Chesapeake Bay Blue Crab Stock Assessment Review 2011

29 – 31 March, 2011, Baltimore, Maryland

On behalf of the Center for Independent Experts (CIE)

Dr Julian Addison
Le Mas du Coq
946 Chemin du Puits des Gavottes
84300 Cavaillon
France

julian@masducoq.com

6 May 2011

Executive Summary

The 2011 assessment for the Chesapeake Bay blue crab fishery uses a sex-specific catch multiple survey analysis (SSCMSA) model. This is a major improvement on the 2005 assessment primarily because of the integration of the reference point calculations within the model using per recruit analyses. Previously, the population model and the calculation of reference points were two separate processes which were not necessarily consistent. The two main differences from the 2005 model are the inclusion of a renewal function in the form of a stock–recruitment relationship, and the incorporation of separate components for each sex, which allows the generation of sex-specific reference points. This is particularly important considering recent management changes which have been geared to conservation of the female stock and which have consequently changed the sex-specific exploitation rates.

The assessment is a two stage (age-0 and age-1+ crabs) model, which incorporates three fishery-independent surveys including the extensive winter dredge survey, which is treated as an absolute index of abundance. Two main concerns with the formulation and robustness of the model were identified. Firstly, the stock-recruitment relationship is based on both males and females, and whilst it is an innovative approach, the way in which the relationship is formulated within the assessment model suggests that some recruitment will always occur even if male population size is zero. Secondly, the assessment model appeared to be extremely sensitive to treating the winter dredge survey as a relative abundance index, an approach conventionally used when incorporating survey data. Such an approach appeared to have significant implications for estimating stock status, reference points and stock status relative to reference points, and therefore it should be a priority to investigate the basis for treating the winter dredge survey as an absolute index given that the best model fit is treating the survey as a relative index of abundance.

The derivation of reference points considers the sex-specific formulation of the stock-recruitment relationship and whilst the use of US Federal and Regional Management Council regulations to define the limit and target reference points is supported, the assessment makes a value judgement about the overfished limit and target reference point rather than give a range of potential values, and the value is not as precautionary as it could be. This is important because the model has been developed as part of a benchmark assessment of the stock and there will not be a new assessment each year. Instead in future years, empirical estimates of exploitation rate and abundance will be calculated from observed catch levels and abundance based on the winter dredge survey and these will be used to evaluate the status of the stock in relation to a control rule based on exploitation rate and abundance reference points. It is important therefore that any inconsistencies in the model structure are ironed out before the control rule is applied to future empirical estimates of exploitation rate and abundance.

A number of other areas were highlighted where the assessment could be improved. The assessment would benefit from the incorporation into the model of additional fishery-independent surveys, the inclusion of the Potomac Fisheries Commission time series of fisheries data, a consistent standardised catch rate dataset that can be used in the assessment, an updated assessment of annual levels and variation in recreational catch, and some quantification of incidental mortality in the fishery. There is also potential scope for refining and improving the model through the incorporation of size classes to more fully represent blue crab biology, time steps of less than one year, i.e. a quarterly or even monthly model, and possibly incorporating the information on soft/peeler crabs alongside the hard crab fishery as separate entities, and the incorporation of spatial structure in to the model.

Background

Blue crab is the most important commercial fishery in Chesapeake Bay. The first Baywide stock assessment (Rugulo *et al*., 1997) was conducted using a length-based approach to estimate exploitation, and an unweighted average of the four principal fishery-independent surveys to determine abundance. Consequently biological reference points were crude. In 2001, the Bi-State Blue Crab Advisory Committee (BBCAC) developed a new management framework that proposed threshold reference points based on maintaining 10% of the virgin spawning potential and on the lowest observed abundance, and a target exploitation rate that would lead to an effective doubling of the spawning stock present in 2001. In 2005 a second Baywide benchmark assessment for blue crab was conducted (Miller *et al*., 2005) and recommended adopting the exploitation fraction, defined as the proportion of crabs available at the beginning of the season that are subsequently harvested, in place of less intuitive measures (F) used in previous assessments. Estimates of exploitation fractions were calculated based on the Baywide winter dredge survey (WDS) and within a modified catch-survey analysis that permitted the use of multiple surveys. The approach used in the 2005 assessment was reviewed by a panel of international scientists with expertise in crustacean fisheries who found that it was a substantial improvement over previous assessments (Haddon, 2005) but also identified issues to be addressed in future assessments.

Since the 2005 assessment, the three management jurisdictions have implemented a range of regulatory changes aimed at attaining the target exploitation rate of 46% of the available stock. In 2009 the NOAA Chesapeake Bay Office initiated and supported development of another benchmark assessment and that assessment (Miller *et al*., 2011) is the basis of this CIE Peer Review.

Description of Review Activities

The CIE review for the Chesapeake Bay blue crab assessment took place in the Sheraton Hotel, Baltimore, Maryland on 29-31 March 2011. In addition on 28 March there was a special session at the National Shellfisheries Association Annual Meeting devoted to the results from the Blue Crab Advanced Research Consortium which provided interesting background to the fishery to the review panel but which was not formally part of the review process. Prior to the review, the NOAA Chesapeake Bay Office in conjunction with the University of Maryland, Chesapeake Biological Laboratory provided all the documents on a website for the Review (Appendix 1). Following a request for the code in a form that could be run in ADMB and R, these were kindly provided to the review panel via the website.

The objective of the CIE Chesapeake Bay blue crab review was to review and discuss input data, assumptions and parameters to the model, the assessment itself and the biological reference points for the population.

The Terms of Reference were:

a) Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.
b) Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

c) Describe and quantify patterns in fishery-independent surveys. Analyses should include an evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time series.
d) Describe and quantify patterns in catch, effort and survey-based estimates of exploitation by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.
e) Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.
f) Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.
g) Characterize scientific uncertainty with respect to assessment inputs and stock status.
h) Evaluate stock status with respect to reference points.

On the first day of the review meeting (29 March), Dr Tom Miller, the lead investigator on the assessment, gave a presentation covering all aspects of the above Terms of Reference (ToR), and there was a wide-ranging discussion between the assessment team, the review panel, other interested scientists and representatives from the relevant management jurisdictions. On the second day (30 March), the review panel discussed the research with the assessment team against each ToR. Throughout the review meeting, the whole assessment team was extremely helpful and flexible in critically discussing all aspects of the assessment and the ToRs of the review, which made the task of the review team immeasurably easier and more enjoyable. Upon request, the assessment team provided the following additional documentation, information and tests:

1. Correction of the lagged plots;
2. Extra residual and predicted versus observed plots;
3. An additional preliminary model run assuming that the winter dredge survey is a relative index of abundance.

The review panel met in private on 31 March to evaluate progress in relation to all the ToRs and in particular to evaluate the assessment and the consequent derivation of biological reference points for the population. The panel also met briefly on 1 April to review some additional material provided by the assessment team and to complete a summary report.

<u>Summary of findings</u>

The findings of the review are structured in line with the eight Terms of Reference (ToRs) provided to the review panel, although many aspects of the assessment discussed at the review related to more than one ToR, and this is reflected in the summary of findings below.

ToR a Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

Over the last decade, significant progress has been made in understanding the life history and vital rates of blue crab in Chesapeake Bay, which has improved the reliability of parameter estimates for input to assessment models. The initial question that is relevant to an assessment of the stock in Chesapeake Bay is whether the Chesapeake Bay population can be considered as a single stock. Genetics studies suggest that it is a reasonable to assume that the

Chesapeake Bay population is a single stock for fisheries management purposes, and whilst larvae studies suggest some limited exchange with other areas, they generally confirm this exchange is not expected to significantly affect population dynamics of blue crab in Chesapeake Bay. In comparison with most fish "stocks", I agree with the conclusion that the Chesapeake Bay stock can reasonably be considered as a separate stock.

In terms of the assessment model, the key biological parameters are the growth rates, reproductive parameters and natural mortality rates of the adults.

Growth rates. This seems to be a key area for the assessment in relation to the method used to allocate crabs in the surveys to the two stages (age-0 and age-1+ crabs) for the population model. In addition, these two stages form the basis of estimates of recruitment, exploitation rate and spawning stock reference points with age-0+ crabs providing exploitation rate reference points, whereas age-1+ crabs provide spawning stock /abundance reference points. An important question is how much in-year dynamics might influence the recruitment of crabs to the fishery, and whilst the assessment team were confident that empirical observations from the surveys show that the cut-off points for age-0 and age-1 crabs are very clear and do not vary significantly from year to year, I would recommend that annual variations are investigated to ensure that assumptions underlying the population model are correct.

Growth information comes from the use of estimates of age based on accumulation of lipofuscin in neural tissue pioneered for the blue crab by Ju *et al*. (1999, 2001) but significant questions concern the use of this technique. Firstly the technique has been criticised as being inappropriate for estimating the age of long-lived animals and a study on freshwater crayfish (Fonseca *et al*., 2003) showed that lipofuscin concentration could under certain circumstances decline with time (age of animal). In the case of the blue crab this should not be a problem because the species is relatively short-lived. In particular, however, there has been considerable disagreement about the methodology of measuring lipofuscin concentration by employing spectrofluorimetry of crude tissue extracts (Harvey *et al*., 2008; Sheehy, 2008). However the recent study by Puckett *et al*. (2008) using known-age blue crabs from a hatchery suggests that greater confidence can now be attached to the estimated growth rates and subsequent partial recruitment curves for blue crabs. Nevertheless there remains some uncertainty around these parameter estimates, and in particular I would recommend some sensitivity analysis of the assessment model outputs to the partial recruitment curve generated from the Puckett *et al*. study.

Reproduction. Little is known about age and size at maturity in blue crabs, but mating and spawning periods are well understood. Female crabs produce up to 8 clutches per breeding season, but viability is lower in later clutches, so most recruitment comes from the initial clutches. Larger female crabs produce larger, but fewer clutches, but no information was presented on size and viability of eggs in relation to the size of female crab. The larvae cycle is well understood and empirical observations suggest that the sex ratio at recruitment is 52% female.

Whilst there is a wealth of knowledge about the reproductive biology of blue crabs, and more is becoming available through a concerted research effort in conjunction with a hatchery and stock enhancement programme in recent years, reproductive parameters are handled relatively simply within the population model. In the population model, spawning stock estimates seem to be based solely on numbers of age-1+ female crabs, and there would appear to be scope to make this spawning stock calculation rather more sophisticated by including fecundity-at-size and hence producing more accurate estimates of egg production. At present my main criticism of the way in which the model is formulated is that it does not appear to adequately accommodate the assumed terminal moult in blue crab females. It is not physiological because the Y-organ does not degenerate, but appears to be functional because females don't appear to moult after being inseminated. At present the population model assumes that all female crabs of age-1+ are identical, when in fact with a terminal moult, females will mate once, not moult again and hence are unlikely to contribute further to the spawning stock/egg production. On that basis, the assumption that spawning stock is proportional to the number of age-1+ females is incorrect because the larger, older females will not be contributing to the spawning stock. I would recommend that the population model is modified to more accurately assess spawning stock.

Whilst there is potential for sperm limitation in blue crab, recent monitoring has shown that all mature females appear to have sufficient sperm. However, some recent unpublished PhD work has suggested lower clutch sizes in females, and so there is clearly a need to investigate further the potential for sperm limitation in blue crabs, particularly in the light of recent new management measures to protect the female stock and which are therefore causing an increase in the male exploitation rates.

Natural mortality rate estimates

The only feature of the adult life cycle that is incorporated into the assessment is an estimate of natural mortality rate (M). Major effort has been put into estimating M since the first assessment in 1997 when M was assumed to be 0.375 (Rugulo *et al*., 1997). The previous stock assessment (Miller *et al*., 2005) used a range of values of M from 0.6 to 1.2, but also included the previous estimate of 0.375.

Since the last assessment, Lipcius and Smith (2011) have analysed returns from a tag-recapture study of mature female crabs using a Brownie-type model (Brownie *et al*., 1985) to estimate average annual survival rates of 0.15 varying from 0.09 to 0.28. Estimates of maximum age were then used with Hoenig's (1983) empirical relationship to give estimates of M ranging from 0.65 to 0.79. (A similar estimate of M could be calculated for the maximum age of crabs identified through the lipofuscin-based ageing approach, but this was not possible because the lipofuscin studies were undertaken at a time when only small, young crabs were available for sampling.) It would be instructive to investigate how the various components of mortality contribute to this estimate of 15% survival. As an initial step, use of the Brownie model with a different parameterisation that includes M and F, or M and catchability and effort would be informative. This would be a better method of calculating M than the method of using the maximum age calculated from the tagging experiment.

In comparison with most fisheries, the assessment team have collated a huge range of information on estimating the natural mortality, which provides added confidence that they are using a plausible range of values for M. The population model does however assume that M is stage and sex-independent and constant. Whilst there is no evidence that blue crabs are susceptible to very high incidences of mortality due to disease such as Hematodinium infection, a paper presented at the special session of the Blue Crab Advanced Research Consortium on 28 March (Kramer *et al*., 2011) showed that 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.

In conclusion, major recent advances have been made on understanding the life cycle of the blue crab and its variation through the Blue Crab Advanced Research Consortium and the hatchery and stock enhancement programme, yet the population model used in the assessment remains relatively simple. For example, there is significant spatial structure of the stock within the whole bay/estuary, both east-west and north-south, in terms of habitat differences, mortality rates, prey and predator abundance, rates of cannibalism etc., and also in terms of movement of crabs around the bay including migratory corridors of mature females. Yet none of that spatial variability is incorporated in the model. I recommend therefore that a spatial component be added to the model or at the very least some spatial analysis of the survey data is undertaken. As with most hatchery and stock enhancement programmes, it remains to be seen whether enhancement could be a viable alternative to conventional fisheries management measures, or at the very least whether enhancement programmes could form part of an integrated management scheme. However, there is no doubt that the hatchery and enhancement programme is making major advances in the understanding of blue crab biology and ecology.

ToR b Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

Reference points for the Chesapeake Bay blue crab fishery have evolved significantly since the first baywide assessment in 1997, and until now reference points have been estimated independently of the analyses used to assess exploitation rates and population abundance. The 2005 benchmark assessment derived reference points using the exploitation fraction (U) defined as the proportion of crabs available at the beginning of the season that are subsequently harvested. The overfishing definition equivalent to maintaining 10% of the virgin spawning stock was estimated as $U_{threshold}$ = 53% of all available crabs, and target exploitation rate (equivalent to maintaining 20% of the virgin spawning stock) was estimated as U_{target} = 46%. These reference points were used for management from 2006 onwards, but recently fishery managers, who were increasingly reliant upon abundance estimates from the winter dredge survey, became uncomfortable about the emphasis on exploitation rates alone, and so they adopted additional reference points of an abundance threshold of 86 million age-1+ crabs and an interim target abundance of 200 million.

However there has been concern that the reference points based on exploitation rate and abundance may not be consistent, so a key element of the latest (2011) assessment is the development of a fully integrated analysis that estimates both reference points and stock status. This is a major advance on the approach taken in previous benchmark assessments and is a real strength of the new assessment model. In addition the model permits the derivation of sex-specific exploitation rates and the authors recommend that all exploitation reference points should be based on age-0+ females, as some of the age-0 crabs will recruit to the fishery during the season. Secondly, all abundance-based reference points should be based on age-1+ females as these represent spawning stock and can be accurately measured during the WDS. This approach is fully supported.

The following reference points were defined from the 2011 assessment -

(1) Overfishing limit is defined as the exploitation rate of age-0+ female crabs that corresponds with MSY

(2) As blue crab is considered a data poor species (because there is no uncertainty attached to this estimate) and in line with regional Fishery Management Councils' regulations, target exploitation rate is set at 0.75 x U_{msy}.
(3) Overfished abundance threshold should be 0.5 x N_{msy}
(4) Target abundance reference point should be the equilibrium abundance when target exploitation rate is achieved, i.e. $N_{0.75 \text{ x } msy}$

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.

I shall leave discussion of the actual values for these reference points to a later section as they sit more easily alongside the results of the assessment under ToR e. I do have some concerns about the way in which the threshold and target abundance reference points have been estimated as the assessment team has made a choice that is normally the role of fisheries managers or policy makers, but again I will leave this discussion to ToR e.

Finally, I note that all reference points are now framed in terms of abundance of females, but with new management measures restricting the exploitation of females, recent increases in male exploitation rate and the increased potential for sperm limitation, it would be worth considering whether some male-based reference points might be appropriate, for example, in relation to abundance of age-1+ male crabs.

ToR c Describe and quantify patterns in fishery-independent surveys. Analyses should include evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time-series.

The blue crab assessment benefits from having a series of long-running fishery-independent surveys. Three of these surveys, the VIMS trawl survey of the southern end of the Bay, the Maryland trawl survey of the north-eastern area of the Bay and the Baywide winter dredge survey (WDS), are used in the assessment, and other shorter time series survey data are available but were not used in the assessment. In the assessment, the abundance of age 1+ crabs in the WDS was assumed to be an absolute index of abundance, and other stages in the WDS and the other two surveys were considered as relative abundance indices.

In response to the peer review of the 2005 assessment, the authors adopted a delta generalised linear modelling approach (GLM) to obtain standardised indices of abundance for the three surveys. For all three surveys, the analysis showed that the best fitting models (lowest AIC) included survey design and environmental factors. For the WDS the best fitting models for males, females and crabs of both sexes included all design and environmental factors. The use therefore of standardised abundance indices is a significant improvement on raw abundance data from the surveys. Standardisation of survey indices using the delta generalised linear model (deltaGLM) method is appropriate. However it is recommended that interaction terms in the second stage of the delta GLM be considered, especially strata and year, temperature and year for the dredge survey. However even after standardising the data, the correlations between abundance of age-0 crabs in year t with age-1+ crabs in year t+1 is relatively poor, suggesting that the allocation of crabs to age classes may not be as good as it might be due to inter-annual variation in growth rates.

It is clear that the fishery-independent surveys are a vital component of the assessment of the Chesapeake Bay blue crab stock. In particular, the WDS is a very important survey because it is an extensive baywide, stratified random survey, it takes place when the temperature is less than 10 degrees C and so crabs are buried in the substrate at these temperatures, year and vessel specific catchability are calculated for each year, and return visits are made to high density stations to estimate winter mortality. I would therefore strongly recommend continuation of the WDS as an essential component of the assessment and the evaluation of stock status relative to the reference points.

The other two fishery-independent surveys are of less importance, but I believe that more could be made of these survey data. For example, the Virginia (VIMS) trawl survey was not standardised and only 3 of the 7 strata were used in the trawl survey Index. The Virginia trawl survey should therefore be standardised and there are two options for making best use of the available data. The four strata that are not currently used could either be incorporated into a single index for the whole geographical area covered by the VIMS survey, or used as a second independent time series. The latter may be preferable because it would place greater emphasis on the winter component of the survey portion and potentially provide information on spawning abundance). Secondly, it should be possible to look at the WDS survey results in Virginia and Maryland separately and see how they correlate with trawl survey indices in the two states. There are also other survey time series, such as the Chesapeake Bay Multispecies Monitoring and assessment Program, which should be examined in more detail and potentially included in the assessment, and I would particularly recommend investigation of an earlier life history phase recruitment index, e.g. larvae or megalopae abundance.

ToR d Describe and quantify patterns in catch and effort by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

The blue crab fisheries in Chesapeake Bay are managed by three separate jurisdictions, Virginia Marine Resources Commission, Maryland Department of Natural Resources and the Potomac River Fisheries Commission and each management jurisdiction has historically evolved slightly different methods for collection of catch and fishing effort data, as well as having varying management legislation. Whilst pots, trot lines and dredges have been used consistently in the past, potting is the principal method of fishing now. Whereas biological and survey data for the fishery are excellent, there are significant doubts about the quality of landings and effort data from all areas.

One of the key problems in obtaining an accurate time series of catch and effort data has been changes in reporting procedures in both Virginia and Maryland over recent years. Following the study of Fogarty and Miller (2004), the current assessment reconstructed commercial landings for Virginia and Maryland using time series analysis. Results show an impact of the 1993 reporting change in Virginia, suggesting that landings were under-reported by a significant amount prior to 1993, and this time series of adjusted landings suggests a significant recent decline in landings. There were also changes in reporting procedure in Maryland in 1981 and 1993 but analysis suggested that only the 1981 intervention was found to be significant. Again re-constructed time series suggests major under-reporting prior to 1981. Time series analysis of catch and effort trends is the standard accepted way of analysis to examine any effects of reporting changes. However the analysis of adjusting the catch to past reporting changes is very influential in the catch series used in the assessment and has a significant impact on the evaluation of the status of the stock. As it is possible that other concurrent events may have led to an increase in landings at the time of the change in reporting procedures, it is important that the scale of these changes be independently verified, even if indirectly and/or anecdotally.

The Potomac Fisheries Commission has the best time series of fisheries data because they have not been compromised by any changes in reporting procedures. However these data have not been analysed as part of the assessment because they are still on old style format discs. I strongly recommend that the assessment team gains access to, and analyses, these data and compares trends across other two states' surveys and the WDS.

There are a number of other areas where I would recommend further detailed investigation in order to provide a better understanding of removals from the fishery and potential areas for incidental mortality due to the fishing process.

First and foremost, I would recommend that accurate and complete catch and effort data are collected from the fishery so that a consistent standardised catch rate dataset can be used in the assessment. Such an index needs only to be representative, and it would be worthwhile investigating whether individual fishers had kept personal diaries of daily records of catch and effort. In addition, it would worthwhile reviewing the collection of size and sex data as sex-specific catches often have to be extrapolated across areas.

The assessment assumes that an additional 8% of commercial landings are due to the recreational sector, but recent studies by Aguilar *et al*. (2011) and Roberts *et al*. (2011) suggest that recreational harvest rates vary across sub-estuaries, and in some areas may be as much as 50%. I would recommend therefore that a baywide survey of recreational catches be undertaken immediately and at regular intervals to quantify the overall landings from the recreational sector. Similarly, it would be helpful to gain some quantitative assessment of under-reporting rates.

Finally, there are a number of aspects of the blue crab fishery that highlight the potential for large incidental mortality. The summer dredge fishery (currently closed) undoubtedly caused significant mortality to crabs, and there may also be high mortality rates during handling in the summer pot fishery. Some quantification both historically and in the current fishery could be included in the assessment calculations. The soft and peeler crab catch does not record catches at the point of removal from the water, but is instead based on landings of animals that have *survived* being kept in a shedding tank, and so the animals that died during this process have not been considered as a component of the catch. This requires either a) that the catch is recorded or b) the length of time they are kept in the tank is recorded and consequently the survival rate of these crabs is estimated. It would also be instructive to obtain quantitative estimates of discard mortality rates and any potential “ghost fishing” due to lost pots.

ToR e Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.

The 2011 assessment for the Chesapeake Bay blue crab fishery uses a sex-specific catch multiple survey analysis (SSCMSA) model. I consider that this is a major improvement on the 2005 assessment which used a catch multiple survey analysis (CMSA) model. The main differences from the 2005 model are the inclusion of a renewal function which allows recruits in year t+1 to be generated from abundance of adults in year t, so management reference points can now be generated in relation to both spawner per recruit and MSY. Previously, the population model and the calculation of reference points were two separate processes which were not necessarily consistent. Secondly, the model incorporates separate components for each sex, which allows the generation of sex-specific reference points and which is particularly important considering recent management changes which have been geared to conservation of the female stock and which have consequently changed the sex-specific exploitation rates.

The model uses standardized time series of fishery independent abundance and is fit to time series of total and sex-specific catches. In general, the model performs relatively well. The model replicated time series of total catch and the time series of the winter dredge survey (WDS), and the Virginia and Maryland trawl surveys. The model assumes that the abundance of age-1+ crabs in the WDS is an absolute index of abundance but catchability is estimated for other surveys. The best model fit estimated catchability of 0.4 for age-0 crabs in the WDS. However the model does not provide a particularly good fit to some of the data. The main areas of “lack of fit” are an inability to capture the high recruitment levels in many years, and secondly the model consistently underestimates the female population size and overestimates the male population size irrespective of input parameters. I would recommend detailed investigation into why the model provides such a poor fit to the sex-specific population sizes.

The model structure is considered to be appropriate (although see reservations below) but is not very well documented – both in terms of parameters, but also in terms of how it deals with life history events. I recommend that a conceptual model is presented before the mathematical components to allow readers to evaluate the assumptions of the model in relation to spawning, recruitment timescales etc.

The most important new aspect of the model is the introduction of a stock-recruitment relationship based on both males and females (equation 6.5). Whilst this is an innovative approach, this equation seems to suggest that some recruitment will always occur even if male abundance is zero! At current stock levels, this two sex SRR is acceptable, but we note that at extreme stock levels, the model does not perform in a realistic manner, and this should be fully documented in the assessment.

In addition there are a number of minor changes required to the formulation of the model identified by the review panel such as 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.

Determination of reference points

The derivation of the reference points estimated from the assessment model have been discussed previously under ToR b, so in this section I will confine myself to discussion of the generation of the specific values below estimated for each of the agreed reference points.

- The overfishing limit is the exploitation rate of age-0+ female crabs that corresponds with MSY, U_{msy} = 0.34 based on a catchability of age-0 crabs of 0.4, and partial recruitment to the fishery of age-0 crabs of 0.6. This estimate of U_{msy} is relatively invariant to the ratio of the sex-specific exploitation rates.
- As the blue crab is considered to be a data-poor species and in line with regional management council regulations, the target exploitation rate is estimated at 0.75 x U_{msy} = 0.255
- The overfished abundance limit/threshold is calculated as 0.5 x N_{msy} =70 million age-1+ females. This figure of 70 million is taken from the lower end of the range of 68-82 million females estimated for varying sex-specific exploitation rates.
- Target abundance is defined as the equilibrium abundance when the target exploitation rate is achieved, i.e. $N_{0.75 \text{ x } msy}$ = 215 million age-1+ females. Again this figure of 215 million is taken from the lower end of the range of 210-340 million dependent on the assumptions about sex-specific exploitation patterns.

As stated under ToR b, I believe that the assessment team has taken a very sensible approach to defining reference points and this approach should be supported. For the abundance limit reference point, Regional Management Council guidance is to set this at 0.5 x N_{msy}. This is appropriate, but where I question the assessment team's approach is the choice of this limit reference point based on the range of sex-specific exploitation patterns assumed within the model. Figure 6.12 shows that N_{msy} occurs at a range of values dependent on the assumption underlying the sex-specific exploitation rates. For the observed range of relative exploitation rates, 0.5 x N_{msy} falls between approximately 68 and 82 million age-1+ crabs. However instead of giving this range of values from which fishery managers could choose their reference point, a more appropriate approach in my opinion, the assessment team have themselves made a decision to accept a value at the lower end of this estimated range at 70 million age-1+ crabs. Whilst it is open to discussion as to whether this is the assessment team's role, my main criticism is that this value of 70 million age-1+ crabs is in fact at the lower end of the predicted range, and therefore cannot be considered precautionary. A fully precautionary value for the overfished abundance threshold would be something around 82 million age-1+ crabs because this would be the estimated value for the highest sex-specific exploitation rates observed previously (a ratio of F_{male} to F_{female} of 1.6). In defining the target abundance, the authors have again chosen a specific value for $N_{0.75 \text{ x } msy}$ instead of providing managers with a range of options dependent on assumptions about the ratio of sex-specific exploitation rates.

Alternative stock assessment models

In addition to the SSCMSA model the assessment team have also provided the results of fitting two alternative assessment models. Production models fitted to time series of abundance and effort data in Working Paper 3 (Miller 2011a) suggest similar yields to the SSCMSA model, but abundance reference points were significantly greater than those from SSCMSA. Re-running the 2005 CMSA model (WP3, Miller 2011b)) did capture the trends in abundance of age-1+ females, but predicted much higher values of the exploitation rate (U). The assessment team should be congratulated on providing alternative assessment models, but the review panel agreed that differences in model structure in that female spawning stock size is not calculable from the production and CMSA models, comparing these with the new reference points is not possible, and concluded that the SSCMSA model was the best of the three models.

Finally, I would recommend that in the future consideration is given to incorporating more biological realism and structure to the model, particularly as so much more information is emerging from the major research programme carried out over the last 10 years. Potential options include incorporating shorter time steps such as quarterly or even monthly time steps to more explicitly incorporate timing of life history events, making the model size-based for better incorporation of fecundity–at-size and providing a better approximation of the reproductive status of females that undergo terminal moult, incorporating a spatial component to the model because of the widely-observed spatial differences, such as the very different nature to the fisheries in Maryland (primarily males) and Virginia (primarily females), and possibly incorporating the soft/peeler crab fishery alongside the hard crab fishery as separate entities.

I note that this is still very much a single species model and ultimately ecosystem-based models should be developed for the Chesapeake Bay blue crab population. However I would not recommend such a development as a priority at the current time.

ToR f Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.

The assessment notes that temporal patterns of abundance (N) and exploitation rate (U= catch (C)/abundance) over time are mirror images of each other suggesting depensatory exploitation. The possibility of depensation was examined in detail in a working paper by Lipcius (2011). This elegant analysis calculated U from the annual estimates of baywide abundance (N) and the landings (C) from the three published Fisheries Commission statistics. His analysis showed that a depensatory relationship is likely with higher values of U at low population abundance, and so fisheries managers may need to consider additional management measures at low abundance.

Other density-dependent processes have been addressed in detail as part of the Blue Crab Advanced Research Program. For example, Johnson *et al*. (2011) observed density dependent movement of larvae, and also density dependent mortality attributed to cannibalism. Such density-dependent processes should be incorporated into the population model.

ToR g Characterize scientific uncertainty with respect to assessment inputs and stock status.

The assessment document provides a minimum of sensitivity analysis of the output in relation to the input parameters M, partial recruitment and sex ratio. I recommend therefore that a more systematic sensitivity analysis is undertaken to understand how robust the model's output is to these and other input parameters. In addition, new information is emerging about variation in some of these parameters. For example, Kramer *et al.*'s (2011) study showed annual, seasonal and size–related variations in natural mortality which should be incorporated in the model.

More importantly, however, I would like to see additional runs to investigate the sensitivity of the model to various aspects of the model's configuration and some of the key assumptions underlying the model. These might include running the model with and without the modified time series of catches, relaxing the assumption that abundance of age-1+ crabs are an absolute estimate of abundance, runs where observation error and process error are not estimated simultaneously, and alternative configurations of the sex-specific stock recruitment relationship.

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, whereas the model estimates catchability of age-0 crabs from this survey. In most stock assessment models, it is conventional to assume that all survey indices are relative rather than absolute indices of abundance, and so the panel requested the assessment team to carry out additional runs where catchability of the age-1+ crabs was estimated from the model rather than assumed to be 1.0. Preliminary runs of the model showed a much better fit to the sex-specific abundances, but still failed to predict the high levels of recruitment. Estimates of catchability appeared to be very different for this model than the baseline run and the catchability estimates may change the interpretation of the reference points and trajectory of the stock. The model is therefore extremely sensitive to treating the WDS as a relative index of abundance in terms of absolute stock status, reference points and stock status relative to reference points. 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.

I recommend therefore that a full investigation of the behaviour of this revised model and its outputs are undertaken immediately.

ToR h Evaluate stock status with respect to reference points.

The SSCMSA model allows an integrated approach to estimating stock status and reference points, and therefore there is now consistency between the exploitation rate and abundance reference points. Looking at the time series of exploitation rate and abundance, the revised control rule (Figure 6.14) shows that the crab fishery was overfished in 1998-2004, but by 2009, the exploitation rate was below the $U_{0.75 \times msy}$ reference point and abundance was above the overfishing threshold, but below target abundance. WDS abundance in 2010 was well above this target abundance.

I think that the presentation of this section of the report could be improved to help the reader through the conclusions drawn on stock status in relation to the estimated reference points. Firstly the points on Figures 6.1 and 6.14 should be denoted by the years to which they apply, as this would add to the description in the text of how the stock status has changed over the last 20 years or so in relation to the calculated reference points. Secondly there are two changes that occur between the calculation of the 2005 reference points and those in the 2010 assessment. The panel agreed that it would be very helpful to have an additional figure as follows – Figure 6.1 shows the old method for standardising the survey data and the old reference points and it would therefore be helpful to have an intermediate figure showing the new standardised data for surveys with the old reference points before moving on to Figure 6.14 which shows the new standardised survey data with the newly calculated reference points. This would provide an interpretation of the changes when moving from the 2005 method to the 2010 method.

I note that the population model has been developed as part of a benchmark assessment of the Chesapeake Bay blue crab stock and that there will not be a new assessment each year. The authors recommend that in between these benchmark assessments, the status of the stock with regard to the reference points such as Umsy can be estimated from the empirical estimates of age-0+ female exploitation based on reported female harvests and the abundance of age0+ females in the winter dredge survey calculated using the SSCMSA estimates of q from the model. It is important therefore that any inconsistencies in the model structure are ironed out before the control rule of Figure 6.14 is applied to future empirical estimates of exploitation rate and abundance. This means that the assessment team should reconsider their recommendation about the threshold and target abundance reference points in relation to variable sex-specific exploitation rates, and in particular need to fully examine the sensitivity of the model to the relaxation of the assumption that the WDS is treated as an absolute abundance index.

This will be particularly important because the stock appears to be responding to female only management measures, such as the termination of the winter dredge fishery, which specifically target conservation of the female part of the stock.

Finally whilst not necessarily the role of the assessment team, I would like to see the development of some harvest control rules which kick in if and when reference points are reached or exceeded. At present, fishery managers have a methodology for evaluating the status of the stocks in relation to reference points, but no guidelines on how to respond if those reference points are breached.

Conclusions and Recommendations

The 2011 assessment for the Chesapeake Bay blue crab fishery uses a sex-specific catch multiple survey analysis (SSCMSA) model. This is a major improvement on the 2005 assessment primarily because of the integration of the reference point calculations within the model using per recruit analyses. Previously, the population model and the calculation of reference points were two separate processes which were not necessarily consistent. Two significant concerns with the model were identified. Firstly the formulation of the stock–recruitment relationship implies that it is possible to maintain recruitment even at male populations size of zero, and secondly the model is extremely sensitive to treating the winter dredge survey as a relative abundance index (as opposed to an absolute index of abundance as assumed in the baseline run). These concerns need to be addressed along with some value judgements about the precautionary nature of the chosen reference points.

The extensive biological knowledge of the species and the wealth of fishery-independent survey data suggest that the fishery is essentially data rich, and there are numerous ways described below where the assessment could be improved in the future.

Specific recommendations

1. 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.
2. Sensitivity analysis of the assessment model outputs to the partial recruitment curve generated from the Puckett *et al*. study should be carried out.
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. The Virginia (VIMS) trawl survey should be standardized, and better use should be made of the additional four strata.
12. WDS survey results for Virginia and Maryland should be evaluated separately to investigate whether they correlate with trawl survey indices in the two states.
13. Other survey time series, such as the Chesapeake Bay Multispecies Monitoring and Assessment Program, should potentially be included in the assessment.
14. Consideration should be given to the development of an early life history phase recruitment index, e.g. larvae or megalopae abundance.
15. 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.
16. The assessment team should gain access to, and analyse, the time series of catch data from the Potomac Fisheries Commission.

17. 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.
18. A baywide survey of recreational catches should be undertaken immediately and at regular intervals to quantify the overall landings from the recreational sector.
19. Some quantitative assessment of under-reporting rates should be obtained.
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.
21. Detailed investigation is required as to why the model provides such a poor fit to the sex-specific population sizes.
22. A conceptual model providing a timeline of life history events such as spawning and recruitment should be incorporated.
23. 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.
24. The chosen values for the overfished abundance threshold and the target abundance reference point should be re-considered.
25. 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.
26. A more systematic sensitivity analysis should be undertaken to understand how robust the model's output is to the various input parameters.
27. 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.
28. 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.
29. 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.
30. Harvest control rules should be developed to agree management actions if reference points are reached or exceeded.

Acknowledgements

I would like to thank Tom Miller and all the blue crab assessment team who were refreshingly open and helpful throughout the review, which made it easy for the review panel to understand the strengths and weaknesses of the 2011 assessment. Particular thanks are due to Tom Miller and Mike Wilberg who carried out some additional requested model runs at very short notice. In addition to the assessment team, the review panel benefited greatly from interaction with representatives of the management jurisdictions, and with a wide range of fisheries biologists working in the Blue Crab Advanced Research Consortium, in particular Yoni Zohar who kindly gave us a tour of the blue crab hatchery.

Derek Orner from the NOAA Chesapeake Bay office coordinated all of the review panel meetings and throughout provided a broad perspective on the science, assessment and management of the blue crab population.

References

Aguilar, R., E.G. Johnson, A.H. Hines, P.M. Roberts, M.R. Goodison and M.A. Kramer. 2011. Spatial dynamics of the blue crab (*Callinectes sapidus*) fishery in Chesapeake Bay: understanding post-maturation processes to improve efficacy of marine stock enhancement. Presentation at the Blue Crab Advanced Research Consortium Symposium, Baltimore, Maryland, 28 March 2011.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson. 1985. Statistical inference from band recovery data: a handbook. 2nd edition. U.S. Fish and Wildlife Service, Resource Publication 156, Washington D.C.

Fogarty, M.J. and T.J. Miller 2004. Impact of a change in reporting systems in the Maryland blue crab fishery. Fisheries Research **68**: 37-43.

Fonseca, D.B., M.R.J. Sheehy, and P.M.J. Shelton. 2003. Unilateral eyestalk ablation reduces neurolipofuscin accumulation rate in the contra-lateral eyestalk of a crustacean, *Pacifastacus leniusculus*. J. Exp. Mar. Ecol. Biol**. 289**: 277-286.

Haddon, M. 2005. Summary review of the Chesapeake Bay blue crab (*Callinectes sapidus*) 2005 stock assessment.

Harvey, R.H., D.H. Secor and S.-J. Ju. 2008. The use of extractable lipofuscin for age determination of crustaceans: reply to Sheehy. Mar. Ecol. Prog. Ser. **353**: 307-311.

Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. Fish. Bull. **81:** 898-903.

Johnson, E., A. Hines, H. Soulen, P. Roberts, M. Kramer, M. Goodison and R. Aguilar. 2011. Evidence for large-scale recruitment limitation of blue crabs in Upper Chesapeake Bay. Presentation at the Blue Crab Advanced Research Consortium Symposium, Baltimore, Maryland, 28 March 2011.

Ju, S.-J., D.H. Secor and H.R. Harvey. 1999. Use of extractable lipofuscin for age determination of blue crabs *Callinectes sapidus*. Mar. Ecol. Prog. Ser. **185**: 171-179.

Ju, S.-J., D.H. Secor and H.R. Harvey. 2001. Growth rate variability and LF accumulation rates in the blue crab *Callinectes sapidus*. Mar. Ecol. Prog. Ser. **224**: 197-205.

Kramer, M. A.H. Hines, and E.G. Johnson. 2011. Annual, seasonal and size-dependent patterns of juvenile blue crab, *Callinectes sapidus*, mortality in Upper Chesapeake Bay. Presentation at the Blue Crab Advanced Research Consortium Symposium, Baltimore, Maryland, 28 March 2011.

Lipcius, R.N. 2011. Density dependent exploitation of blue crab in Chesapeake Bay. Assessment Working Paper 2 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Lipcius, R.N. and A. Smith. 2011. Survival, longevity and natural mortality of mature female blue crabs. Assessment Working Paper 1 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Miller, T.J. 2011a. Development and application of simple production models to the Chesapeake Bay blue crab fishery. Assessment Working Paper 3 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Miller, T.J. 2011b. Application and update of a catch, multiple survey model to the Chesapeake Bay blue crab fishery. Assessment Working Paper 4 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Miller, T.J., S.J.D. Martell, D.B. Bunnell, G. Davis, L.A. Fegley, A.F. Sharov, C.F. Bonzek, D.A. Hewitt, J.M. Hoenig and R.N. Lipcius. 2005. Stock assessment of Blue Crab in Chesapeake Bay: 2005. Final Report Ref: [UMCES] CBL 05-077. UMCES Tech. Ser. No. TS-487-05-CBL., University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD.

Miller, T.J., M.J. Wilberg, A.R. Colton, G.R. Davies, A. Sharov, R.N. Lipcius, G.M. Ralph, E.G. Johnson, and A.G. Kaufman. 2011. Stock assessment of the blue crab in Chesapeake Bay 2011. UMCES Tech. Ser. No. TS-614-11-CBL., University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD.

Puckett, B.J., D. H. Secor and J-S Ju. 2008. Validation and application of lipofuscin-based age determination for Chesapeake Bay blue crabs (*Callinectes sapidus*). Trans. Am. Fish. Soc. **137**: 1637-1649.

Roberts, P.M., R. Aguilar, E.G. Johnson, A.H. Hines, M.R. Goodison, and M. A. Kramer. 2011 Population and fishery dynamics of blue crab (*Callinectes sapidus*) in a sub-estuary of Chesapeake Bay. Presentation at the Blue Crab Advanced Research Consortium Symposium, Baltimore, Maryland, 28 March 2011.

Rugulo, L.J, K. Knotts, A. Lange, V. Crecco, M. Terceiro, C.F. Bonzek, C. Stagg, R. O'Reilly, and D.S. Vaughan. 1997. Stock assessment of Chesapeake Bay blue crab (*Callinectes sapidus*). National Oceanic and Atmospheric Administration: 267p.

Sheehy, M.R.J. 2008. Questioning the use of biochemical extraction to measure lipofuscin for age determination of crabs: Comment on Ju et al. (1999, 2001). Mar. Ecol. Prog Ser. **353**: 303-306.

Appendix 1 – Bibliography of materials provided for review

Prior to the meeting of the review panel in Baltimore, panel members were provided with password protected access to the University of Maryland, Chesapeake Biological Laboratory website as follows:

http://hjort.cbl.umces.edu/crabs/Assessment.html

This website provided access to the full 2011 assessment document, four associated Working Papers, files containing the catch data, control rule calculations, ADMB code and R scripts and associated data, links to the 2005 assessment and the associated peer review, and other previous assessment papers.

Full assessment report

Miller, T.J., M.J. Wilberg, A.R. Colton, G.R. Davies, A. Sharov, R.N. Lipcius, G.M. Ralph, E.G. Johnson, and A.G. Kaufman. 2011. Stock assessment of the blue crab in Chesapeake Bay 2011. UMCES Tech. Ser. No. TS-614-11-CBL., University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD.

Working papers

Lipcius, R.N. and A. Smith. 2011. Survival, longevity and natural mortality of mature female blue crabs. Assessment Working Paper 1 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Lipcius, R.N. 2011. Density dependent exploitation of blue crab in Chesapeake Bay. Assessment Working Paper 2 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Miller, T.J. 2011a. Development and application of simple production models to the Chesapeake Bay blue crab fishery. Assessment Working Paper 3 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Miller, T.J. 2011b. Application and update of a catch, multiple survey model to the Chesapeake Bay blue crab fishery. Assessment Working Paper 4 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Appendix 2 CIE Statement of Work – Dr. Julian Addison

Statement of Work for Dr. Julian Addison

External Independent Peer Review by the Center for Independent Experts

Blue Crab Benchmark Stock Assessment - 2010

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: The blue crab stock has been subject to Baywide stock assessments on two previous occasions. In the years between benchmark assessments, updates on the stock status are provided by the NOAA Chesapeake Bay Office's Chesapeake Bay Stock Assessment Committee. The most recent update concluded that the stock was not overfished and was not then experiencing overfishing. Since the 2005 assessment, the three management jurisdictions have implemented a range of regulatory changes aimed at attaining the target exploitation rate of 46% of the available stock. Thus, it is appropriate that another, Baywide benchmark assessment be conducted. The blue crab resource, specifically for soft and peeler crabs, in Chesapeake Bay has recently been declared a fisheries resource disaster by the Secretary Commerce. In 2009 and 2010, annual updates (not peer-reviewed) have shown slight improvements in the resource. Blue crab is the most important commercial fishery in Chesapeake Bay with annual Baywide landings recently as low as 50 million pounds – roughly 25 million pounds below the long-term average. 2010 predicted landings (if fished at the target exploitation level) could top 100 million pounds. This is obviously a large fluctuation in landings and thus value of the resource to the Bay community.

The first Baywide stock assessment was conducted using a length-based approach to estimate exploitation, and an unweighted average of the four principal fishery-independent surveys to determine abundance. Consequently biological reference points were crude.

In 2001, the technical subcommittee of the Bi-State Blue Crab Advisory Committee (BBCAC) developed a new management framework that relied on exploitation and biomass threshold and target reference points. Threshold reference points were proposed based on a maintaining 10% of the virgin spawning potential and on the lowest observed abundance in the surveys. A target exploitation rate that would lead to an effective doubling of the spawning stock present in 2001 was also selected. The most recent Baywide benchmark assessment for blue crab in the Chesapeake Bay was conducted in 2005. This assessment critically evaluated and revised estimates of the natural mortality rate, the impact of reporting changes on landings estimates, and spawning potential ratio reference points. The 2005 assessment, using data through 2003, recommended adopting the exploitation fraction, defined as the proportion of crabs available at the beginning of the season that are subsequently harvested, in place of less intuitive measures (F) used in previous assessments. Estimates of exploitation fractions were calculated based on the Baywide winter dredge survey (WDS) and within a modified catch-survey analysis that permitted the use of multiple surveys. The approach used in the 2005 assessment was reviewed by a panel of international scientists with expertise in crustacean fisheries who found that it was a substantial improvement over previous assessments. However, the panel also identified issues to be addressed in future assessments. In particular, the panel recommended exploration of the impact of density-dependent processes in life history traits, improvements to the fishery-independent surveys, particularly with regard to catchability, the possibility of developing a sex-specific assessment model and reference points, and a fuller analysis of the impacts of uncertainty on all aspects of the assessment.

The 2010 assessment and targeted research program is a highly collaborative and integrated program to address specific concerns raised by the international review panel from 2005. The assessment activities are divided into eight specific Terms of Reference (TOR) that were developed based on the review comments received from panel of experts convened to review the 2005 assessment, and from extensive discussion with managers from MDNR, the Potomac River Fisheries Commission and the Virginia Marine Resources Commission, the three relevant management jurisdictions.

NOAA Fisheries is playing a significant role in coordinating disaster assistance to Maryland and Virginia to ensure a sustainable blue crab fishery in Chesapeake Bay. This 2010 Benchmark assessment and research program represents a large investment by NOAA and the state management agencies and should be reviewed internationally.

The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment and crustacean fisheries. Each CIE reviewer’s duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Baltimore, Maryland during the tentative dates of 29-31 March 2011.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in the Baltimore, Maryland during the tentative dates of 29-31 March 2011, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 14 April 2011, submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

22 February 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
15 March 2011	NMFS Project Contact sends the CIE Reviewers the pre-review documents
29-31 March 2011	Each reviewer participates and conducts an independent peer review during the panel review meeting
14 April 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
28 April 2011	CIE submits CIE independent peer review reports to the COTR
5 May 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each CIE report shall address each ToR as specified in **Annex 2**,
(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

Derek M. Orner
NOAA Chesapeake Bay Office, 410 Severn Avenue, Suite 107A, Annapolis, MD 21403
Derek.orner@noaa.gov Office: (410) 267-5676 Cell: (410) 570-2268

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Blue Crab Benchmark Stock Assessment - 2010

The stock assessment review has the following eight specific terms of reference:

a) Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

b) Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

c) Describe and quantify patterns in fishery-independent surveys. Analyses should include an evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time series.

d) Describe and quantify patterns in catch, effort and survey-based estimates of exploitation by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

e) Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.

f) Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.

g) Characterize scientific uncertainty with respect to assessment inputs and stock status.

h) Evaluate stock status with respect to reference points.

Annex 3: Tentative Agenda

2010
Blue Crab Stock Assessment Review
Sheraton Baltimore City Center Hotel
101 West Fayette St., Baltimore, MD
March 29-31, 2011

March 29, 2011

12:30 Welcome & Introductions Orner
- Stock Assessment Committee
- Review Panel

12:45 Presentation of the 2010 Blue Crab Stock Assessment
Miller

4:00 General / Open Question Period Orner
- Public Comment
- Review Panel

5:30 Adjourn

March 30, 2011

8:30 Term of Reference Review and Discussion

I. Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

II. Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

III. Describe and quantify patterns in fishery-independent surveys. Analyses should include evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time-series.

IV. Describe and quantify patterns in catch and effort by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

12:30 Lunch

2010
Blue Crab Stock Assessment Review
Sheraton Baltimore City Center Hotel
101 West Fayette St., Baltimore, MD
March 29-31, 2011

March 30, 2011 (cont.)

1:30 Term of Reference Review and Discussion (continued)

V. Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.

VI. Examine density-dependent exploitation patterns derives from survey-based and model-based approaches.

VII. Characterize scientific uncertainty with respect to assessment inputs and stock status.

VIII. Evaluate stock status with respect to reference points.

5:15 Adjourn

March 31, 2011

9:00 Review Session *[closed-door]*

- Review Panel to discuss assessment methodologies and develop individual opinions.

- Initiate development of summary documents

12:00 Lunch

1:15 Review Session (continued)
4:30 Adjourn

Appendix 3 - Panel Membership and Review Meeting

The review panel consisted of Dr Julian Addison, an independent fisheries consultant based in France, Dr Cathy Dichmont of CSIRO Marine and Atmospheric Research, Queensland, Australia and Dr Billy Ernst, Department of Oceanography, University of Concepción, Chile. The CIE did not formally appoint a chair of the review panel. The meeting was co-ordinated by Dr Derek Orner of NOAA on behalf of the CIE review team. The panel met formally from 29-31 March at the Sheraton Hotel, Baltimore, Maryland following the agenda outlined in Appendix 2, Annex 3.. The panel met briefly again on 1 April to consider some additional model runs provided by the assessment team, and a summary review was agreed representing the view solely of the CIE panel members.

The Review process worked extremely well. The whole assessment team was extremely helpful and open about the strengths and weaknesses of the new assessment and were very keen to provide any new model runs requested by the review panel. The panel benefitted greatly from having a NOAA Coordinator present throughout the panel meetings to provide broad expertise in the science and management of the fishery. The only minor criticism of the process that I have was that the identity of the panel was not known to panelists or assessment team until the start of the review process.