North Pacific Fishery Management Council Crab Modeling Workshop

February 26 – March 1, 2013
Hilton Hotel, Anchorage

Summary

For more information contact:
Diana L. Stram
NPFMC
605 West 4th Ave
Anchorage, AK 99501
(907) 271-2809
diana.stram@noaa.gov

A technical crab modeling workshop took place from February 26 – March 1, 2013, 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), members of the Scientific and Statistical Committee (SSC), 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 third NPFMC workshop to review models for Bering Sea and Aleutian Islands (BSAI) crab stocks, which are currently under development or have not been subject to a previous in-depth review. The main objectives of the workshop were to provide the assessment authors with feedback and recommendations on model development. The January 2012 workshop focused on Aleutian Islands golden king crab and eastern Bering Sea Tanner crab, while this workshop focused on Aleutian Islands golden king crab and Norton Sound red king crab.

Assessment models for Aleutian Islands golden king crab and Norton Sound red king crab were presented to the workshop. The workshop followed the following format to facilitate real-time model development: each model was considered in turn, with additional work requested for clarification or development and then shifting to the next model while the suggested new work on the first model was being completed. Model code and documentation were provided for review a week prior to the workshop. Discussions about the data, assumptions, assessment models, and interpretation of the results took place during workshop. 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 stock below. The workshop also reviewed results of an ADF&G research pot survey cooperatively conducted with the fishing industry, and received a presentation of the progress developing a generic crab model.

The key consensus conclusions and recommendations were:

General comments

1. Although considerable work was undertaken prior to and during the workshop to standardize catch and effort data to remove factors unrelated to abundance, the workshop noted that there was no guarantee that the resulting standardized indexes would be proportional to abundance. For example, it is never possible to be sure that the standardization has removed all of the effects of changes in fishing practices or that the pots sample the full range of the vulnerable population of Aleutian Islands golden king crab.
2. The stock assessment for Norton Sound red king crab included data from ADFG mark-recapture estimates of abundance, which could not be reconstructed from the primary (raw) data. Furthermore, when the NMFS trawl survey was reconstructed during the workshop, substantially different abundance indices were obtained. The workshop recommended that base-case assessments should only use indices which can be replicated from primary data, and that the underlying calculations supporting those indices should be checked before use.

3. The problems identified with model coding and convergence during the workshop highlight the need for thorough simulation testing of all assessment models. This has been done to some extent for eastern Bering Sea snow crab, but not for the stock assessment models reviewed during this workshop or any other BSAI crab stocks with approved assessments.

Aleutian Islands golden king crab

The workshop focused on two key aspects of the stock assessment: (a) the standardization of the catch and effort data, and (b) the model on which the stock assessment is based.

In relation to CPUE standardization, the workshop noted that catch and effort data from fish tickets and from observer records are available, and that both data sets can be used to create standardized CPUE series for the commercially retained catch. Only the observer data set can be used to create a standardized CPUE index series for the discarded component of the catch. The workshop examined whether the trends (in particular the large increase in standardized retained CPUE since rationalization) were attributable to an increase in abundance or can be attributed to changes in the fishery dynamics (e.g., fewer boats and longer soak times). The workshop concluded that soak time affected the catch rate per potlift and hence should be included in the CPUE standardizations. Consequently, the fish ticket data could not be used to develop a catch-rate index because soak times linked to the appropriate effort are not available from this data source. It is known that the industry is modifying its fishing practices to minimize the catch of sub-legal crab, which means that an index based on the catch of sub-legals is not likely to reflect sub-legal abundance because of a trend of reduced vulnerability in this component of the catch, and a CPUE index for sub-legal crab consequently should not be included in the base-case assessment. The workshop recommended that, for the CPUE standardization, which is to be presented to the May 2013 CPT meeting: (a) the catch-rate indices for the pre- and post-rationalization period should be treated as separate series, and (b) soak time should be forced into the CPUE standardization in the same way that year is forced.

In relation to the stock assessment model, the workshop identified discrepancies between the model description and the code, which implemented the model (e.g., how growth was modeled and whether old and new shell crab were represented separately in the model), as well as problems with how the model was coded. The workshop noted that several of the recommendations from the 2012 model review have not yet been implemented. The workshop concluded that the model was not ready to be used as the basis for providing management advice. It provided detailed recommendations for future model development.

Norton Sound red king crab

The workshop focused on the following aspects of the stock assessment: (a) the standardization of the catch and effort data, (b) the survey data which forms a key input to the assessment, (c) the reason(s) that the model is unable to mimic the first two survey estimates of abundance in previous assessments, and (d) aspects of the structure of the model.

In relation to the catch rate data, the workshop agreed that (a) the analyses should ideally not be based on imputed data, (b) the analysis for the years 1978-92 should not include interactions, only main effects, and (c) the model for the years 1993-2012 should account for interactions between year and other factors.

In relation to the assessment model, the workshop found that the survey abundance estimates for the NMFS surveys were not based on the raw data (unlike the case for other crab assessments). However, a preliminary re-analysis of the data led to estimates which are substantially different from the current estimates. The assessment author should therefore produce model runs using the current and revised NMFS estimates. The workshop found that several model runs did not converge; this needs to be resolved before May 2013 assessment. In addition, the CPUE data were found to be overweighted.

Generic crab model

Athol Whitten introduced the Generic Crab Model project. The workshop supported this effort, noting that standardization and testing of code should improve efficiency in model development and reduce errors. The workshop suggested that a side-by-side comparison of the generic model and one or two of the current approved assessments be presented in one year's time. The Bristol Bay and Norton Sound red king crab assessment authors volunteered for this comparison. Comment from the SSC was sought to inform long-range plans for crab assessment, and the workshop noted that the next step after the tandem comparisons would be to use the generic model to conduct assessments for other crab stocks.

In closing the workshop, the Chair recognized the considerable work undertaken by the analysts (Gretchen Bishop, Bob Foy, Toshihiro Hamachan, Doug Pengilly, Shareef Siddeek), by responding to the numerous requests for additional analyses. He also noted the excellent work undertaken by the rapporteurs (Cody Szuwalski and Athol Whitten) in assembling the draft report.

A. Aleutian Islands golden king crab

A.1 ADF&G pot survey

Doug Pengilly presented results from a joint government/industry research project in which research pots (i.e., pots with smaller mesh sizes and no escape panels/rings, designed to retain all crab that enter) were placed within strings of commercial pots (which have larger mesh and escape panels) as they were fished by a single commercial vessel over two fishing trips. The CPUE and length frequencies of crab captured by the research pots were compared to the adjacent commercial pots with the objectives of: (1) comparing the relative size classes of crabs captured by the two pot designs; and (2) evaluating the capacity of cooperative research to gather data useful to management. It is known that length frequencies have been shifting toward higher length classes in the commercial fisheries since the mid-1990s. Several hypotheses, some of which were in conflict, could explain this observation: either recruitment had dropped or fishers were setting their gear in such a way that minimized the catch of smaller fish (either through changing pot design or pot location). The objective of this project was to determine if smaller crab were available in the locations that were being commercially fished, thus reducing the likelihood of the recruitment failure hypothesis. It was shown that length frequencies from the research pots were similar to many of the summarized length frequencies from the pre-rationalization fishery. Doug and industry members noted that it is clear that sub-legal animals exist in the area where the fishery is prosecuted.

The workshop discussed how the data gathered from the research pots interspersed with commercial pots could be used in the assessment and to inform management. The consensus of the workshop was that these data could: (1) be used to estimate gear selectivity; and (2) form the basis for an index of abundance which covered a broader range of sizes than the CPUE for the retained catch. However, in relation to (1), such a selectivity function would only measure the retention by the gear of the crabs available at the time of fishing (selectivity), but would not include availability, which relates to that part of the population not available to the pots. The latter information is required in addition to the selectivity information to describe total vulnerability in an assessment model (see the snow crab BSFRF studies for an example). In relation to (2), data collected from research pots interspersed with commercial pots would need to be standardized in the same way commercial CPUE data are standardized if they were to be used to develop an index of abundance. The workshop agreed that joint cooperative research efforts were promising, but required planning, both in terms of experimental design and how the information gathered would be used.

A.2 CPUE standardization

Siddeek Shareef presented a range of standardized CPUE indices for AIGKC. The workshop requested a more in-depth description of the data and how the data were ‘groomed’. Industry representatives noted that observers collect data at the level of individual potlifts, which contain more associated information, but do not monitor every potlift made by the fleet, while fishery ticket data comprise all the catch and potlifts made by the fleet, but only in a form which is summarized by trip and statistical area. This came out in discussion with industry when responding to questions related to the reliability of the estimates of “week of catch” and “statistical area of catch” that are reported on fish tickets. There was discussion in the workshop about the protocol used by observers to differentiate between “legal”, “sub-legal”, “retained” and “discarded” crabs. There was concern that the observers would diverge from the equivalent definitions used by industry when sorting the pots for these categories. Doug Pengilly affirmed that the observers used the same definition of “legal” crab (based on carapace width) as does the industry, while the observers also measured crab size as carapace length, which is how “retained” and “discarded” were defined in the analyses. There was also discussion about the nature of the vessels included in the core data set compared to the vessels which were excluded from the standardization analysis. However, that discussion was rendered moot when it was determined to be an issue of data representation in plots rather than a real problem with the analysis.

The workshop noted the lack of soak time data in the fish ticket data and the increasing trend in soak time observed in the observer data. The workshop particularly noted the strong increase in total soak time that was accompanied by a rapid drop in vessel participation after rationalization in 1995/96. The workshop concluded that this increase would be interpreted as an increase in abundance when the fish ticket CPUE data were standardized without the benefit of soak time information. In addition, interactions between year and other variables and multi-co-linearity among variables, when there were a large number of factors, may result in bias for the year coefficients. The workshop recommended that alternative error distributions should be explored for the potlift data because these data are discrete and not continuous as assumed by the adopted lognormal distribution.

CPUE analyses based on the fish ticket data presented to the workshop explained substantially more variation than those presented at earlier workshops. Siddeek noted that this was because the data had been "groomed", removing spurious catch records.

Requests from 26 February 2013

A. Request: Plot the normal scores for the raw annual CPUEs from the fish tickets and the observer data for vessels that span the period over which rationalization took place; compute Pearson's correlation.
Rationale: Although there are more fish ticket data, data on, for example, soak time are not available for the fish ticket data. The workshop wished to understand whether this might lead to problems with the use of the fish ticket data as the basis for a CPUE index.
Response: CPUEs (catch divided by number of pots) by trip from the observer data are very similar to CPUEs by trip from fish tickets (r^2 > 0.9 for all presented examples: Figures 1-3). Observers on catcher-processors do not sample as many pots as on catcher vessels; consequently the data for catcher-processors are more noisy (r^2 = 0.9). Request J was generated.

B. Request: Conduct a CPUE standardization in which the covariates "EastVessel" and "WestVessel" are the number of vessels than operated in the fishery (by region) during the year and not the number in the core set of vessels. See if it possible to construct a covariate to replace these covariates which, for each record, are based on the number of real days of fishing for the specific week, statistical area and gear.
Rationale: The aim of this covariate is to quantify the "competition" of effort. However, competition would occur among all vessels (not just those in the core set)
Response: There was general confusion as to what the presented plots actually meant. A GLM was not produced using the revised data, and another model formulation was requested (request K). Influence plots were requested to address this issue, and a table of coefficients in normal space for the model was needed (request L).

C. Request: Plot the distribution of soak times (without truncating the top 5% of values) and repeat the CPUE standardization including records for which soak time is in the top 5%.
Rationale: The distribution for soak time appears to be bimodal
Response: The plot was produced for western Aleutian golden king crab (WAG) (Figure 4), and it has very long tails. It is not clear what functional form (e.g., linear, asympototic, dome-shaped) should be assumed for the relationship between CPUE and soak time. The tail for eastern Aleutian golden king crab (EAG) is not as long as for WAG, but displays a similar shift to higher soak times after rationalization (Figure 4 from Part II Model Assessment on the website).

D. Request: Explain the lack of perfect correlation between week and month.
Rationale: One would expect these two covariates to be perfectly correlated
Response: There are errors in the estimates of week. However, the extent to which these errors are pervasive is not known, and the aberrant data were not excluded from the analysis.

E. Request: Identify which types of data points appear in the tails of the q-q plots for the fish ticket data.
Rationale: These data points may reflect outliers.
Response: This request could not be addressed in the time available during the workshop.

F. Request: Plot the proportion of zero records in the observer data by year and the annual probability of zero catches from the binomial component of the model.
Rationale: The document provided to the workshop included considerable information on the non-zero catches, but very little information on the zero catches.
Response: The trend over time in the proportion of zero catches appears to be linear and decreasing (Figures 5-6). The lack of an obvious breakpoint at rationalization was surprising, and could imply an increase in abundance and the impact of changes to the number of vessels. This result prompted request K.

G. Request: Repeat the analyses of the observer data using an alternative distribution (such as the negative binomial).
Rationale: Plots such as Figure 14 (Part I: CPUE standardization) suggest that some of the catches may only be one crab; the current assumption of a log-normal distribution gives equal weight to catches of 1 and 100+.
Response: This request could not be addressed in the time available during the workshop.

H. Request: Repeat the CPUE standardization including a year:month interaction
Rationale: The workshop was interested to see whether there are different trends by month given the changes in the timing of the fishery.
Response: Year:month interactions were not selected during the stepwise GLM.

I. Request: Plot a histogram of soak times every five years by region.
Rationale: The workshop wished to understand how soak times had changed over time.
Response: This plot (Figure 7) further strengthened the hypothesis that a large change in soak time occurred post-rationalization.

Requests from February 27, 2013

J. Request: Plot the observer catch per 'true' unit effort, where effort is defined by both pots lifted and time spent soaking. Compare these trends to the CPUE from the fish ticket data.
Rationale: "Effort" is currently poorly defined, and the proportion of zeros decreases over time. (This leads to request K).
Response: Catch per pot lift increased after rationalization, but so does soak time (Figure 8). Consequently, soak time and year are confounded, and this makes standardization across pre- and post-rationalization periods very difficult. The workshop concluded that the CPUE time series should be broken at rationalization and fit separately in the assessment. Only observer data should be used because it has associated soak time data. The trend in the relationship between soak time and catch should be addressed in the pre-rationalization period, but the post-rationalization relationship between soak time and catch appears to be primarily on the part of the curve for which increased soak time does not result in higher catch.

K. Request: Predict catch using soak time and pot lifts (and other covariates) as predictors rather than predicting CPUE.
Rationale: It is unclear which covariates reflect effort, and much of the variance introduced by changes in soak time may be captured by the year effect.
Response: This request could not be addressed in the time available during the workshop.

L. Request: Produce 'influence plots' for each standardization model and tables of estimated coefficients.
Rationale: The workshop wished to understand how different variables influence the model predictions, and needed coefficients in normal space to make sense of them.

Response: The desired influence plots were not produced. Tables of estimated coefficients support the results from request J in that the effects of soak time are confounded with 'year' in the standardization.

M. Request: Produce a plot of the predicted proportion of zeros vs. the observed proportions of zeros for both WAG and EAG. (This is the denominator of equation 6, not only 'y_{binom}'.)
Rationale: The produced plot (from request F) was not easy to interpret.
Response: This request could not be addressed in the time available during the workshop.

Requests from February 28, 2013

N. Request: Produce standardization models for CPUE from observer data that: 1) are forced to include soak time; and 2) include both soak time and year for only the pre-rationalization period.
Rationale: The influence of soak time and year are confounded, and this exercise should demonstrate the respective influences.
Response: Forcing the models to incorporate soak time changes the standardized CPUE indices (Figure 9).

Major recommendations and conclusions

Only the observer CPUE data should be standardized and used in assessments because these data contain associated soak time information. Observer CPUE should be considered to be two time series, with rationalization marking the break point, and soak time forced into the model without considering its significance.

Future work

For the May CPT meeting, provide a revised CPUE standardization which:

- analyzes the observer data pre- and post-rationalization separately for both the EAG and WAG;
- forces soak time as a covariate in all models (binomial and log-normal);
- uses standard model selection diagnostics (e.g., improvement in R^2) to select other covariates; and
- includes all diagnostics (including influence plots) in the document for both the binomial and log-normal components of the analysis;
- calculates a CPUE index which combines the lognormal and binomial series; and
- show the relationship between all lognormal, binomial and combined series.

The longer-term tasks related to CPUE standardization are:

- consider a model for the observer CPUE where soak time and depth are treated as continuous covariates;
- identify the data points which appear in the tails of the q-q plots for the fish ticket data (and remove aberrant data points);
- repeat the analyses of the observer data using an alternative distribution (such as the negative binomial); and
- construct a competition covariate which, which, for each record, is based on the total number of days of fishing or the number of pots set in a day for the week and statistical area .

A.3 Stock assessment model

Siddeek Shareef presented the current draft assessment model for AIGKC, with a focus on the model structure, rather than results. This model has not yet been approved for providing management advice and the stock currently remains in Tier 5, which stipulates that the calculation of OFLs will be based upon average catch calculations. The treatment of the EAG and WAG as separate stocks was questioned, given the similar trends in CPUE, and the potential utility of combining the data into a single assessment. The separation was based primarily on the footprint of the fishery, which leaves a large unfished region between the two areas. Discussion with industry revealed that this area of no fishing between the WAG and EAG existed because the density of crabs was not at a commercial level, but that there was a continuous distribution of crabs between the two areas, albeit at lower density. Natural mortality (*M*) was specified as $0.18yr^{-1}$ in the model, based on estimates of *M* for Bristol Bay red king crab. However, golden king crab are a deeper water species and may be longer-lived.

Many issues were identified when discussing the model structure line by line (Appendix A of the stock assessment document). First, some documentation did not match the code (e.g., how growth was modeled and whether old and new shell crab were represented separately in the model). Growth increments and molting probability, estimated outside of the model, were only informed by tagging data, and would be confounded in the current configuration of the model. In response to a question whether it would be preferable to analyze the tagging data within or outside of the population model, the workshop noted that it was common practice to analyze all data simultaneously in assessments for lobster in Australia and New Zealand and for prawn fisheries in Australia. This is done because (a) the range of sizes tagged and recaptured can be quite small compared to lengths of animals caught in the fishery; (b) the fishery and survey size-composition data also provide information on growth, and (c) Bayesian formulations of these models correctly incorporate the uncertainty in the growth sub-model.

The proportion of legal crab which are expected to be retained, but are actually discarded (parameter QQ; Eq. 7) was only estimated for a sub-set of years without providing justification for the choice of years. Moreover, the assumed proportion of legal crab discarded for years for which QQ is not estimated and for which there were no discard data (pre-1990) was left at 1. The workshop recommended that (a) QQ be estimated for all years after 1989 (if the estimate should be 1, it will be estimated to be 1); and (b) QQ for the years before 1990 be set to the average after 1989.

Weighting for likelihood components should be performed in ‘sigma-space’ for interpretability. Odd model behavior (e.g., CPUE goes negative because retained CPUE is higher than total CPUE) arises and “if” statements have been (inappropriately) used to address this problem. Inconsistent likelihoods (e.g., eq 9 in Part II stock assessment documentation) may be influencing the output of the model. Steve Martel provided a document (Appendix C) outlining potential flaws when coding models using AD Model Builder, identifying the consequences of such flaws and providing best-practice coding solutions to address the flaws in the code.

Requests from February 27, 2013

Request: Include figures summarizing the tagging data and diagnostic plots related to the fits of the model to the tagging data
Rationale: It is difficult for reviewers to assess the model without this information.
Response: These were not produced because the work required on the model would change the answers.

Major recommendations and conclusions

The model is not ready for adoption in its current form and there is a significant amount of work that needs to be done before adoption, including implementation of the standards outlined in Appendix C. Many of the presented diagnostics indicate the model does not fit the data well. Recommendations from the last workshop (e.g., adopting a more generalized modeling framework, reporting standard deviations of normalized residuals, and using a non-robust formulation of the multinomial during the early phases o of the estimation) have not been implemented, and must be addressed before adoption of the model. Ideally, the EAG and WAG assessments should be based on a single piece of code (i.e., one TPL file) rather than the two presently being used as this will reduce the chances for coding errors and ensure that both areas are being treated equivalently. Differences between areas can be controlled through the data input files. The base model for future assessments should be fit to the tagging data (to estimate the size-transition matrix) and the assessment for the WAG should use the EAG tagging data, given the lack of growth information in the WAG. Given the concerns with the model, AIGKC should remain in tier 5 for this assessment round.

Future work (before the model is next reviewed)

- Remove the "old shell" component from the model.
- Recode the model so that CPUE is never negative and so that there are no "if" statements which involve differentiable variables.
- Reformulate the transition matrix to avoid modeling molting probability unless this is supported by residual diagnostics (this is to reduce potential confounding among parameters).
- Explore the sensitivity of the results to the value assumed for natural mortality (in particular lower values).
- Ensure the documentation matches the code and use names for estimated parameters in tables that describe what they actually are.
- If QQ is not estimated for all years, show the data on which the choice of years for which QQ is estimated is based. Use the expectation for the proportion of legal crab discarded based on the years for which data are available for the years in which there are no data (rather than using 1).
- Remove penalties on fishing mortality during the final phase of the estimation—they introduce biases and can exert strong influence on the outcome of the assessment.
- Always present fits to length comps as bubble plots.
- Address issues with effective sample sizes and check that residuals are as random as possible.
- Explore different selectivity curves pre- and post- rationalization to respond to the shift in length frequencies.

B. Norton Sound red king crab

B.1 CPUE Standardization

Gretchen Bishop provided an overview of the methods used to standardize the summer fishery CPUE data for NSRKC. CPUE had not previously been standardized. This analysis was undertaken given concerns that raw CPUE may not be linearly related to abundance. Gretchen noted there were large management changes in the fishery in 1993 that are likely to affect CPUE standardization attempts if not properly considered.

Data preparation involved four steps: grouping of statistical areas, combining of data from fish tickets with those from the Commercial Fisheries Entry Commission (CFEC), constructing a continuous variable to represent the difference between landing date and start of season, and finally, using filtering criteria to subset the data to obtain a more balanced design. After applying stepwise model selection criteria, CPUE was modeled as a function of: Year, Modified Statistical Area (MSA), Week of the Year (WOY), a polynomial function of Length Overall (LOA), and the Permit Fishery (PF).

The Workshop discussed the use of various factors in the GLM to explain changes in CPUE, including 'day of season' as opposed to actual day of the year. Gretchen explained that the day of season factor was intended to be independent of actual time, which might capture, for example, depletion during the season, given the season does not start on the same date each year. The workshop also discussed the assumption that all data come from a single continuous stock, and the use of fish tickets to determine statistical areas.

The workshop considered issues with having vessel as a factor in the GLMs because vessel might be a proxy for other factors such as MSA. The discussion of these and other related concerns led to several requests, mostly to revise GLM analyses using alternative explanatory factors.

Requests from 26 February 2013

A. Request: Provide a map that shows the assumed area of the various red king crab stocks near Norton Sound
Rationale: Large catches of red king crab have historically been taken outside of the area now considered to be the core of the stock. It is possible that these earlier catches may have come from a different stock.

Response: The workshop examined plots of survey catch rates (1976 onwards). It was noted that red king crab can be found outside of Norton Sound, but the catch rates are low. Industry representatives noted that red crab are found in inshore waters north of Norton Sound, but these crab aggregations are not fished.

B. Request: Provide the means and standard errors for the coefficients related to the main factors in the GLM, and plot the effect of LOA against LOA.
Rationale: The workshop wished to understand the size of the various factors, and was concerned that the cubic function applied to the LOA covariate might exhibit anomalous behavior.
Response: Gretchen noted that the means and standard error estimates for the coefficients in the GLM are provided in Appendix C of the draft stock assessment report. Requests I and N are follow-up requests.

C. Request: Plot the CPUE indices from the GLM for the full time series, that for 1978-1992, and that for 1993 onwards. Scale the series for 1978-1992 and 1993-2012 so that their means match the means for the CPUE indices for the full time series.
Rationale: The plots in the document suggest that the trends in standardized CPUE for the analyses based on subsets of the data exhibit different trends than those from the GLM based on the full time-series, perhaps because of an unbalanced design.
Response: The patterns for 1993 onwards were very similar between the analyses based on data for 1992-2012 and the full data set (Figure 10). However, the estimate for 1978 from the full (1978-2012) model is lower than that from 1978-93 model, while the estimates for 1979-82 from the full model are higher than those from the 1979-92 model.

D. Request: Extend tables 5, 6 and 7 by adding consideration of covariates which are interactions of factors; consider interactions involving variables not indicated as main effects in the model.
Rationale: Gretchen mentioned that inclusion of some interactions led to a large increase in R^2, but that interactions were ignored in the final model selection.
Response: The tables presented (Table 1) indicated that there were many interactions which involved year. The number of data points was low for the 1978-92 model. Even so, inclusion of the interaction between Year & MSA and Year & Week improved the R^2. The 1993-2012 analysis was based on far more data points, and the interaction between Year & Vessel was identified as important. However, it is possible this interaction is mimicking other effects, such as MSA. This issue is explored further in requests F, G and H.

E. Request: Repeat the GLM ignoring any data points for which catch and effort were imputed.
Rationale: GLM standardizations should not be conducted using imputed data
Response: The results of this analysis were qualitatively similar to those for the analyses which used the imputed data. However, only ANOVA tables were presented. Gretchen noted that the results were less stable when the imputed data were ignored. In general, the workshop preferred not to impute missing data. Request J provides a follow-up request.

Requests from 27 February 2013

F. **Request:** (1) Repeat the analyses from request E, except start the analysis in 1996. (2) Repeat the analysis in request E for the years 1978-1992 and 1993-2012, except ignore vessel as a possible factor (and therefore not include year*vessel interactions).
Rationale: Vessel can be a surrogate for area or time, because sometimes particular vessels go to the same areas repeatedly, or are active at particular times of the year. This can confound other factors when standardizing the CPUE data.
Response: Substantially less of the variance was explained when vessel was excluded from the model and other interactions entered the model.

G. Request: Repeat Request E except replace 'week' by 'month'.
Rationale: The week*year interaction could be due to short-term effects of random variation.
Response: Month of year did enter the model, and the explained variance dropped.

H. Request: Choose the best GLM model using model selection criteria, and provide time series information for each statistical area. Plot year*(X interaction): show how much variation in the CPUE is explained, and plot the trends by year and 'X interaction'. Show significant trends in plots. Only show these results for a best model for 1978-92 and one for 1993-2012 of those considered.
Rationale: The workshop was interested to understand whether the year-interactions were perhaps proxies for other effects.
Response: This request was not completed in the time available (but see Request M).

I. Request: Plot the polynomial function of LOA.
Rationale: The coefficients in the report are not sufficient to understand the effect of LOA on CPUE.
Response: This request was not completed in the time available during the workshop.

Requests from 28 February 2013
J. Request: Provide the output from the GLMs which ignore imputed data in the form of time-series of standardized CPUE.
Rationale: The results from request E did not include plots of standardized CPUE for the analyses with and without imputed data.
Response: This request was not completed in the time available during the workshop.

K. Request: Repeat the CPUE standardizations with vessel as a main effect, but do not include vessel-interactions.
Rationale: Vessel is clearly an important factor in the GLM, but the implications of year*vessel interactions are hard to understand.
Response: This request was not completed in the time available during the workshop.

L. Request: Confirm that the apparent mislabeling of the outputs from some analyses are indeed mislabels.
Rationale: Some of the results appear to have been mislabeled.
Response: This was confirmed.

M. Request: Plot CPUE trajectories by vessel and by year for the analysis based on data for 1993-2012.
Rationale: This is an important interaction and explains a considerable amount of information, but needs to be understood.
Response: Figure 11 shows the interaction plots.

N. Request: Plot the polynomial function of LOA.
Rationale: The coefficients in the report as not sufficient to understand the effect of LOA on CPUE.
Response: This request was not completed in the time available, but code to conduct the analyses was developed (Appendix D).

O. Request: Explore whether it is possible to start the CPUE standardization in 1977.
Rationale: Fishery catch statistics are available for 1977, but it was unclear to the workshop whether raw data were available for these years.
Response: This request was not completed in the time available during the workshop.

Major recommendations and conclusions

The workshop agreed that including interactions in the model for 1978-1992 did not improve the fit sufficiently to justify their inclusion, especially given the small number of data points for these years. The interactions between year*vessel, and year*week in the model for 1993-2012 should either be treated as random efforts or filtering methods applied that might eliminate the need for these interactions. The workshop established a steering group to work with Gretchen to get the CPUE standardization finalized for Hamachan to complete the assessment. Gretchen will include André, Martin, Jack, and Buck on correspondence, and get feedback and input over the coming weeks. Paul may be able to help as well, especially with regard to the influence plots (he can help find code).

Future work

For the May CPT meeting, provide a revised CPUE standardization which:

- ignores any imputed data;
- checks whether the data for 1977 could be included in the analysis;
- analyzes the data for 1978-92 and 1993-2012 separately;
- includes only main effects for the 1978-92 model;
- addresses the year*vessel, and year*week interactions for the 1993-2012 model either by treating them as random effects or by applying additional filtering criteria that remove the need for these interactions;
- addresses the year*MSA interaction for the 1993-2012 model by combining the outer and outer*North MSA strata; and
- provides full diagnostics (including influence plots and a plot of the impact of LOA on CPUE).

B.2 Assessment model

Hamachan provided an overview of the methods used in the current stock assessment model. The panel discussed the pot survey data and raised concerns about the collection, treatment, and assumptions concerning this information. The workshop did not review the detailed mathematical specifications of the model as the assessment had been approved previously by the SSC and CPT.

Requests from 26 February 2013

A. Request: Provide a likelihood profile for survey catchability (Q); consider a range from 0.1 to 2 for survey Q; retain the output files in case the workshop wishes to see whether some parameters (e.g. those which define survey selectivity change with survey Q).
Rationale: Survey catchability is assumed to be 1 for all years and survey types. It is unclear whether other data sources in the model are in conflict with this assumption.
Response: Hamachan presented the results of a likelihood profile analysis on Q (Figure 12). The workshop was satisfied that the results revealed no major problems with the assumption Q=1, but noted that this conclusion is predicated on the data and assumptions of the analysis. Some of the analyses clearly did not converge.

B. Request: Show the fit of the model to all of the survey indices (those collected during pot and trawl surveys) on a single plot.
Rationale: The fit to the pot indices is not included in past documents.
Response: Hamachan provided the plot as requested, and it will be provided in future versions of the stock assessment document. The model fitted the pot indices well given the high CV (0.34) assumed for these data.

C. Request: Explain how the data collected during the tagging study were used to compute estimates of absolute abundance for 1985.
Rationale: The methodology for computing these abundance estimates is not available.

Response: The workshop reviewed the report of the survey for 1985 (Brannian, 1987) and a report for some of the earlier surveys (Powell *et al*., 1983). The methodology used for the 1985 surveys appears standard, but that for earlier surveys less so. See Request D for further considerations of the pot survey data.

D. Request: Explain how the survey estimates for the early surveys (which pertain to 100mm+ animals) were used to derive the survey estimates used in the model (which pertain to 74mm+ animals).
Rationale: The reported survey estimates of abundance and those used in the assessment differ markedly for some years
Response: Hamachan explained that the pot survey indices were scaled by the ratio of the legal to sublegal crab for the 1980, 1981, and 1982 surveys, but were based on the reported estimates of sub-legal crab for the 1985 survey. The approach for 1985 is more appropriate. See Requests E and I for follow-up analyses

E. Request: Conduct a sensitivity test in which the "pot survey" indices are dropped from the assessment.
Rationale: It is not clear whether these indices are comparable with the indices from the trawl surveys. Also, the basis for the assumed CV of 0.34 was unclear.
Response: Hamachan provided two plots: one showing the fit of the model to all survey indices, and another showing the fit to the data when the pot surveys are ignored. The results suggest that the results are insensitive to the pot survey data: the model with the data in and the model with the data out differ negligibly in terms of fit to the survey index. No rationale can be provided for the assumed CV of 0.34.

Requests from 27 February 2013

F. Request: Replace the current selectivity curve by a logistic curve (over all size bins) and therefore estimate two parameters (instead of fixing four parameters and estimating two). Provide comparative results.
Rationale: The current approach (fixing selectivity for the last four size-classes to 1 and assuming a logistic selectivity pattern for the first two size-classes) is unusual for crab assessments and the results (all crab are fully selected) unexpected given the results for other crab stocks.
Response: The selectivity pattern for the trawl survey was still flat, but there were differences in selectivity for some other fleets.

G. Request: Provide results for a model with a fixed value of Q=1 for NMFS surveys and an estimated Q value for the ADFG surveys.
Rationale: The data for the various surveys are not inconsistent with the assumption Q=1 for all surveys. However, the gear used in the surveys (and the areas covered during the surveys differ between the survey types)
Response: The fits to the early survey estimates were essentially unchanged, but the Q for the ADFG surveys was 1.5. Information was not available to assess how the estimate of Q for the ADFG surveys impacted the fits to the other data.

H. Request: (1) Define a common survey grid for each of the NMFS and ADFG surveys: standardize the areas, to provide comparable indices. (2) Provide a summary paragraph describing the methodology that was applied to the pot survey data and how this has changed over time and among different agencies (NMFS and ADFG).
Rationale: Survey designs have changed over time and are different, at different times, between agencies (NMFS and ADFG).
Response: Bob recomputed the survey estimates for 1976, 79, 82, 85, 88 and 89 using the data included in the NMFS database and using a 20x20 grid. The estimates were substantially larger (3-4 times for

some years) than those used in the current assessment (Table 2; Figure 13). If correct, this changes the qualitative understanding of the dynamics of this population. See Request P.

I. Request: Provide the results of a sensitivity test in which the approach used to correct the 1980-82 pot survey estimates is also applied to the 1985 data.
Rationale: Data are available on sublegal crab for 1985, but not for 1980-82. Applying the method used to estimate sub-legal crab to the 1980-82 data to the data for 1985 will help to understand the behaviour of the method applied to the 1980-82 data.
Response: The estimate of abundance for 1985 was 1960.5 compared to 2320.4. The workshop noted its general discomfort with using survey estimates which cannot be reproduced from the raw data.

J. Request: Provide the results of a likelihood profile analysis on the initial abundance (InitPop). Do this for Q = 1.
Rationale: The Workshop would like to understand which information is driving the estimate of InitPop, other than catch.
Response: The profiles were provided, but there was clear evidence for non-convergence for several values for InitPop. Also, the size of the contribution of the CPUE data to the objective was surprisingly large (larger in some cases than the contribution of the trawl survey data to the objective function). This was subsequently confirmed to be due to the way the CVs for the CPUE data were set.

K. Request: Provide the results of a likelihood profile analysis on the initial abundance (InitPop) as above, but with Q=0.5.
Rationale: The fits to earlier abundance indices may be better if Q=0.5.
Response: See above.

L. Request: Provide the results of a sensitivity test in which the initial size-structure is estimated.
Rationale: The initial size-structure was set to the size-structure from the 1976 trawl survey in the model presented to the workshop, but the 1976 trawl survey data are subject to observation error.
Response: The results were essentially unchanged by estimating the initial size-structure.

M. Request: Provide the results of a sensitivity test in which the CVs on the first two survey estimates of abundance are equal to 0.01. Do this for survey Q=1, and then repeat for estimated survey Q.
Rationale: This test should help the workshop to understand what is forcing the model not to fit these earlier data.
Response: The model was able to fit the 1976 abundance estimate almost perfectly, but the fit to the 1979 estimate was not exact. The workshop agreed this sensitivity test confirmed that the inability to the fit the 1976 data point was likely not because of the high catches. The results for Q estimated should have been better than those for Q=1, but this was not the case. This is indicative of a convergence problem. The workshop noted that using a “PIN” file could help address this issue. See request P.

N. Request: Provide the results of a sensitivity test by dropping the CPUE data altogether. Just fit to the survey data.
Rationale: The workshop wished to understand the relative impact of the CPUE and survey data on the results of the analyses
Response: Ignoring the CPUE data led to a better fit to the 1976 abundance estimate, but poorer fits to the abundance data for 1979-1982. The reason for this was unclear.

Requests from 28 February 2013

O. Request: Repeat the base case analysis, but halve the weight on the size data

Rationale: The workshop wished to understand the impact of the weight assigned to the size data on the ability to fit the survey data. The workshop looked at the plots of input and effective sample sizes and these appeared adequate
Response: The fits to the abundance data were better

P. Request: Repeat request M, but estimate Q for the ADFG surveys in the last phase.
Rationale: There appeared to be convergence problems with the earlier analyses.
Response: There was still evidence for convergence problems

Q. Request: Repeat the requests presented to the workshop today when there is an "additional variance" term in the CPUE likelihood and an "offset" is removed the contributions by the size data to the objective function. Show the likelihood component for each run, and have all the outputs from the model available to enable residual plots to be produced.
Rationale: The CVs assumed for the CPUE data are those from the GLM (0.175 for all years) which is too small given the actual residuals. Also, removing the offset from the size data makes it easier to interpret the sizes of the component of the overall objective function.
Response: The offset was added as requested. However, the allowance for additional variance was incorrectly coded.

R. Request: Go back to the original raw data for the NMFS surveys and check the data against the information in the database. Also, recompute abundance estimates using the same grid (10x10) as for the ADFG surveys.
Rationale: There was concern that the estimates of abundance may be wrong. While the use of different gear in the NMFS and ADFG surveys will impact survey Q, the magnitude of the effect seems very large.
Response: This request could not be completed during the workshop.

Major recommendations and conclusions

The workshop noted that several of the abundance estimates included in the assessment cannot be replicated from the primary (raw) data. It was agreed that the base model for future assessments should be use only abundance estimates (and associated size information) which can be replicated. The CPUE data appear to be overweighted, but attempts during the workshop to correct this were unsuccessful. Some progress was made identifying the causes for the inability to fit the early NMFS abundance data, but the impact of lack of model convergence and the inability to allow for additional variance when fitting the abundance estimate meant that no definitive conclusions could be drawn. The revised NMFS estimates of abundance for 1976-91 differ markedly from those used in current assessments, which requires checking.

Future work

For the May CPT meeting:

- The NMFS estimates of abundance should be checked. Analyses should be conducted using the current and any revised estimates. Check whether the NMFS survey gear changed during the years 1976-1991. If so, consider treating the data from each gear type as a separate index. Consider sensitivity to the grid (10x10 or 20x20) used to estimate abundance.
- Base case analyses should only include data which can be replicated from the primary source (which may mean that the 1980, 1981 and 1982 pot surveys may need to be dropped from the base-case analysis).
- The rationale for a CV of 0.34 for the pot survey estimates must be provided.
- The code should be modified to include the ability to estimate additional variation for the CPUE indices.

- Efforts should be made to check that the model has converged (e.g., by using different starting value and using a PAR file as a PIN file).
- Include a set of model configurations based on the requests by the workshop for each of (a) the old / revised NMFS survey data, (b) inclusion / exclusion of the CPUE data, and (c) setting Q=1 for all survey / setting Q=1 for the survey for Q is estimated to be highest.

Longer term:

- Integrate the tagging data into the assessment.
- Consider sensitivity to different choices for the number of size-classes.

C. Introduction to Cstar and the Generic Crab Model

Athol Whitten noted that each crab species and associated fishery is unique. Thus, the Generic Crab Modelling (GCM) project[1] aims to provide a flexible and extensible framework for development of fishery-specific crab models, rather than just a 'one size fits all solution'. Generalized modelling frameworks have been applied to many of the world's fisheries, and have proven very effective in improving model development and therefore management of a wide range of species. Modelling frameworks such as Stock Synthesis have enabled fisheries modellers to avoid coding errors, standardize model diagnostics and reporting, and to focus less on problems with models and more on data analysis.

He noted further that the GCM will be the first implementation of a library of mathematical functions under development for ADMB: The Common Stock Assessment Routine (Cstar) library[2]. The library will become available as a contributed package to ADMB[3], and therefore available as an open source repository. The library will follow strict coding standards set out by the ADMB foundation and will be documented online using the Doxygen[4] system.

He noted that collaboration is key: the Cstar and GCM projects will actively seek to collaborate with intended end users of the GCM software at the ADFG and NMFS. Through collaboration it is hoped that both projects will evolve into ongoing open source projects with multiple contributors and therefore available for use for many years to come.

In discussion, it was noted that standardized plotting, input, and output, the potential for compulsory simulation testing and version control are all benefits of a common repository of assessment code. Common banks of code are exposed to a larger 'audience' to identify coding issues. Concerns about the open source nature of the code were raised by the workshop—if anyone can edit the code, how does one know if the code is still correct? Version control is one answer, but standardized batteries of tests designed to ensure changes in the common code do not influence the performance of the code can also be developed and used. Generic code will allow more time to be spent on data issues rather than coding issues.

The workshop suggested that a side by side comparison of the generic model and the currently used assessments be conducted in one year's time. Bristol Bay and Norton Sound red king crab authors volunteered for the comparisons. There should be many models in development and compared at each evaluation of the assessments to attempt to account for model uncertainty. Comment from the SSC is sought to inform long-range plans for crab assessments, and the next step after tandem comparisons is to see how quickly the data from the other assessments can be formatted for use in the generic model.

References

Brannian, L.K. 1987. Population assessment survey for red king crab (*Paralithodes camtschatica*) in North Sound, Alaska, 1985. ADF&G Technical Data Report No. 214 (61pp).

[1] https://code.google.com/p/generic-crab-model/
[2] https://code.google.com/p/admb-cstar/
[3] http://www.admb-project.org/
[4] http://www.doxygen.org

Powell, G.C., Peterson, R. and Schwartz, L. 1983. The red king crab *Paralithodes camtschatica* (Tilesius) in Norton Sound, Alaska: History of biological research and resource utilization through 1982. Information Leaflet No. 222 (116pp).

Table 1a. Analysis of deviance for stepwise lognormal model selection for Norton Sound red king crab. The response variable is log(CPUE) in terms of numbers/pot. The explanatory variables Year, Vessel, Modified Statistical Area, Week of Year and their second-order interactions were offered to the model. Data is from vessels having three deliveries for three years over the time series 1978–1992. The forward stepwise selection process used a stopping point of R^2 difference > 0.01. Notation from the open source programming language R is used.

	Difference from null in				
Variable	**Null deviance**	**Null df**	**Residual deviance**	**Residual df**	**R^2**
Year			-25.5	149	0.51493
+Vessel	-9.953	-5	-35.4	144	0.56241
+Modified Statistical Area	-5.959	-3	-41.4	141	0.60295
+Year:Modified Statistical Area	-15.956	-8	-57.4	133	0.64738
+Week of Year	-15.957	-8	-73.3	125	0.69040
+Vessel:Week of Year	-45.913	-23	-119.2	102	0.77734
+Modified Statistical Area:Week of Year	-7.975	-4	-127.2	98	0.80238
+Year:Week of Year	-9.985	-5	-137.2	93	0.81716
- Year:Modified Statistical Area	3.992	2	-133.2	95	0.80935
+Vessel:Modified Statistical Area	-7.990	-4	-141.2	91	0.81960

Table 1b. Analysis of deviance for stepwise lognormal model selection for Norton Sound red king crab. The response variable is log(CPUE) in terms of numbers/pot. No imputations were conducted. The explanatory variables Year, Vessel, Week of Year, Modified Statistical Area and their second order interactions were offered to the model. Data is from vessels having three deliveries for three years over the time series from 1993–2012. The forward stepwise selection process used a stopping point of R^2 difference > 0.01. Notation from the open source programming language R is used.

	Difference from null in				
Variable/Formula	**Null deviance**	**Null df**	**Residual deviance**	**Residual df**	**R^2**
Year			-39.8	4255	0.23446
+Vessel	-111.9	-56	-151.6	4199	0.38301
+Year:Vessel	-641.8	-321	-793.5	3878	0.54141
+Week of Year	-30.0	-15	-823.44	3863	0.56054
+Vessel:Week of Year	-1019.89	-510	-1843.33	3353	0.66864
+Year:Week of Year	-284.0	-142	-2127.29	3211	0.71032

* Coefficients undefined because of singularities

Table 2. Current and revised survey estimates of abundance for Norton Sound red king crab.

	Current		Revised	
Year	Abundance	CV	Abundance	CV
1976	4219.294	0.163	5837.825	0.161
1979	901.000[1]	0.233	1403.364	0.214
1982	2323.379	0.256	9864.138	0.151
1985	3195.535	0.263	7142.310	0.141
1988	3035.621	0.298	10499.191	0.141
1991	3092.794	0.350	9115.345	0.184

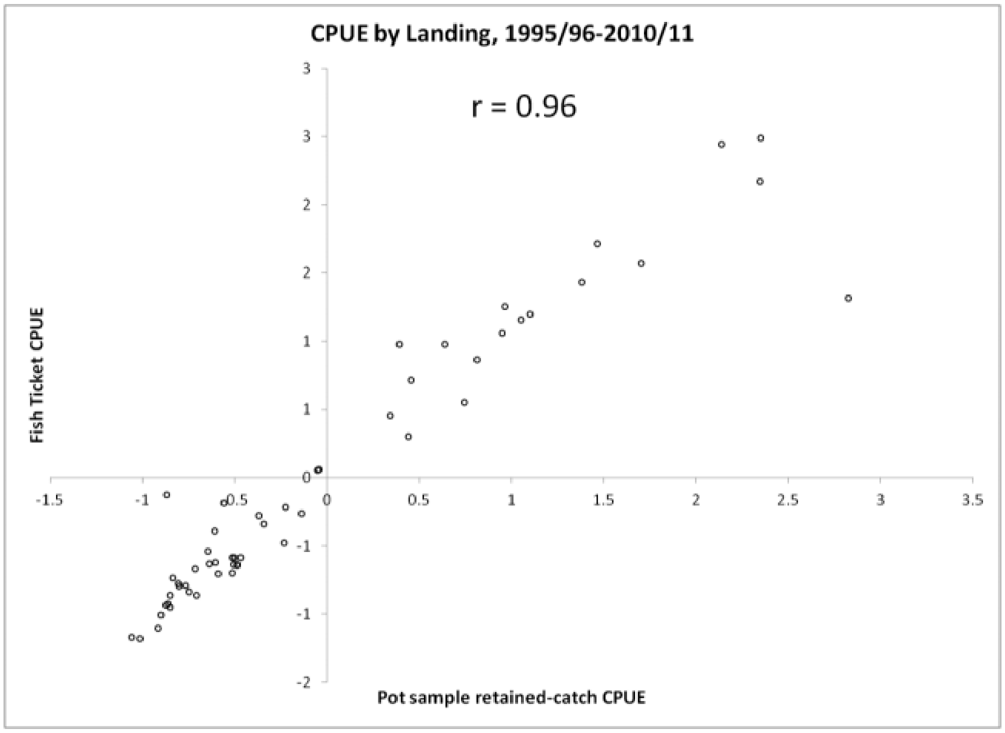

Figure 1. Fish ticket CPUE vs. observer-sampled CPUE for a single vessel.

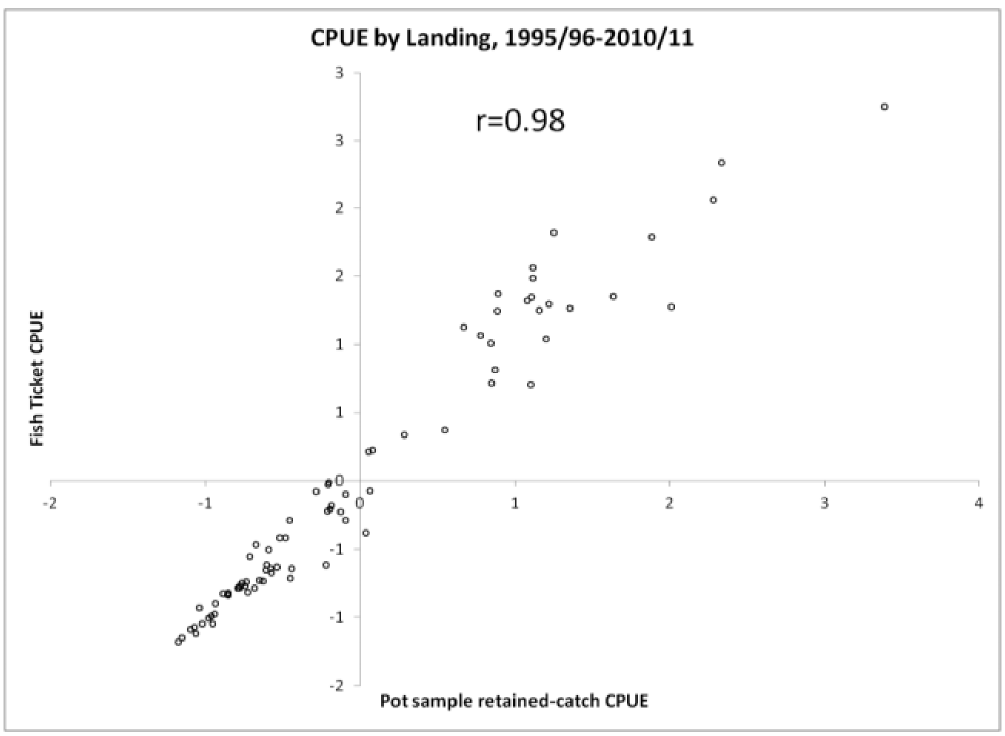

Figure 2. Fish ticket CPUE vs. observer-sampled CPUE for a single catcher vessel.

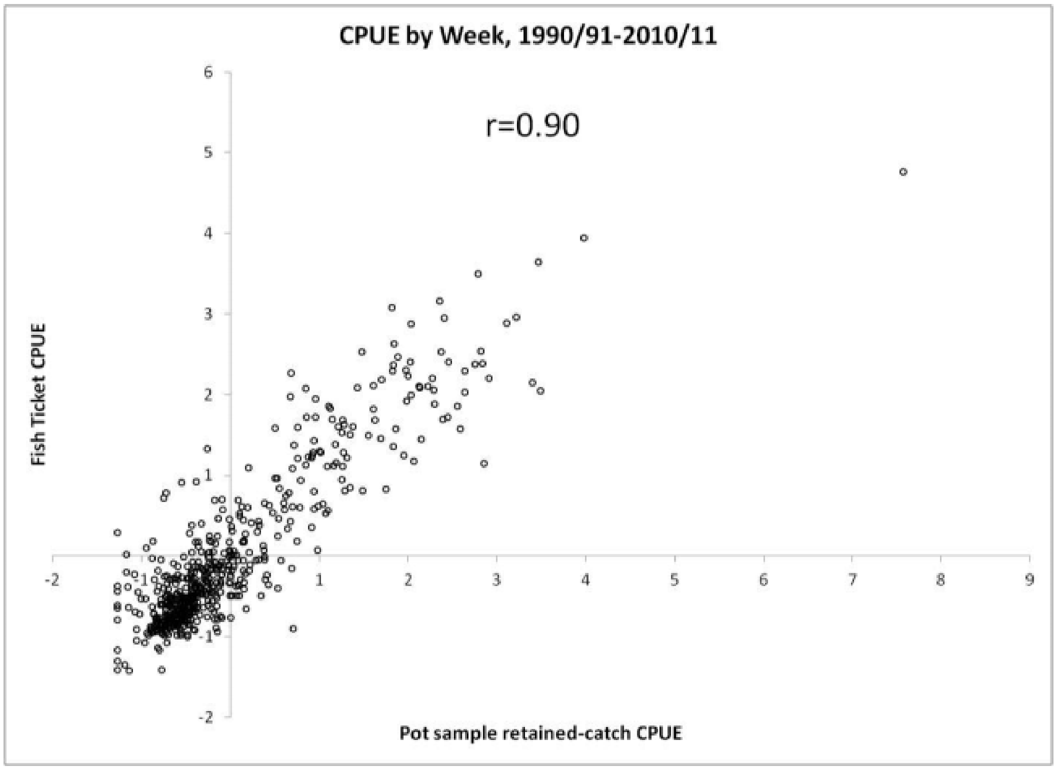

Figure 3. Fish ticket CPUE vs. observer-sampled CPUE for a single catcher-processor vessel.

1995	0.1783
1996	0.2289
1997	0.1885
1998	0.1004
1999	0.1765
2000	0.2007
2001	0.1681
2002	0.1585
2003	0.1248
2004	0.1133
2005	0.0674
2006	0.0706
2007	0.046
2008	0.0265
2009	0.0305
2010	0.0531

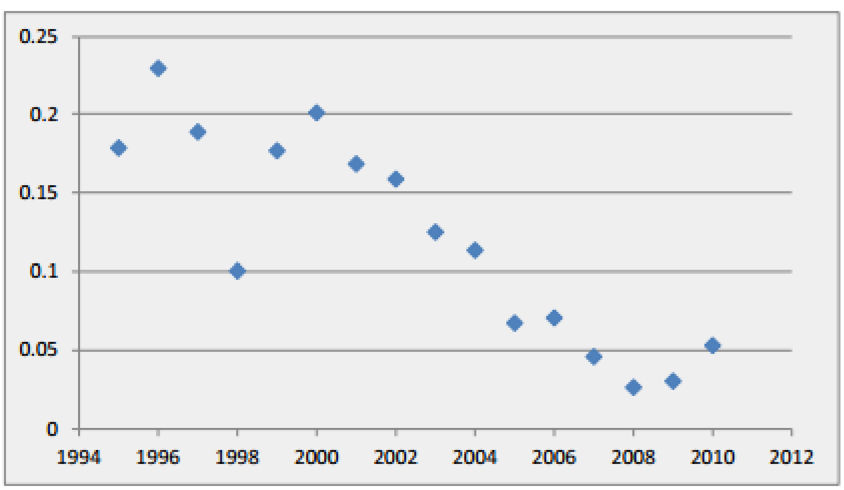

Figure 4. Proportion of zeros in catch data for the WAG.

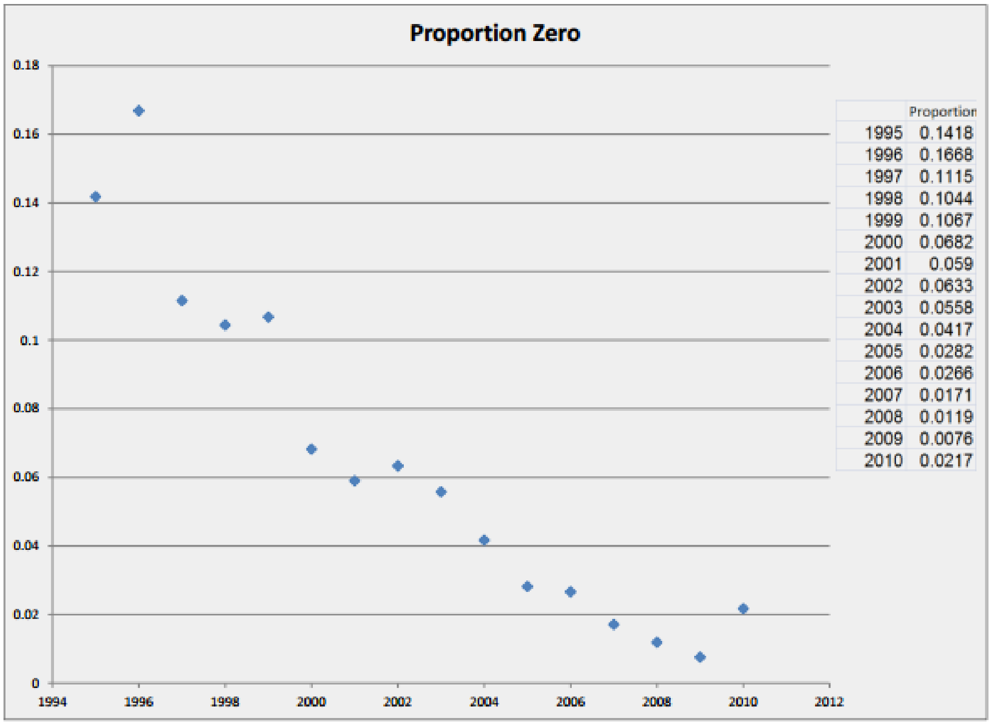

Figure 5. Proportion of zeros in catch data for the EAG.

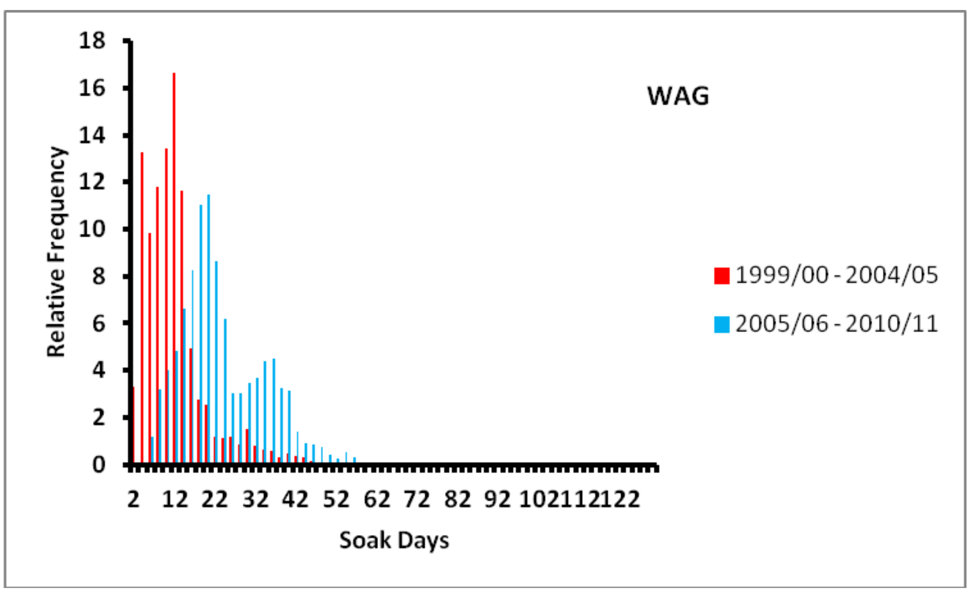

Figure 6. Non-truncated histograms of soak days pre- and post-rationalization.

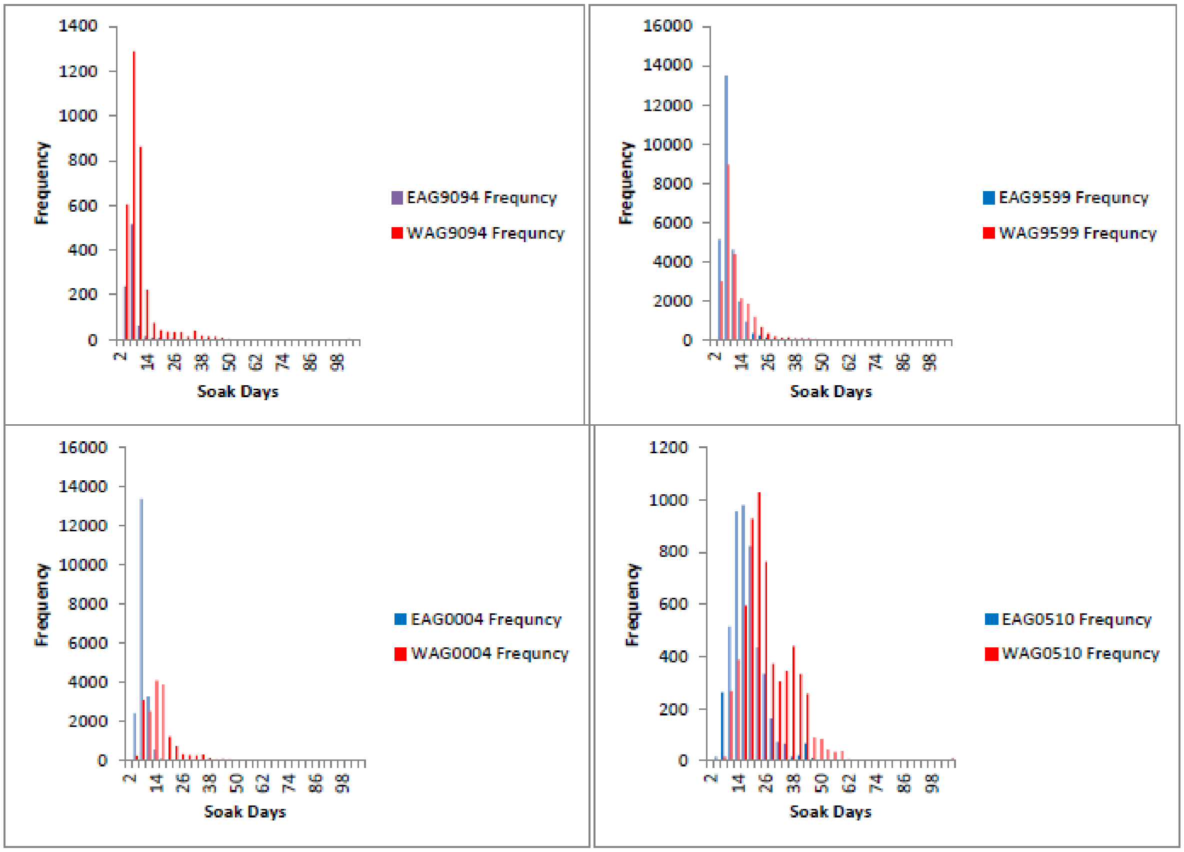

Figure 7. Histograms of soak times by area for five year periods. Note that the y-axes are different scales.

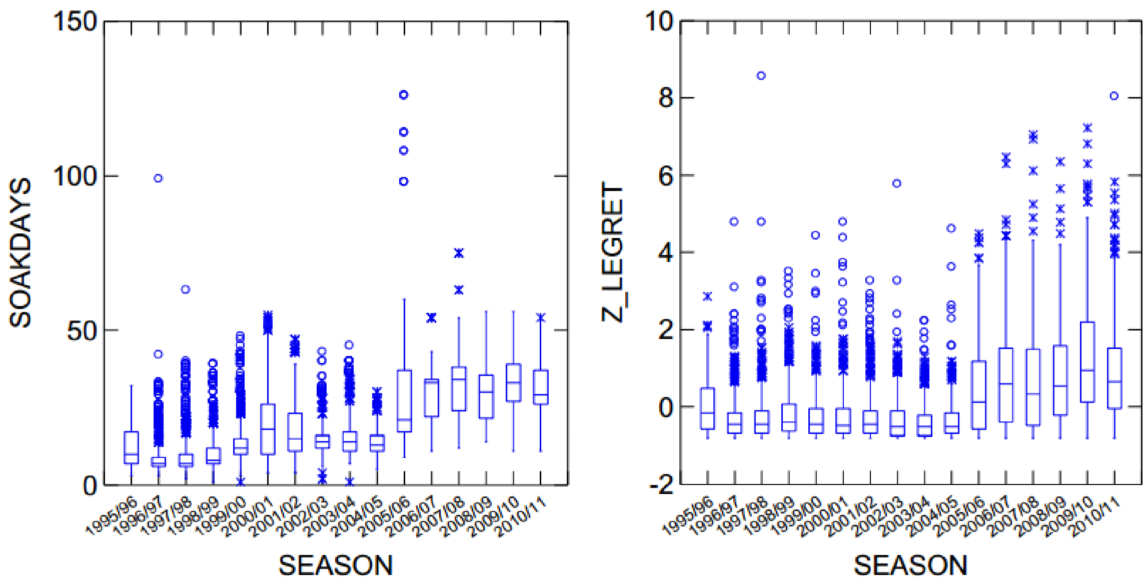

Figure 8. Soak time and standardized retained catch for an individual vessel spanning both pre- and post-rationalization seasons.

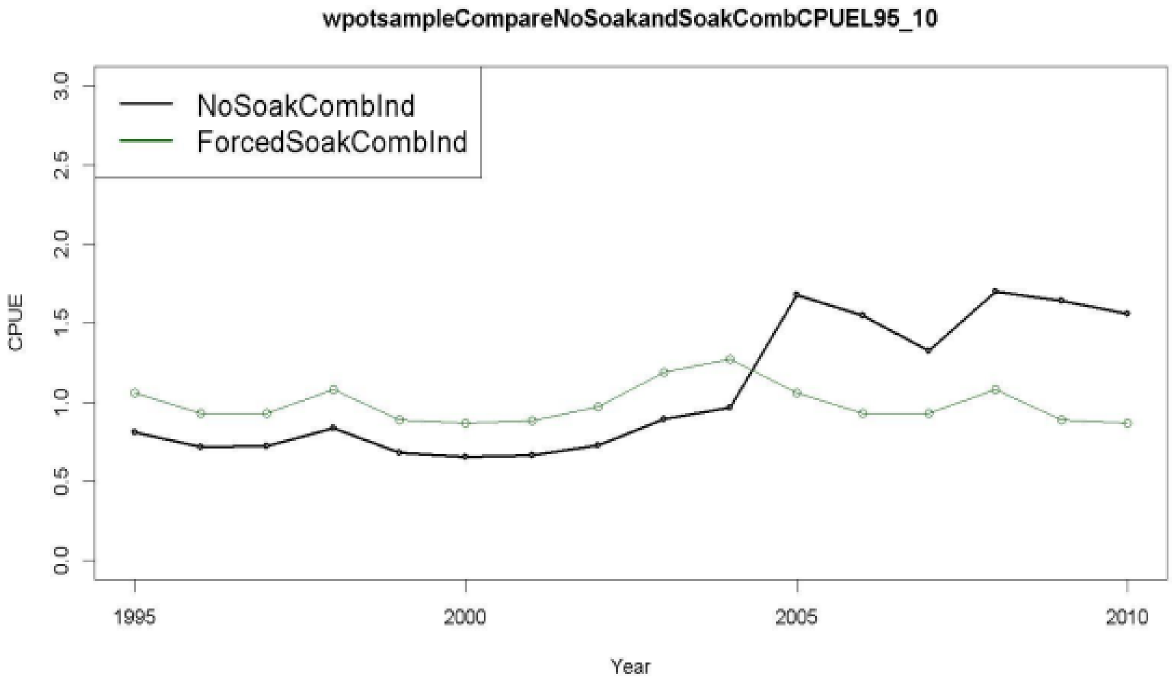

Figure 9. Changes in WAG standardized CPUE when soak time not selected and when it is forced into the model.

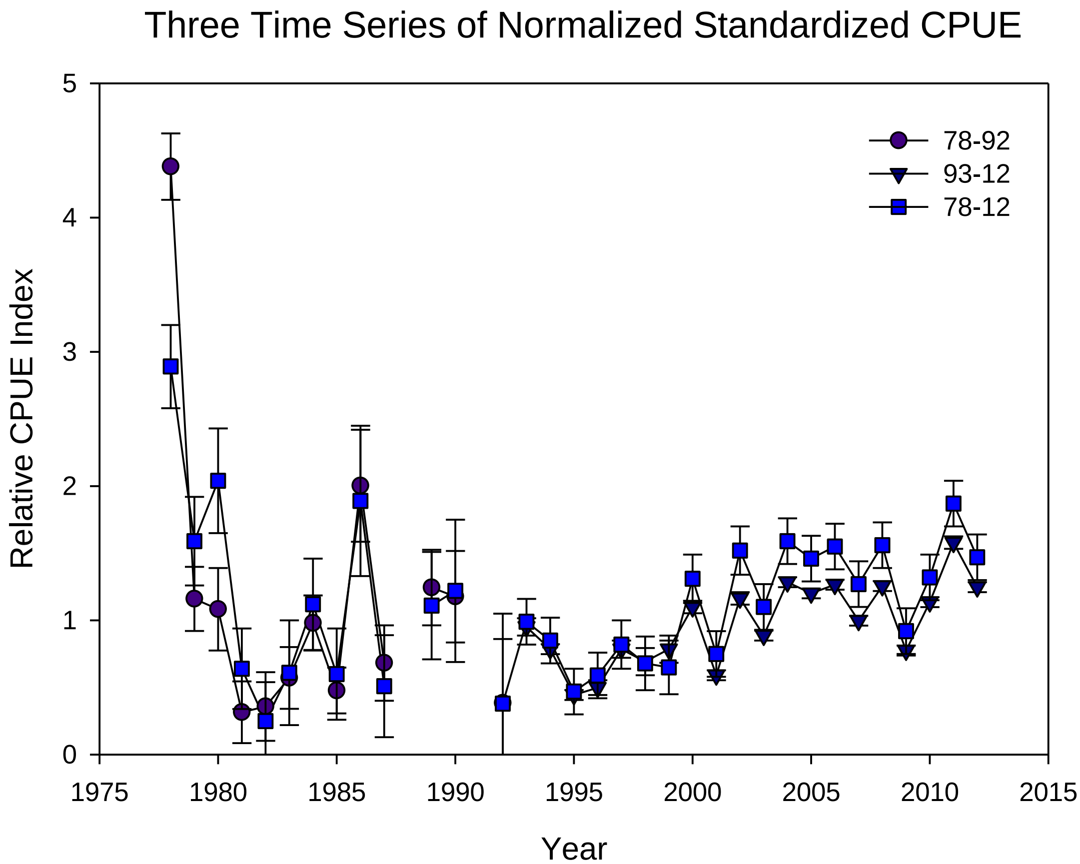

Figure 10. Trends in standardized (ST.CPUE), catch per unit effort (CPUE) in numbers per pot and their standard errors from generalized linear models for three time series: 1978–1992 (log(*CPUE*)~*Year*+ *Vessel*+*Modified Statistical Area*+*Week of Year*), 1993–2011 (log(*CPUE*)~*Year*+*Vessel*+*Week of Year*+*Modified Statistical Area*+*Permit Fishery*.), and 1978–2012 (log(*CPUE*)~*Year*+*Modified Statistical Area*+*Week of Year*+poly(*Vessel Length Overall*,3)+*Permit Fishery*). CPUE for 1978–1992 and 1993–2011 were normalized to the 1978–2012 mean. For each of the respective time series, data is from the summer commercial fishery for red king crab in Norton Sound for vessels having three deliveries for three years. The forward stepwise selection process used a stopping point of R^2 difference > 0.01.

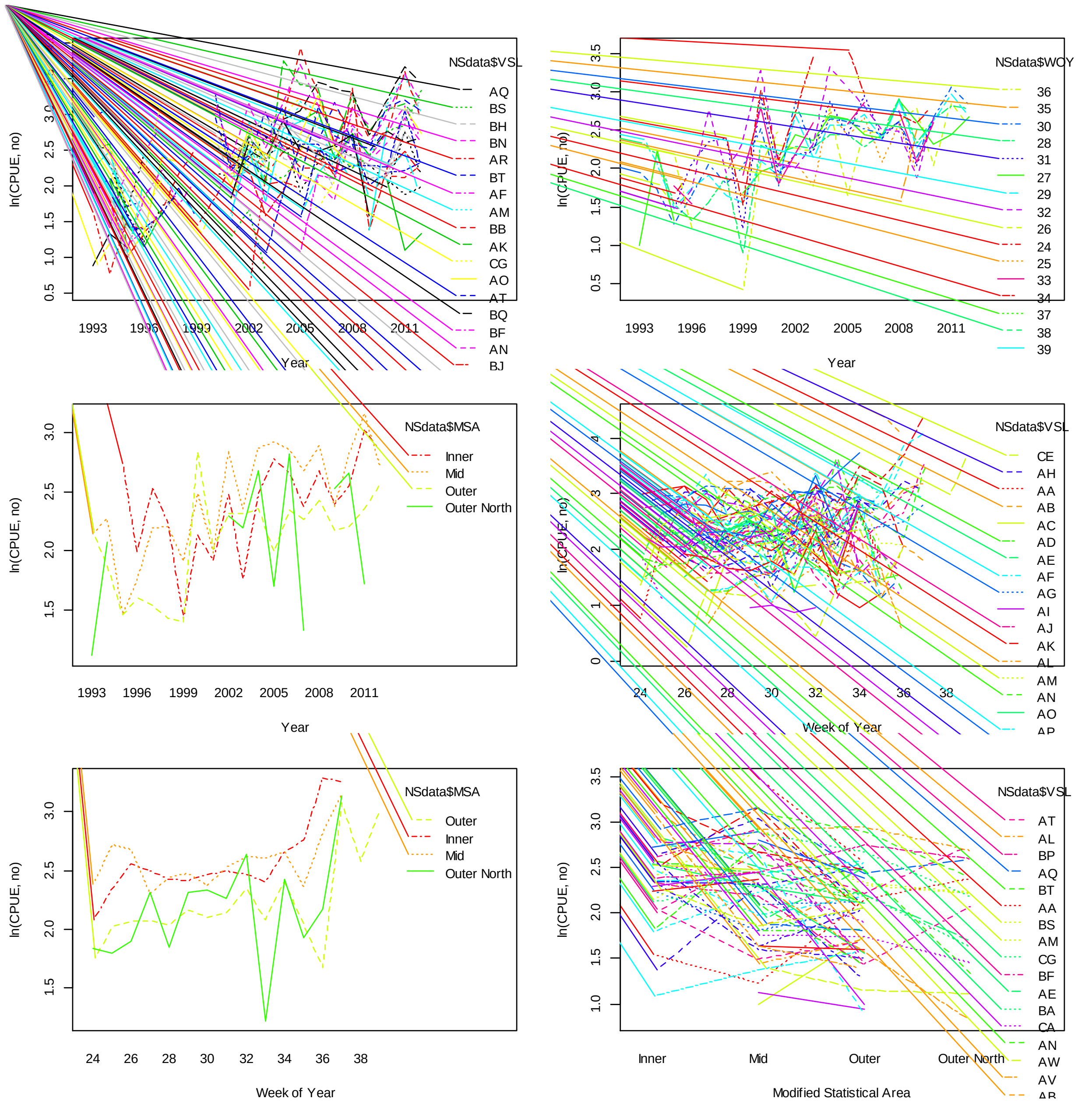

Figure 11. Interaction plots for the GLM from Table 5b above from the glm of LN.CPUE.NO~YR+VSL+YR:VSL+WOY+VSL:WOY+YR:WOY. Data is for the time series 1993–2012 from the summer commercial fishery for red king crab in Norton Sound for vessels having three deliveries for three years. The forward stepwise selection process used a stopping point of R^2 difference > 0.01.

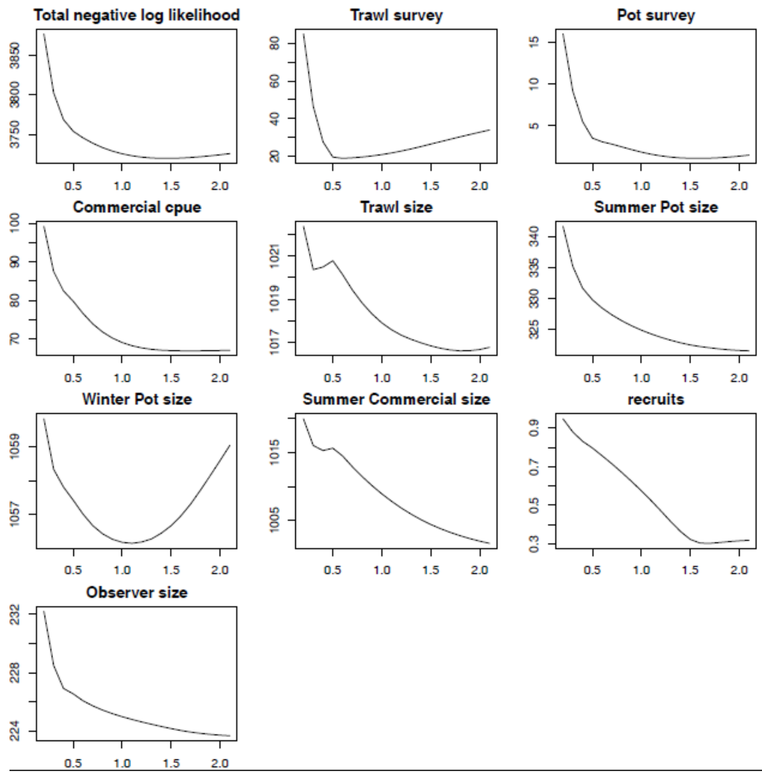

Figure 12. Likelihood profiles for survey Q for NSRKC.

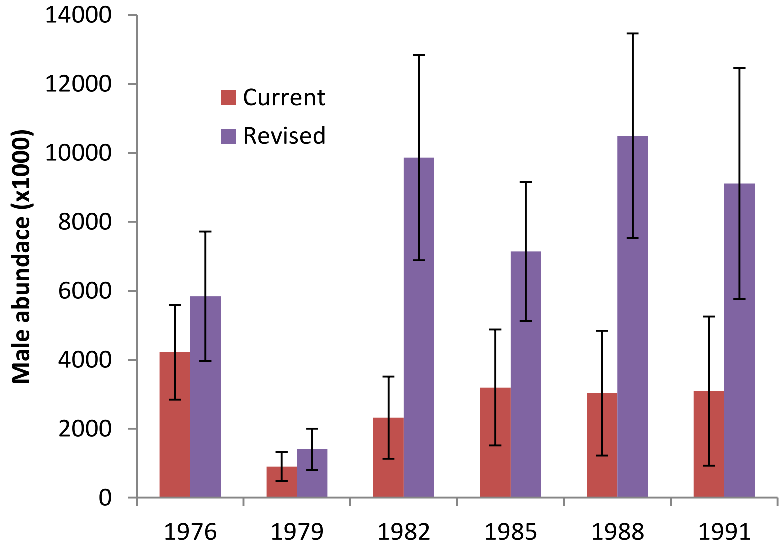

Figure 13. Comparison of the current NMFS survey estimates of abundance and the recomputed values.

Appendix A
Workshop participants and attendees

André Punt	UW (Chair; CPT)
Diana Stram	NPFMC (CPT)
Robert Foy	AFSC Kodiak (CPT)
Siddeek Shareef	ADFG Juneau (CPT; analyst)
Martin Dorn	ASFC Seattle (CPT)
Jack Turnock	AFSC Seattle (CPT)
Bob Foy	AFSC Kodiak (CPT)
Jason Gasper	NMFS RO Juneau (CPT)
Doug Pengilly	ADFG Kodiak (CPT)
Buck Stockhausen	AFSC Seattle (CPT)
Steve Martell	IPHC (SSC)
Terry Quinn	UAF (SSC)
Gordon Kruse	UAF (SSC)
Hamachan Hamazaki	ADFG (analyst)
Athol Whitten	UW (rapporteur)
Cody Szuwalski	UW (rapporteur)

Paul Starr	New Zealand
Jie Zheng	ADFG Juneau
Dick Powell	
Denby Lloyd	Aleutian king crab research foundation
Wes Jones	NSEDC
Linda Kozak	
Rip Carlton	GKC Research Foundation
Vicki Vanek	ADFG
Jim Ianelli	AFSC*
Gretchen Bishop	ADFG Juneau*
Anne Hollowed	AFSC *
Chris Siddon	ADFG Juneau*
Charlie Lean	NSEDC
Dick Tremaine	Aleutian No. 1
Kevin Keith *	

*participated by webex

Appendix B
Primary Documents Reviewed

Documents were made available for review prior to and during the workshop at:

www.tinyurl.com/ak-crab

Appendix C
Notes on golden king crab models
Steve Martell

1 The objective function must be differentiable & continuous

There are a number of occurrences of conditional statements (e.g., if, max, min) that create a discontinuity in the objective function. For example, the following line of code changes the value of a differentiable variable to a fixed constant if the following condition is true:

```
if(preddiscdcpue(t,l)<=0.0) preddiscdcpue(t,l) = 0.00001;
```

In this case the conditional likelihood of prediscdcpue(t,l) = -100 is the same as prediscdcpue(t,l) = -0.01 because in both cases prediscdcpue(t,l) is a fixed constant. Great efforts should be made to avoid such conditional statements. Two suggestions, recode the model such that preddiscdcpue can never be less than or equal to 0, or if this is unavoidable use posfun to ensure the objective function remains continuous. Without such a change, the assessment becomes conditional on the assumed constant, in this case 0.00001, and can have profound effects on model estimates (e.g., Figure 1A, left panel). An alternative would be to use posfun to ensure the objective function is continuous:

```
preddiscdcpue(t,l) = posfun(preddiscdcpue(t,l),1.e-5,fpen);
```

This minor change results in a slight difference in the estimate of mature male biomass (figure App.CA right panel).

1.1 A simple test

If you are using posfun, there is a possibility that model predictions could be based on the small constant used in posfun, rather than the model-based prediction. During development of the model the following statements can be used to validate that objective function is based entirely on model parameters, and not the assumed constant:

```
if(fpen >=0) cout<<"Fpen = "<<fpen<<endl;
if(fpen ==0 && last_phase()) exit(1);
```

This statement can be placed just before the calculation of the final objective function. If the code does not stop before calculating the hessian matrix, then the user should be aware that the likelihood calculation is comparing the data with an assumed constant, not the model predictions based on model parameters. If the model stops, then all is well in the last phase of parameter estimation. If you are seeing nonzero values for fpen during the hessian calculation, then the model results are partially based on the value of the constant used in posfun.

2 Molting probability

There is no information in either of the data sets to estimate the molting probabilities (i.e., newshell oldshell data). This was pointed out in last year's modeling workshop. As a result the a and log b parameters are not estimable. Yet the code attempts to estimate them in the Dutch Harbor model. If you simply change the upper bound of log b for example, the model no longer converges and you get a very different answer.

In the AI case, the molting probabilities from the Dutch Harbor model are then used as the molting probability for this model (i.e., dmolt in the input data). The growth transition matrix is also used, but this is the intended assumption is to assume that growth transitions are the same in the two areas and based on the growth increment data from the tagging study in DH.

The population dynamics accounting is still based on newshell and oldshell animals. This should be changed in the code such that the model results for both areas are not sensitive to the assumed prior distributions (in this case bounded uniform priors) for molting probabilities.

3 Size transition matrix
The log_beta parameter is sitting on its upper bound of -0.6. The bounds for this parameter are extremely small (-0.696, -0.600); likely due to the confounding with molting probability. If the molting probability is removed from model and accounting is based on old shell and newshell combined, then the length-transition matrix model will have to be re-parameterized to include the expectation and variation in the gamma distribution (or alpha and beta coefficients).

4 Simulation testing & diagnostics.
Before this model can be used for management advice, simulation testing must be performed. In addition to simualtion testing, a few additional diagnostic tests should also be performed. For example if posfun is being used to improve numerical stability, e.g.;

```
double tiny = 1.e-10;
predretcpue(t,l) = posfun(predretcpue(t,l),tiny,fpen);
```

then the value of fpen should be printed out to ensure that it is zero.

At a bare minimum, the stock assessment model should be used to simulate the observations during the PRELIMINARY_CALCS_SECTION, and then allowed to estimate the model parameters that were used to generate the fake data. The Generic model that is available at https://code.google.com/p/generic-crab-model/ has a built in command line option -sim that will overwrite the existing data with simulated data.

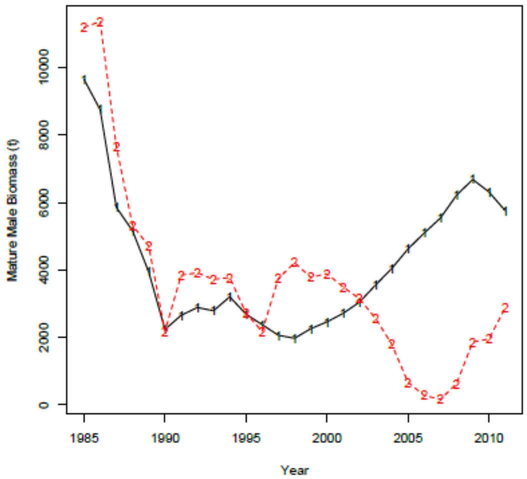

Figure App.C.A. Estimates of mature male biomass with a conditional constant set at 0.1 (label 2) versus 0.0001(label 1) in the Aleutian Island golden king crab model (Left panel). In the right panel, label 3 corresponds to a change in the code that uses “posfun" to ensure the objective function remains continuous.

Appendix D
Code to plot polynomial interactions

```
set.seed(98010)
Ndata <- 20

VSL <- factor(c("A","B","C"))
Year <- factor(c(1,2,3,4,5))
YearV <- c(1,4,9,16,25)
VSLv <- c(1,2,5)
LOAs <- runif(Ndata,0.2,5)
LOAv <- 1-20*LOAs+5*LOAs^2
print(LOA)
print(VSL[1])
design <- data.frame(YEAR=sample(Year,Ndata,rep=T),VSL=sample(VSL,Ndata,rep=T),LOAs=LOAs)
true <- YearV[design$YEAR] + VSLv[design$VSL] + LOAv
print(cbind(design,true))

# Gretchen, compare these two outputs to see the funky way R handled polynomials!
M1 <- lm(true~YEAR+VSL+poly(LOAs,2),data=design)
print(summary(M1))
M1 <- lm(true~YEAR+VSL+LOAs+I(LOAs^2),data=design)
print(summary(M1))

# This is the bit you care about
LOAOut <- seq(from=0.2,to=5,by=0.1)
getpolydesign <-
data.frame(YEAR=rep("1",length(LOAOut)),VSL=rep("A",length(LOAOut)),LOAs=LOAOut)
#print(getpolydesign)
# Do the projection
LOAPred <- predict(M1,getpolydesign)
#print(LOAPred)
truepoly <- 1 - 20*LOAOut+ 5*LOAOut^2
par(mfrow=c(2,2))
plot(LOAOut,truepoly)
# This is what we want
plot(LOAOut,LOAPred)
plot(truepoly,LOAPred)
```