Center for Independent Experts (CIE) Independent Peer Review Report

Bering Sea Snow Crab Stock Assessment Review

Seattle, Washington. January 21-24, 2014

Dr. Noel Cadigan

Centre for Fisheries Ecosystems Research
Marine Institute of Memorial University of Newfoundland
St. John's, NL. Canada

Executive Summary

The review panel focused on the east Bering Sea (EBS) snow crab assessment model and not the harvest control rule or reference points. The EBS model integrates a substantial amount of information/data on stock productivity to provide management advice. The assessment model has a considerable number of parameters that are necessary to account for the complex biology of snow crab; however, the parameters are mostly well-identified by the assessment model. The model is not well documented and the statistical inferences (i.e. standard errors and confidence intervals) it produces may not reliable. The population dynamics of the stock are poorly understood and more research is required to improve our understanding of how to optimally and sustainably harvest this stock. Important elements are: 1) natural mortality rate of commercial sizes males, 2) the contribution of large males to reproductive potential, and 3) factors governing recruitment to the exploited population.

Background

The peer review meeting was held at the Alaska Fisheries Science Center (AFSC) in Seattle, Washington, during the dates of January 21-24, 2014. The purpose of the meeting was to review the Bering Sea trawl survey data for snow crab, additional survey data used in estimation of catchability, the stock assessment model structure, assumptions, life history data, and harvest control rule. More specific terms of reference (ToRs) are provided in Appendix 2.

The Panel was composed of three independently appointed Center for Independent Experts (CIE) reviewers (Dr. N. Cadigan, Canada; Dr. N. Hall, Australia; Dr. B. Ernst, Chile) and the chair, (Dr. M. Dorn, AFSC). Assessment documents were prepared and presented by Mr. J. Turnock (AFSC). A description of the Bering Sea trawl survey program and data for snow crab was presented by Dr. B. Foy (AFSC). Presentations were provided by Dr. C. Szuwalski on MSE and environmental effects on recruitment, and Dr. A. Whitten on the generalized crab assessment model. The support of all of these scientists to the review process is gratefully acknowledged.

The CIE reviewers were tasked with conducting an impartial and independent peer review in accordance with the SoW and ToRs herein. The reviewers were required to be active and engaged participants throughout panel discussions and to voice concerns, suggestions, and improvements while respectfully interacting with other review panel members, advisors, and stock assessment technical teams. The CIE reviewers were required to have expertise in conducting stock assessments for fisheries management, and to be thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, size-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language. Familiarity with invertebrate stock assessment, knowledge of crab life history and biology, harvest strategy development was desirable. Each CIE reviewer's duties were not to exceed a maximum of 14 days for conducting pre-review preparations with document review, participation in the panel review meeting, and completion of the CIE independent peer review report in accordance with the ToR and Schedule of Milestones and Deliverables.

Role of reviewer

A working paper (WP) describing the assessment model and some supporting materials were emailed to the Panel two weeks before the meeting. These documents are listed in Appendix 1. I reviewed the backgrounds documents I was provided. I attended the entire review panel meeting in Seattle, Washington, January 21-24, 2014. I reviewed presentations and reports and participated in the discussion of these documents, in accordance with the SoW and ToRs (see Appendix 2). This report is structured according to my interpretation of the required format and content described in Annex 1 of Appendix 2.

At the beginning of the meeting the ToRs were discussed. It was decided not to review the harvest control rule or the reference points. A consideration of reference points was felt to have implications for other stocks.

Summary of findings

ToR 1. A statement of the strengths and weaknesses of the Bering Sea snow crab stock assessment and stock projection models.

Strengths

Unlike snow crab stocks in eastern Canada, a model is used for EBS snow crab to integrate a substantial amount of information/data on stock productivity to provide management advice. The model is size-based and accounts for some of the rather complicated life history characteristics of snow crab, including a long juvenile stage and a terminal molt at the morphological mature stage after which the crab stop growing. The highly size and sex selective nature of the fishery is included in the model.

Fishery landings statistics seem reliable. Observer coverage seemed good since 1992 for sampling sex and size frequencies, and catch per unit effort. Spatial information on the fishery seemed readily available.

There is a fairly long time-series (1978-2013) of NMFS survey information available that covers most of the area where mature crab are found. Time-series in eastern Canada are shorter or have more limited biological sampling of crab early on.

There catchability of the survey has been examined in two studies (2009-2010).

The assessment model has a considerable number of parameters that are necessary to account for the complex biology of snow crab; however, the assessment lead provided the correlated matrix of the estimated model parameters and confounding amongst parameters did not seem to be an issue. Only those parameters that were expected to be partially confounded (e.g. intercept and slopes of logistic selectivity models) had high correlations. This provided evidence to the review panel that the model results may be robust.

The following text on weakness will be more in-depth but I find that overall the EBS snow crab assessment model is more advanced than for snow crab stocks in eastern Canada that I am familiar with.

Weakness

There is substantial spatial structure in the size/sex/maturity distribution of snow crab in the Bering Sea. A substantial fraction of the immature population, and some mature females, can be found north of the main EBS survey area. I did not understand what contribution these immature crabs have towards the mature population in the EBS survey area. However, whenever a large part of a stock exists outside of a survey area then the reliability of survey indices of stock size is usually much diminished, and this comment applies to <u>immature</u> EBS snow crab.

The assessment model was not well documented and technical descriptions (i.e. model equations) were often poor. For example, the equation describing the recruitment model had an "l" superscript that was confusing, because the "l" subscript indexed recruitment to various length bins in the model. The equation for the population dynamics model (pg. 18 in WP) does not seem correct. As I understand the model, the sum in this equation should range from the minimum model length to "l", and the sum is over $l^{'}$, because crab of different lengths (i.e. $l^{'}$s) in year t-1 growth to length l in year t. Some other equations (e.g. recruitment deviation likelihood equation) seemed to be based on an earlier version of the model and did not accurately describe the current model formulation. A highly unusual penalty was used for estimating the commercial sized fishing mortality rate for males. The document provided little information about this and the assessment lead was not clear on why this was used, and whether it was appropriate to use. Line types seemed to be mis-labelled in Figure 4.

Maturity determination was not well described in the assessment WP. My understanding is that female maturity status is visually determined at sea but male maturity status is not. The latter is based on a chela height relationship for the proportion mature. This relationship was assumed to be constant over time. One maturity curve for males was estimated using the average fraction mature based on chela height data and applied to all years of survey data to estimate mature survey numbers. The probability of maturing was estimated to match the observed fraction mature for all mature males and females observed in the survey data. The assessment WP was not clear on how this was done. I could not find a likelihood equation for the probability of maturing. In the presentation by the assessment lead (but not in document) a figure was shown for the annual proportions mature versus carapace widths, for 1989-200 and 2007. It was not clear if this was the only data used to estimate probability of maturing. Also, there was evidence of considerable inter-annual variation in proportions mature. This may have important consequences for the assessment model because it affects the size distributions of males and females. The document needs a clearer description of how maturity is estimated. Consideration should be given to allowing the probability of maturing to change over time.

The model could not reliably estimate the natural mortality rate (M) of commercial sized males, and females of all sizes. A major source of uncertainty in this assessment is the appropriate value for M. This is common problem in stock assessment models so this is not really a criticism of this snow crab model. The model used a prior on commercial sized male M (mean M=0.23 with a se = 0.054). This was based on an assumption that in a virgin population of snow crab, longevity would be at least 20 years. This age is a little higher than considered on the east coast of Canada, where the longevity of adult males (after reaching the terminal molt) is approximately 5-7.7 years (Hébert et al., 2011; DFO, 2013). Hence, if snow crab in the Bering Sea terminally molt in 8-10 years then longevity would be in the range of 13-18 years. Also, M probably increases with age after terminal molt as shell condition deteriorates and crab lose legs. The posterior mean of adult male M from the model was 0.261 which indicates that the model and data favors a higher value than the prior indicates. There are alternative model configurations that may provide additional information on the appropriate value for M. M=0.23 was used for all female crab in the model and I have no better suggestion. M for immature crabs, both males and females, was freely estimated at 0.386 and this value seems reasonable.

A weakness of the current model is that it does not use stage-structure (i.e. shell condition) information for terminally molted crab. This is considered further under the next ToR.

The model uses external weighting of likelihood components. This is a fairly common procedure in US stock assessment models. However, confidence intervals have a fairly arbitrary interpretation. If one were to multiply all weights by 100, so that the relative weighting of likelihood components is the same then, although parameter estimates will remain the same, hessian-based standard errors and confidence intervals for parameters will decrease by a factor of 10. This model basically fixes the variance parameters of data inputs and other structural components. There did not seem to be procedure used (e.g. Francis, 2010) to make sure that likelihood components received appropriate weight. The issue is analogous to deriving a confidence interval from a simple random sample of measurements. A confidence interval based on an assumed variance (say variance = 1) will not be informative if the sample variance is much different than one.

Survey indices are externally weighted by their survey CV's, although the likelihood equation on Pg. 24 of the assessment model WP indicated that there is an extra weighting component. I am not sure what value was used for this λ. However, the survey CV is only the measurement error component on the survey variability. This is also a type of process error component (e.g. Francis, 2011) that should be considered.

Aspects of the projection model were discussed – in particular, the lognormal bias correction and how auto-correlated process errors were modelled. It seemed odd that the recruitment model on Pg. 17 of the WP did not have a lognormal bias correction term while the projection model on Pg. 29 did. A verbal explanation was given and this should be included in the report. The equation on Pg. 29 for including temporal autocorrelation in the projected recruitment deviations does not seem right. If $\eta_t \sim N(0, \sigma_R^2)$ then the stationary variance of $\epsilon_t = \rho\epsilon_{t-1} + \eta_t$ is $\lim_{t\to\infty} Var(\epsilon_t) = \sigma_R^2/(1-\rho^2)$. If one wants the stationary distribution of AR(1) recruitment errors to have variance equal to σ_R^2 then the generating equation should be $\epsilon_t = \rho\epsilon_{t-1} + \sqrt{1-\rho^2}\eta_t$. This is different than the equation on Pg. 29.

ToR 2. Recommend for alternative model configurations or formulations if required.

ToR 3. Recommendations of alternative model assumptions and estimators if required.

These ToRs are very similar and are treated as one.

The model currently is fitted to commercial catch statistics (landings, and discards), survey indices of mature and immature biomass, separately for males and females, and their corresponding size compositions. Mature and immature male biomass indices are inferred using an annually constant chela size dependent maturity relationship. A better approach is to continue fitting to the commercial catch statistics and female survey information, but fit only to the survey male biomass and size composition information and fit the probability of maturing for males directly in the model using a likelihood component based on the chela height data.

Growth data for Bering Sea snow crab are limited. The Base model fitted a sex-specific linear model to the growth data reported by Somerton (2013). Sample sizes were 17 for males and 18 for females. The linear models did not fit the data well. There were substantial patterns in residuals. Somerton (2013) used a segmented linear model to fit the data; however, the assessment model did not converge when fitting the growth data using a sex-specific two segment model. Few details were provided on this, but ADMB will have problems fitting a segmented model depending on how the model is configured. An alternative is to consider a smoothed version of the segmented linear model. For example, if the two regression equations are $f_i(x) = a_i + b_i x$, i=1 or 2, but constrained to be equal at a breakpoint δ, e.g. $a_2 = a_1 + (b_1 - b_2)\delta$, then a smoothed segmented model could be $f(x) = f_1(x)\left\{1 - \Phi\left(\frac{x-\delta}{s}\right)\right\} + f_2(x)\Phi\left(\frac{x-\delta}{s}\right)$, where Φ is the cumulative distribution function for a standard normal random variable and is used to provide a mixture distribution for the two linear regression equations. s is a scale parameter governing how smooth the transition is between f_1 and f_2. If $x << \delta$ then $\Phi\left(\frac{x-\delta}{s}\right)$ will be close to zero and $f(x) \approx f_1(x)$. Conversely, if $x >> \delta$ then $\Phi\left(\frac{x-\delta}{s}\right)$ will be close to one and $f(x) \approx f_2(x)$. I suggest fixing s and estimating a_1, b_1, b_2, and δ. The cumd_norm function in ADMB can be used to compute $\Phi\left(\frac{x-\delta}{s}\right)$.

The growth data in Somerton et al. (2013) seemed very reliable because of similarity with growth information from eastern Canadian studies. If the snow crab model is unable to fit the data well, then I suspect this may be related to a mis-specification of some other model assumption rather than problems with the growth data.

The issue of discard mortality was considered in detail by the assessment group. A 30% discard mortality rate was used, and this was thought to be an upper bound on the real value. However, it was also indicated that discard mortality has probably decreased since the early 2000's. Some consideration should be given to how much discard mortality rates may have changed over time, and whether the stock assessment is sensitive to this.

The model is only length structured which may be a serious deficiency for snow crab because of the terminal molt characteristic of this species. When snow crab terminally molt and become morphometrically mature they stop growing in size, although weight, in terms of meat yield, may increase for approximately a year after terminal molt. The lack of size or age structure information for terminally molted crab makes it difficult to estimate M and I suspect F, which may be why the model has to use an F penalty towards 1.15. However, stage-structure information about shell-condition from surveys may provide additional information about mortality rates of terminally molted males. As snow crab age after terminal molt their shell condition (i.e. appearance) changes. The assessment lead indicated that this information was used in earlier versions of the model but it was difficult to use because of uncertainty in the number of years that it takes for a crabs shell condition to change. I suggest that this could be addressed by using a distribution of years for the various shell conditions that can be reliably be differentiated.

For illustration purposes I consider three shell conditions: new, intermediate, old. Let S be a random variable for shell condition, and let A be a random variable representing the age distribution of commercial sized mature males (CSMMs). I assume all CSMMs regardless of shell condition are fully selected by the survey. Age is in years since terminal molt. Prob(A) will depend on total mortality rates (Z) and recruitment. For simplicity I assume the recruitment rate into CSMM is constant, in which case Prob(A) depends only on Z and is the steady-state age distribution. I projected for 30 ages and pooled ages 7+. If we know something about Prob(S|A) (e.g. from radiometric or eye-stalk aging, or tagging) then we can compute Prob(S=s) = Σ_aProb(S=s|A=a)Prob(A=a) for various levels of Z and compare this with the survey distribution of shell condition to infer more likely values of Z. For example, if Prob(S|A) is like Figure 1 (left hand panel) then Prob(S) is shown in the right-hand panel of this Figure for several values of Z.

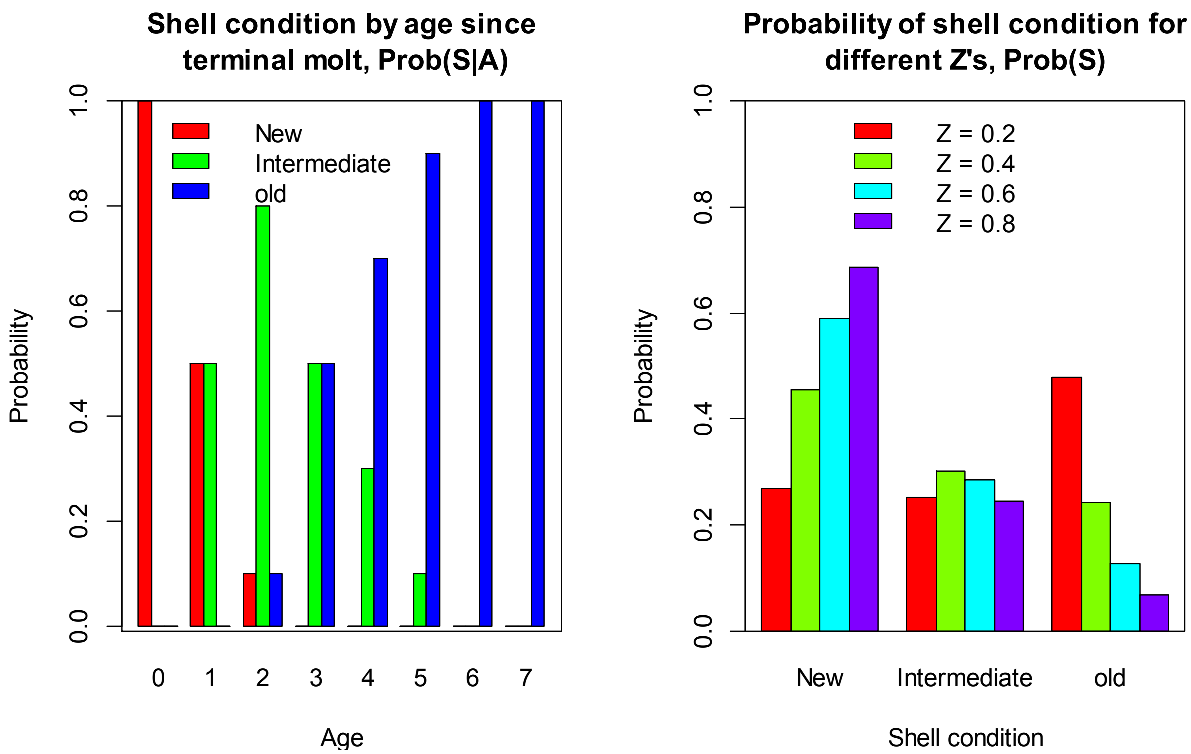

Figure 1. Scenario 1: left: Prob(S|A). Right: corresponding Prob(S) for several values of Z.

Another scenario in which shell condition is more age-specific is shown in Figure 2 and illustrates that Prob(S) is more sensitive to the value of Z than Prob(S|A). The base run of the assessment model (Figure 3 in WP) indicated that recent harvest rates for SCMMs were about 30%, implying Z = 0.62 based on M=0.261. Hence one would expect a distribution of shell conditions similar to the light blue bars in the right-hand panels of Figure 1 and 2 (i.e. 50-60% new shell, 13-14% old shell). However, there may be other reasonable choices for Prob(S|A) that could have a larger effect on Prob(S). These figures are only provided for illustration purposes only.

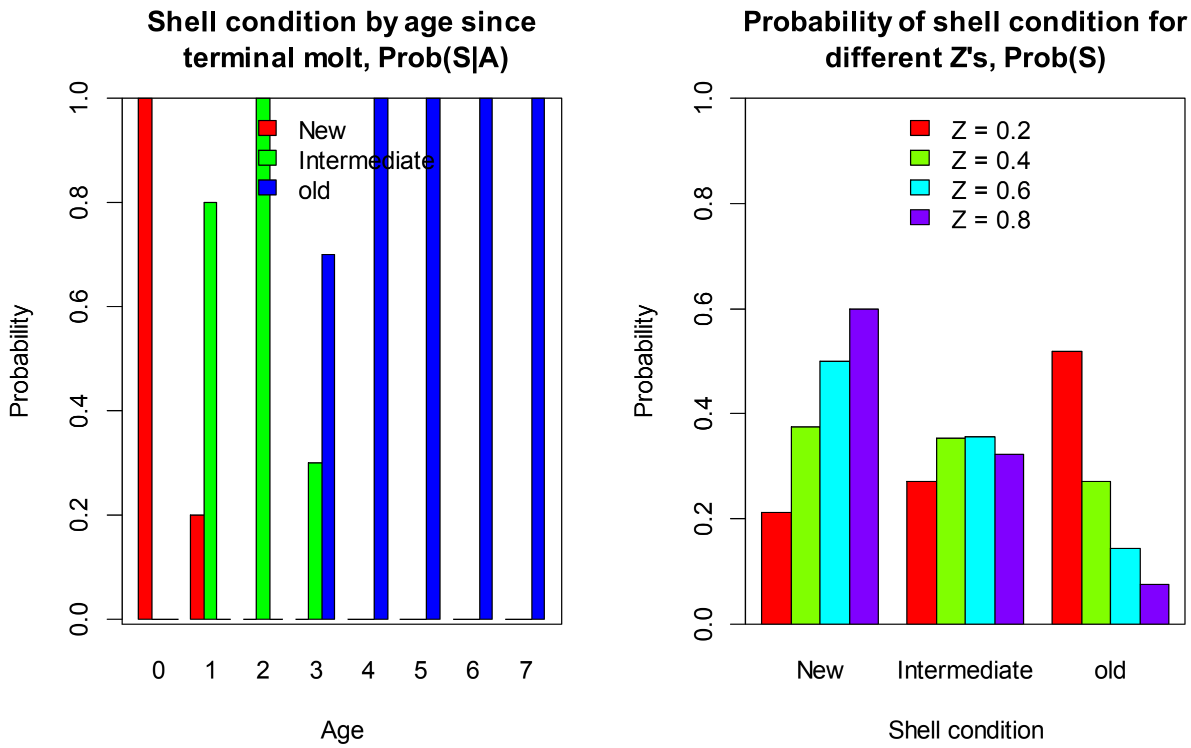

Figure 2. Scenario 2: left: Prob(S|A). Right: corresponding Prob(S) for several values of Z.

If recruitment is not constant then this will also affect Prob(S) and would need to be accounted for. A good way to do this is to change the assessment model to keep track of abundance at age since terminal molt (i.e. a length- and age-based model), include a likelihood component for data on Prob(S|A) and a likelihood component for the annual survey shell condition distribution to provide additional information on Z and consequently F and M. If there is no reliable information on Prob(S|A) then scenarios could be investigated, and the results in Figures 1 and 2 suggest that results may not be very sensitive to a range of assumptions for Prob(S|A).

M probably increases with age since terminal molt, and this is another factor that will affect Prob(S) as well as the stock assessment model and fishery management reference points.

The assessment model assumed a length dependent selectivity function for males and females that was constant over time; however, commercial pots were modified in 1997 and 2001 to increase escapement of small crab. One would expect this to impact fishery and discard selectivity. While the assessment document recognized this issue, additional consideration is required. This could involve a sensitivity analysis.

The Bering Sea Fisheries Research Foundation (BSFRF) conducted a survey of 108 tows in a portion of the Bering Sea in summer 2009. A larger area was survey in 2010 and included paired tow comparative fishing with the NMFS survey vessel. Separate indices (total biomass and size composition) were developed from these surveys and included in the stock assessment model. NMFS abundance (i.e. numbers) indices were also developed for the BSFRF 2009 and 2010 study areas and included in the model. The purpose of including these indices was to provide information to estimate the catchability of the EBS NMFS survey biomass index. The assessment WP does not describe very well how this was achieved. On Pg. 27 it gives a description of the model value of NMFS study area survey indices but does not give the corresponding equation got the MBFRF survey.

This approach of developing indices for the study area is more reasonable for the 2009 BSFRF survey than the 2010 survey. It does not utilize the pairing of tows in 2010. A better approach for the 2010 data is to include a likelihood component for the conditional distribution of the NMFS study area catches given the total of the NMFS and BSFRF catches. References on what the conditional distribution should be are included for ToR 4. There was little description of the 2009 BSFRF study. A Somerton (2010) reference was given but not included in the background papers or in the assessment WP reference list. However, I gather that the BSFRF survey completed four random tows in 27 survey stations/grids, and the NMFS survey completed one tow in each of the 27 stations. A better to estimate the relative efficiency of the NMFS survey compared to the BSFRF survey is to develop a likelihood component that includes a station effect and a survey effect; that is, $E(C_{slij}) = \rho_{sl}\lambda_{il}$, where C_{slij} is the catch of length l snow crab during the jth tow of survey s at station i, λ_{il} is the length distribution of crab at station i (which is assumed to be homogeneous throughout the station area) and ρ_{sl} is the survey effect which is assumed to be one for all lengths for the BSFRF survey. Hence, ρ_{sl} is the relative efficiency of the NMFS survey compared to the BFRFR survey. The ρ_{sl} term is the same in the likelihood terms for the 2009 and 2010 surveys results. A similar model is considered in Benoît and Cadigan (2014).

ToR 4. A review of fishery dependent and fishery independent data inputs to the stock assessment.

The assessment lead expressed confidence in the reliability of the fishery catch statistics. Discard statistics seem more uncertain and were at time high in the catch time-series, but this does not seem too important because of the relatively low discard mortality used in the model.

I reviewed the background paper by Somerton et al. (2013) on the catchability of snow crab in the eastern Bering Sea (EBS) bottom trawl survey. It involved a paired tow comparison with a vessel fishing using a *Nephrops* trawl that was assumed to have an efficiency of one. I have several criticisms of this paper that are relevant to the snow crab assessment model:

1. Somerton et al. (2013) did not seem to adequately account for Binomial over-dispersion when fitting paired-tow catches. With this experimental design it is virtually impossible

to ensure that exactly the same densities are fished by each vessel. The best one can hope for is that differences are small and random from pair to pair. Failure to account for this variation can result in spurious estimates, and I suspect this is the issue with some of the curious catchability patterns in Somerton et al. (2013), particularly that the EBS survey catchability for large males increased substantially with small changes in carapace width. Between pair variability has been addressed recently by Miller (2013) who recommended a conditional beta-binomial approach, and Cadigan and Dowden (2010) used a binomial generalized linear mixed model. Cadigan and Bataineh (2012) concluded that when there is uncertainty about the type of Poisson over-dispersion present in paired-tow catches then a Binomial random effects model is a good choice.

2. Somerton et al. (2013) concluded that the size-dependent catchability (q) of the EBS survey also varied with depth and bottom sediment size. This seems reasonable. To calculate an overall survey q Somerton et al. (2013) calculated a catch weighted average of station specific q's predicted by their GAM model. I think their weighted-average procedure is wrong for reasons outlined in Appendix 3.

The assumption that the *Nephrops* trawl has an efficiency of one may not be entirely accurate. In the southern Gulf of St. Lawrence there is dedicated fishery-independent bottom-trawl survey conducted annually since 1988. The survey gear is a Nephrops trawl (20 m Bigouden trawl net) and the target fishing procedure at each site is a 4-6 minute tow at an average speed of approximately 2 knots. Three chartered vessels have been used to conduct the survey since 1988. In addition, there is a multi-species survey conducted at the same time that provides synoptic estimates of snow crab abundance and distribution for a consistently sampled area that largely overlaps with, and is larger than, the area covered by the crab survey. Benoît and Cadigan (2014) analyzed these data and found some significant differences in the catchabilities when crab survey vessels changed. Reasons for these differences are poorly understood; however, the results demonstrate that the assumption of 100% efficiency of a *Nephrops* trawl for snow crab may not be accurate.

The area surveyed in the annual NMFS EBS bottom trawl survey has changed several times including more sampling stations farther north starting in 1989. Juvenile crabs are much more common in these more northerly stations. This suggests the potential that the size selectivity of the EBS survey has changed. This was addressed by including EBS survey selectivity blocks in the assessment model. However, this could have unintended consequences. I suggest that if the spatial distribution of snow crab does not change substantially from year to year then an alternative approach is to conduct a spatio-temporal analysis of survey catches with a focus on producing a standardized index for the maximal area surveyed. An example of such an approach is given in Cadigan (2012); he used a simple model to fill in historical gaps in spatial survey coverage related mostly to extensions in the survey area over time. Some type of spatio-temporal analysis is probably better than simply "blocking" survey selectivity in an assessment model because the spatio-temporal analysis uses more information to fill "gaps" than a non-spatial assessment model. Note that the approach of Cadigan (2012) was not pursued further because the approach of Benoît and Cadigan (2014), which utilizes more information to fill gaps, was considered to be better.

Benoît and Cadigan (2014) also considered how to include multiple tows at a site when deriving a survey abundance index. They basically included a random site effect which accommodates the correlation one anticipates from multiple tows. However, this approach will not be appropriate if there is local depletion of crab densities due to the multiple tows.

In the EBS NMFS survey there have been occasional multiple tows at crab “hot spots”. This is a type of two-phase sampling that is known to produce an estimate of trawlable abundance that is biased low (e.g. Francis, 1984). However, the hot spots were not based on snow crab densities and so they should not theoretically cause a bias. However, one would expect correlation in such multiple tows in a model-based survey analysis.

In the survey presentation by Bob Foy, he mentioned that NMFS is re-evaluating how to derive survey indices of stock size from the EBS survey. There are a number of issues that should be addressed, and I have outlined some of them in this report. I suggest that a purely design-based approach will not be appropriate. A joint model and design based approach will probably be better. A review and application of such as approach is given by Chen et al. (2004). The re-evaluation of the NMFS survey estimation methods should be done in a workshop or working group format with external experts in the analysis if fisheries survey data.

ToR 5. Recommendation on research needs that would reduce uncertainty in key parameters used or estimated in the assessment.

The mature male natural mortality rate has been demonstrated to be an important parameter in the stock assessment. Better age and shell condition information could be used to improve estimation of mature male M, as outlined under ToR 3.

ToR 6. Suggested research priorities to improve the stock assessment.

More research is needed on the spatio-temporal stock dynamics of EBS snow crab. This should include:

1. What is the contribution of immature crab to the north of the EBS survey area to the mature stock?
2. A better measure of the reproductive potential of the stock, including an improved understanding of the reproductive value of large mature males.

3. I concur with the research need outlined in the assessment WP that there is a need for age information of terminally molted crab, and the relationship between age and shell condition.
4. Better information is required on growth increments and skip-spawning, and potential spatial variability of these processes.
5. Within season changes in commercial CPUE may provide some information on exploitable stock size (via depletion). However, this depends on the within season spatial dynamics of the fleet. This should be considered for future stock assessments.

There is a need to perform sensitivity analyses to better understand the robustness of the stock assessment model. A common task in most stock assessment meetings and reviews is to perform sensitivity analyses of various model inputs and assumptions in an attempt to find the ones that have large effects on advice. There is little point in refining data and assumptions that have little impact on the important results of the assessment. In their presentation to the review panel, Szuwalski and Punt (also see Szuwalski and Punt, 2012) demonstrated that the OFL derived from an earlier version of the snow crab assessment model was sensitive to the value of M for mature males and also the amount of information used to model the growth curve. In the review we did not ask for sensitivity analyses. A problem with providing ad hoc sensitivity runs is that important sensitivities may be missed and less important sources may be over-emphasized to give a false impression of model robustness. More objective procedures were discussed in which model inputs and assumptions are changed (i.e. perturbed) in a systematic and objective manner, and then the impacts of these perturbations on important model outputs are examined. Such high-dimensional sensitivity analyses can be computationally prohibitive. Cadigan and Farrell (2002) presented an approach that is computationally more tractable.

The growth data is sparse and does not seem sufficient to accurately define the variability or range of the growth transition matrix in the model. It is useful to assess how sensitive the assessment model is to width of growth transition probabilities. If important results are sensitive to this, then efforts should be made to collect more growth data.

Conclusions and Recommendations

ToR 1. A statement of the strengths and weaknesses of the Bering Sea snow crab stock assessment and stock projection models.

Conclusions

- The EBS snow crab assessment model integrates a substantial amount of information/data on stock productivity to provide management advice.
- Fishery landings and size composition information seems reliable.

- The assessment model has a considerable number of parameters that are necessary to account for the complex biology of snow crab; however, the parameters are mostly well-identified by the assessment model.
- A substantial fraction of the immature snow crab population can be found north of the main EBS survey area. This suggests that survey indices of immature stock size may be less reliable than indices of mature stock size.
- The assessment model was not well documented and technical descriptions (i.e. model equations) were often poor.
- There was evidence of considerable inter-annual variation in proportions mature. This may have important consequences for the assessment model because it affects the size distributions of males and females.
- The model could not reliably estimate the natural mortality rate (M) of commercial sized males, and females of all sizes. A major source of uncertainty in this assessment is the appropriate value for M.
- The model uses external weighting of likelihood components. Hence, confidence intervals may have a fairly arbitrary interpretation.

Recommendations

1. The contribution to the mature population of immature snow crab north of the main EBS survey area needs to be better understood.
2. Provide more accurate and complete documentation on the assessment model.
3. Consideration should be given to allowing the probability of maturing to change over time in the assessment model.
4. Consider alternative weighting schemes (i.e. inverse variance) so that statistical inferences may be relevant.

ToR 2. Recommend for alternative model configurations or formulations if required.

ToR 3. Recommendations of alternative model assumptions and estimators if required.

These Tors are very similar and are treated as one.

Conclusions

- The base assessment model fitted a sex-specific linear model to the growth data reported by Somerton (2013). The linear models did not fit the data well for either males or females.
- The current assessment model lacks size or age structure information for terminally molted crabs, and this makes it difficult to estimate M and F for these crabs.
- The assessment model includes survey information from a trawl gear that is assumed to catch all crab encountered, and comparisons with the NMFS EBS survey catches allows

for the estimation of the catchability of this survey. These catchability estimates seem reasonable.

Recommendations

1. Continue fitting the assessment model to the commercial catch statistics and female survey information, but fit only to the survey male biomass and size composition information and fit the probability of maturing for males directly in the model using a likelihood component based on the chela height data.
2. Fitting a segmented regression model to growth data is more complicated in ADMB because of the autodiff functionality of this software. Consider fitting a smoothed version of the segmented linear model (see proposal in the Summary section of this report).
3. Include stage-structure information about shell-condition from surveys to provide additional information about mortality rates of terminally molted males (see proposal in the Summary section of this report).
4. A better approach for including the 2010 comparative fishing data is to include a likelihood component for the conditional distribution of the NMFS study area catches given the total of the NMFS and BSFRF catches. The goal is to better control for extraneous sources of variation affecting the trawl catches.

ToR 4. A review of fishery dependent and fishery independent data inputs to the stock assessment.

Conclusions

- The assumption that the *Nephrops* trawl has an efficiency of one may not be entirely accurate.
- The area surveyed in the annual NMFS EBS survey has changed several times including more sampling stations farther north starting in 1989. Juvenile crabs are much more common in these more northerly stations. This suggests the potential that the size selectivity of the EBS survey has changed.
- In the EBS NMFS survey there have been occasional multiple tows at crab “hot spots”. This is a type of two-phase sampling that is known to produce an estimate of trawlable abundance that is biased low. However, the hot spots were not based on snow crab densities and so they should not theoretically cause a bias.
- The external estimation of NMFS EBS survey catchability in Somerton et al. (2013) can be improved as outlined in Summary of Findings section.

Recommendations

1. Changes in the area surveyed in the annual eastern Bering Sea (EBS) bottom trawl survey were addressed in the assessment model by estimating catchability and survey selectivity in time-blocks corresponding to changes in survey coverage. Some type of spatio-temporal analysis of the survey catches is probably better than simply "blocking" survey selectivity in an assessment model because the spatio-temporal analysis uses more information to fill "gaps" than a non-spatial assessment model. A reference was given to illustrate such an application to Southern Gulf of St. Lawrence snow crab. However, I cannot make more specific recommendations on a spatio-temporal model based on the information that was available during the review. Note that this issue is probably outside the scope of this review.
2. Multiple tows at hot spots should be included in the survey index, using an appropriate statistical analysis that accounts for the anticipated correlation among these tows. However, if local depletion is suspected then only the first tow at a hot spot should be used.

ToR 5. Recommendation on research needs that would reduce uncertainty in key parameters used or estimated in the assessment.

Conclusions

- The mature male natural mortality rate has been demonstrated to be an important parameter in the stock assessment.

Recommendations

1. Collect age and shell condition information, and use this information in the assessment model to improve estimation of mature male M, as outlined under ToR 3 Summary of Findings.

ToR 6. Suggested research priorities to improve the stock assessment.

Recommendations

1. Determine the contribution of the immature crab to the north of the EBS survey area to the mature stock?
2. Better quantify the reproductive value of large mature males.
3. Conduct objective sensitivity analyses to better understand the robustness of stock assessment models.

Appendix 1: Bibliography of materials provided for review

Materials reviewed prior to meeting

September 2013 Crab Plan Team Report. C-3 supplemental.

Somerton, D. A., and R. S. Otto. "Net efficiency of a survey trawl for snow crab, Chionoecetes opilio, and Tanner crab, C. bairdi." Fishery Bulletin 97.3 (1999): 617-625.

David Somerton , Scott Goodman , Robert Foy , Lou Rugolo & Laura Slater (2013): Growth per Molt of Snow Crab in the Eastern Bering Sea, North American Journal of Fisheries Management, 33:1, 140-147.

Somerton, David A., Kenneth L. Weinberg, and Scott E. Goodman. "Catchability of snow crab (Chionoecetes opilio) by the eastern Bering Sea bottom trawl survey estimated using a catch comparison experiment." Canadian Journal of Fisheries and Aquatic Sciences 70.12 (2013): 1699-1708.

Benjamin J. Turnock and Louis J. Rugolo . Stock Assessment of eastern Bering Sea snow crab. National Marine Fisheries Service. September 19, 2013.

DRAFT REPORT of the SCIENTIFIC AND STATISTICAL COMMITTEE to the NORTH PACIFIC FISHERY MANAGEMENT COUNCIL. September 30th – October 1st, 2013.

Additional materials reviewed for this report

Benoît, H.P., and Cadigan, N. 2014. Model-based estimation of commercial-sized snow crab (*Chionoecetes opilio*) abundance in the southern Gulf of St. Lawrence, 1980-2012, using data from two bottom trawl surveys. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/114. v + 46 p.

Cadigan, N. G. and Farrell, P. J. 2002. Generalized local influence with applications to fish stock cohort analysis. Appl. Statist. 51: 1-15.

Cadigan, Noel G., and Jeff J. Dowden. "Statistical inference about the relative efficiency of a new survey protocol, based on paired-tow survey calibration data." (2010).

Cadigan, N. G., and O. M. Bataineh. "Inference about the ratio of means from Negative Binomial paired count data." Environmental and Ecological Statistics 19.2 (2012): 269-293.

Cadigan, N. G. 2012. Preliminary investigations into standardizing the 4T snow crab (*Chionoecetes opilio*) bottom-trawl survey index. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/016. iv + 48 p.

Chen, Jiahua, Mary E. Thompson, and Changbao Wu. "Estimation of fish abundance indices based on scientific research trawl surveys." Biometrics 60.1 (2004): 116-123.

DFO, 2013. Assessment of Newfoundland and Labrador Snow Crab. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2013/013.

Francis, R. I. C. C. "An adaptive strategy for stratified random trawl surveys." New Zealand Journal of Marine and Freshwater Research 18.1 (1984): 59-71.

Francis, RIC Chris. "Data weighting in statistical fisheries stock assessment models." Canadian Journal of Fisheries and Aquatic Sciences 68.6 (2011): 1124-1138.

Hébert, M., Wade, E., Biron, M., DeGrâce, P., Landry, J.-F., and Moriyasu, M. 2011. The 2010 assessment of snow crab, *Chionoecetes opilio*, stocks in the southern Gulf of St. Lawrence (Areas 12, 19, E and F). DFO Can. Sci. Advis. Sec. Res. Doc. 2011/082.

Miller, Timothy J. "A comparison of hierarchical models for relative catch efficiency based on paired-gear data for US Northwest Atlantic fish stocks." Canadian Journal of Fisheries and Aquatic Sciences ja (2013).

Szuwalski, Cody, and André E. Punt. "Identifying research priorities for management under uncertainty: The estimation ability of the stock assessment method used for eastern Bering Sea snow crab (*Chionoecetes opilio*)." Fisheries Research (2012).

Appendix 2: CIE Statement of Work

Statement of Work

External Independent Peer Review by the Center for Independent Experts

Bering Sea Snow Crab Stock Assessment Review

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 Representative (COR), 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 snow crab assessment was last reviewed by the CIE 2008. Since that time, the analyst has made a number of improvements to the model. These changes should be reviewed by an independent panel. The snow crab assessment is a high profile assessment which has undergone significant change in results due to incorporation of data on catchability of the survey net and estimation of natural mortality. This review would encompass the Bering Sea trawl survey data, additional survey data used in estimation of catchability, the stock assessment model structure, assumptions, life history data, and harvest control rule. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the tasks and ToRs described in the SoW herein. The CIE reviewers shall have expertise in conducting stock assessments for fisheries management, and be thoroughly familiar with various subject areas involved in stock assessment, including population dynamics, size-structured models, harvest strategies, survey methodology, and the AD Model Builder programming language to complete the tasks of the scientific peer-review described herein. Familiarity with invertebrate stock assessment, knowledge of crab life history and biology, harvest strategy development is desirable. Each CIE reviewer is requested to conduct an impartial and independent peer review in accordance with the ToRs herein. The CIE reviewer's duties shall not exceed a maximum of 14 days conducting pre-review preparations with document review, participation in the panel review meeting, and completion of the CIE independent peer review report in accordance with the ToR and Schedule of Milestones and Deliverables.

Location of Peer Review: Each CIE reviewer shall participate and conduct an independent peer review during the panel review meeting scheduled at the Alaska Fisheries Science Center (AFSC) in Seattle, Washington during the tentative dates of January 21-24, 2014.

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

Tasks prior to the meeting: The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review in accordance with the tasks and ToRs within the SoW. Upon completion of the independent reviewer selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, and FAX number) to the contractor officer's representative (COR), who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment ToRs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be 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 COR prior to the commencement of the peer review.

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (not by email) the 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/.

Pre-review Background Documents: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the COR the necessary background information and reports (i.e., working papers) for the reviewers to conduct the peer review, and the COR will forward these to the contractor. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

Tasks during the panel review meeting: Each reviewer shall conduct the independent peer review in accordance with the SoW and stock assessment ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and contractor.** Each 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 stock assessment ToRs as specified herein. The NMFS Project Contact will be responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact will also be responsible for ensuring that the Chair understands the contractual role of the reviewers as specified herein. The contractor can contact the COR and NMFS Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Tasks after the panel review meeting: Each reviewer shall prepare an independent peer review report, and the report shall be formatted as described in **Annex 1**. . If any existing Biological Reference Point or their proxies are considered inappropriate, or if an inappropriate model formulation is identified, the report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report shall indicate that the existing BRPs are the best available at this time. Additional questions and pertinent information related to the assessment review addressed during the meetings that were not in the ToRs may be included in a separate section at the end of an independent peer review report.

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.

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 at Seattle, Washington during FeJanuary 21through January 24, 2014.
3) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than February 7, 2014, each CIE reviewer shall 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 CIE Regional Coordinator, Dr. David Die via email to ddie@rsmas.miami.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.

17 December 2013	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
23 December 2013	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers.
21-24 January 2014	Each reviewer shall conduct an independent peer review during the panel review meeting in Seattle, Washington
7 February 2014	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
21 February 2014	CIE submits CIE independent peer review reports to the COR
28 February 2014	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:

(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,

(2) The 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 COR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COR. The COR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COR

NMFS Office of Science and Technology

1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910

William.Michaels@noaa.gov Phone: 301-427-8155

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:

Jack Turnock

Resource Ecology and Fisheries Management Division

Alaska Fisheries Science Center

7600 Sand Point Way NE

Seattle, WA 98115

jack.turnock@noaa.gov Phone: 206-526-6549

Douglas Demaster, Science Director

Alaska Fisheries Science Center

17109 Pt Lena Loop Rd.

Juneau, AK 99801

douglas.demaster@noaa.gov Phone: 907-789-6517

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. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed. The CIE independent report shall be an independent peer review of each ToRs.

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

Annex 2: Terms of Reference for the Peer Review

Bering Sea Snow Crab Stock Assessment Review

The report generated by the consultant should include:

1. A statement of the strengths and weaknesses of the Bering Sea snow crab stock assessment and stock projection models;
2. Recommend for alternative model configurations or formulations if required;
3. Recommendations of alternative model assumptions and estimators if required.
4. A review of fishery dependent and fishery independent data inputs to the stock assessment.
5. Recommendation on research needs that would reduce uncertainty in key parameters used or estimated in the assessment.
6. Suggested research priorities to improve the stock assessment.

Annex 3: Agenda for Panel Review Meeting

Bering Sea Snow Crab Stock Assessment Review

NOAA Alaska Fisheries Science Center

7600 Sand Point Way NE, Seattle, WA 98115

Contact: jack.turnock@noaa.gov Phone: 206-526-6549

21-24 January 2014

Tuesday, January 21

09:00 Welcome and Introductions

09:15 Role of chair and reviewers, terms of reference

09:30 Overview (species, surveys, fishery, catch levels, bycatch)

10:00 Biology (growth, natural mortality, diets, spawning areas, nursery areas, maturity curves, mating, molting frequency)

11:00 Field experiments on escapement, discard mortality, tagging

11:30 Age Determination, shell condition

12:00 Lunch

13:00 Biology continued

14:00 Harvest control rules and overfishing definition

15:00 Ecosystem considerations - Predation, prey

16:00 Summary of on-going research

Egg viability

Migrations and movement

Larval drift

Spatial modeling

Management Strategy Evaluation

Wednesday, January 22

09:00 Survey methodology and analysis

12:00 Lunch

13:00 Description of stock assessment and projection model

Thursday and Friday, January 23-24

Reviewer discussions with assessment authors. Review of requested model runs if required.

Appendix 3: Addition review information on Somerton et al. (2013)

Let t be the total number of sampled sites and let T denote the total number of possible tow sites in the survey area (T >> t). The EBS NMFS survey is a systematic survey but I will assume it is simple random sampling for estimation my purposes here. Let r_w denote the survey selectivity for crab of width w. Somerton et al. (2013) indicated that selectivity is a complex function of carapace width, depth, and sediment type. Hence, survey selectivity is not constant throughout the stock area; it is location dependent. Let r_{wi} denote the selectivity at site i. Assume that the survey catch at site i, denoted as I_{wi}, is an unbiased estimate for trawlable abundance at site i; that is $E(I_{wi}) = r_{wi}N_{wi}$, where N_{wi} is the number of crabs in width category w at site i.

With simple random sampling the average survey catch ($\bar{I}_w = t^{-1} \sum_{i=1}^{t} I_{wi}$) is an unbiased estimator of the average catch at all possible tow sites,

$$E(\bar{I}_w) = T^{-1} \sum_{i=1}^{T} r_{wi} N_{wi} = \bar{r}_w \bar{N}_w,$$

where $\bar{N}_w = T^{-1} \sum_{i=1}^{T} N_{wi}$ and $\bar{r}_w = (\sum_{i=1}^{T} r_{wi} N_{wi}) / \sum_{i=1}^{T} N_{wi}$.

The issue for stock assessment purposes is how to estimate $\bar{r}_w$ given estimates of r_{wi}, denoted as $\hat{r}_{wi}$. A "plug-in" estimator is

$$\hat{r}_w = \frac{\sum_{i=1}^{t} \hat{r}_{wi} I_{wi}/\hat{r}_{wi}}{\sum_{i=1}^{t} I_{wi}/\hat{r}_{wi}} = \frac{\sum_{i=1}^{t} I_{wi}}{\sum_{i=1}^{t} I_{wi}/\hat{r}_{wi}}$$

This is different from the approach proposed by Somerton et al. (2013), which I think is

$$\tilde{r}_w = \frac{\sum_{i=1}^{t} I_{wi} \hat{r}_{wi}}{\sum_{i=1}^{t} I_{wi}}.$$

It seems they are using an average weighted by a biased estimate of N_{wi}.