Stock Assessment of Blue Crab in Chesapeake Bay
2011: Response to Reviewers

The 2011 Stock Assessment of Blue Crab in Chesapeake Bay was reviewed by a CIE panel on March 29-30, 2011 in Baltimore, MD. The review meeting involved a presentation by the assessment team to the panel on the first morning, followed by a day and half of detailed questions and answers involving the both the review team and the panel. In many cases the assessment team responded directly to questions raised by the panel by conducting additional simulation runs or analyses.

Each of the reviewers fully supported the approach taken in the assessment. They recognize that the assessment provides the "best available" science to support management of the blue crab fishery. There are no fatal flaws identified by any of the reviewers in their comments. We note that because of the nature of this specific review panel, there was no review panel chair and thus there is not a summary report. Yet, Dr. Addison sums up the reviewers comments when he writes,

"I would fully support the approach taken in the assessment. The derivation of the reference point calculations considers the sex-specific nature of the stock-recruitment relationship and is implemented correctly in the model. The exploitation and abundance reference points are meaningful indices which are understood readily by all stakeholders. In particular, the exploitation fraction (U) is a much simpler concept to understand than reference points based on fishing mortality (F). This is particularly important if management action is required when thresholds are exceeded or if targets are reached and agreement is needed on how to achieve the management objectives. I support the use of the US Federal definitions for the limit reference points and the use of the New England and Mid-Atlantic Fishery Management Councils' system to define target reference points. These reference points may seem a little arbitrary, but by their very nature, most reference points have a degree of subjectivity, and the authors have sensibly chosen to define their reference points along the lines of widely-agreed approaches."

However, each of reviewers does make important recommendations for immediate changes to the assessment document and recommendations for research that should be undertaken to improve future assessments. Here we respond to the specific points identified by the each individual review panel member in their CIE review. We address reviewer comments alphabetically in the order raised by each reviewer.

Julian Addison

Dr. Addison provided a 31-page review of the assessment. His recommendations are outlined in the review narrative and identified on a point by point basis in a list of 30 recommendations. There are no recommendations in the narrative that are not identified in the list of 30 recommendations.

A1. *Annual variations in growth rate in relation to the allocation of crabs from the surveys to age-0 and age-1+ age classes should be investigated to ensure that assumptions underlying the population model are correct.*

The review team recognizes that the allocation of crabs to either the age-0 or age-1+ category based on size introduces questions regarding the reliability of the size-age key. However, this issue affects the MD and VIMS surveys more than it does the winter dredge survey – for which the bimodal size distribution is remarkably constant from year to year. We do not believe this is an issue that affects the reliability of the current assessment.

The assessment team recommends that new research is conducted to explore the dynamics of blue crab size structure inter-annually as it relates to environmental parameters and population size.

A2. Sensitivity analysis of the assessment model outputs to the partial recruitment curve generated from the Puckett et al. study should be carried out.

The partial recruitment into the fishery likely varies inter-annually as a function of the growth of blue crabs during their first full summer when individual crabs grow rapidly from a carapace width of < 60 mm on emergence in April to attaining maturity by the autumn. In the base model we assumed that the partial recruitment of these age-1 crabs was 0.6. We addressed concerns raised by this reviewer by conducting sensitivity runs that are fully documented in the assessment. The three sensitivity runs include two runs with fixed partial recruitments of 0.3 and 0.9 which together with the base case fully bracket the range suggested in the Puckett et al. paper. We conducted an additional run in which we allowed the partial recruitment to be estimated. This simulation still used a constant partial recruitment value – but the specific value was estimated by the model. Model results from this run suggest the model performed best with a partial recruitment close to 1. We caution that this should not be taken to indicate that the true partial recruitment =1. We maintain that the base case assessment model with the partial recruitment =0.6 still represents the best available information for the true value of the partial recruitment.

The assessment team recommends that additional research to evaluate the utility of including a variable partial recruitment parameter be conducted.

A3. The population model should be modified to take into account the terminal moult in female crabs replacing the current assumption that all age-1+females contribute to the spawning stock.

The assessment team is not fully certain of the intent of this comment, but we believe that it relates to the fact that there is no age-specific decline in fecundity in the

assessment model. As currently implemented, the model assumes that were a female to survive to an age of 6 or 7, that female would have exactly the same fecundity as she experienced with her first brood. Empirical evidence from experience of raising crabs in the NOAA-funded crab hatchery in Baltimore and from field studies in other states suggest this is not the case.

The assessment team recognizes this deficiency and suggests that it could only be addressed by a division of the age-1+ category in the model into two separate stages – first time spawners and experienced spawners. Although this would indeed provide additional resolution of age-specific fecundity, it would do so at a cost of compromising our ability to identify age categories in the survey data. Given the high mortality rate of females, however, a very small proportion of females is expected to survive long enough to experience reduced fecundity.

The assessment team believes that a more resolved stage- or age-structured model should be a goal for a future assessment.

A4. The potential for sperm limitation in blue crabs should be revisited particularly in the light of recent new management measures to protect the female stock which have caused an increase in the male exploitation rates.

Sperm limitation is known to occur in several crustacean species and has been explored experimentally and empirically for blue crab in Chesapeake Bay. It is clear from these studies that there is the potential for sperm limitation for occur and the presence of sperm limitation should be a concern for managers responsible for the blue crab fishery. However, evidence suggests that sperm limitation is currently not limiting population productivity in the Chesapeake Bay system. Provided that managers follow the guidelines in the assessment of maintaining the sex ratio in the population within observed bounds, concerns over sperm limitation should not restrict the freedom of managers to act.

The assessment team recommends that resources be invested to understand the scope for sperm limitation in blue crab in the Chesapeake Bay. Research could take advantage for example of the natural trend in sex ratios up and down the bay to assess the impact of a reduction in male abundance on female reproductive success.

A5. Use of the Brownie model with a different parameterisation that includes M and F, or M and catchability and effort to analyse tag-recapture returns would be informative.

The assessment team recognizes the important role that the tag-recapture data can play in informing future assessments. Such data could provide a time series of estimates of M to which a future model could be fit, or more simply could inform and improve estimates of M. The assessment team fully supports this recommendation.

A6. Natural mortality rate (M) can vary annually, seasonally and with crab size. It is recommended that an analysis is carried out to assess the sensitivity of the model's output to variable M.

It is highly likely that the assessment will be sensitive to the value of M used or to the existence of interannual variability in M. However, we currently have few data that inform us as to the magnitude and pattern of that variability. Accordingly the assessment team recommends that research to address this recommendation be tied directly to the previous recommendation to document patterns in the magnitude and variability in M.

A7. A spatial component should be added to the population model or at the very least some spatial analysis of the survey data should be undertaken.

The assessment team considered development of a spatially-explicit model early on in the assessment process, but rejected it as a viable alternative because we have insufficient knowledge regarding the movement of blue crab. Lacking such information, inclusion of a spatial component risked becoming a model fitting exercise, allowing the model much more freedom to fit observed pattern, but unconstrained by any information that would allow us to evaluate the reliability of the movement patterns implied from fits to a spatially-explicit model.

The assessment team continues to believe that a spatially-explicit model should be a goal for future assessments. Spatially-explicit data are available from fishery-independent surveys, and increasingly from fishery-dependent sources. We recommend that simulation modeling be undertaken to explore trade-offs in changing the spatial and temporal resolution of the existing model.

A8. Recent increases in male exploitation rate and the increased potential for sperm limitation, suggest that male-based reference points might be appropriate, for example, in relation to abundance of age-1+ male crabs.

No explicit male-specific reference points were brought forward in the model. However, male-specific reference points are implicit in the model through our strong caution that the ratio of sex-specific exploitation rates be maintained within the envelope of values that have been observed empirically. Thus, in setting a female-specific reference point and in maintaining observed ratios of sex-specific exploitation rates, managers are implicitly managing to a male-specific reference point.

The assessment team notes that male-specific reference points could be formally developed if so desired. But in keeping with points made by this same reviewer in his full narrative we believe that such reference points should come out of discussions between managers and stakeholders rather than from the assessment process alone.

A9. When standardising survey indices using the delta generalised linear model (deltaGLM) method, interaction terms in the second stage of the delta GLM should be considered, especially strata and year, temperature and year for the dredge survey.

The assessment recognizes that an interaction term would be appropriate in the standardization approach. The Delta-GLM package that is widely used within NMFS for stock assessments only permits first order effects to be included in the model. This does not excuse the oversight.

This was the first time that any of the fishery-independent indices have been standardized for a blue crab assessment. We believe that the standardization effort did improve the information provided to the assessment model. The standardized indices appear to describe the data well and we do not believe that inclusion of the interaction term would induce a substantial change in our conclusions. However, we do recognize that the inclusion of interactions may have improved the information content of the surveys even more. Therefore, the assessment team fully endorses the continued use of standardization analyses for future assessments. In addition, the assessment team recommends that future assessments should continue to include interaction terms in any effort to standardize fishery-independent surveys.

A10. The winter dredge survey (WDS) is an essential component of the assessment and the evaluation of stock status relative to the reference points and its continuation is strongly recommended.

The assessment team fully supports this recommendation. The winter dredge survey is the most accurate and comprehensive index we have available for this population and it plays an essential role in managing the population in between benchmark assessments.

A11. The Virginia (VIMS) trawl survey should be standardized, and better use should be made of the additional four strata.

The VIMS trawl survey is the longest single fishery-independent survey available and as such provides invaluable information to managers and assessment scientists. For this assessment, we did not make as full use of this index as perhaps we should have done – and we accept the criticism implied in this comment.

We note that the VIMS trawl survey is longer and its design has evolved and is accordingly more complex than the other surveys. The assessment team recommends that a high priority future research objective should be the comprehensive analysis of the full VIMS trawl blue crab database. To our knowledge this has not been done since Chris Bonzek (VIMS) re-evaluated the size-age class categorization as a part of the Rugolo et al. assessment. The application of modern standardization approaches to these data, and the use of these data to re-evaluate the size-age convention (points A1 and A2 above) would be of high utility to future modeling efforts.

A12. WDS survey results for Virginia and Maryland should be evaluated separately to investigate whether they correlate with trawl survey indices in the two states.

Such analyses have been done informally, but were not presented in the assessment. In general, the age-1+ abundances from the WDS correlate well with VIMS trawl survey indices and less well for the MD survey. Age-0 abundances from the WDS do not correlate well with recruitment indices from either state – further reinforcing the conclusion in the assessment regarding the low information content of these surveys for recruits. Additionally analyses have also correlated the mature female abundance from the winter dredge survey with the abundance of mature females on the spawning grounds the following spring and summer from the VIMS trawl survey. There are discrepancies that are apparent in this comparison that suggest processes may be occurring between the time when the females are surveyed in the winter dredge survey in the winter and when the females move onto the spawning grounds later in the year.

The assessment team does not believe there is much to be gained from the direct comparison of WDS and state surveys as recommended by the reviewer. However, the assessment team does suggest that continued evaluation of the relationship between mature female abundance in the dredge survey and in the VIMS trawl survey is warranted.

A13. Other survey time series, such as the Chesapeake Bay Multispecies Monitoring and Assessment Program, should potentially be included in the assessment.

The exclusion of the CHESMAPP survey data by the assessment team was a regrettable error. These data should have been evaluated. We chose not to evaluate them because we believed, *a priori*, that the time series was too short. However, this is a weak excuse for the lack of inclusion of these data and they should have been evaluated.

Since the assessment meeting we have obtained, courtesy of Dr. Rob Latour (VIMS), the full CHESMAPP blue crab data set for use in a separate production modeling exercise for blue crab (Colton, A. R. 2011. An Evaluation of the Synchronization in the Dynamics of Blue Crab Populations in the Western Atlantic. MS Thesis. University of Maryland College Park, College Park, MD.). Colton used a similar assessment model to that used here but included four fishery-independent indices for the Chesapeake (WDS, VIMS, MD Trawl and CHESMAPP). Her findings suggest that the utility of the CHESMAPP survey data for blue crab assessments is currently low, but we expect it will increase.

The assessment team recommends that any future assessment of blue crab explore and to the extent possible include the blue crab data from the CHESMAPP survey.

A14. Consideration should be given to the development of an early life history phase recruitment index, e.g. larvae or megalopae abundance.

The assessment team recognizes that only the WDS appears to provide reliable information on the abundance of recruits and thus, a survey specifically designed to sample blue crab megalopae as they return to the bay would be useful.

The assessment team notes that under one aspect of the funding obtained for the assessment, new work is being undertaken to improve the reliability of the WDS as an index of abundance. In this research, the WDS is being expanded into shallow water using a smaller dredge. Options of the design of such a survey, and preliminary estimates of the catchability of the gear are being considered now.

A15. The scale of changes to the time series of catch data following re-calibration to take into account changes in reporting procedures should be independently verified, even if indirectly and/or anecdotally.

The assessment team appreciates the spirit in which this comment is made, but believes that such a direct assessment would be fraught with difficulties. In considering review comments provided by all three reviewers, we believe a better approach would be to have used the time series analyses only to identify whether reporting changes were necessary, but to use the assessment model itself to estimate the magnitude of the reporting change by simply adding a parameter to the model itself in which the magnitude of the landings is scaled by a common factor before and after a prospective reporting change.

We propose that future assessments should implement the catch correction within the assessment model rather than external to it. We believe this would be an improvement because the time series approach used here was in itself a form of simple stock assessment – in that it assumed something about the relationship between catch and abundance. Using the output from such a model in a second assessment model does not permit for the error structure in the time series model to be brought forward into the assessment model

A16. The assessment team should gain access to, and analyse, the time series of catch data from the Potomac Fisheries Commission.

The assessment used the Potomac River data as the data from the Potomac is partitioned out to the two states. However, the reviewer is correct to the extent that the spatial and temporal resolution of the catch data from the PRFC is much greater than in either of the states. To the extent we did not do this, this was an opportunity lost. In particular, because the PRFC maintains good catch and effort data, we did indeed loose the opportunity of developing a commercial CPUE index that could have been used as an additional input to the model.

The assessment team recommends that future assessments explore data from the Potomac River to ensure catch and effort data available from this system are fully utilized.

A17. Accurate and complete catch and effort data should be collected from the fishery so that a consistent standardised catch rate dataset can be used in the assessment. The potential existence of personal diaries of daily catch and effort data should be investigated.

There have been considerable improvements to the quality of catch and effort data since the first assessment by Rugolo et al. in 1997. One of the ToRs of the project that funded the assessment was to analyze the existing effort data to determine if there was sufficient information in the data to form the basis of a CPUE index. We failed to complete this ToR.

18. A baywide survey of recreational catches should be undertaken immediately and at regular intervals to quantify the overall landings from the recreational sector.

The assessment team agrees that surveys of recreational removals should be a high research priority. We believe that such data would greatly improve the data available to managers and stakeholders to make informed decisions.

19. Some quantitative assessment of under-reporting rates should be obtained.

The extent of under-reporting is not quantified. Apocryphal tales suggests the magnitude of the problem may be substantial in some areas and at some times. The assessment team agrees that this should be a high priority research item.

20. Incidental mortality and discard mortality rates should be ascertained for the pot, trot line and dredge fisheries and survival rates for the peeler/soft crabs held in tanks should be qunatified.

Although poorly quantified for all sectors, the assessment team does not believe that the discard mortality rate in this fishery is substantial. However, we do support improvements to our understanding of the dynamic of the soft and peeler fishery.

A21. Detailed investigation is required as to why the model provides such a poor fit to the sex-specific population sizes.

We held long discussions with the review panel over the tension in the model introduced by the empirical patterns in the sex ratios in the surveys and the catches. We have tried numerous model runs with different parameterizations to try to account for this pattern in model results. The bias could be resolved by setting substantially natural mortality rates for males and females, or by establishing a sex-ratio at recruitment far different to those observed empirically. Both runs were not considered realistic. Overall, we developed no deep insight into either what are the biases inherent

in the data, or whether there is an important interaction among processes that occurs in the field but that is not fully captured in the model. We have no reason to believe that there is any coding error in the model that causes this pattern. Currently, we believe the discrepancy may result from patterns introduced into the data because of the catchability corrections for the winter dredge survey. Accordingly we believe this is a data issue more than it is a model issue. The assessment team strongly supports the research recommendation proposed that a detailed investigation be undertaken to understand the sources of the discrepancy.

A22. A conceptual model providing a timeline of life history events such as spawning and recruitment should be incorporated.

It was clear from the beginnings of our discussions with the assessment review panel that we had not done an adequate job of explaining the timing of key life history events or of key fishery activities. For the assessment presentation we developed a conceptual diagram of the model which helped greatly in resolving some of these issues. This is now included in the assessment document as Figure 6.3.

A23. A number of minor changes are required to the formulation of the model. These include defining the initial conditions of the model, decrementing M by time and F by k (kappa) in equations 6.7 and 6.11, incorporating a bias correction and incorporating a compensatory function into the stock-recruitment relationship.

The review panel and assessment teams had numerous exchanges during the Q & A period related to these points. The assessment team has formulated the two equations in question under the assumption that the temporal pattern of fishing mortality and natural mortality are identical and constant. If we accept this assumption, then the changes are not required.

The changes recommended by the review panel imply that the temporal trends in natural and fishing mortality are different. The assessment team accepts that this may be the case. But we would argue we do not have an *a priori* expectation for how M and F differ that would be required to follow the review panel's advice. We recommend viewing this question as a research recommendation rather than a short term change needed for this assessment.

The assessment team is not quite certain over the final two suggestions of the bias correction and the compensatory function. We believe the reviewer is misguided in his suggestion of the need for a bias correction in estimates. This is normally done when estimates from the assessment model are used in a separate framework to estimate reference points. In this case our reference points are estimated within the assessment model itself and thus we do not believe bias correction is necessary (i.e., the bias correction to calculate the mean of the alpha parameter of the stock-recruitment relationship from the median is done in the reference point calculations).

We believe the request for a compensation term in the stock recruitment relationship relates to the observation that the current stock recruitment model predicts maximum stock productivity at a zero male abundance level – something that is clearly biologically absurd. We have implemented this in additional sensitivity runs included in the assessment document and Assessment Working Paper III. We note that the uMSY reference point from this combined sex stock recruitment model is identical to that developed in the sex-specific base model. We further note that the non sex-specific stock recruitment model recommended by the review panel is not free of absurd implications in the face of extreme sex ratios. Specifically, because the sex composition of the spawning stock is not specified in the non sex-specific stock recruitment model, the implicit implication that average recruitment will be observed from stocks comprised of a normal sex ratio, an all female population and an **ALL** male population.

The assessment team continues to believe that the sex-specific stock recruitment model used in the assessment is more appropriate for blue crab than the non sex-specific model. To address the review panel's concerns we are currently working on a modified version of the sex-specific model that imposes a recruitment penalty on increasing female dominated populations to reflect sperm limitation. However this refinement of the model is not yet fully developed.

A24. The chosen values for the overfished abundance threshold and the target abundance reference point should be re-considered.

It is not clear to the assessment team exactly what the review is asking in this comment. We did have extensive discussions with the review team over assessment philosophy and whether it was the responsibility of the assessment team to offer managers with a range of options for reference points or a single value. Following federal guidelines, the assessment team believes that it should provide strong guidance with regard to limit reference points – albeit recognizing that even here there are no absolute standards and that any recommendation reflects the assessment history and practice for individual stocks. However, we recognized more flexibility with regard to target reference points. Paraphrasing John Hoeing (VIMS), "Science tells you want you can't have, not what you want." In this light, threshold reference points provide scientifically-based guidance as to limits for abundance and exploitation. The target reference point is simply one suggestion, following federal guidelines of what society might "want," but it is not the sole possible value.

The assessment team is happy to work with managers and stakeholders to explore a wider range of target reference points if so desired.

A25. Consideration should be given to incorporating more biological realism and structure into the model including a size-based approach, shorter time steps, and a spatial component.

The assessment team fully supports this recommendation for any future assessments. However, we would argue that we have taken a substantial step in developing sex-specific reference points in this assessment. We believe it is currently premature to go to a fully spatially resolved, high temporal resolution model.

A26. A more systematic sensitivity analysis should be undertaken to understand how robust the model's output is to the various input parameters.

Although the additional simulation runs required to address this review comment were not in the initial draft of the assessment made available to the review panel, such simulations were conducted during the review meeting and made available to the review panel. The results of these simulations are integrated into Table 6.2 and are available in Assessment Working Paper 3.

A27. The model should be run with and without the modified time series of catches, when observation error and process error are not estimated simultaneously, and with alternative configurations of the sex-specific stock recruitment relationship.

We have addressed aspects of this comment above in A15. We believe a better way in the future to assess the likely impact of the reporting change is to do so internally to the model. However, we do not believe it would be appropriate to do so here because it would substantially change the reviewed document and assessment.

We do now present an alternative run of the model with a traditional, non-sex specific stock recruitment model. We continue to argue that although traditional, the non-sex specific stock recruitment model is not free of assumptions about stock productivity at extreme sex ratios. The sex-specific model makes the claim that reproduction can occur in the absence of males – whereas the traditional model makes the claim that it can occur in the absence of females or males.

A28. One of the key assumptions of the model is that the abundance of age-1+ crabs from the WDS is an absolute estimate of abundance. As the assessment model is not robust to this assumption, the panel recommended that a sensitivity test is run that uses the raw winter dredge survey indices and that the derivation of the catchability of the winter dredge survey index of abundance should be investigated and possible sources of this difference (both assessment and survey) should be investigated.

We had extensive discussions with the review panel over the claim that the WDS is an index of absolute abundance. The review panel is correct in noting that surveys that are indices of absolute abundance are rare. Although the panel is certainly correct that such surveys are rare, we do feel that a good case can be made that the survey catchability based on the number of stations and coverage for age-1+ crabs is near unity. A simulation model run in which the catchability for the winter dredge survey was estimated (Table 6.2 – Est WDSq simulation) resulted in estimates of uMSY

exploitation of almost twice that of the other model runs. We believe adopting this as the base run to be a non-precautionary approach.

However, the assessment team fully supports additional studies to characterize the catchability of age-1+ adult crabs in the winter dredge survey. An additional benefit that might accrue from such research would be insights into the sex-ratio at recruitment knowledge of which might help us explain the failure to achieve sex-specific sex ratios in the data. We note that one aspect of the research funded alongside the assessment was to do just that.

A29. Additional interpretation should be provided about the two changes that occur between the calculation of the 2005 reference points and those in the 2010 assessment.

We believe that the reviewer is requesting that we run the 2005 assessment model with standardized indices and the 2010 assessment with the raw indices (as used in the 2005 assessment model) to evaluate whether the change in the assessment model or the change in indices is most responsible for the change in reference points. We believe such simulations are not necessary. Given the close agreement between the simple survey means and the standardized index values for most age classes and surveys, it is clear to the assessment team that the change in the assessment model is most responsible for the change in the reference points.

A30. Harvest control rules should be developed to agree management actions if reference points are reached or exceeded.

We agree with the review panel that development of a management control rule would be a helpful addition to the framework for managing blue crab in the Chesapeake.

Dr. Cathy Dichmont

Dr. Dichmont provided a detailed review of the assessment. In her review, she identified two categories of recommendations: future work and recommendations. It is clear that the two recommendations in the future work section relates to medium- to long-term investments that may provide additional flexibility to assessment analysts working on blue crab in the future. However, while many of the forty recommendations relate to specific actions that she recommends are addressed in the current assessment, the assessment team believes that some of these also are best addressed as targeted research initiatives to be conducted over the next five years to prepare the ground for the next assessment. In addition there is considerable overlap between comments from Drs. Dichmont and Addison. Where appropriate we refer the reader back to comments made in response to Dr. Addison's points.

We comment on each below

Future work beyond recommendations

1. *This assessment would benefit from being size---based with a shorter time step – this would internalize the growth rate assumptions, allow in---season changes to be accommodated and also incorporate different gear types. Adding some spatial structure to the model should also be considered.*

   This recommendation derives directly from Dr. Dichmont's experience of decapod assessments in Australia. In some of her work, she utilizes a size-based model that runs on a monthly time step to provide management advice. The assessment team recognizes the spirit in which this recommendation is made and believes that it will be of use to explore such a model for the Chesapeake Bay. However, the assessment team also notes that model complexity is in itself not necessary just because it can be done – the Australian example uses in season management tools and thus requires a shorter time step. Were the three Chesapeake jurisdictions to consider in season management for blue crab, then such a model would potentially be required here. However, until that time, the higher resolution assessment model that Dr. Dichmont recommends may not be necessary.

2. *There are clear spatial---temporal changes in the data that have not been fully captured in the assessment. Other crab studies have shown the value of detailed spatial and temporal mining of the data. There would be value in undertaking a detailed and cohesive analysis of spatial and temporal dynamics of different population components on different surveys and their relationship with abiotic factors.*

   The assessment team fully supports this recommendation, but notes that some of these analyses are currently underway through the research awards that supported the assessment. Not specifically mentioned in Dr. Dichmont's recommendation, but one which we believe is also relevant is to conduct similar analyses of the commercial catch and effort data. Resolution of fine scale fishery-independent patterns will only be of use if they are matched with similar scale analyses of the fishery-dependent data. Both categories of analyses would be necessary for assessment model referenced above.

Recommendations

D1. It is unclear whether these different inputs of implicit age are internally consistent. Some evidence of this should be provided.

This review comment was largely addressed during the Q&A session with the review panel. It related specifically to concerns prior to the assessment meeting on behalf of Dr. Dichmont and the other reviewers over the reliability of the size-age cut off for blue crab. We provided detailed size distributions from the winter dredge survey that

allayed their concerns. This figure is now provided as Fig 3.1 in the assessment document

D2. A possible change in fecundity with respect to size should be monitored in terms of its impact on the assessment.

The assessment team fully support this recommendation. We note that a recent PhD dissertation at ODU provides evidence of such a change and continued monitoring of changes in stock productivity should be monitored.

D3. In the assessment model, fecundity is not weighted by age, whereas it is likely to be a consideration and quite a major factor within the assessment where the category age 1+ implies that all animals greater than age 1 are equally fecund even though 2+ animals are unlikely to breed. Splitting the model into three age classes should be considered.

This is a similar recommendation to A3 above and we refer back to our response to that comment in general.

D4. A research task for future assessments is to study age related fecundity and other related fecundity patterns for Chesapeake Bay.

The assessment team fully supports this recommendation. Given that the population size of blue crab in the Bay has experienced considerable variation over time, the potential exists for density-dependent changes in stock productivity. These should be monitored.

D5. In future, use the tagging data in the assessment to estimate female fishing mortality and natural mortality.

This issue was discussed in depth during the Q&A session. The review panel were made aware that we had indeed tried to include the time series of M derived from tagging directly into the assessment. However, the relatively short duration of this time series and the contradictory nature of the pattern in the data meant that these attempts were not successful – the assessment model effectively averaged through the time series and did not permit interannual variation to be expressed.

However, if the tagging efforts for mature female blue crab are continued, the time series will be very useful in future assessments. The assessment team fully support this recommendation.

D6. In the interim, modify the Brownie model to estimate survivorship and natural mortality through a parameterization that includes M and F, or M and catchability and effort.

The assessment team fully support this recommendation and will work to modify the application of the Brownie model to develop partitioned estimates of vital rates (see comment A5).

D7. Include the Baywide Multispecies Monitoring and Assessment Program in the assessment as a survey index of abundance.

This comment parallels A13 above and we refer back to answer to that recommendation.

D8. Incorporate interaction terms in the second stage of the delta GLM, especially with respect to strata and year, temperature and year for the winter dredge survey. This comment parallels A9 above and we refer back to answer to that recommendation.

D9. Include in the assessment model either using the four other Virginia trawl survey strata as an additional index of abundance or (better option) analyse the Virginia trawl survey and standardize into a single index.

This comment parallels A11 above and we refer back to answer to that recommendation. Dr. Dichmont provides additional specific guidance as to how the VIMS trawl survey data may be standardized.

D10. Conduct more detailed analyses of the dredge survey catchability estimates, especially with regard to availability or else undertake a spatial model of the survey.

This comment parallels A28 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D11. Consider applying consistent methods across jurisdictions.

This research recommendation relates to the operation of the winter dredge survey. The survey is conducted by two different agencies – MDNR and VIMS and slightly different approaches are used to how data are adjusted for catchability. This recommendation is a request for more uniformity in how data are collected and analyzed and the assessment team fully supports this recommendation.

D12. The two-stage component of the dredge survey is an extremely valuable component of the analysis and should be continued.

The assessment team supports this recommendation – the continued stratified random survey design with adaptive sampling in areas of high initial abundance should be continued to provide estimates of overwinter mortality. As the time series of relative overwinter mortality expands, it will be of increasing utility in the assessment.

D 13. The dredge survey is essential and should be continued.

The assessment team strongly supports the continuation of the winter dredge survey. It is an essential management tool.

D 14. Run a sensitivity test that uses the raw standardised winter dredge survey indices (not converted to absolute).

See D15 below

D15. Review the derivation of the absolute index of abundance in detail and investigate possible sources of the difference between the absolute and relative model runs (both in terms of the assessment and the survey). This is the highest short---term priority.

During the review, there were several discussions about potential for the expansion of the WDS catches to abundance to cause biases in estimates of sex-specific abundance (see D11). The assessment team fully supports this recommendation and sees it as a potential way to reconcile differences between the sex-specific catch and WDS indices.

D16. Use the daily Potomac data in the assessment, which seem to be very rich and useful.

This comment parallels A16 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D17 Independently substantiate these adjusted changes to the catch data through interviews of dealers/fishers or investigate whether there is a discrepancy within the spatial landings data as the fishery effort moves over time.

This comment parallels A15 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D18. Undertake model sensitivity tests with unadjusted catch data and an intermediate value of the reporting change factor.

This comment parallels A15 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D19. Estimate the reporting change factor (or intervention parameter) internal to the model (best option). If the model is unable to estimate this parameter, then an alternative would be to include the reporting change parameter (with variances) as an input value within the model so sensitivity tests of this parameter can be undertaken and the error in this variable could be included internal to the model.

This comment parallels A15 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D20. Investigate the potential of dividing the catch data into gear type or at least undertake a gear analysis.

The assessment team fully supports this research recommendation. Representing the different fishery sectors as separate fleets in the model has a lot of advantages. For example, we could reflect the different spatial and temporal dynamics of each fleet. However, we believe that such an expansion of the number of fleets in the model must be accompanied by a shift to the higher spatial, finer temporal resolution model recommended in Dr. Dichmont's long term research plan. We do not believe it is appropriate to undertake these analyses for the current assessment.

D21. Investigation should be carried out to establish whether it is possible to produce a standardised catch rate dataset that can be used in the assessment, as the index only needs to be representative rather than complete. Also, investigate the use of effort data within the assessment.

This comment parallels A17 above and we refer back to answer to that recommendation. The assessment team fully supports this research recommendation.

D22. Given the value of this fishery in the region, there should be more emphasis placed on collecting accurate and complete catch and effort data that are well aligned with the needs of the assessment.

There have been substantial improvements in the quality of information coming from fishery dependent sources. As noted in response to A17 above, the current assessment effort was amiss in not completing evaluation of the potential of the effort data that is currently available to provide a suitable CPUE index. It is a high priority item for the upcoming period. The assessment team fully supports this research recommendation.

D23. Either record the catch of the soft and peeler crab at the point of landing or record the length of time the crabs are kept in the tank and estimate the survival rate of these crabs.

The research team fully supports this research recommendation. Such data will be vital if the assessment moves toward the high resolution models recommended by this reviewer. Sources of catch and discard mortality in this sector of the fishery are poorly described and improvement in the information available would be beneficial.

D24. The conversion of catch in weight (bushels) to numbers should use the mean weight from the catch for each year both for the past and the future. This means that the mean size in the catch by year is required.

Dr. Dichmont is quite correct in this point, that the approach we took requires information on the mean size of crabs available in the catch. Such data are available in Maryland through their sentinel fishery program. Similar data are not routinely

collected in Virginia. We note that the data available for Maryland are not available for all fishery sectors and may not adequately reflect variation in the sizes of crabs collected in different regions of the Bay. Thus if we are to move to the high resolution assessment model as proposed by this reviewer, it is important that an adequate baywide sampling program be established to monitor the biological characteristics of the harvest. The assessment team fully supports this recommendation.

D 25. Undertake a baywide recreational survey as the resource is increasing and recreational effort may be increasing. This is a priority.

This comment parallels A18 above and we refer back to that recommendation. The assessment team fully supports this research recommendation.

D 26. Undertake sensitivity tests for larger recreational catches in the assessment.

The model utilized an 8% correction in landings to reflect the magnitude of the recreational catch. As noted in Section 5.4.2, the magnitude of the recreational catch is poorly described. As such, it could be argued that we should have explored the sensitivity of the model to this source of uncertainty. However, because individual fishery sectors are not recognized in the model – that is harvest occurs as a result of the actions of single composite fishery, the assessment team believes that sensitivity analyses to specifically examine the consequence of uncertainty in the recreational fishery as a single sector are not warranted. Because the model includes recreational catch as a constant proportion of the commercial catch, the effect of changing the assumed 8% is likely a proportional change in abundance (as is often seen in other assessment models).

We believe that a broader analyses of the sensitivity of the model to misreporting generally would be more appropriate and we did not undertake such an analysis. The assessment team recommend that analyses of the potential impact of misreporting on reference points and stock status be a high priority.

D 27. Undertake full sensitivity tests of the stock---recruitment relationship (following this into the reference point calculations) and also develop a function that captures the principle of what is presently formulated but does not optimise at zero males i.e. incorporate a compensatory function in the stock---recruitment relationship.

There remains considerable disagreement between the assessment team and the review team regarding the stock-recruitment model used in the assessment. The review team point out, quite correctly, that when taken to the extreme the stock recruitment model used in the assessment predicts continued recruitment when there are no males in the populations. As Dr. Dichmont notes, the model predicts the best thing to do is to catch all the males! This is clearly not biologically reasonable. We revised the assessment report to strongly caution that managers should implement policies to stay

within the empirically observed sex ratios – and that pursuing policies that cause the population to move to more extreme sex ratios is fraught with potential danger of overfishing males and thus reducing the stock's productivity.

In its place, the review panel recommended we implement a standard non sex-specific stock recruitment model. We have implemented this in additional sensitivity runs included in the assessment document and Assessment Working Paper III. We note that the uMSY reference point from this combined sex stock recruitment model is identical to that developed in the sex-specific base model. We further note that the non sex-specific stock recruitment model recommended by the review panel is not free of absurd implications in the face of extreme sex ratios. Specifically, because the sex composition of the spawning stock is not specified in the non sex-specific stock recruitment model, the implicit implication that average recruitment will be observed from stocks comprised of a normal sex ratio, an all female population and an **ALL** male population.

The assessment team continues to believe that the sex-specific stock recruitment model used in the assessment is more appropriate for blue crab than the non sex-specific model. To address the review panel's concerns we are currently working on a modified version of the sex-specific model that imposes a recruitment penalty on increasing female dominated populations to reflect sperm limitation. However this refinement of the model is not yet fully developed.

D28. Develop a conceptual model which would provide a much better description of the biological timeline and how this connects to the assessment.

This comment parallels A22 above and we refer back to answer to that recommendation. We have included an additional figure to reflect these changes.

D29. The decrementation in, for example, equations 7 and 11 etc. should index M with time of year and F by effort.

This comment parallels A23 above and we refer back to that recommendation.

D30. Split the partial recruitment parameter (that ties selectivity and growth) to include gear specific selectivity. The model should also overtly include selectivity of the gear (including how it changed over time).

The assessment team notes that the current assessment model does not represent separate fishery fleets in the model, but rather models removals as occurring from a single composite fishery. As a result this recommendation relates to Dr. Dichmont's long term recommendation that a size-based, high spatial resolution, fine temporal resolution assessment model be developed. Were such a model to be developed, then this research recommendation becomes not only relevant but imperative.

D31. For consistency, set the initial conditions correctly allied with the assessment implementation i.e. from 2 to 4 parameters.

This comment refers to assumptions made in the model about abundance in the four sex-age categories in the first year. The model currently estimates one parameter for the initial number of total recruits and another for the initial number of males and females. The total recruits are partitioned into males and females based on the observed sex ratio from the WDS (the same assumption that is used to partition recruitment for all later years). In contrast, no difference in sex ratio of fishing mortality rates is applied for abundance in the age-1+ category. Thus, the model assumes that abundance of age-1+ males and females is equal in the first year. The assessment team agrees that it would make more sense to estimate separate parameters for initial age-1+ abundance of males and females and that this should be pursued in the next assessment.

D32. Include the calculation of effective sample size for the winter dredge survey in the likelihood of the assessment.

We believe that this recommendation refers to the methods used to obtain the weights applied to the different fishery independent data sources in the model. The model uses a weighting approach where the fit of each index of abundance is weighted inversely with its standard deviation (SD; or log-scale SD). The SDs of the likelihood components of the three surveys used in the assessment were estimated differently. The likelihood components for the MD and VIMS trawls were estimated by iteratively changing the assumed SDs until the input SDs were approximately equal to the output SDs for each survey. The approach of estimating the SDs of the indices of abundance is often used when estimates of precision from the indices are not thought to fully represent the uncertainty in the index. However, the WDS was treated differently. Because of the large number of samples, the full sampling of the stock's range, and the high degree of interannual consistency in the distributions of catch each year, we used the design-based estimate of the variance in the likelihood function for the winter dredge survey. This essentially assumes that the survey design fully describes the uncertainty in estimated abundance. This does not require any correction for effective sampling size. The assessment team recommends that in the future simulations be conducted to assessment potential differences between results from using both approaches.

D33. Include the bias correction in the stock---recruitment relationship.

This comment parallels A23 above and we refer back to that recommendation.

D34. There would be some benefit in running sensitivity tests where the model is formulated as either process or observation error driven.

The model currently includes both observation error in the indices of abundance and process error in the catch data. The assessment team agrees that this is an important research recommendation that will help us understand more fully the tension that model experiences in trying to fit the survey and catch time series.

D35. Undertake a sensitivity test of the influence of the different indices as well as implement the dredge survey as a relative index.

The assessment team has undertaken specific additional sensitivity analyses that are included in the assessment document (Table 6.2) and in Assessment Working Paper III. These sensitivity analyses explored the impact of treating the winter dredge survey estimates of abundance as a relative index, rather than as an absolute index as currently used. The overall effect of assuming that the winter dredge is a relative index is to suggest higher reference points than are developed in the base model. We have included these results in the report, but have retained the base model because we believe it is a more accurate representation of the WDS and because it is more precautionary than the models that assume the winter dredge survey is a relative index.

D36. Run sensitivity tests of different input data and parameters and its effects on reference points. Consider systematic approaches such as FAST or a designed experiment.

The assessment team agrees that additional exploration of the sensitivity of recommended reference points to uncertainty in model inputs is required. It was an original goal of the assessment team to bring forward not only point estimates of the reference points, but also an indication of the uncertainty associated with each. We are committed to continuing work in this area and fully support the research recommendation.

D37. Provide more detailed fit statistics like q---q plots and residual plots for all indices.

We have added an additional index to the assessment document that now provides all outputs from the base run of the model. We provide individual annual estimates for each of the time series that were fit in the model and for key population parameters.

D38. Provide more choices of reference points rather than providing single values for each.

This comment parallels A24 above and we refer back to answer to that recommendation. In order for managers to make use of information on multiple reference points, additional research would need to be conducted to characterize the tradeoffs inherent in adopting a set of reference points and a harvest control rule. The assessment team fully supports this recommendation.

D39. Provide the option for more precautionary limit reference points.

The assessment team and the review panel spent considerable time in the Q&A session over the management philosophies in different regions. There are clearly no absolutes when it comes to reference points – with the exception perhaps of MSY – are to an extent arbitrary. For example SPR reference points are "more guidelines than a code." The assessment team followed the tradition that has developed in the Chesapeake region and in the mid-Atlantic generally of using MSY and 0.5*MSY as reference points. The review panel is quite correct to point out that there is no a priori reason why these management points should be selected. Indeed the assessment team strongly believes that target reference points should be negotiated among managers and stakeholders to reflect societal decisions regarding the best use of the resource. Specifically target reference points are not a scientific choice, but a socio-political one. Science can provide values following guidelines, but should not be the sole arbiter of the decision making.

The assessment team fully supports this recommendation.

D 40. The assessment document should be updated to include various corrections and clarifications as described in the review report.

The assessment team has made all of the changes recommended by Dr. Dichmont that relate to information not provided, or incorrect references within the document.

Dr. Billy Ernst.

Dr. Ernest provided 24 recommendations that were of both a short term and a longer term nature. Solely because Dr. Ernst's review is the third one we discuss, many of his comments have already been raised by the two previous reviewers. We detail our response to each of his comments below.

1. There is an appropriate description and mathematical representation of the stock assessment model in the main document; nevertheless it is necessary that the authors include some modifications in the final document:
a. A conceptual model should be included, which would provide a better description of the components, biological timeline and how this connects to the assessment, especially with regard to survey data. Main model assumptions should also be addressed.

We have addressed this concern with additional text in the assessment document.

b. The stock recruitment function should include the bias correction factor to avoid misinterpretation of recruitment parameters estimates.

This comment parallels A23 and D33 above. We refer back to these responses in addressing this comment.

c. Initial conditions of the population dynamic model should be clearly specified in the document.

We have addressed this concern with additional text in the assessment document (see also comment D31).

d. Appropriate documentation should be provided of the non-precautionary nature of the stock-recruitment function, particularly at low population stock size.

We have included additional text regarding the reviewers concerns over the stock recruitment model. We continue to defend robustly our choice of stock-recruitment function, but have including cautionary notes to managers in the assessment document regarding the risk of exceeding the observed ratios of sex-specific exploitation rates.

e. A better description is needed of the sensitivity tests presented in the document, including estimated parameter values and standard errors, likelihood values associated to each data component, and uncertainty around the main model outputs.
f. Sensitivity tests undertaken during the review should be described.

All sensitivity runs are now fully documented in Table 6.2 and in Assessment Working Paper III. Additionally we provide all output of the base simulation in Appendix V to the assessment document

g. Data points on the reference point figures (6.1 and 6.14) should be annotated with the years, as this would link better with the text.

We have changed the two reference point figures as requested.

h. A composite reference point figure should be provided, including: a) the 2005 method for standardizing the survey data and the 2005 reference points, b) the 2010 standardized data with the 2005 reference points and c) the 2010 standardized data with the 2010 reference points. This figure should be accompanied with an interpretation of the changes when moving from the 2005 to the 2010 assessment.

We have changed the assessment document as requested.

2. In many ways this fishery can be considered data rich and from all available data the winter dredge survey is the most important piece of information. It is used in the assessment and the evaluation of stock status relative to the reference points. It is important that this survey is continued.

The assessment team fully supports this recommendation. The WDS is essential to the ongoing management of the blue crab fishery in the Chesapeake Bay. This survey is the single highest priority for funding.

The following four points need no response

3. The structure of model (Age-0/Age-1+ and sex) seems appropriate considering the nature of the historical data. It also allows one to keep track of meaningful population statistics (spawning abundance by sex), which is of special relevance under recent changes to the management regime in the fishery.

4. The calculation of reference points is incorporated within the assessment model. This is a major advantage over previous models, because it assures consistency between statistics.

5. One inconvenience of collapsing the entire adult age structure into one age group is losing track of female fecundity at age. The implicit assumption is that female fecundity remains constant, which is probably rather unlikely to happen. Additionally it is unclear how this would influence the spawning stock size at MSY calculations.

6. In the assessment report there was an important effort to assess model structure uncertainty (results of two additional models were presented). Despite the intrinsic value of this exercise, a direct comparison of reference points between the models is not possible, because female spawning stock is not available for the other models. Nevertheless of the three assessment models provided, the best model is the SSCMSA.

7. On the other hand the parameter uncertainty of the proposed model was not fully characterized through the documented sensitivity tests, nor was the robustness of reference points estimates adequately presented. For example, test runs during the review highlight that the assessment model output is extremely sensitive to treating the winter dredge survey as a relative index of abundance. The fit was much better than the base case presented in the report. This is of great concern, as traditionally the default assumption is to treat surveys as a relative index of abundance. It is recommended that:

a) A sensitivity test is run using the raw winter dredge survey indices.
b) The entire calculation of survey catchability should be checked in detail and possible sources of difference should be investigated.

The assessment team has completed the additional sensitivity run requested and it is now included in the assessment document.

The assessment team remains concerned over the estimation of survey catchability for the winter dredge survey. There remain differences in survey protocols between the two jurisdictions. This is particularly unfortunate given that the ecology of blue crab is believed to differ along the north-south axis of the Chesapeake Bay – thus determining to what extent survey inferred difference relate to differences in the way the survey is conducted and the data are handled between the two states or to differences in the underlying ecology is difficult. The assessment team strongly recommends that an interjurisdictional team (possibly CBSAC) be assembled to review the operation of the WDS, the theoretical and practical foundation for the calculation of survey catchabilities and intergration of the two data streams. The assessment team believes this is of the highest possible priority.

8. The reference point calculations seem appropriate based on the structure of the new model. The authors followed the Federal reference points system to define the limit reference points and the New England and Mid-Atlantic Fishery Management Councils' system to define target reference points. This seems appropriate, but is in the end a policy decision. With respect to stock based reference points there is a value judgment about how precautionary the overfished limit and target reference point needs to be. Different values arise using different male to female F ratios, and coincidentally the one chosen in the report is not the most precautionary one.

Future work

9. This assessment lacks a catch per unit effort index of abundance, an index that captures the condition of the stock from the fisheries perspective. Future assessments should investigate the feasibility of generating a standardized catch rate index. This would ultimately be of importance in future economic analyses.

This recommendation parallels similar recommendations from the other reviewers (A17 and D21, D22). The assessment team fully supports this recommendation. We believe that several potential approaches might be fruitful in this regard:

a) analysis of catch and effort from the Potomac River Fisheries Commission – the PRFC has maintained accurate, spatially resolved catch and effort data in the Potomac River since 1963. CPUEs are routinely calculated from these data and provided to the Commission and its workgroups. However, these data have not been utilized within the assessment as they should have been.

b) Both states are now collecting mandatory daily reports from individual waterman. Although the time series is, at present, of insufficient duration to be of direct use in the assessment currently, that is not likely to be the case when the next assessment is conducted. Efforts should begin now to assemble and review these data

c) While catch and effort data from the two states may be too short to be of direct use in an assessment currently, they do provide extensive spatial information on the pattern of fishing on a daily and seasonal basis. It is likely that, were these data combined with data from fishery-independent surveys considerable insight could be gained in the response of the fishery to changes in the underlying distribution of blue crab.

The remaining 14 points are fully covered by responses to the comments of the two other reviewers – in our responses we refer back to our specific comments

10. The substantial difference between the model and estimated catchability coefficient of the winter dredge survey requires immediate attention. This is a high priority task.

This comment parallels A9 above and we refer back to answer to that recommendation.

11. The spatial component seems to play an important role in the population dynamics of different ontogenetic stages of Chesapeake Bay blue crab. A detailed data driven analysis (combining all of the pieces of information) of spatial and temporal dynamics of different stock components should be undertaken. Environmental covariates should also be factored in.

This comment parallels A7 and Dr. Dichmont's second long term priority above and we refer back to answer to that recommendation.

12. More work on the natural mortality parameter estimation should be undertaken using the available mark recapture data. Alternative Brownie model parameterization should be considered in the analysis, to explicitly model natural and instantaneous fishing mortality rates. Ultimately, the tagging data could be directly incorporated into a more integrated stock assessment estimation model.

This comment parallels A5 and D6 above and we refer back to answer to that recommendation.

13. Only a nominal index of abundance was constructed from the Virginia trawl survey data during this assessment. Index standardization should be undertaken. A second index of abundance can be potentially developed based on the winter portion of the Virginia trawl survey, which could provide information on the female spawning stock.

This comment parallels A11 and D9 above and we refer back to answer to that recommendation.

14. Interaction terms were omitted in the standardization of survey indices. It is recommended to develop additional models with interaction terms in the second stage of the delta GLM analysis.

This comment parallels A9 and D8 above and we refer back to answer to that recommendation.

15. In recent years some areas in the Chesapeake Bay seem to be experiencing higher recreational fishing mortality than reported in this analysis. It is recommended that a new baywide recreational survey be undertaken.

This comment parallels A18 and D25 above and we refer back to answer to that recommendation.

16. Despite the data-rich nature of this assessment, there is a major need for accurate and complete catch and effort information. Some of these factors are presented in the following list:

(a) the conversion of the catch in weight to numbers should be based on measured average weight in the catch, rather than using a constant factor. This means that mean size in the catch by year is required;

(b) a better and more direct estimate of soft and peeler crab mortality should be developed for the assessment;
(c) the adjustment of past changes in catch reporting requires some independent verification. The catch data is influential to the assessment results; (d) Anecdotal information suggests that incidental winter dredge and summer pot fishery mortality might be important. It is important to quantify these additional sources of mortality.

This comment parallels D22-24 above and we refer back to answer to that recommendation.

17. Additional model runs should be undertaken to assess the robustness of the results to:
(a) using the modified/original time series of catch;
(b) estimating the catchability coefficient of winter dredge survey for adult crab;
(c) using different pieces of information;
(d) not estimating observation and process error simultaneously;
(e) to free up the assessment model to allow for a better fit of extreme values. A sensitivity test might be organized in a systematic and comprehensive way, using appropriate techniques (i.e. FAST, designed experiments).

This comment parallels A26-28 and D34-36 above and we refer back to answer to that recommendation.

18. Include additional sources of information into the assessment such as Baywide Multispecies Monitoring and Assessment Program and Potomac daily catch data.

This comment parallels A13 and 9 above and we refer back to answer to that recommendation.

19. Adjust the reporting changes in the catch internal to the assessment model and input the intervention parameter so that this acknowledges some of the uncertainty.

This comment parallels A9 above and we refer back to answer to that recommendation.

20. Incorporate formally the different fisheries into the assessment. In addition reparameterize the partial recruitment parameter into formal selectivity and growth components.

This comment parallels A2 above and we refer back to answer to that recommendation.

21. The initial conditions of the model should have 4 as opposed to 2 parameters. If initial conditions stay in a 2-parameter configuration, the implicit assumption for the other 2 parameters should be consistent with the rest of the model configuration.

This comment parallels A9 above and we refer back to answer to that recommendation.

22. In equation 7 and 11, kappa should be multiplied by only F and M should be scaled by time.

This comment parallels A9 above and we refer back to answer to that recommendation.

23. Incorporate some kind of compensatory component into the stock recruitment relationship to avoid non-precautionary behavior of the model (i.e., depleting male population, while maintaining high recruitment rate).

This comment parallels A9 above and we refer back to answer to that recommendation.

24. A more complex version of the model might include:

(a) size structure;
(b) monthly time step to allow for the evaluation of in-season regulations; and
(c) spatial structure.
The model development could be frameworked into a MSE type of configuration.

This comment parallels Dr. Dichmont's first long term priority recommendation above and we refer back to answer to that recommendation.