
E. Document how the reweighting of the model was done (including changes in effective Ns for the age and length data and extra CVs for the abundance indices)
F. Add the recommendations from the September 2010 SSC CPS Subcommittee review and the November 2010 SSC review to the recommendation list from the 2009 STAR Panel (see 2010 assessment document, p 135+).
G. Include profiles and prospective and retrospective analyses for the final base model and the full range of sensitivity tests, including those in which the age and length data are downweighted, and each survey is assumed to be an absolute index of abundance, in the final report.

Reference

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 54: 284–300.

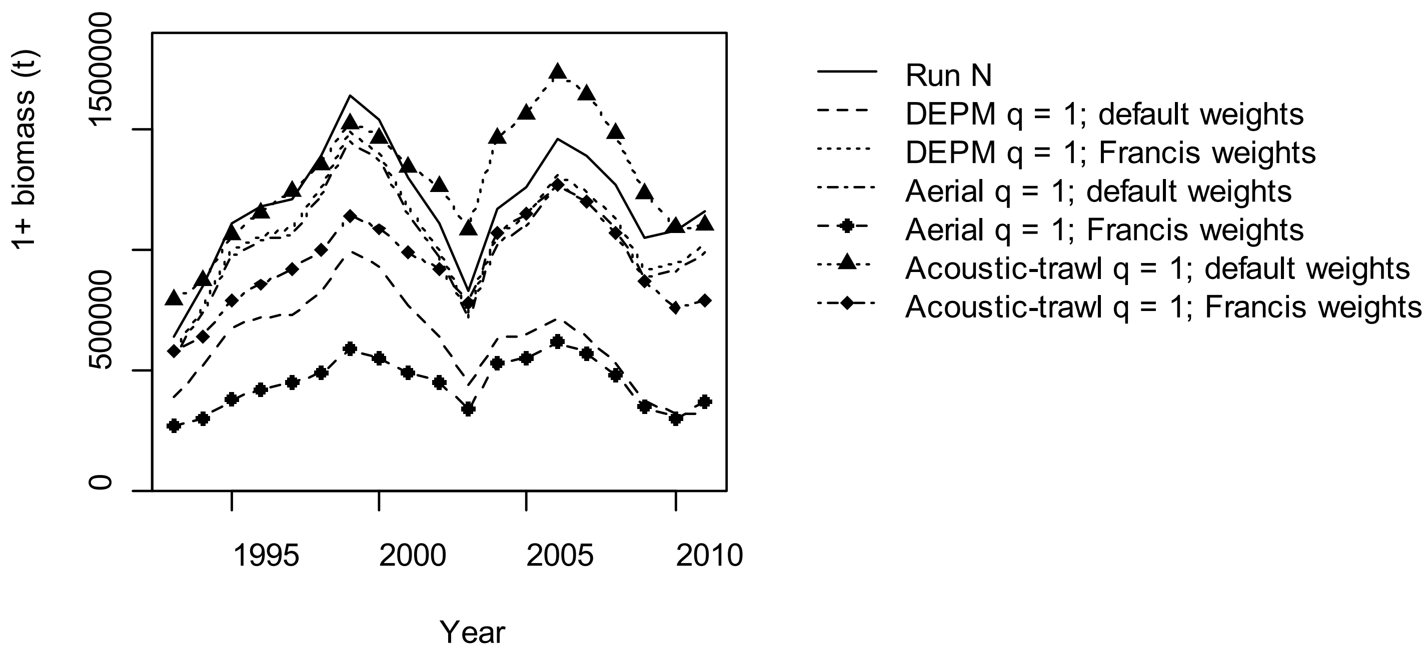

Figure 1. Time-trajectories of 1+ biomass from run N and six variants of this run in which each of three survey series are assumed to be absolute indices of abundance and the weights assigned to the age and length data are set to the default values and reduced as in run X.

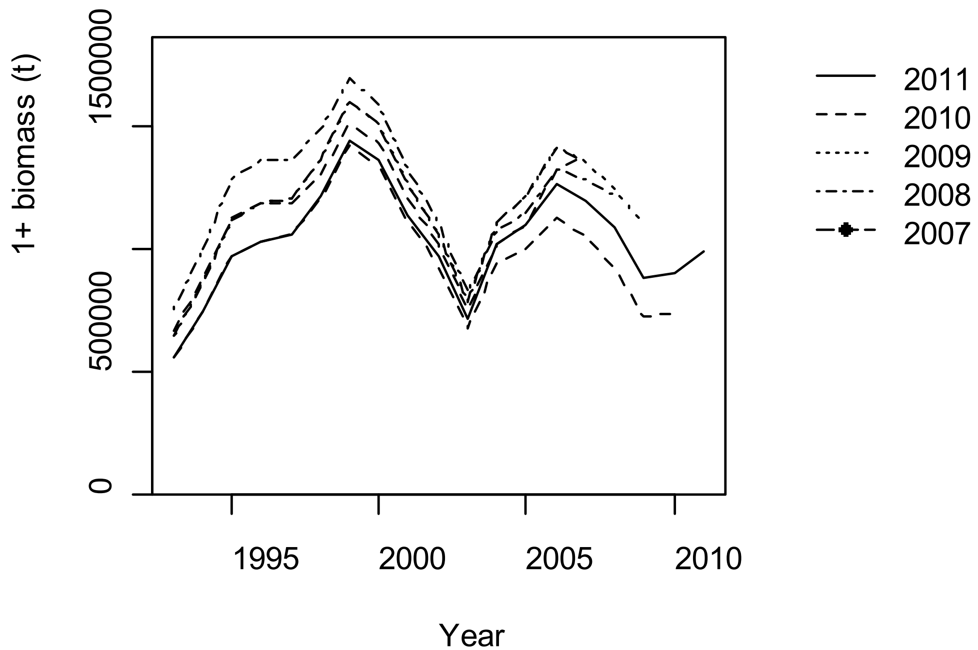

Figure 2. Results of the retrospective analysis based on the final base model.

Appendix 1
2011 Pacific Sardine STAR Panel Meeting Attendees

STAR Panel Members
André Punt (Chair), University of Washington
Ray Conser, NOAA Southwest Fisheries Science Center
Chris Francis, New Zealand National Institute of Water & Atmospheric Research
Larry Jacobson, NOAA Northeast Fisheries Science Center

Other Attendees
Mike Okoniewski, CPSAS Rep to STAR Panel
Lorna Wargo, CPSMT Rep to STAR Panel
Kevin Hill, NOAA Southwest Fisheries Science Center (SWFSC)
Kerry Griffin, Council Staff
Jenny McDaniel, SWFSC
Nancy Lo, SWFSC
Beverly Macewicz, SWFSC
Paul Crone, SWFSC
David Demer, SWFSC
Greg Krutzikowsky, ODFW
Steve Marx, Pew Charitable Trusts
Piera Carpi, UMass, Dartmouth
Sandy McFarlane, Canadian DFO & Canadian Pacific Sardine Association
Linnea Flostrand, Canadian DFO
Bob Seidel, Commercial fishing
Kirk Lynn, CDFG
Jerry Thon, Northwest Aerial Sardine Survey (NWSS)
Tom Jagielo, NWSS
Dale Sweetnam, SWFSC
Erin Reed, SWFSC
Sam Herrick, SWFSC
Diane Pleschner-Steele, CA Wetfish Producers Association
Ryan Howe, NWSS
Richard Carroll, Ocean Gold Seafood
Ed Weber, SWFSC
David Haworth, Commercial fishing
Fabio Campanella, SWFSC
Josh Lindsay, NMFS SWR
Christina Show, SWFSC
Russ Vetter, SWFSC
Emmanis Dorval, SWFSC
Kristen Koch, SWFSC
Briana Brady, CPSMT

Appendix 2
Comments on Weighting of Composition Data
Chris Francis

The composition data in many stock assessment models are given too much weight because most approaches to assigning weight to this type of data ignore the strong correlations in these data (and also in the associated residuals). A useful way to highlight this problem is to plot observed and expected mean lengths (or ages), as in done in Figure 1 for the base model length comps. The fact that the expected mean lengths in this plot are often outside the confidence intervals for the observations indicates that the data are over-weighted. Down-weighting these data (by decreasing the multinomial sample sizes) would increase the width of the plotted confidence intervals.

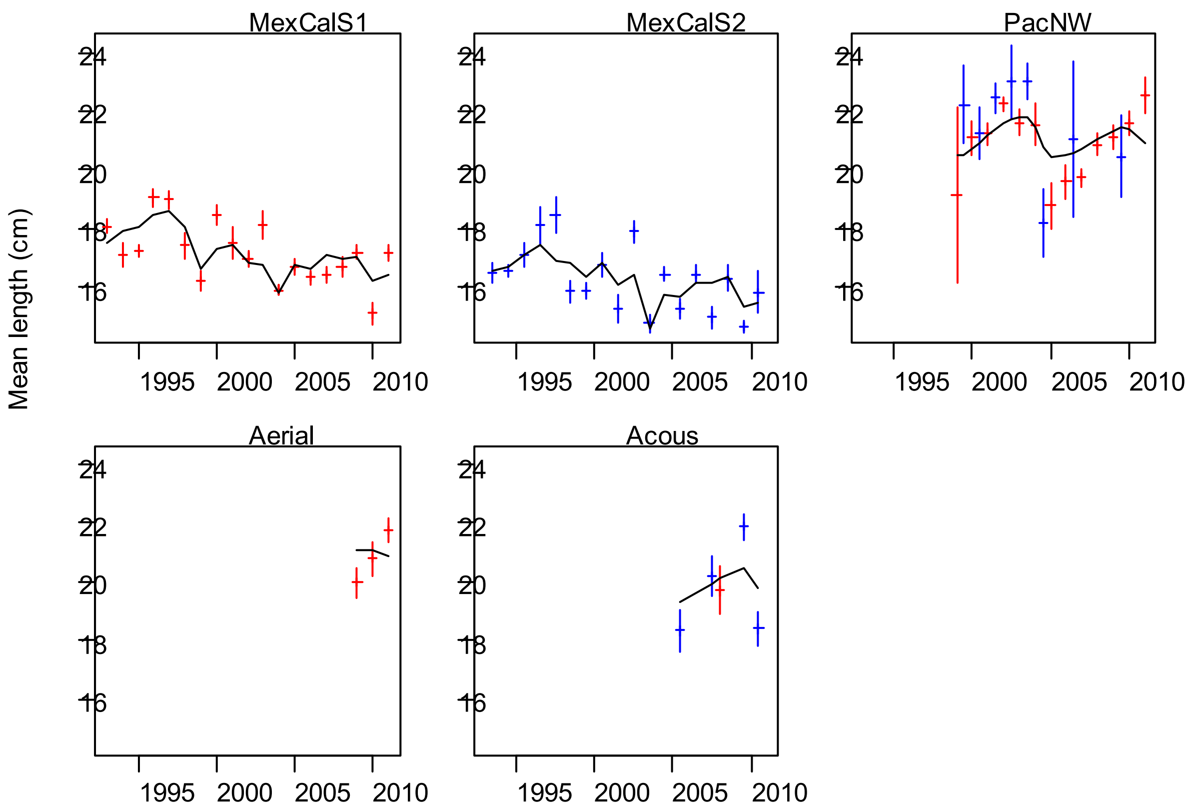

Figure 1: Observed ('+', with 95% confidence intervals shown as vertical lines) and expected (lines) mean lengths for all length composition data in the base model. The plotting colour of the observed values indicates the semester (red for semester 1, blue for semester 2). The confidence intervals were calculated using the multinomial sample sizes assumed for the base model (i.e., the products of the initial sample sizes and effN_mult_Lencomp values in Tables 4 and 9 of the assessment report).

The method of iteratively reweighting composition data in Stock Synthesis implicitly assumes that the residuals associated with one length (or age) bin are uncorrelated with those in another bin. In fact, correlations between composition residuals are often strong, and show a characteristic pattern like that in Figure 2.

One way of avoiding over-weighting composition data (by ignoring these correlations) is to base the re-weighting calculations on the residuals of mean length (or age), rather than on residuals of individual proportions. When this was done for the length composition data in the base model it suggested that the multinomial sample sizes for these data should be smaller by a factor of 0.06 – 0.1 (Table 1).

Full details about this method of re-weighting composition data are given in Francis (2011) [see method TA1.8 in Table A1; the w_j in that table is the same as the N_multipler in Table 1 below].

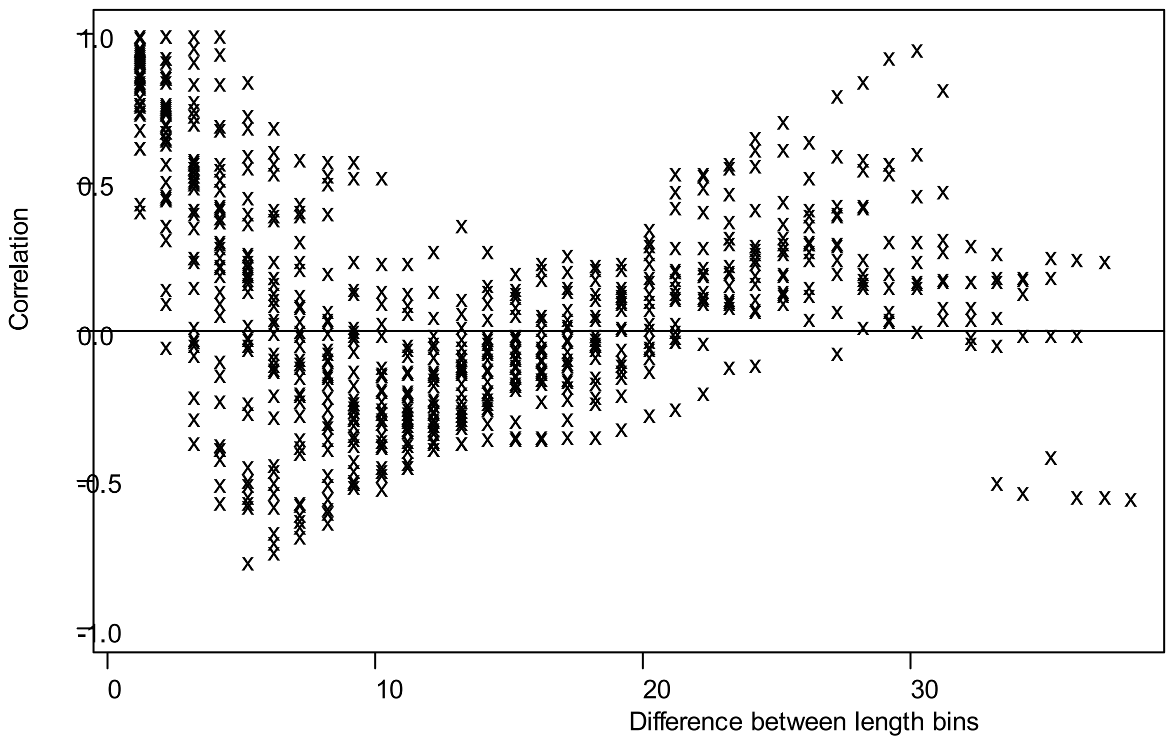

Figure 2: Correlations amongst the residuals from the MexCal_S1 length comps in the base model. Each plotted point represents a correlation between the vector of residuals for one length bin and that for a different length bin; the x-axis shows the difference (number of bins) between the two length bins.

Table 1: Suggested reweighting of the length composition data from the base model, showing the median sample sizes assumed for each data set in the base model (N_base), an N multiplier calculated from the mean length residuals, and the suggested median sample sizes (N_new), which are the product of N_base and the multiplier. Because of small sample sizes (i.e., few years of observations), the N_multiplier for the aerial and acoustic-trawl surveys was calculated by combing these two series.

Data set	Median N_base	N_multiplier	Median N_new
MexCalS1	135.9	0.058	7.9
MexCalS2	117.7	0.061	7.2
PacNW	40.9	0.104	4.3
Aerial	14.8	0.067	1.0
Acous	43.5	0.067	2.9

Reference

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

Appendix 3
West coast of Vancouver Island sardine trawl survey

Provided by L. Flostrand and J. Schweigert
Department of Fisheries & Oceans Canada
Pacific Biological Station, 3190 Hammond Bay Rd. Nanaimo, BC V9T 6N7

Summer surveys directed at collecting information on sardines off the West coast of Vancouver Island (WCVI) started in 1997. Fishing is conducted in surface waters ($\leq$ 30 m) using a mid water trawl towed at average speeds approximating 4-5 knots. Since 2006, sampling has been conducted at night. Biomass estimates are based on extrapolating the average sardine catch density (metric ton /km^3) by stratum over an estimate of the stratum's spatial size (km^3) and then summing across strata. The core area of the survey region is approximately 16,740 km^2 and catch densities are assumed to represent sardine distributions in the top 30m of the region, therefore the region's surface volume is estimated at ~ 502.2 km^3 (see Figure below). Recent regional estimates of sardine catch density and seasonal biomass in the WCVI core survey region from night sampling in 2006 and 2008 to 2010 (no survey was conducted in 2007) show a declining trend, whereas the 2011 estimates are approximately double the 2010 estimates (see Table below).

The current Canadian harvest control rule is based on the U.S. assessment of coastwide adult biomass and the migration rate of sardines into Canadian waters (Ware 1999, Schweigert et al 2009, DFO 2009), upon which a harvest rate equivalent to the U.S. rate is established (a 15% harvest rate has been in place since 2002; DFO 2010). More information on the provision of science advice and the harvest control rule is reported in the 2011 Science Advisory Report on the *Evaluation of Pacific sardine stock assessment and harvest guidelines in British Columbia* (http://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2011/2011_016-eng.pdf, DFO 2011)

Table. Summary information and statistics associated with West Coast Vancouver Island (WCVI) trawl survey sardine catch densities and biomass estimates. For 95% confidence interval, LL= lower limit and UL= upper limit.

YEAR	2006	2008	2009	2010	2011 *
WCVI SAMPLING					
Tows with sardines / total number of tows	42/45	44/71	53/109	40/72	41/68
Core survey region					
Tows with sardines/ total number of tows	41/44	40/60	47/95	37/57	41/68
SARDINE DENSITY (mt/km^3)					
Mean	759.9	420	378.3	163.2	~300.0
95% LL	461.6	196.5	220.2	57.6	Not available
95% UL	1,105.60	736.4	557.8	309.7	Not available
CV **	0.23	0.33	0.23	0.39	~0.28
BIOMASS (mt)					
Mean	381,617	210,924	189,977	81,964	~150,000
95% LL	231,816	98,682	110,589	28,927	Not available
95% UL	555,232	369,820	280,127	155,541	Not available

* 2011 estimates are preliminary and have not been reviewed
** CVs presented above have been corrected from previously reported estimates (reported to have ranged from ~ 1-3).

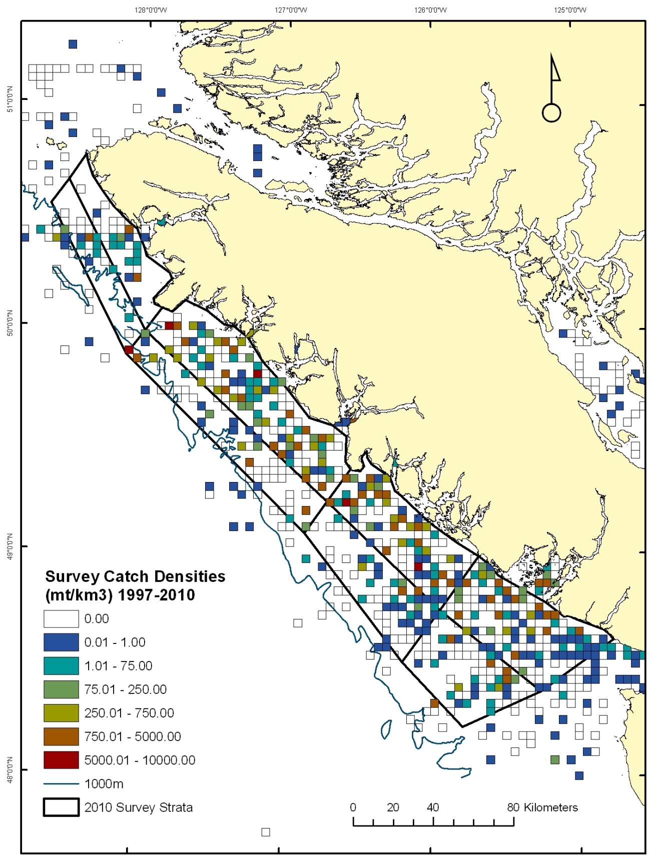

Figure. Mean sardine densities for all 1997-2010 sardine survey trawl tows based on 4x4 km sized grid cells. Outer boundaries define the core WCVI survey region. Also shown are sub-regional boundaries as they pertain to future work interests for stratification schemes.

REFERENCES

DFO. 2009. Proceedings of the Pacific Scientific Advice Review Committee (PSARC) meeting for the assessment of scientific information to estimate Pacific sardine seasonal migration into Canadian waters. DFO Can.Sci. Advis.Sec. Proceed. Ser. 2009/034.

DFO. 2010. Pacific Sardine Integrated Fisheries Management Plan 2010/2011. Government of Canada.

DFO. 2011. Evaluation of Pacific sardine (*Sardinops sagax*) stock assessment and harvest guidelines in British Columbia. DFO Can. Sci. Advis. Sec. Science Advisory Report. 2011/016.

Schweigert, J., McFarlane, G.A., and Hodes, V. 2009. Pacific sardine (*Sardinops sagax*) biomass and migration rates in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/088. 14p.

Ware, D.M. 1999. Life history of Pacific sardine and a suggested framework for determining a B.C. catch quota. DFO Can. Sci. Advis. Sec. Res. Doc. 1999/204.