NPFMC Crab Modeling Workshop

Hilton Hotel, Anchorage

January 14-17, 2014

Summary

A technical crab modeling workshop took place January 14-17, 2014, at the Hilton Hotel in Anchorage, AK. The workshop was chaired by André Punt (University of Washington), and was attended by members of the Crab Plan Team (CPT), the authors of crab and groundfish stock assessment models, outside technical stock assessment experts, and the general public (see Appendix A for a list of participants). This workshop was the fourth NPFMC workshop to review and improve stock assessment models for Bering Sea and Aleutian Islands (BSAI) crab stocks. The workshop focused on a review of the generic crab model which is currently under development, the assessment for Norton Sound red king crab, and how raw data are assembled for use in crab assessments. A generic model was also applied to data for Bristol Bay red king crab.

A variety of presentations were made to the workshop (see http://www.afsc.noaa.gov/REFM/stocks/Plan_Team/crab/draft_assessments.htm). The workshop followed a format intended to facilitate real-time model development and review. Presentations were made to the meeting, with additional work requested for clarification or model development. The meeting then moved to a different topic to allow analyses and model changes to be made. In the context of the above format, the workshop made a number of requests for additional analyses to the analysts. These requests, their rationale and the outcomes from the work conducted are reported for each topic below.

The key consensus, conclusions, and recommendations were:

Generalized crab model (Gmacs)

- Gmacs represents a worthwhile approach that should, in time, provide a unifying framework to the current suite of assessment models that have been developed for Alaskan crab stocks. However, Gmacs is currently a "work in progress" and more work needs to be done before it can be used as an alternative to any of the current assessment models. The workshop does not anticipate that it will be possible to use Gmacs for the September 2014 assessments, but this should be possible for the September 2015 assessments.
- The model code and documentation will be updated based on the detailed recommendations of the workshop (see Section A for details).
- Athol Whitten and Jim Ianelli will work with Jie Zheng and Hamachan Hamazaki to achieve credible assessments for NSRKC and BBRKC using Gmacs.
- There is a need to consider how the software will eventually be handed over to assessment authors and be updated over time.

Norton Sound red king crab assessment

- The next assessment of Norton Sound red king crab should set the extent of variation in recruitment (σ_R) to 0.5. The previous use of a weight of 0.01 implied a much higher σ_R (7.0), which meant that the model was following "noise" in the observer length-frequency data. Using σ_R to specify recruitment variation ensures that this parameter is set to a reasonable value.
- The model is likely overly complex, estimating more parameters than can be supported by the assessment information that is available for the stock. The workshop evaluated several ways to simplify the model, but further simplification is warranted. Approaches that were evaluated included:
  a. forcing selectivity to equal 0.999 (or some similar high value) at the largest size bin to reduce the number of parameters and also reduce poor convergence behavior;
  b. assuming selectivity for the ADFG and NMFS trawl surveys are the same;
  c. modeling growth and molting processes using a single transition matrix; and
  d. removing the winter pot survey data.
- There is a fundamental conflict between the tagging data used for growth and the assumed longevity of Norton Sound red king crab. This conflict is still present, despite attempts at simpler modelling approaches. One possible way forward would be to include the tagging data in the assessment model to estimate growth within the model. This would allow the trade-offs between fitting different data sources to be made in a statistically consistent way.

Data for use in crab assessments

- Many crab assessment include 'legacy' data, the origins of which are uncertain. This is partially a result of changes in analysts over time, as well as the length of some of the data time-series.
- The data used to construct historical information on bycatch (groundfish and crab fisheries) are of varying quality. Early data should be considered highly uncertain. In some cases (e.g. groundfish discards), it is no longer possible to reconstruct the historical data used in assessments.
- Data preparation procedures were reviewed for each of the main data categories within each assessed crab stock to check uniformity in data preparation procedures among assessments. In particular, the preparation of historical data was reviewed to better understand how these data were prepared and to ensure that the best available procedures had been followed.
- The survey data (estimates, CVs and length-composition data) for Bristol Bay red king crab and St. Mathews blue king crab should be based on an extract by Bob Foy to ensure that the same methodology is used over time. The survey data for all stocks should be updated once a new approach for constructing indices is selected by the CPT.

General modeling recommendations

- Input data should be verified, well documented, and reproducible.
- Survey size-composition data should be accompanied by an index of population abundance.
- Use tagging data (if available) for direct estimation of growth by incorporating the observed length-increment information into the model and estimating the model growth parameters simultaneously with the other model parameters.

- Avoid using dome-shaped selectivity.

Actions arising for the May 2014 CPT meetings

- Bob to provide alternative survey time-series, and the CPT to recommend which approach to use in the future.
- Bill Gaeuman to provide all assessment authors updated crab fishery discard information (total numbers discarded and length-frequency of discards and the total observed catch).
- Bill Gaeuman to provide a comparison of the "simple averaging" method for estimating discard in the crab fisheries and the "design-based" approach, for the years for which comparisons can be made.
- Jack Turnock to obtain the early groundfish trawl length-frequency data (pre-1991) and include it in the assessment of snow crab.
- Norton Sound red king crab:
    - A full assessment should be conducted with a range of suggested scenarios so the May CPT can recommend an OFL and an ABC for the 2014-15 management cycle. The assessment will need to be revised again for the September 2014 CPT meeting and the September specification cycle.
    - Provide alternative model runs where selectivity for the ADFG and NMFS trawl surveys are assumed the same, and where different selectivity patterns are estimated for each survey.
    - Provide alternative model runs in which the growth transition matrix incorporates both growth and molting probabilities. If possible, develop a model that incorporates the growth data, and the growth transition matrix is estimated.
    - Provide alternative model runs in which: 1) the size-composition data from the winter pot survey are excluded, and 2) the CPUE data for the winter pot survey are included.
    - Review the data used to calculate the growth transition matrix, and provide an overview of the new tagging program and the data that it is expected to provide.
- Further review the progress on applying Gmacs to NSRKC and BBRKC. Athol should provide an updated version of the documentation of the model as well as diagnostic statistics and model outputs for the two stocks.
- All assessment authors should provide model scenarios which mimic the September 2013 assessments by replacing the bycatch data in the crab fisheries with updated data from Bill Gaeuman using the "simple averaging" method and by replacing the NMFS survey data with recalculated series based on updated methodologies so the CPT can evaluate the implications of these changes to the data.
- The May 2014 CPT may need to be 4-5 days long (rather than 3 days), given the need to review the NSRKC and BBRKC assessments as implemented in Gmacs as well as reviewing a revised assessment for Aleutian Islands golden king crab.

In closing the workshop, the Chair recognized the considerable work undertaken by the analysts (Toshihide (Hamachan) Hamazaki and Athol Whitten), by responding to the requests for additional analyses. He also noted the excellent work undertaken by the rapporteurs (Martin Dorn, Paul Starr, and Buck Stockhausen) in assembling the draft report. Finally, he recognized the considerable effort made by the Anchorage-based participants (particularly Hamachan and Diana) by attending the meeting sessions given the less-than-desirable weather conditions.

A. Gmacs overview and technical description

Athol Whitten, a post-doctoral researcher at the University of Washington, provided both a written and an oral overview related to version 1.0 of the size-structured assessment model framework called Gmacs, *Generalized Modeling for Alaskan Crab Stocks*. He also provided technical details of Gmacs to workshop participants in both the writeup and presentation. Athol is developing Gmacs in collaboration with André Punt (UW) and Jim Ianelli (NOAA/NMFS/AFSC). In addition to developing this modeling framework, Athol is also constructing two demonstration assessments using the Gmacs framework: one for Bristol Bay red king crab and one for Norton Sound red king crab. The project to develop Gmacs is funded by NOAA and the Bering Sea Fisheries Research Foundation (BSFRF). Mark Maunder and Steve Martell are also collaborators on the development.

Gmacs is a statistical size-structured population modeling and stock assessment framework that is designed to be flexible, scalable, useful for both data-rich and data-poor stocks, and applicable to Alaskan crab stocks. Assessment models based on Gmacs will be able to incorporate multiple data types from a variety of fisheries and surveys by combining all data in the form of an integrated analysis. This is important for Alaskan crab stocks, as most stocks are represented by numerous data sources—some of which have short or incomplete time series.

Version 1.0 of Gmacs is currently under development. It is based on ADMB (Auto-Differentiation Model Builder; Fournier et al., 2012) code, as well as a library of generic ADMB-based functions called Cstar (Common Stock Assessment Routines) that Athol is developing concurrently. Gmacs is being written in a fashion that allows contributions from multiple developers, as well as facilitating documentation. Source code for Gmacs is being hosted on the web-accessible site GitHub at https://github.com/awhitten/gmacs, from which it can be downloaded as a zip file. Alternatively, researchers interested in contributing to the development of Gmacs can create a local repository of the Gmacs project using "git", edit existing files and create new ones, and (once reviewed by Athol and Jim) update the project hosted on GitHub. Gmacs source code aspires to be well-commented and uses Doxygen (http://www.stack.nl/~dimitri/doxygen/) to provide easily updated project documentation via a simple but powerful markup language for comments in the source code.

Workshop participants discussed the merits of the Gmacs framework, reviewed the technical details of the framework and made suggestions for enhancing the generality, flexibility, and applicability of the framework to Alaskan crab stocks. Participants also provided editorial comments regarding the technical description made available to the workshop.

In general, workshop participants thought that Gmacs represented a worthwhile approach that should, in time, provide a unifying framework to the current suite of assessment models that have been developed for Alaskan crab stocks. However, they also noted that Gmacs is currently a "work in progress" and required further work before it could be used as an alternative to any of the current assessment models.

The generalized population dynamics in Gmacs are governed by a size-structured model that accounts for changes over time to numbers-at-sex, -maturity, and –size class. Currently, numbers are not tracked by shell condition, but the ability to do so is under development. There was discussion whether it would be worthwhile extending the model to track numbers by age class as well. André Punt noted that he has an NPRB-funded project to develop age- and age/size-structured models for crab stocks (http://project.nprb.org/view.jsp?id=a3b426b7-7d16-46e2-b352-d88547d05807). André felt that it would probably be more useful to incorporate spatiallyexplicit considerations in Gmacs rather than incorporating age explicitly within the framework at this time.

The model accounts for changes in population numbers due to growth, maturation, natural mortality, fishery removals and recruitment. Model parameters can be specified using time blocks to implement temporal variation in parameter values. Time blocks are specified in a model control file, rather than being hard-wired in the model code (as is the case for some current assessment models).

Recommendations regarding the model formulation and dynamics:

1. implement mean total (as opposed to sex-specific) recruitment deviations, and (possibly) time-varying sex ratios to describe time-varying recruitment, rather than using sex-specific means and deviations;
2. consider implementing spline functions as an alternative to "piecewise linear" functional descriptions of model processes (e.g. probability of maturing at size);
3. reformulate estimation of fishing mortality rates when effort is available, but catches are missing (section D.4);
4. reformulate estimation of initial population structure by building up recruitments over time prior to the "start" of the model, rather than directly estimating initial numbers-at-sex/maturity/shell condition/size class (sections C.9 and D.5(2));
5. consider using a cumulative normal distribution (and/or other functions) rather than a gamma function to describe size-at-recruitment and/or growth increment to reduce computational time;
6. implement additional survey variance to represent year-to-year process error given that the survey standard errors underestimate the true extent of difference between the survey estimates and actual population abundance.;
7. incorporate the potential to use historical tagging data to estimate growth (but not population abundance) within the model;
8. implement time-varying recruitment using deviations from a stock-recruitment curve, as alternative to using deviations from a mean;
9. implement the capacity to change the functional form for selectivity over time in blocks;
10. implement the capacity for penalties to be placed on the extent to which parameters change over time when, for example, selectivity is modeled using time blocks;
11. allow natural mortality (*M*) to depend on maturity state and length as well as being time-varying (as implemented in current Tanner and Norton Sound red king crab models);
12. implement time-specific, flexible retention curves (e.g., double normal) and eliminate the "high-grading" parameter;
13. formulate documented likelihood equations using complete functional forms (i.e., include constants);
14. consider alternative choices for the likelihood function for the length-composition data such as the robust normal for proportions as well as the new approach proposed by Chris Francis (Francis, in press);
15. turn off penalty E.3 in final estimation phase;
16. implement the 'hybrid method' as an option for calculating annual fishing mortality rates (Methot and Wetzel, 2013) – the results from the hybrid method could be used as starting values if the annual fishing mortality rates are treated as estimable parameters;
17. as a future consideration, consider alternatives to MMB as proxies for fecundity;

18. consider developing a likelihood for the composition data which fits the composition data by sex/shell/maturity state separately and the proportion of the total catch which is in each sex/shell/maturity state;
19. include molting probability (which may change over time) into the calculation of the growth transition matrix; and
20. add code so that retrospective analyses can be conducted automatically.

Recommendations regarding model input, outputs and diagnostics:

1. create duplicate versions of the input files as the data are read into the model so that they can be checked and also so that they can be used as the basis for simulation tests;
2. make the number of size-classes in the input data fleet-specific (e.g. you may enter 20 size-classes for survey data but only 10 for fishery data) and reformat the data internally to a consistent number of length bins;
3. add an option to 'jitter' the starting values for all parameters to help assist with testing whether convergence has been achieved;
4. implement Francis' iterative re-weighting (Francis 2011) as an alternative to the current Ianelli/McAllister weighting for size composition likelihoods;
5. report residual plots, Q-Q plots, plots of fits to the length-frequency data, etc. (see Punt and Kinzey [2009] for examples of plots and Section E for additional recommendations related to diagnostics);
6. add confidence intervals to plots of data points which are considered uncertain (effort and abundance indices);
7. list in an output file the estimated parameters, their starting values, and final estimates, their bounds and phases, as well as their asymptotic standard errors – automatically highlight cases where parameters are on or close to bounds in the output.

Recommendations regarding the model description and nomenclature:

1. standardize the numbering of model equations so that equation numbers correspond to the paragraph sequence within the document (sub)sections (e.g., the first two equations in section C.1 would be labeled C.1.1 and C.1.2 rather than 1 and 2 or C.1 and C.2);
2. revise equation C.2 – there should be a plus sign instead of a minus sign in front of "D";
3. distinguish "fleets" from 'fisheries"; at present, the term fleet is used for the internal representation of data types as well as for actual fleets which can be confusing – one suggestion is to use the terms 'data components' and 'fleets';
4. make sure subscripts/superscripts in all equations are consistent and correct;
5. retention as a logistic function should be an ascending (not descending) function of size (eq. C.8 and prior description);
6. check consistency in eq. C.9 (α_R^s and β_R^s in eq. C.9 are sex-specific but PR_l is not);
7. don't use "R" when referring to both retention and recruitment;
8. replace the term "piecewise linear" with a more appropriate description;
9. maturation should be described by an "ascending" (not descending) logistic function (section C.5);
10. refer to 'abundance indices' rather than surveys because some crab assessments include fishery CPUE data
11. the notation in section D.1 should be improved and made consistent with section C (i.e.., use "C" for retained catch and "D" for discards in eq. D.1);

12. equations need to be provided in section D.2; and
13. document handling of 'sub-surveys' (e.g., the approach used in the snow crab assessments to make use of the BSFRF survey data to estimate survey selectivity).

B. Example Applications of the Generic Model to Bristol Bay Red King Crab and Norton Sound Red King Crab

Athol provided the meeting with an initial Gmacs version of a Bristol Bay red king crab assessment. The model on which this application was based did not include shell condition, and is preliminary. Athol and Jim Ianelli worked on the model code and its implementation for BBRKC and NSRKC during the workshop. The workshop made several requests for additional analyses. The analyses conducted during the meeting led to identifying several areas where the code provided to the meeting needed to be generalized further (see Section A of this report).

1. Request: Provide additional residual diagnostics for BBRKC. Show the fits to the length data as well as residual plots. Also, add confidence intervals to the fits to the abundance data.
Rationale: The diagnostics provided were inadequate to fully evaluate the model.
Response: The example application to Bristol Bay red king crab generally fitted the data adequately (Figure 1), but the model was unable to fit the high catches in the late 1970s, tended to overestimate the survey index in the early years of the time-series, and was unable to mimic the length compositions for red king crab caught during the Tanner crab fishery. Moving to a dome-shaped selection pattern for the Tanner crab fishery should resolve the latter problem. However, Gmacs fitted the abundance estimates for recent years better than the currently-adopted assessment.

2. Request: Conduct a preliminary assessment of Norton Sound red king crab.
Rationale: Preliminary results were only shown for NSRKC.
Response: The model fit the data very poorly. In particular, it was unable to mimic the catch data, which was unexpected. Further investigation revealed that the length composition data were down weighted in the NSRKC assessment, but not in the example application.

3. Request: Evaluate why the model does not fit the trawl by-catch (there are positive catches for some years for which the observed data appear to be zero catches).
Rationale: It would be expected that the model would mimic the catch data very well.
Response: This was a plotting error.

4. Request: Report how many parameters are at or near bounds (and ideally create a table which included standard errors for the parameters).
Rationale: The meeting wished to understand how often parameters were being estimated close to bounds.
Response: This could not be completed during the meeting, but will be available for the May 2014 CPT meeting.

5. Request: Repeat this assessment with the same size groupings (including the smaller size groups) that are used in the current Bristol Bay red king crab assessment.
Rationale: only the largest size bins were used in the current example.
Response: To be presented to the May 2014 CPT meeting.

C. Norton Sound Red King Crab Model

The NSRKC assessment was not accepted by the SSC in September 2013 because the SSC did not agree with the approach used by the CPT to specify the 2013 harvest information. The CPT recommended a model that excluded summer 2013 observer data. The summer 2013 observer data, which had not been "finalized," included a mode of small crab. The model interpreted this mode as a very strong 2013 year class and calculated a substantial increase to the 2014 OFL and ABC. The SSC recommended that the winter fishery be managed using the ABC and OFL from the model adopted at the June Council meeting. This approach delayed moving the NSRKC to an October-September assessment cycle, and will require that assessment model recommendations be brought forward at both the June and September 2014 Council meetings. The June model should provide the ABC and OFL for the July-June cycle, and will incorporate all recommended changes to the model from this workshop and earlier CPT meetings.

One concern the CPT had about moving to an October-September assessment cycle was whether information from the summer fishery would be available in time to be used for the assessment. The experience in 2013 suggested that this is unlikely to be a problem. The 2013 summer fishery for NSRKC extended later than usual, yet the observer data from the summer fishery were available in preliminary form.

Hamachan presented likelihood profiles for natural mortality (*M*) and recruitment variability (σ_R), which were requested at the September CPT meeting. The likelihood profile for recruitment variability was generated for σ_R in the range of 0.01 to 0.5. The profile for the total log likelihood and most data components was lower for higher values of σ_R. **The use of σ_R = 0.5 was considered appropriate for the next assessment.**

The likelihood profile on *M* assumed constant *M* for all length intervals, unlike the accepted assessment model which assumes that *M* for the final length interval is 3.6 times the nominal *M*. The model showed the best overall fits for *M* in the range 0.4-0.5yr^{-1}, but there was contrast among data components. The weights given to each data component will determine the minimum in the total log likelihood profile. It was noted that the model failed to converge when *M* was 0.36yr^{-1}, which generated considerable discussion among workshop participants. Analyses conducted during the meeting showed that the lack of convergence for M = 0.36yr^{-1} was the result of a parameter converging on its minimum bound. Widening the bounds for the parameters removed the convergence problems. It became apparent that many of the selectivity parameters were being estimated inconsistently, primarily because selectivity appeared to be high and without much curvature for both surveys (NMFS and ADFG) and in the fishery.

1. Request: Redo the likelihood profile for natural mortality for NSRKC. NMFS survey selectivity, ADFG survey selectivity, and fishery selectivity should be specified so that selection is 0.99 for the largest size category, and only the slope parameter is estimated. Record the max gradient value from the output parameter file.

Rationale: Examination of selectivity parameter values over the profile suggested they were outside the range where they could be informed by the data.

Response: Selectivity was estimated as constant for all lengths for the NMFS and ADFG surveys. There was a convergence problem for the M = 0.14yr^{-1} model run. The maximum gradients were large, indicating that model had not converged. This run needs to be checked further (i.e., using the derivative checker).

2. Request: Apply the ADMB derivative checker to identify where the code for the model is non-differentiable.
Rationale: Non-differentiability may be the cause of the convergence problems.
Response: Hamachan reported that the ADMB derivative checker did not indicate any problems with non-differentiability.

3. Request: Redo the likelihood profile for natural mortality for NSRKC by reducing the weight assigned to the summer length composition data to 10% of the original weights.
Rationale: It appeared that the summer length composition data contributed substantially to the overall likelihood.
Response: The run provided to the meeting was based on reducing the weight assigned to the length data by 90%. However, the changes made for Request 1 also reduce the weight on the summer length composition data.

4. Request: Provide results for model runs with values of σ_R of 0.5 and 7.0, and the September (preliminary) and December (final) versions of the 2013 summer length composition observer data, resulting in four runs covering all combinations.
Rationale: This request attempted to identify the reason for the very large recruitment estimate in 2013 by evaluating two hypotheses: incomplete observer data and/or the use of a weak constraint on recruitment deviations.
Response: Both model runs using σ_R of 0.5 eliminated the very large recruitment estimate in 2013 and gave an acceptable result (Figure 2). Results were not sensitive to whether or not the finalized observer data were used.

5. Request: Tabulate all size transition data, and the molting probability for crab that were at liberty for one year using the available tagging data. Document the method used to produce the current size transition matrix.
Rationale: To investigate the growth transition matrix used in the assessment model
Response: Some results were provided. However a full response to this request will require additional work and should be provided to the CPT in May 2014.

6. Request: Provide a model run which estimates a multiplier on the transition probabilities between both the 4^{th} and 5^{th} size classes and the final size class. Transition probabilities should be renormalized to sum to 1.0. Assume selectivity is flat, $M = 0.18yr^{-1}$ and $\sigma_R = 0.5$.
Rationale: To evaluate whether the changes in transition probabilities to the final size bin could account for the lack of crab in the final size bin (instead of increased natural mortality as in the base model).
Response: Results were very similar to those for the base model. This model included estimated selectivity for the winter pot survey with a decrease for the last size bin, which was not intended in the request. Therefore, request 7 was made.

7. Request: Rerun the model which estimates a multiplier on the transition probabilities between both the 4^{th} and 5^{th} size classes and the final size class. Transition probabilities should be renormalized to sum to 1.0. Assume selectivity is flat (**including the winter pot survey**), $M = 0.18yr^{-1}$ and $\sigma_R = 0.5$.

Rationale: To evaluate whether the changes in transition probabilities to the final size bin could better account for the lack of crab in the final size bin (instead of increased natural mortality as in the base model).
Response: Results were very similar to the base model.

8. Request: Provide a model run which estimates a new growth transition matrix based on the raw tagging data and fixes the molting probabilities to 1.0. This transition matrix should include only crab at liberty for one molting cycle, and include the processes of both molting and growth.
Rationale: This simplifies the model by combining molting and growth into a single transition matrix, and evaluates a transition matrix that was constructed from raw tagging data.
Response: Model results appeared plausible, although fits to the size composition data were poorer. This model still included higher natural mortality for the largest size bin, which was not intended in the request. Therefore, request 9 was made.

9. Request: Provide a model run which estimates a new growth transition matrix based on the raw tagging data and fixes the molting probabilities to 1.0. This transition matrix should include only crab at liberty for one molting cycle and includes the processes of both molting and growth, **Fix M=0.18yr^{-1} for all length intervals**.
Rationale: To obtain a model that satisfies the intent of request 8.
Response: Removing the higher natural mortality for the final size bin had a very strong influence on model results. Stock size was lower in the initial years relative to models which had a high *M* in the final length interval, then declined to low levels, followed by a strong increasing trend. Model fits to the data were much poorer than for the base model,

10. Request: Provide a model run with the winter pot survey data removed. Use old base model, but fix M=0.18yr^{-1} for all lengths.
Rationale: To evaluate the value of including the winter pot survey data in the model. Some workshop participants found little rationale for including survey size composition data when survey indices were not considered reliable.
Response: Removing the winter pot survey data did not have a strong effect on model results. The winter pot survey length composition data do not appear to contribute much to the assessment model results.

11. Request: Provide results for two models, both of which exclude the winter pot survey length data and input a new growth transition matrix based on the raw tagging data and fixes the moult probabilities to 1.0. One model should use a constant M=0.18yr^{-1} for all lengths; the other model should include the higher *M* for oldest length bin.
Rationale: To see if a model with the new growth transition matrix and the winter pot survey data removed still shows the conflict between the growth data and assumed natural mortality of 0.18yr^{-1}.
Response: Conflict between growth data and assumed longevity is still present with these changes to model. Elimination of the winter pot survey data and use of simpler approach to modeling growth and molting probability seem to be improvements to the model, but do not resolve the fundamental problem.

D. Assembling Data for Use in Crab Assessments

The workshop conducted a review of the major data preparation procedures for each of the crab stock assessments. The purpose of this review was to ensure that data preparation procedures were as uniform as possible among the various stock assessments, and that they followed accepted practices. In particular, the CPT wanted to review historical data inputs to ensure that these values had been prepared acceptably, given the availability of information. The CPT also wanted to compare the historical data that have been used in past stock assessments with the equivalent estimates after being re-analysed based on currently available information.

D.1 Survey data

Bob Foy presented re-analyses of the Tanner crab data collected during the NMFS Bering Sea bottom trawl survey. These analyses included a consideration of the survey design and how it has been executed since 1975, the handling of crab for which biological information (sex, length) is missing, and the collection of additional data to characterize male Tanner crab maturity. The workshop was presented with a number of options regarding how to use this information to create a comparable time series of abundance indices from this survey.

The workshop made a number of recommendations arising from this presentation:

- The CPT should consider standardizing how survey data are analysed (rather than basing, for example, selection of survey stations so that the calculated abundance estimates match those produced historically). The workshop noted that changes to survey design have crept into the execution of this survey over its long history. These modifications affect the comparability of the abundance indices when included in a stock assessment model. Bob provided the workshop with several options for selecting stations and tows, such as including only a single tow per grid, dropping all corner stations, and dropping all "hot spot" tows. The workshop agreed that a reasonable option would be to accept multiple tows within a stratum, as long as the tow density was uniform across all grids within that stratum. The strata do not need to be the same between years, although the total spatial coverage needs to be as equivalent as possible for comparability. The workshop also agreed that all strata should consist of contiguous grids and that "re-tows", done for Bristol Bay red king crab, should not be included in the abundance indices for crab species other than Bristol Bay red king crab. The workshop requested that Bob report abundance estimates (and length compositions) by stock, along with a summary of survey coverage, for several 'options', with the aim to base survey inputs on a suite of tows by survey year that represent uniform sampling density across each stratum. This can be achieved by dropping additional tows, particularly "hot spot" tows, which do not conform to best statistical design. One option specifically requested for Tanner crab is the following:
  1. select one "standard haul" (HAUL_TYPE=3) per station for each station at which standard hauls were conducted,
  2. assign stations to the same standard strata for all years (Prib's MTCA, St. Matt's MTCA, rest of west 166W, east 166W),
  3. calculate average CPUE (numbers and weight) and standard errors by year over all sampled stations for each stratum,
  4. report the area in each valid stratum by year (i.e., the sum of girds associated with sampled stations),

5. provide results (average CPUE in numbers and weight) by year, stratum, and sex, as well as by year, stratum, sex, shell condition, maturity state (for females) and size (1mm bins).

- Unmeasured (“crushed”) crabs have been encountered in the survey in some years. Available biological sample information should be used to characterize these catches, preferably from the same tow. If no crabs were measured in a tow, adjacent similar tows can be used. The file of data supplied to assessment authors should clearly indicate which crab were actually measured and which ones were ‘inferred’.
- Biological information should be included in the file provided to assessment authors with sufficient information that the sum of the measured crabs, after accounting for the sampling fraction, would match the observed haul weight for the species. The information provided in the past did not always satisfy this specification: for example, when the haul was not completely sampled, there were “crushed” crabs in the tow, or the length-weight regression differed from that assumed.
- Representative maturity information for Tanner and snow crab should be collected. Mature male Tanner crabs are identified based on chela height using a functional relationship that assumes fixed proportions of mature and immature males by size category across years. This is unlikely to be correct, given annual variation in recruitment. The workshop identified that separating mature and immature males by size category was analogous to the use of an age-length key in groundfish surveys and noted that such keys are never applied across years. Consequently, the current practice, which summarises sampling over a 30 year period, is not correct. The workshop recognises that changing this method should not be implemented without careful consideration of the consequences. The workshop recommended that a future modeling workshop consider how the Tanner and snow crab models should be modified to use the information on the proportion mature by length over time.

In discussion, the workshop recommended that Bob provide Jie and Bill with revised abundance and biomass indices and length-composition data for the entire period considered in the assessments.

D.2 Bycatch data

Bob also reviewed the preparation of estimates of the bycatch of commercial crab species in the Bering Sea groundfish trawl fishery. These estimates are of varied provenance over the history of these fisheries and cannot be validated before 1991. Following that year, estimates have been assembled from a variety of sources to provide a usable series of catch estimates. Between 1991 and 2002, catch estimates were derived from a combination of industry reports and observer estimates. From 2003 onwards, crab bycatch estimates have been derived from ratio estimators based on observer data, with varying levels of spatial resolution, depending on the year. Estimates by species prepared after 2008 are available by Alaska state statistical area. The workshop emphasized that assessment authors consider the quality of the available data when selecting weights for the groundfish biomass and length composition data.

D.3 Reports for individual species

Bill Gaeuman presented how catch estimates were obtained from the directed St. Matthew Island blue king crab fishery. These estimates are derived from dockside monitoring of landings, which record the landed totals by weight and sub-sample each landing for length distribution. These landings consist entirely of legal male crabs. Estimates of discards from this fishery are derived from observer data. These observer data are summed by length bin, sex and shell state to provide estimates of the distribution of crabs entering the commercial pots. It was determined that it was not possible to weight these observer samples to represent the sampled catch (either by string or by trip) because of the lack of linkage between the observer and catch/effort databases (except for Aleutian Islands golden king crab).

Bill elaborated on how he provides estimates of discarded crab bycatch in the directed commercial fisheries operated in state waters. For the assessment authors, he uses a simple estimator derived from all observed pots and is based on the observer determination of discarded crabs. The estimate of discarded catch is the sum of all discarded males divided by the total number of observed pots multiplied by the total fleet effort (derived from the fish ticket information). A more elaborate estimation procedure is done at the end of the season (e.g., Gaeuman 2013) which takes into account differing observer sampling fractions among vessel types (catcher and catcher/processors). However, these estimates are not used in any of the crab stock assessments and are only available from 2007/08 onward. The methodology used for these estimates is documented in Appendix A of Gaueman (2013).

Jie Zheng noted that the data are assembled for Bristol Bay red king crab using the same methods as outlined by Bob and Bill.

Shareef Siddeek described how he assembles retained and discarded catch estimates for the Eastern (EAG) and Western (WAG) Aleutian Islands golden king crab fisheries. Annual retained catch estimates and dockside length frequencies are estimated by number using the same method described by Bill for the St Matthew Island blue king crab fishery. Siddeek described two methods used to estimate discarded male catch. Method 1 is the same as used by Bill for other directed commercial crab fisheries and is based on the determination of discard status by the observers. These estimates are constrained to males greater than 100 mm carapace length to conform to the assumptions of the AIGKC model, and are only available from 1995/96 because the observer discard estimates are considered unreliable before that year. Method 2 estimates the total (retained plus discarded) CPUE of crab landed from the observed length distributions. This CPUE is converted to a value for the total fleet based on the total effort for each fishery derived from the fish ticket information and is also constrained to males greater than 100 mm carapace length as in Method 1. Discarded catch is determined by subtracting the retained dockside catch estimates from the fleet total. Estimates using this method are available from 1991/92 because they are not based on the observer determination of discard status. Siddeek computed weighted relative length frequencies where the weightings were the sampled vessel’s catch, for distributing total, retained, and discard catches into size bins.

André Punt described a multi-stock (species) spatially-discrete model which is intended to explore the effects of ocean acidification on crab populations and the impact of the snow crab fishery on the Tanner crab fishery by accounting for the spatial structure of the two populations.

E. Additional Diagnostics

Paul Starr summarized the diagnostics plots used in recent assessments of rock lobster in New Zealand. The workshop recommends that assessment authors report the following additional diagnostic statistics:

- Index data: Q-Q plots of standardized residuals and standardized residuals versus time and the predicted index values (Figure 3).

- Length-composition data: Q-Q plots of standardized residuals and standardized residuals (as box-and-whisker plots) versus time, standardized residuals versus predicted proportions (Figures 4 and5).
- Observed vs predicted proportions by maturity class (Figure 6).

References

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

Francis, R.I.C.C. In press. Replacing the Multinomial in Stock Assessment Models: a First Step. *Fisheries Research* 00: 00-00.

Gaeuman, W.B. 2013. Summary of the 2012/2013 mandatory crab observer program database for the Bering Sea/Aleutian Islands commercial crab fisheries. Alaska Department of Fish and Game, Fishery Data Series No. 13-54, Anchorage. 74 p.

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

Punt, A.E. and D. Kinzey. 2009. Report of the Alaska Crab Stock Assessment Workshop http://www.npfmc.org/wp.content/PDFdocuments/resources/SAFE/Appendix_CrabWKSHPreport909.pdf

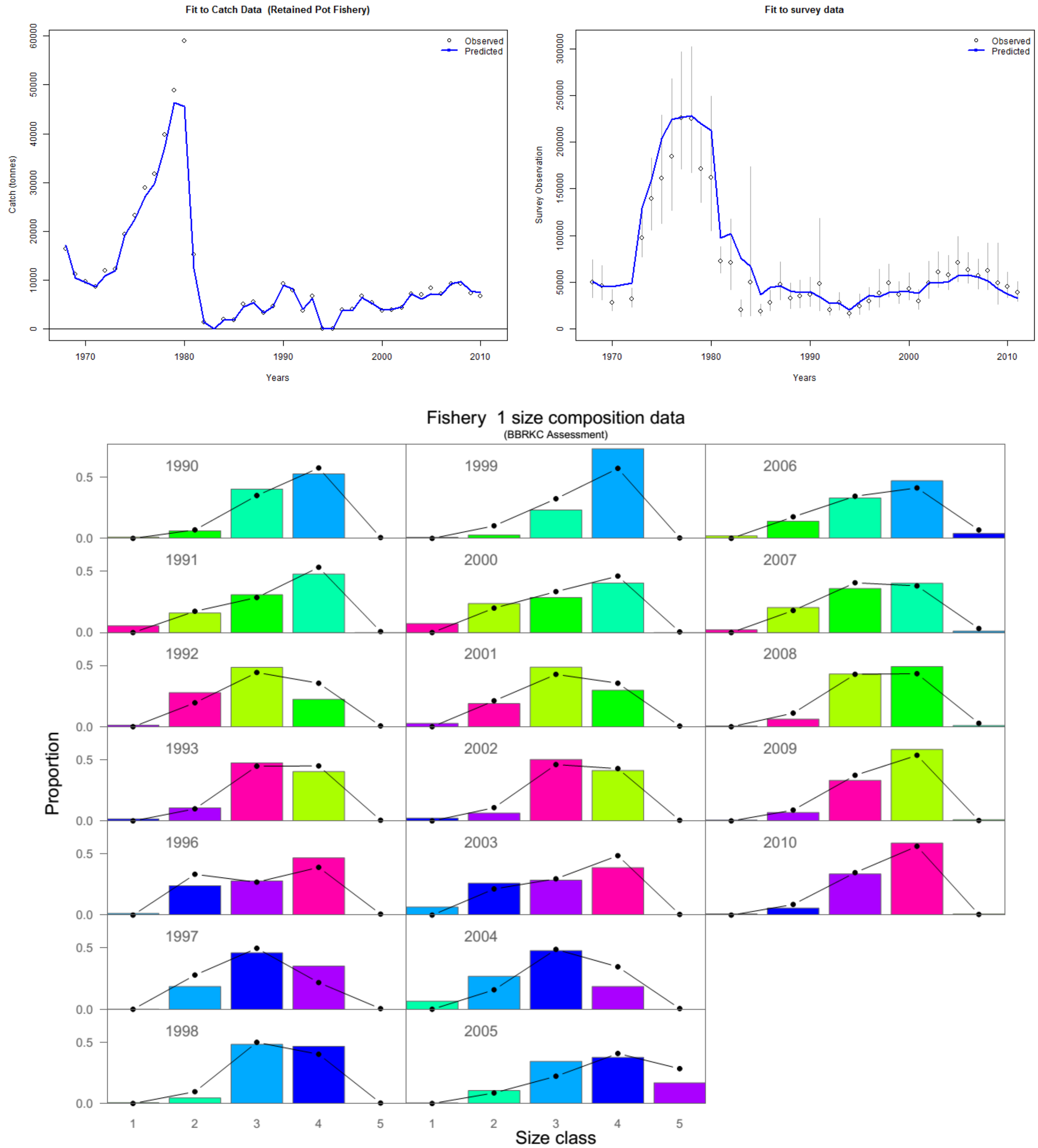

Figure 1. Fits of Gmacs to the directed fishery catches, the survey biomass indices, and the discard length composition for BBRKC.

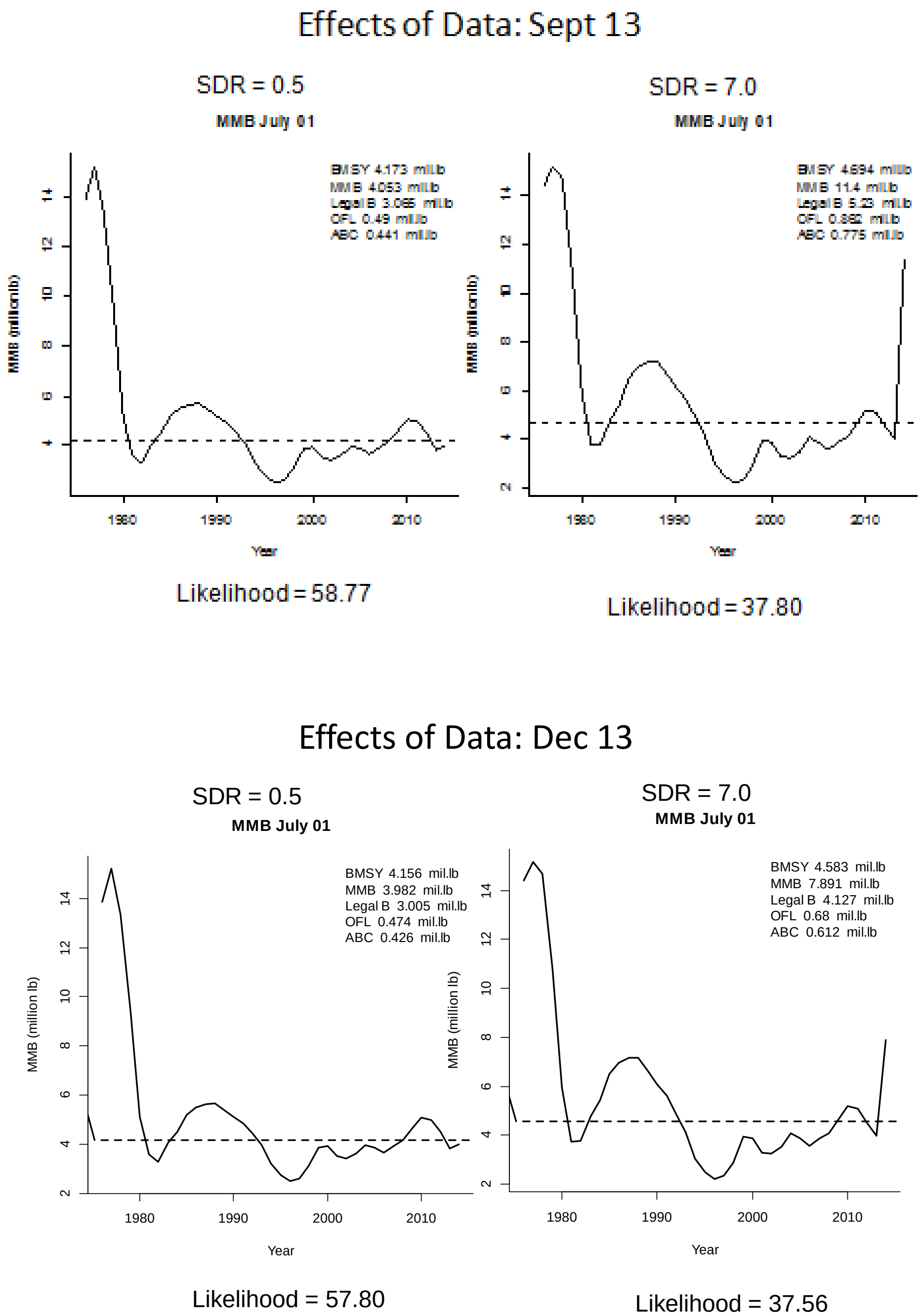

Figure 2. Time-trajectories of MMB as a function of the data set (September 2013 – upper panels; December 2013 – lower panels) and the value for σ_R (0.5 left panels and 7.0 – equivalent a weight of 0.01 – right panels)

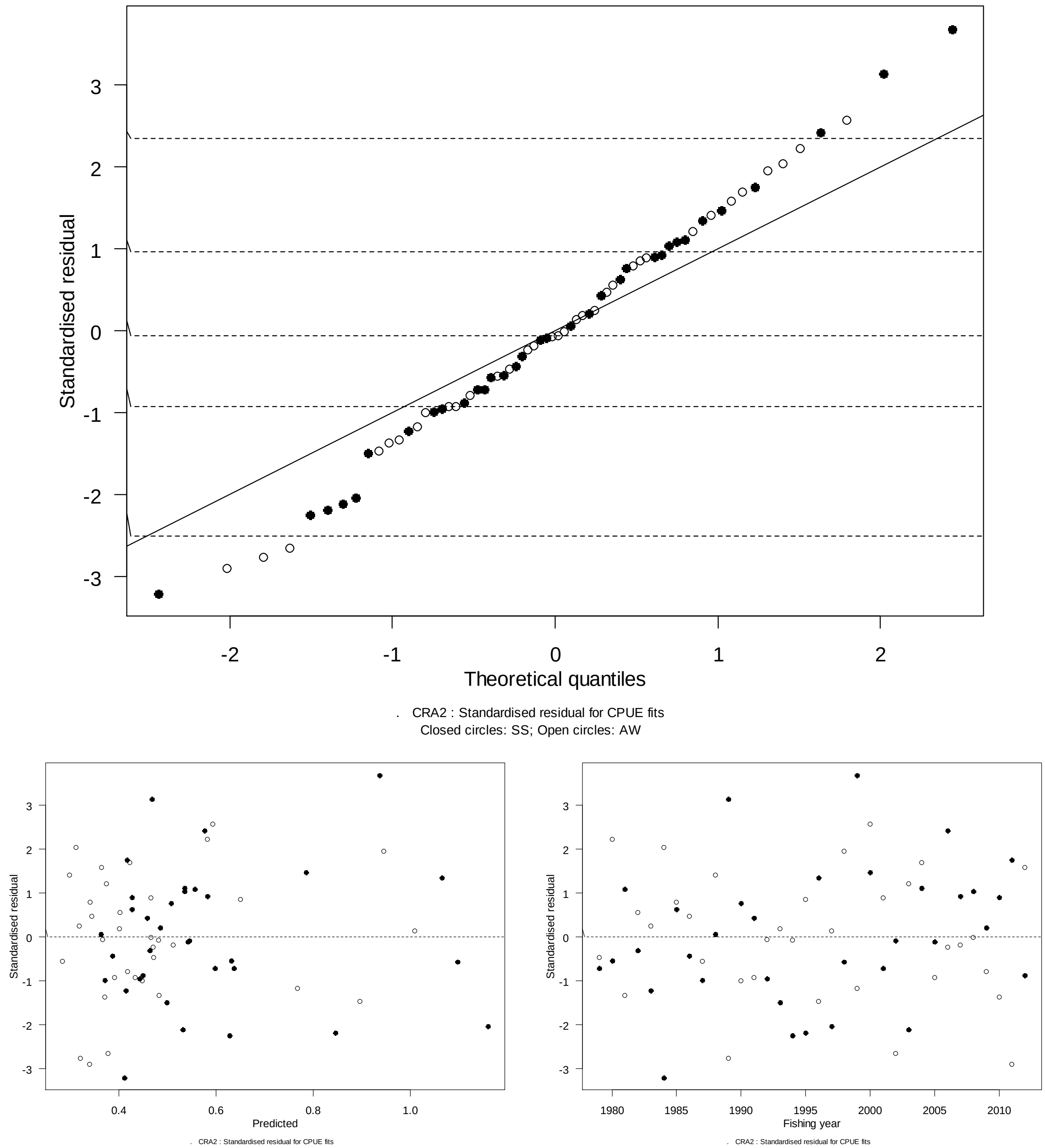

Figure 3. Examples of a Q-Q plot for the index data and plots of standardized CPUE indices vs predicted values and time.

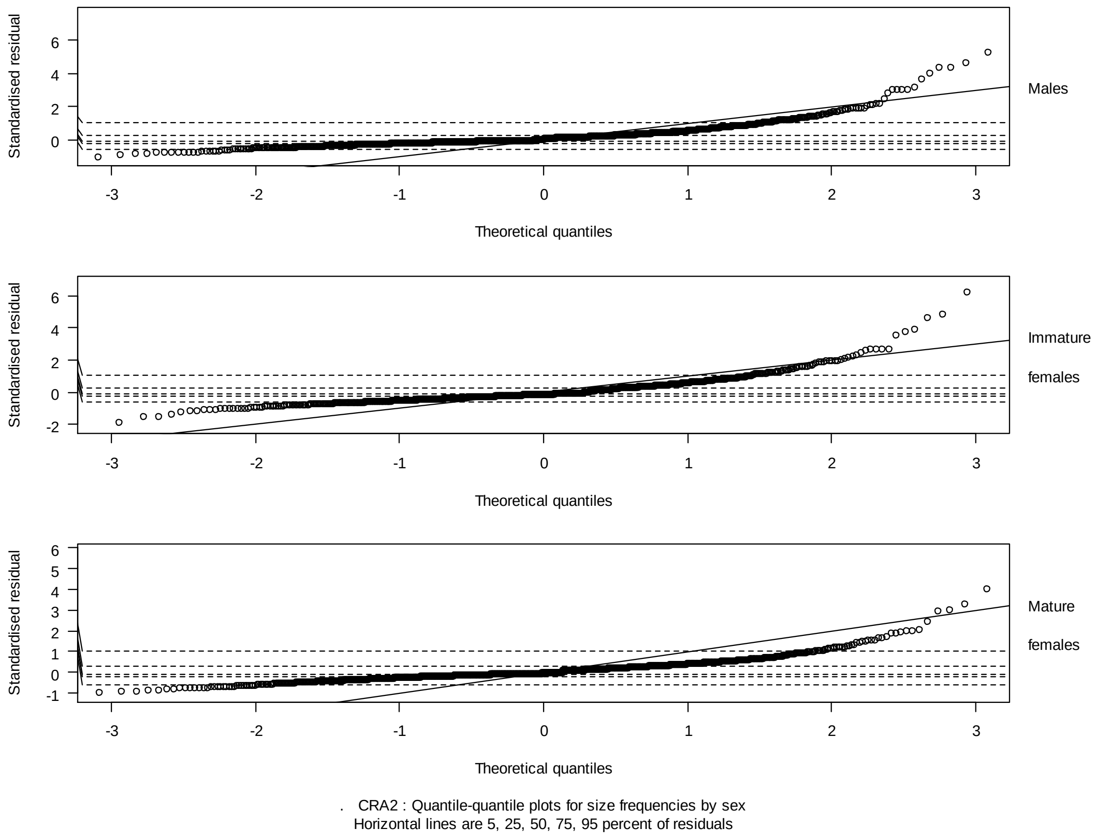

Figure 4. Examples of a Q-Q plot for the standardized length-composition residuals.

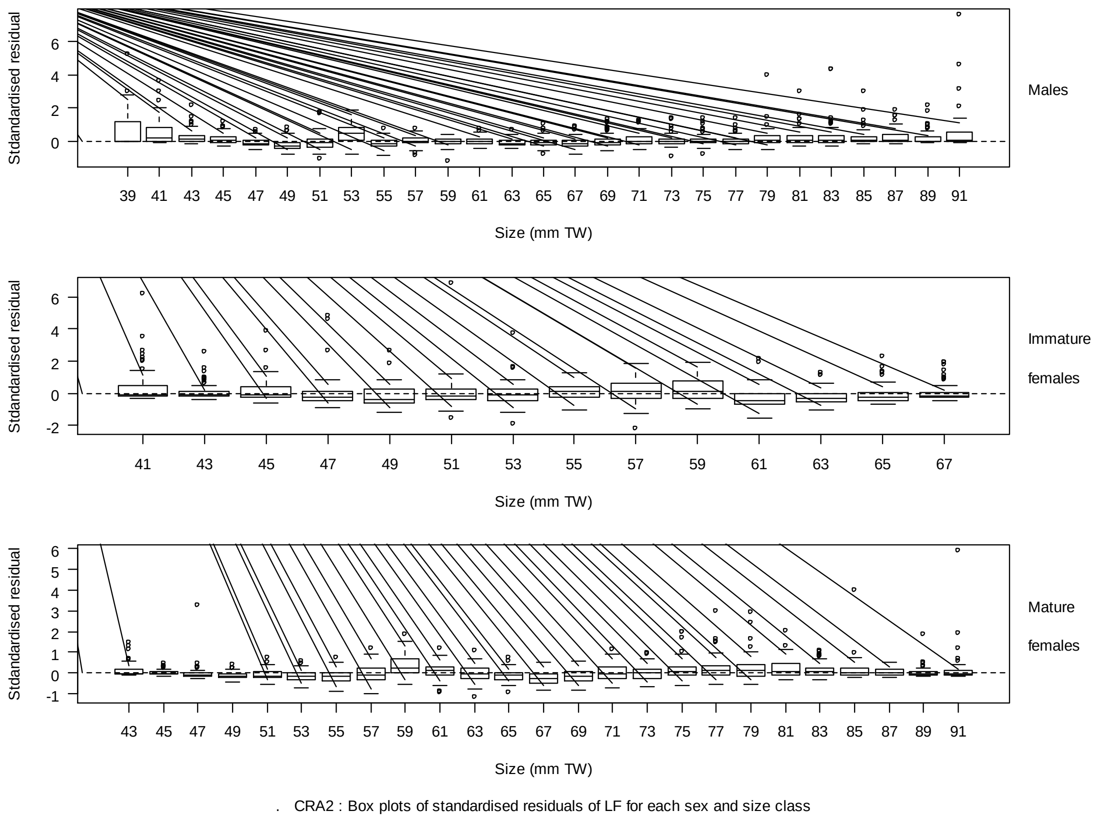

Figure 5. Box-and-whisker plots of (standardized) length-composition residuals (over years) by size-class.

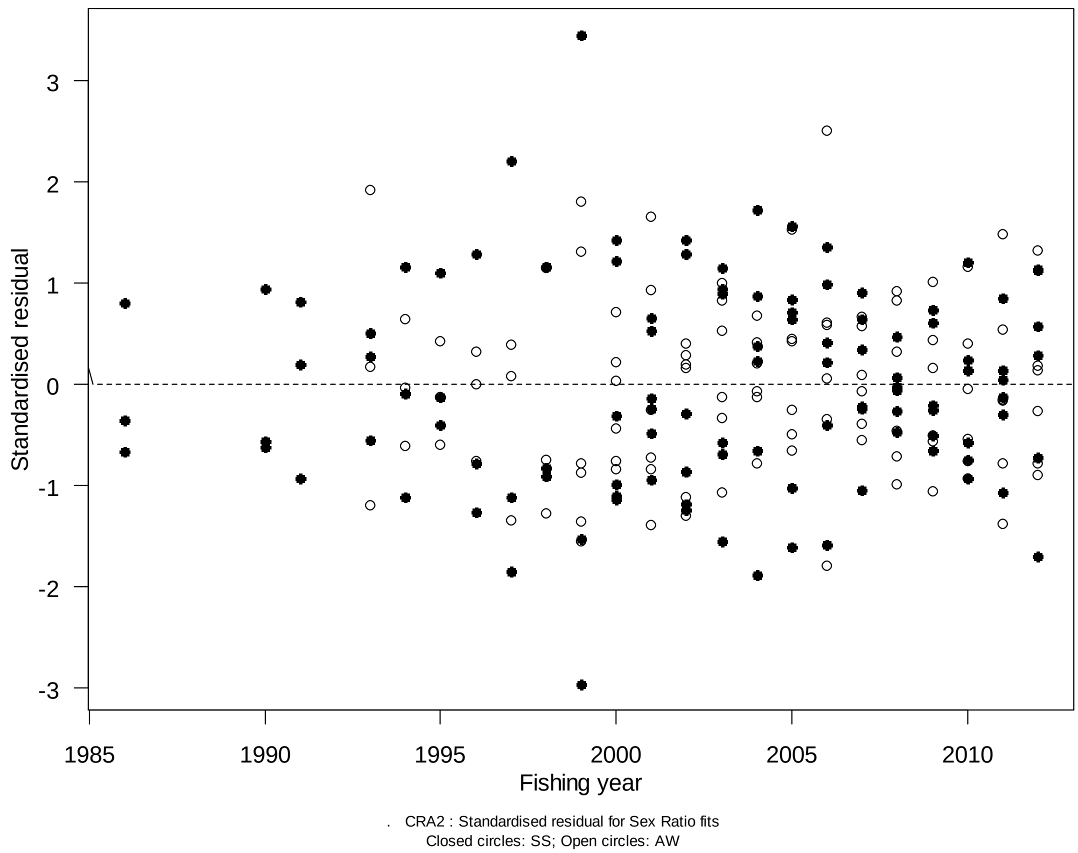

Figure 6. Standardized residuals about the proportion of the catch by maturity class.

Appendix A
Participants and Attendees

Andre Punt UW (Chair, CPT)	Paul Starr New Zealand
Diana Stram NPFMC (CPT)	Ernie Weiss (AEB)
Robert Foy AFSC Kodiak (CPT)	Jae Houng Yang
Martin Dorn AFSC Seattle (CPT)	John Hilsiger
William Stockhausen AFSC Seattle (CPT)	Chris Siddon (ADF&G)
Jack Turnock AFSC Seattle (CPT)	Jaebong Lee
Siddeek Shareef ADFG (CPT, analyst)	Dick Powell
Doug Pengilly ADFG (CPT)	Heather Fitch (ADF&G)*
Athol Whitten UW (Analyst)	Ruth Christianson*
Jie Zheng ADFG (Analyst)	
Hamachan Hamazaki ADFG (Analyst)	
Jim Ianelli AFSC Seattle (Analyst)*	
Bill Gaemann (ADF&G)	

* Participated by webex

Appendix B
Primary Documents Reviewed

Documents were made available for review prior to and during the workshop at:
http://www.afsc.noaa.gov/REFM/stocks/Plan_Team/crab/draft_assessments.htm