UM Independent System for Peer Reviews
Consultant Report on:

Blue Crab Stock Assessment Review

9-11 August 2005 Annapolis, Maryland

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paul.medley@virgin.net

Contents

Executive Summary

The blue crab stock assessment appears to be developing well with marked improvements on the previous assessment in 1997. In particular, considerable work has been put into defining natural mortality to a more realistic and acceptable level, and estimates of catches appear to have been made more accurate. Progress has also been made in the other areas addressed, but the assessment has not yet been developed enough to have full confidence in the results.

This report concentrates on recommendations which are intended to help guide future development of the stock assessment. They address those areas where I believe that there is greatest uncertainty. The most important recommendations are:

- It is apparent that most of the fishery independent survey indices could be improved using generalized linear models. This would provide corrections for changes in gear, weather, habitat and survey design that might have occurred during the years. In particular, a linear modelling approach to standardise the winter dredge survey using the depletion experiments should result in better index estimates.
- The most useful additional abundance index would be a recruitment index as early in the life history as possible. This will increase the predictive power of any stock assessment model and improve advice and management planning. The best candidate for this would be the current megalopal survey.
- Given the high estimated mortality rates, it is necessary to switch to shorter time steps than a year to allow a reasonable chance of estimating fishing mortality. The first step would be to adapt the current annual stock assessment to a within-year model. This will mean describing within season processors, such as the time of arrival of recruits and generating data based on the shorter time step (e.g. catches in each month). The shorter time step will help model the population dynamics of an animal with such a short life history and measure the impact of fishing.
- To set decision rules, an interim stock recruitment relationship is required. This could only be based on historical recruitment, but nevertheless should be used to define the point at which recruitment overfishing occurs.
- Decision rules should be developed based on total catch. As the winter dredge survey (WDS) is strongly correlated to catch, management decisions can be based either on each WDS result or on some simple derivative, such as a moving average value.

Background

The blue crab (*Callinectes sapidus*) is an important commercial fishery for the Chesapeake Bay region and is one of the most valuable fisheries in the Bay. Ecologically, blue crab is an important component of the Chesapeake Bay ecosystem. The first bay wide assessment for blue crab was undertaken in 1995 and completed in 1997. It concluded that the stock was moderately to fully exploited and at average levels of abundance. Subsequent to this assessment, concerns over the continuing status of blue crab were raised because of declines in abundance and harvests. In response to concerns from stakeholders, a Bi-State Blue Crab Advisory Committee was established in 1996. Work by this committee lead to the establishment in 2001 of biomass and exploitation thresholds and an exploitation target reference point. Since 2001, the status of the blue crab stock has been updated annually and its status determined relative to the reference points.

A new stock assessment was completed in 2005. The assessment re-evaluated estimates of natural mortality rates, applied a new time series analysis to adjust historical landings for reporting changes, developed a new assessment model and developed an individual-based spawning potential per recruit model to estimate reference points based on the fraction of the vulnerable population that was harvested each year.

This report reviews the new 2005 Stock Assessment of Blue Crab in Chesapeake Bay and makes recommendations for future directions in data collection, analysis and research to improve the assessment.

Review Activities

The stock assessment documents were received prior to the meeting and consisted of the printed stock assessments and considerable background material on compact disk (See Annex I Bibliography). The stock assessment meeting was held at Radisson Hotel in Annapolis, Maryland, from 9-11 August 2005, where a presentation was given and discussions took place.

The CIE panel consisted of four external experts. The meeting was held in public and other scientists and stakeholders were present to observe and comment on the assessment. This report is based upon a review of the documents received, some analyses of data and discussions at the meeting.

Life History

Summary of Findings

The current analysis has provided improved life history parameter estimates and considerable progress has been made in this area. In particular, the estimate for natural mortality appears to be more firmly grounded on the available life history information. However, blue crab has certain characteristics, such as the terminal female moult and variable growth, which may make the species difficult to assess reliably.

The sex ratio in the population is probably equal, although interpretation depends on equal selectivity. Catches in the winter dredge are split approximately evenly between the sexes, although the catch favours females. Females may be more susceptible to fishing as males can be used as bait. Reporting could also favour females as males are sorted and sold separately to preferential outlets, which may be monitored less carefully. Males and females appear to share the same growth rates, but females have a terminal growth immediately after breeding. Given moulting is a significant source of mortality and females have a terminal moult, it is quite also possible natural mortality is higher for older males, and they decline in the population as they continue to grow.

The Von Bertalanffy growth model is likely to be reasonable as a description of average growth for a cohort even if individual animals exhibit discrete moult growth. Problems are more likely to arise from varying growth rates from year to year with stock density, food availability and temperature. Further modal analysis may give insight into interannual variability in growth which has yet to be explored.

Although an interesting and innovative way to measure age, the lipofuscin growth estimates are probably relatively inaccurate compared to other methods. Lipofuscin can build up in tissues for reasons other than age. In this case, however, there appear to be few options for improving growth estimates.

Tagging is probably the best alternative. Internal coded wire tags are being used on hatchery-released crabs, but recapture presents the greatest problem. To be cost effective a significant proportion of the landings would need to be monitored. For estimating growth only relatively few returns are needed, but they should cover a wide number of periods of release. Notably, it may be possible to obtain estimates of density dependent effects which could have a significant impact on the assessment. Information could include, for example, density dependent growth and mortality effects.

Conclusions and Recommendations

The main improvement to the estimates may be derived from modelling density dependent mortality and length specific mortality. A significant source of mortality is thought to be due to cannibalism. Other mortality may be a direct result of moulting errors or disease. Density dependent effects should make the population more resilient to exploitation.

The natural mortality estimate, in comparison with the 1997 assessment, appears to be more reasonable. It is based on growth rates, the latest aging and tagging work. The general method, to compare a number of different ways to estimate M, is in my opinion the best approach. The preferred estimates of natural mortality lie within the limits from all methods, but are mainly supported by a Delaware study and tagging studies of mature females.

Tagging is only likely to provide good information if catches are more closely monitored. Coded wire tags would be the ideal option, but would require a significant proportion of the catch being monitored for tag returns. This might be done at larger scale processing facilities as well as spot checks at landing and marketing points.

Catch monitoring in general should be improved. Catch monitoring to obtain catch composition and detect tagged crabs would appear to be a necessary step to improve life history parameter estimates. It is highly likely that there are density-dependent and size effects on mortality and tagging would probably be the only way these might be understood.

Fishery Independent Surveys

Winter Dredge Survey

Summary of Findings

The winter dredge survey (WDS) is probably the most accurate biomass survey available. It can be used to estimate next season's catch, and may help estimate natural mortality. Its main problems are the relatively short time series and the need to correct for dredge efficiency and catchability effects.

The catchability (*q*) is estimated from depletion experiments which then allow the surveys to estimate overall biomass as well as correct for changes in gear and vessels. These depletion experiments are a major source of uncertainty. No detailed documentation of the estimation method was available, but it appears that the experiments are analysed separately which may not be the most efficient way to standardise these indices.

The winter dredge survey appears to be able to give estimates of total biomass as well as relative spatial density estimates and relative abundance of different components of the stock, such as size composition, sex and maturity. There is considerable scope for work on these indices, both in improving them and studying the changes they indicate in the population.

Conclusions and Recommendations

In standardising the survey, the assessment scientists face many of the same problems that have occurred in standardising the dredge used to survey the scallop stock over the Georges Bank and elsewhere (contact Dr Deborah Hart or Dr Larry Jacobson, NEFSC, Woods Hole, MA). This scallop survey has used depletion experiments to calibrate the dredge, but also found video surveys more useful in estimating the proportion of scallops which are within the path of the dredge but not captured (dredge efficiency). It is not clear whether video surveys would be effective in this case due to water turbidity and the fact that crabs bury themselves in sediments, but it would be worth a trial. Efficiency estimates can also be improved by recording exact tracks through GPS and accounting for track overlap in the depletion experiments.

While estimating catchability for the gear can help correct estimates, estimating catchability will always be a significant source of error where it varies considerably from survey to survey. Where possible, surveys should be conducted with the same vessels and gear at the same time each year. Minimising the variation in covariates will minimise errors in applying corrections.

A generalized linear model (GLM) can be used as a method to correct for sampling bias and standardise indices. A general approach for dealing with missing survey data and estimating overall biomass is to use generalised linear models (see Annex IV). The GLM can be simultaneously fitted to the depletion experiments used to calibrate the survey as well as the survey data itself. However, an important part of applying corrections is to build an appropriate set of covariates which can be linked to the survey data.

Covariates should include habitat type, weather conditions as well as the gear and vessel characteristics. These covariates could be built retrospectively, assuming benthic habitat (e.g. particle size, depth, salinity) has not changed over time and standard time series indices such as wind speed and water temperature can be used to represent general weather conditions in which the dredge operates.

The depletion experiments used to calibrate the survey should also be used to estimate selectivity. Unfortunately sample sizes can become small as the catch is broken down into categories. A generalized linear model approach becomes more important in these circumstances. This would allow a more sophisticated modelling approach to the survey and should yield more information about the stock.

Other Surveys

Summary of Findings

There are two trawl surveys and a trap fishery survey. The two trawl surveys cover separately the Maryland and Virginia parts of the Chesapeake Bay. The Calvert Cliffs trap survey is located off a power plant and is limited to a particular area.

The Virginia (VIMS) survey is juvenile finfish trawl survey, and crabs are only taken as a by-catch. A non-targeted survey can create problems where catchability for the species of interest is very low. Nevertheless, in this case it appears to be a useful survey and a valuable source of information in fitting a population model (see Annex IV).

The VIMS survey has changed over the years, expanding in scope to the lower bay and undergoing some gear changes, with significant alteration in 1973 and 1979. Catchability is most likely lowest for smaller (age-0) crabs and therefore the recruitment index may be poor. As a long time series of relative abundance it appears overall to be a very useful index.

The Calvert cliffs survey appears to have set up to monitor the potential environmental impact of the power plant. The trap method relies on crab movement, maybe affected by other pots set locally. However the gear has been consistent, and survey has a long time series.

The Maryland Department of Natural Resources survey like the VIMS survey, is a trawl survey, but appears to have been carried less consistently with spatial and temporal gaps in the data series. The implication is that the series is less reliable and noisier.

The analysis of the survey indices was conducted using time series of the standardised as z-scores. This treatment seems inconsistent with the treatment of the same data as log-normal indices in the stock assessment model.

Simple cross correlations between all the log-survey series suggest coherence is low and the series are either noisy or not responding to the same signals. This raises the issue of an appropriate weighting scheme for the indices when they are used together to indicate changes in the stock size.

It is not clear that size divisions are allocating crabs to the correct age groups. Blue crabs show very variable growth and middle size categories probably have indeterminate ages. The poor correlations between “Age 0” category and “Age 1” category one year later are not particularly convincing evidence that the indices are tracking cohorts effectively.

Table 1 Cross-correlation between log-transformed abundance indices at zero lag. High coherence between series should result in high positive correlations between compatible indices (in bold).

	VIMS Age-0	*VIMS Age-1+*	*VIMS Mat Fem*	*MD Age-0*	*MD Age-1+*	*MD Mat Fem*	*CC Age-1+*	*CC Mat Fem*
VIMS Age-0	1.00							
VIMS Age-1+	0.40	1.00						
VIMS Mat Fem	0.30	0.69	1.00					
MD Age-0	0.32	-0.05	-0.11	1.00				
MD Age-1+	0.07	0.03	0.10	0.48	1.00			
MD Mat Fem	-0.03	0.05	-0.09	0.22	0.72	1.00		
CC Age-1+	0.08	0.36	0.26	0.08	0.33	0.27	1.00	
CC Mat Fem	-0.06	0.22	0.11	0.01	0.22	0.13	0.78	1.00

Conclusions and Recommendations

The fishery independent indices appear to contain adequate information on stock size on which to build a stock assessment. Progress has been made in their analysis, but there are a number of statistical issues which could be addressed that could lead to an improved assessment.

The same GLM procedure suggested for the winter dredge survey can be applied to the trawl surveys. Unfortunately there would be no depletion experiment data for calibration, but nevertheless standardisation may improve the indices. Where possible, a retrospective covariate data set should be constructed, including variables, such vessel characteristics, gear type, weather conditions, water temperature and salinity, to correct survey data.

Standardisation may improve coherence between indices and give estimates of within-index variance which may be used as a basis for weighting the indices. However, a GLM may also be used to combine indices, particularly the VIMS and MD DNR trawl surveys which do not overlap spatially. This could be used to explore alternative schemes where indices are additive, as they monitor separate parts of the same stock, rather than being independent.

It may be better to try allocating size compositions to age classes dynamically in each year based on modal analysis rather than using fixed size categories in each year. This not only may improve allocation, but could also be directly justified rather than relying on indirect time series cross-correlations.

The most useful additional index would be a recruitment index. The fishery is predominantly recruitment driven and such an index would explain more of the variation in the stock size. It could be used to predict future catches and plan for a minimum escapement policy, for example, by varying the length of the fishing season. There is already a “megalopal survey” (a general zooplankton survey) that could be tested and developed as a recruitment index. Also collectors have been used and have shown that wind has been a component affecting settlement. The moon phase may also be a factor.

Fishery Dependent Data (Catch and Effort)

Summary of Findings

There is currently no reliable catch and effort data collection. Catches are obtained from various sources, but mostly depend upon fishermen's reported harvest. The observer data for Maryland indicates underreporting and that underreporting is regional. It may be possible to use observer data in future to correct CPUE and improve estimates of catch where total effort is known. Virginia has very limited data collection and co-operation between Virginia and Maryland is poor.

The landings data have been corrected using a time series analysis for accounting for reporting changes. These corrections seem reasonable and should have improved the catch data for use in stock assessment.

Information on catch composition is not available. "Cull-rings" are used to allow undersize crabs to escape, but it appears not to be enforced and as a result it is likely that selectivity will vary from fisherman to fisherman. In addition, because of the different markets, landings of males are more likely to be underreported than female landings.

Conclusions and Recommendations

Catch information does not appear to be very reliable, although adequate for stock assessment. Recording and reporting catches is an area where there could be considerable improvement.

Measures taken to improve enforcement of legal reporting requirements together with educational "outreach" to fishers could lead to better catch reporting and co-operation in scientific research and monitoring. Giving feedback to fishermen on their reported data through either personal reports or public reports in local newspapers, for example, can greatly increase compliance as fishermen begin to understand why they are required to provide data.

Random trip interviews could be conducted at landing sites. Accurate catch and effort measurement and biological sampling would yield better estimates at relatively low cost. There would also be considerable benefit from harmonising the Virginia and Maryland data collection for both the mandatory reporting and random trip interviews.

Stock Assessment Model

Summary of Findings

The model is very simple, assuming a constant natural mortality rate and two variable processes, catch and recruitment, driving population changes. This is appropriate considering the lack of information in catch composition. The model does not account for errors in catch estimates explicitly, although the process error may model a similar effect.

The representation of the process error could be improved. The current log-normal random process appears to be arbitrary. Normally distributed mortality rates could produce a log-normal error, although the error would need to be bounded to avoid increases in population as opposed to just mortality. It is currently not clear how much likelihood is being allocated to increasing population in the model. This effect could be avoided either by keeping process error low or by bounding the process likelihood to ensure the model only accounts for ageing animals dying and not increasing. Increases would imply the population is not closed.

Fishing mortality is negatively correlated with population size. This is probably a result of the indices fluctuating with greater variance than the catches. As the population varies with the indices, relatively constant catches will be explained by changing the effective catchability to compensate. There is no obvious reason why catchability would vary in this way (catchability increasing as crab density falls). It is possible that the population model is fitting to spurious signals in the index, and that the population has not fluctuated as much as the indices indicate. It is also possible that the population model is simply inadequate and these results indicate a model with density dependent mortality could fit the data much better.

Fishing has no measurable impact in the annual model as mortality rates are too high for a significant proportion of the stock to survive from year to year, making an annual assessment suspect. Mortality rates are high enough to detect a within-season depletion. Within-year timings of indices are probably important to management of the fishery. An annual model can only cope with this approximately, but a within-season model would be able to account for seasonal movement, mortality and growth explicitly.

Conclusions and Recommendations

An intra-seasonal model would provide better estimates of fishing mortality. Currently, the modelling approach has been to describe the annual dynamics, but mortality rates are so high that very few adults make it through to their second and third years after recruitment. As there is no stock-recruitment relationship, the model can only hope to explain a very small amount of the variation in stock size and this leads to poor parameter estimates. An alternative, which should lead to better estimates of mortality, is to use the depletion occurring within each year to estimate fishing mortality. If it is possible to generate the necessary within-season data, the fishing mortality should be estimated much more reliably.

The main problem with intra-seasonal models is having adequate data to estimate the relevant parameters. It is unlikely to be possible to separate data into less than two week intervals, and even monthly data may be difficult. Catches and survey indices will need to be specified for the months in which they are obtained. Catches in particular may be difficult to allocate accurately.

The same model as that presented could be used for the intra-seasonal assessment except the time step would be a month rather than a year. The model can be used to estimate the recruitment and survival in each month. Some simplifications may be possible, as the recruitment model can be parameterised to specify the amount of annual recruitment and distribute that recruitment in the same months each year. Each recruitment would then be specified with a parameter for each year and two or three parameters specifying the proportion arriving in each month.

Supporting evidence to indicate changes in catchability would benefit the assessment. Estimated fishing mortality fluctuates as catch (and effort) variation is lower than that observed in the survey indices. Therefore to explain these indices, models will tend to change catchability estimates if catchability is allowed to vary. This may incorrectly explain observation error as changes in catchability or effort. It would be useful to incorporate limitations on catchability or effort changes, through, for example, a conditional random walk term indicating very large changes as unlikely.

I would recommend considering a model that deals with a catch observation error rather than assuming it is known exactly. As catches are assumed to be consistently underestimated, the sensitivity of the results to such a bias can be tested.

It would be sensible to look at a number of assessment models as well as the one presented. Simpler models, such as a recruitment index model, can be fitted to the data (Annex IV). Albeit, this model does not fit as well, it is more parsimonious and it needs to be made clear what advantages are gained from less parsimonious models. Advantages, if any, need to be clearly identified. This would be obtained from a set of alternative models presented for comparison.

Control Rules

Summary of Findings

Although limit and precautionary reference points are being developed, it is not clear how they would be implemented. There appear to be limited options for fishery controls. An explicit decision rule and recovery plan needs to be put in place.

The "overfished" definition depends on a potentially unreliable survey point for 1968. A reference point based on more data points would be desirable. The obvious way to proceed would be to use the residuals from the assessment model allow for error and choose indices which would clearly indicate that the stock biomass has fallen below the chosen threshold.

The assessment does not explicitly deal with a stock recruitment relationship. While the spawners-per-recruit method that has been developed is adequate, the reference points themselves appear arbitrary and are not properly justified.

Conclusions and Recommendations

The winter dredge survey is the most reliable index and is directly correlated with catches. This makes this survey index the best candidate on which to directly base a control rule. The control rules should be explicitly set to maintain or rebuild catches to levels required by fishery managers. This is broadly the approach that has been adopted.

Given the short life history, a stock-recruitment relationship is required to set a meaningful control rule that could address management concerns. SPR reference points will not address total catches and be difficult for managers to support. Therefore a primary objective for the stock assessment should be to develop a workable stock-recruitment relationship that will allow setting reasonable thresholds that will maintain recruitment and therefore catch around some long term agreed average.

Control rules should take into account statistical errors in the index. This can be as simple as using a moving average of the WDS survey instead of relying on point estimates, or based on a more complicated decision analysis approach.

Currently the stock assessment model is not used in setting the control rules. The central aim of further development of the stock assessment model would be to provide relevant information for setting the control rule and testing it through simulations.

Annex I Bibliography

A compact disk was provided with considerable background material, including spreadsheets containing data used in the assessment, AD Model Builder code and supporting articles published in the scientific literature. However, the main documents subject to review were the following:

Rugolo, L., Knotts, K., Lange, A., Crecco, A., Terceiro, M., Bonzek, C., Stagg, C., O'Reilly, R., and Vaughan, D. Stock Assessment of Chesapeake Bay Blue Crab (Callinectes sapidus). April 1997.

Miller, T.J., Martell, S.J., Bunnell, D. B., Davis, G., Fegley, L., Sharov, A., Bonzek, C., Hewitt, D., Hoenig, J., Lipcius, R.N. Stock Assessment of Blue Crab in Chesapeake Bay. 2005. Final Report

Annex II Statement of Work

STATEMENT OF WORK

Consulting Agreement between the University of Miami and Dr. Paul Medley

July 21, 2005

Background

The blue crab supports the most important commercial fishery in the Chesapeake Bay. Commercial landings have exceeded 100 million pounds historically (1993) with more recent average landings reaching approximately 72 million pounds. The total impact of the blue crab fishery to the Chesapeake region exceeds $200 million annually.

Sound management of this resource requires accurate information on the status and trends of the blue crab population and on the dynamics of the fisheries that exploit the stock. There have been two recent stock assessments completed for the blue crab (1997, 1998) and the NOAA Chesapeake Bay Office (NCBO) has produced annual 'Advisory Reports' for blue crab to assist resource managers in the decision making process. Seeing the need for an updated assessment, the NCBO supported the development of a full blue crab stock assessment utilizing FY2003 funds.

This assessment was initiated in October 2003. Due to the political nature of any decision regarding fisheries in Chesapeake Bay, especially blue crab, an independent and expert review of the science is necessary for management of this important fisheries resource. The Habitat Conservation Office is requesting that the Center for Independent Experts (CIE) conduct a review for the NOAA Chesapeake Bay Office's Blue Crab Stock Assessment.

The review workshop for the Chesapeake Bay blue crab assessment will take place in Annapolis, Maryland on August 9-11, 2005. The NOAA Chesapeake Bay Office will provide the following documents prior to the Chesapeake Bay blue crab stock assessment review meeting:

- 2005 Chesapeake Bay blue crab assessment report;
- 1997 and 1998 blue crab stock assessments;
- Annual blue crab advisory reports;
- Adopted management strategies establishing targets and thresholds;
- Chesapeake Bay Fishery Management Plan (1997); and
- Other key publications as necessary.

Objectives of the CIE Review

The Blue Crab Assessment Review Panel will evaluate the Chesapeake Bay blue crab stock assessment, including input data, assessment methods, and model results. The following are the main terms of reference for the review:

1) Evaluate the adequacy and appropriateness of all data used in the assessment, including the following:
   - Life history and vital rates of blue crab in Chesapeake Bay.
   - Patterns in fishery-independent surveys.
   - Patterns in catch and effort by sector and region.
2) Evaluate the adequacy, appropriateness, and application of the assessment models used for the Chesapeake Bay blue crab fisheries and characterize the uncertainty in the assessment.
3) Evaluate the scientific basis for the control rule for the Chesapeake Bay blue crab fishery.
4) Develop recommendations for future research for improving data collection and the Chesapeake Bay blue crab assessment.

The Assessment Review Panel's primary duty is to review the assessment presented. In the course of this review, the Chair may request a reasonable number of sensitivity runs, additional details of the existing assessments, or similar items from technical staff. However, the Review Panel is not authorized to conduct an alternative assessment or to request an alternative assessment from the technical staff present. The Review Panel should outline in its report any remedial measures that the Panel proposes to rectify shortcomings in the assessment.

Specific Activities and Responsibilities

The CIE shall provide a Chair and three Review Panelists to conduct the review of the Chesapeake Bay blue crab stock assessment.

Tasks

Each panelist's duties shall occupy a maximum of 14 workdays (i.e., a few days prior to the meeting for document review; the review meeting; and a few days following the meeting to prepare a Review Report). The Panelist Review Reports will be provided to the Review Panel Chair, who will produce the Summary Report based on the individual Review Reports.

Roles and responsibilities:

(1) Prior to the meeting: review the Chesapeake Bay blue crab assessment report and other relevant documentation in support of this review.

(2) During the meeting: participate, as a peer, in panel discussions on assessment validity, results, recommendations, and conclusions especially with respect to the adequacy of the assessment in serving as a basis for providing scientific advice to management.

(3) After the meeting: prepare individual Review Reports, each of which provides an executive summary, a review of activities and a summary of findings and recommendations, all in the context of responsiveness to the terms of

reference. Advice on additional questions that are directly related to the assessment and are raised during the meeting should be included in the report text. See Annex 1 for further details on report contents and milestone table below for details on schedule. No later than August 25, 2005, these reports shall be submitted to the CIE for review[1] and to the Chair for summarization. The CIE reports shall be addressed to "University of Miami Independent System for Peer Review," and sent to Dr. David Sampson, via e-mail to David.Sampson@oregonstate.edu and to Mr. Manoj Shivlani via e-mail to mshivlani@rsmas.miami.edu.

Milestones or Report Delivery Dates

The following table provides the milestones and delivery dates for conducting the panel review of the Chesapeake Bay blue crab stock assessment.

Milestone	**Date**
Panel review meeting in Annapolis, MD	August 9-11, 2005
Individual panelists provide their draft reports to CIE for review and to Chair for initiating development of the Summary Report	August 25, 2005
CIE provides reviewed individual panelist reports to NMFS COTR for approval	September 1, 2005
COTR notifies CIE of approval of individual panelist reports	September 8, 2005
CIE provides final individual panelist reports to COTR (with signed cover letter) and to Chair to complete Summary Report	September 13, 2005
Chair provides CIE with draft Summary Report for review	September 20, 2005
CIE provides reviewed Summary Report to COTR for approval	September 27, 2005
COTR notifies CIE of approval of Summary Report	September 30, 2005
CIE provides final Summary Report with signed cover letter to COTR	October 5, 2005
COTR provides final Summary Report to NEFSC contact	October 7, 2005

No consensus opinion among the CIE reviewers is sought, and all reports will be the product of the individual CIE reviewer or chairperson.

NOAA Contact person:
Derek Orner, NOAA Chesapeake Bay Office, 410 Severn Avenue, Annapolis, MD 21403; Derek.orner@noaa.gov

[1] All reports will undergo an internal CIE review before they are considered final.

Contents of Panelist Report

1. The report shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report shall consist of a background, description of review activities, summary of findings, conclusions/recommendations, and references.

3. The report shall also include as separate appendices the bibliography of all materials provided during the review meeting and any papers cited in the Panelist's Report, along with a copy of the statement of work.

Annex III GLM for Depletion Experiments

Using an analysis of covariance approach, the x-variables are the cumulative catches within each catch category.

$$U_{it} = a_{it} - q_i \sum_{s=0}^{t-1} C_{is}$$

Both a_{it} and q_i must be fitted within each i category. The intercept and the slope are related as q occurs in the intercept. Where the q is not varied based on common covariates, the intercept will still need to change between experiments. Categories would be defined based on shared covariates, such as weather conditions, benthic habitat vessel and gear. The advantage of this approach is that the error term is shared and the significance of parameters can be tested leading to greater parsimony.

Keeping the model to a strict linear format prevents improved modelling. It is possible to restate the model in a more flexible near-linear form. The model can be defined as:

$$N_{it} = N_{i0} - \sum_{s=0}^{t-1} C_{is}$$

which gives the population size within each experiment. The catchability and selectivity can then be modelled with any number of covariates using a standard GLM. For example, a log-linear model could have the form:

$$Ln(\mu_i) = Ln(N_{it}) + lp_i$$

where lp_i is the appropriate linear predictor for category i and μ_i is the expected CPUE or catch for the survey point as appropriate. The fitting follows an iterative two phase process. Iteratively, fitting the N_{i0} in the population model can be done first, then the GLM can be fitted in the normal way. Because of the near linearity of the model, fitting is straightforward using standard regression techniques. The same technique can also be applied with several populations, split in size groups for example, to estimate simultaneously catchability and selectivity curves.

The error model should be critically evaluated. It is quite possible that any of a normal, Poisson, binomial, gamma or log-normal may provide the best results and all should be evaluated particularly with respect to the variance-mean relationship. The log-normal and gamma models should generally be avoided if possible as they are undefined where catches are zero and delta-models, where a separate binomial likelihood is estimated for zero catches, are more complex. Zero catches in catch categories become a distinct possibility where the categories are broken down. Using a Poisson or binomial error under quasi-likelihood assumptions would be preferable.

Further improvements could be made to the model by considering more detailed spatial models of depletion, if GPS information is available. This would allow a direct estimate of dredge efficiency, that is the proportion of crabs that are captured which are in the path of the dredge. This allow a better direct estimate of density and thereby a better direct estimate of biomass.

Annex IV Assessment Model Analysis

Objective of the Analysis

The objective of this analysis was to review the modelling approach adopted by the assessment and in particular the coherence among the various indices available to monitor the population size. This is not a new assessment, but allowed more detailed exploration issues regarding the use of the indices. Three models were fitted for comparison. The same Collie-Sissenwine population model was used as in the assessment, although the fitting procedure is slightly different. An extension of this model with a density dependent term modelling the effect of adult cannibalistic mortality on recruits was also applied. Finally, a simpler recruitment index model was fitted for comparison. In all cases I used an alternative and simpler minimising routine (Solver in MS Excel), assuming process error was zero. This latter change simplifies the fitting process while still allowing the indices to be assessed.

Fitted Models

The recruitment index model is very similar to the Collie-Sissenwine model, except a recruitment index is assumed to be exactly proportional to the recruitment in each season.

$$N_{t+1} = \left(N_t + \lambda r_t - C_t\right)e^{-M}$$

where N_t = population at time t, λ = parameter raising the index to the population size, M = natural mortality and C_t = catches taken in time period t. The model assumes the recruitment index has no observation error. The longest time series index for age 0 animals, the VIMS trawl index, was found to give the best results and was used in the final model.

The Collie-Sissenwine model allows each recruitment to be fitted as a separate value.

$$N_{t+1} = \left(N_t + R_t - C_t\right)e^{-M}$$

The R_t were fitted parameters to the model, as well as the N_0. Fitting the recruitment time series did not require recruitment to be exactly proportion to the recruitment indices, which allows for observation error. Because the model is deterministic (no process error), the remaining population time series values (N_t) would be fixed by these parameters.

Finally, as an exploratory analysis, a population model was fitted which attempted, somewhat crudely, to capture density-dependent cannibalistic effects of adults on recruits. In this case, survival of recruits to the next year depends upon the number of adults.

$$N_{t+1} = \left(N_t + R_t\, e^{-\alpha N_t} - C_t\right)e^{-M}$$

The parameters in each model were fitted to the indexes by minimising the squared difference between the expected and observed log-indices.

$$\hat{I}_{it} = Ln\left(q_i N_t\right)$$

$$Minimise \quad \sum\left(I_{it} - \hat{I}_{it}\right)^2$$

Eight indices as well as the catch time series were used. All indices were transformed to natural logarithms of the original data and weighted equally. The appropriateness of the log-transform was not explored.

The q coefficients could be estimated directed through linear regression, so only the R_t and N_0 parameters were estimated using the non-linear minimizer. This is broadly the same approach as used in the CMS model presented in the assessment.

Results and Discussion

The model failed to find a reasonable least squares solution unless one scaling of the (q_i) parameters was fixed. I used the winter dredge survey biomass estimates to estimate the biomass directly (fixed q parameter for this index to 1.0). This was equivalent to using the exploitation fractions in the fitting process in the original assessment.

The general results follow the same pattern as those obtained from the fitted model in the 2005 assessment. However, in all cases, the estimated fishing mortality is lower than that obtained from the 2005 assessment. The model consistently estimated higher abundance than indicated by the WDS biomass estimate.

The better fit of the density dependent model indicated that there may be significant improvements in assessment models as the population processes are better modelled. The low fishing mortality in this case is partly an artefact of higher recruitment with its associated higher mortality. An appropriate selectivity model and within season mortality could lead to more realistic fishing mortality estimates.

In all the models, the residuals were still significantly autocorrelated, suggesting that more of the time series behaviour could be captured by a population model. It was also found that the indices are not coherent and do not all possess the same signal (Fig. 1). As well as improving the standardisation of the indices using GLMs, appropriate transforms and weights need to be critically examined to try to improve the use of these indices. This would include considering the timing and distribution of the stock relative to when and where the surveys are conducted. It may be that improvements in the indices will help choose between population models.

The results also show a hint of a workable stock-recruitment relationship (Figure 2). Improvements in the data and modelling would hopefully produce a more reliable functional relationship. As is usual with these relationships, the observation and process error predominates. However, such a relationship can be used as the default hypothesis on which to build a workable control rule that the managers and fishermen should find acceptable.

Table 2 Summary of results from fitting the three models to the catches and available indices. The degrees of freedom (number of data points – number of parameters) remains the almost same for each model. Natural mortality was estimated for the density dependent Collie-Sissenwine model, but otherwise fixed at 0.9 year^{-1}. The density dependent Collie-Sissenwine model fitted the indices best, but exhibits the lowest impact from the fishery with the lowest F.

Model	Degrees of Freedom	Goodness of Fit (χ^2)	Natural Mortality (M)	Fishing Mortality (F) (geometric mean)
Recruitment Index	160	12.27	0.9	0.16
Collie-Sissenwine (No process error)	160	6.88	0.9	0.27
Collie-Sissenwine with density dependent mortality	159	5.85	0.7	0.06

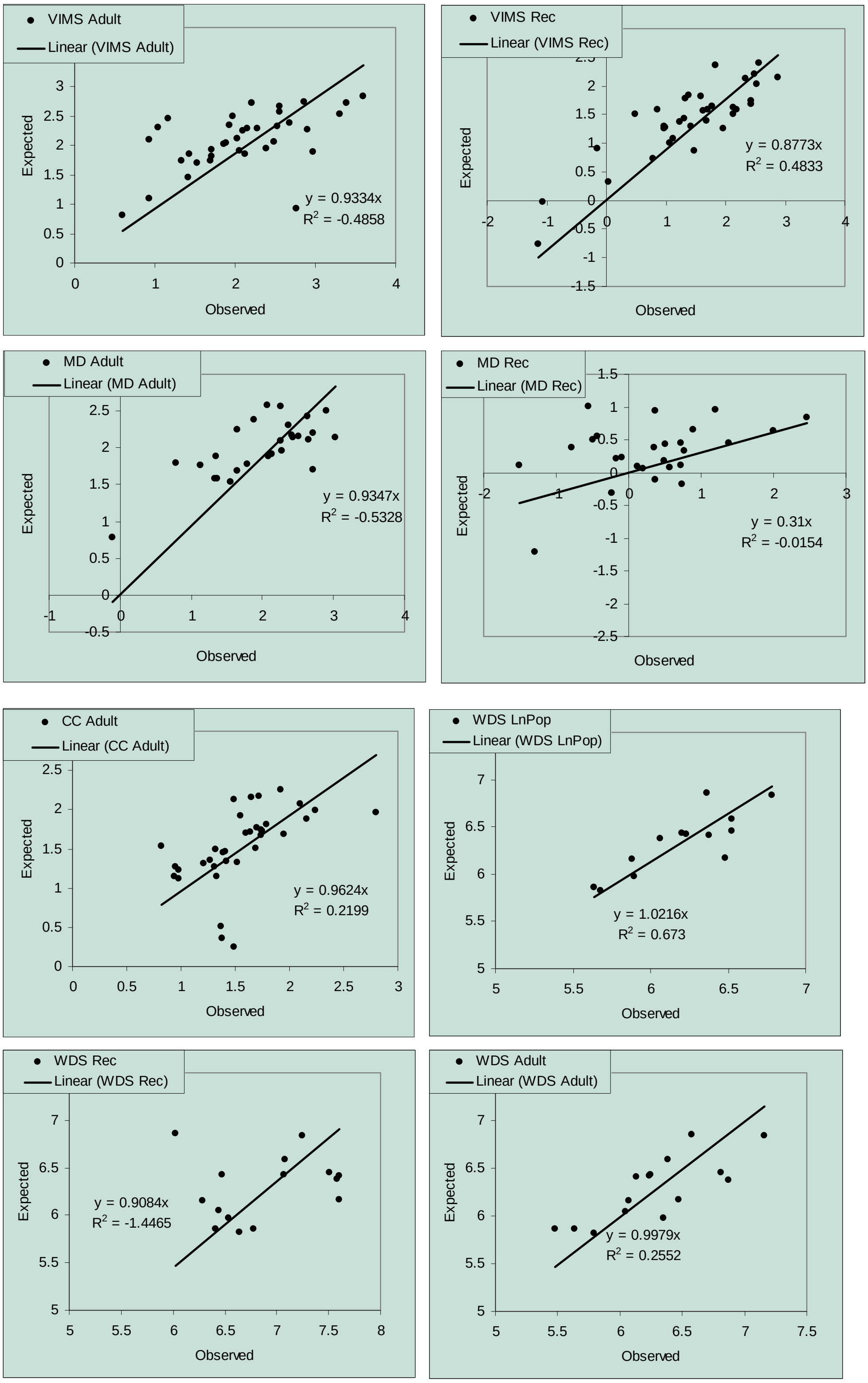

Figure 1 Observed-expected plots for VIMS and MD trawl indices the Collie-Sissenwine model with observation error only fitted in MS Excel. Where positive, the R^2 statistic indicates the proportion of variance explained by the model. A slope close to 1.0 indicates low bias and general good fit. In general, the VIMS trawl, Winter dredge and Calvert Cliffs (CC) indices seemed more coherent, so that variation could be explained by simple changes in population size in these indices. The Maryland trawl index (MD) generally seemed to be the worst behaved agreeing with comments made in the 2005 assessment report.

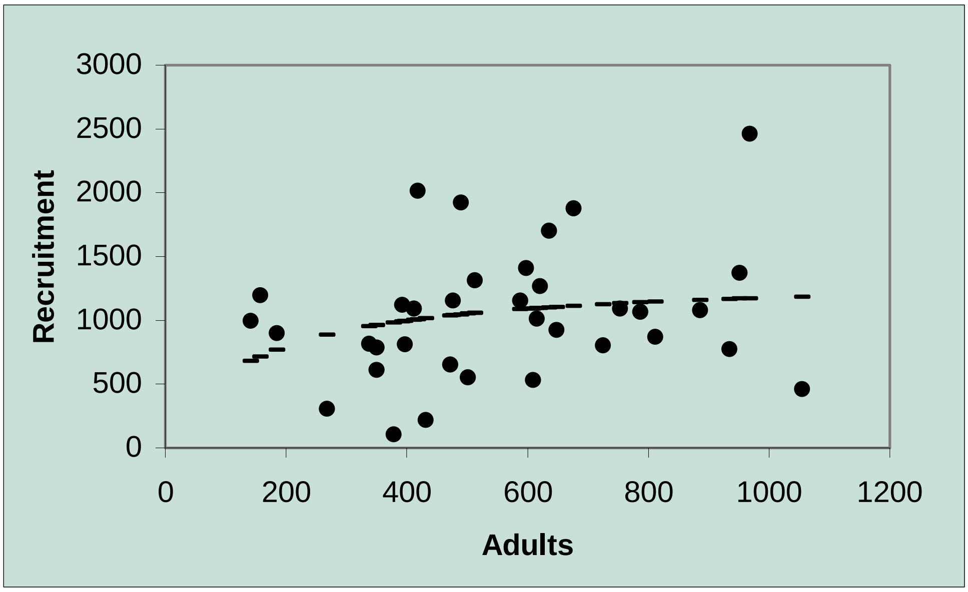

Figure 2 Beverton and Holt stock recruitment relationship fitted to estimated recruits and spawning stock number from the Collie-Sissenwine model.