Center for Independent Experts (CIE) Independent Peer Review of the Generalized Model for Alaskan Crab Stocks (GMACS)

Malcolm Haddon

CSIRO Marine and Atmospheric Research
Castray Esplanade
Hobart 7001
Tasmania
Australia

Malcolm.Haddon@csiro.au

Executive Summary

The June 2015 version of GMACs, the Generalized Modeling for Alaskan Crab stocks software and assessment framework, was reviewed, particularly with respect to its use with the Bristol Bay red king crab assessment. The review included the software and its documentation, its ease of use, and its potential for future use and development.

Given the current state of development of the code base there are limits to the conclusions that can be drawn in relation to some of the terms of reference. The framework within which the code is being developed is certainly fully mature, and is being used to excellent effect so that the collaborative development of code and documentation is now proceeding very well. The use of GitHub and the Wiki for documenting the model simplifies both the dissemination of the material and understanding its content. The structural framework of the modeled dynamics is already clear and well set-up but some of the fundamental aspects of the model (continuous-*F* versus within-year seasonal dynamics) need final design decisions and implementation before complete comparisons can be made between a GMACs assessment and the current BBRKC model assessment.

The great potential of the GMACs software is clearly apparent, but the current functionality belies its future capacity. The GMACs software already has advantages over the current BBRKC assessment model in its readily accessible documentation (although this is still incomplete, but again the potential is obvious), in its use of an easily used R package to display model output results in a repeatable, potentially flexible, and clear manner, and in its ability to continue development and maintenance into the future independently of particular individuals. However, until the developments required to produce the necessary management outputs that are currently output from the BBRKC model it does not yet form a replacement for that software. This could change very rapidly given just a few extra developments.

It is recommended that:

- development of the GMACs software suite continue as rapidly as possible so that its potential can be fully realized (in particular implementation of the within-year dynamics proceed and the on-going generalization of the code to open it up to other species be progressed with priority). Sufficient time needs to be given to the software authors to permit these developments.
- other development possibilities include considering alternative growth descriptions, examining alternative interpretations of how the size at maturity has changed through time, and generalizing many currently fixed constants, such as the size structure used, so that it becomes user selected.
- a formal committee is formed or a schedule of formal meetings is arranged to facilitate the interaction between the GMACs development team and, at least, the Crab Planning Team (plus any additional individuals with applicable expertise), so that design details and directions can be agreed upon, and there is a rise in confidence that the new assessment model will produce what is required by industry and the managers. Some suggestions are listed under ToR 1 and below.
- A core development team be identified to act as final arbiters over the accepted design to be implemented. This should include the main software authors, but also some of the potential users (perhaps from the Crab Planning Team and or the Scientific and Statistical Committee charged with the care of the Alaskan crab fisheries).

Background

Most stock assessments for more valuable fished stocks are conducted using a formal stock assessment model. These are used to produce estimates of the biological reference points that form the basis of a harvest control rule or rules that are used to provide management advice for a fishery. By their nature a fully developed stock assessment model is complex, especially for a fishery which has multiple fleets catching the species concerned, possibly in multiple areas with multiple methods. Such assessment models and their related software often take a number of years to develop and implement fully, and because they are often put together by fishery scientists rather than professional software programmers they are often not necessarily well documented. Because they have been designed with a particular species and fishery in mind, it is also often the case that they are highly specialized to a particular fishery and hence not particularly flexible, so that applying them to a different species can be difficult without having to introduce an array of changes to match differences between the species and fishery.

The preceding paragraph should not be construed as a criticism of stock assessment software authors who have taken on this difficult task, which often continues for years as changes are requested and extra diagnostics added. It is important to understand that writing such software to a level that it can produce results which are defensible in relation to the management of sometimes extremely valuable fisheries is difficult and time consuming.

There has been a move to using standardized software with examples such as Stock Synthesis (Methot and Wetzel, 2013) and CASAL (Bull et al., 2005) being developed. Such software is centralized and usually freely available, often has a development team rather than an individual (though often an individual drives its development) and is widely tested and examined. It is designed to be flexible with multiple options available with respect to how it can handle multiple fleets, different fishing gear selectivity, and spatial issues (along with many other options). There are numerous advantages to using such standardized code:

- There are more fisheries scientists in the world who can use sophisticated software than can write sophisticated software, so more species should be able to be formally assessed when using generic assessment software.
- Such generic software is examined and used by many people so the chances of discovering the more subtle bugs that can occur in complex analytical code is greater or at least quicker than is likely to occur when using custom built software for each species.
- It means a consistent approach can be applied across a wide range of different species. Having output statistics and data plots with the same format aids in the communication of complex outputs across an array of different stock assessment groups.
- Generic software can become institutionalized, meaning that it need no longer be solely reliant on a particular individual but rather can be further developed and maintained through having public documentation and possibly a core development team whose members can change through time. This insures the accepted stock assessment process against personal accident, illness, or staff movements occurring with the people involved with custom software.
- Changes or additions to the code or the outputs requested or required through time can be shared across all assessments that use the software rather than have each implement them separately.

Alaskan Crab Fisheries

The Alaskan crab fisheries include some very valuable stocks with the Red King Crab stock in Bristol Bay, Alaska being the basis for one of the United States most valuable fisheries (Zheng and Siddeek, 2015) and it is currently assessed using a custom built stock assessment model (Zheng et al., 1995, and later). In Alaska there are 10 crab species that have been targeted (three fisheries are currently closed) but only a few have well developed formal assessments. It would be beneficial to have a stock assessment model suited to assessing not only the Bristol Bay red king crab, but also the eastern Bering Sea snow crab, the eastern Bering Sea Tanner crab, and the Norton Sound red king crab. If it were capable of providing assessments for the six other king crab species (Aleutian Island golden, the Pribilof Island red, blue, and golden, the St Matthew blue, and the Adak red king crab) some of which are relatively data-poor, that would also be a valuable advance. One complication is that most development of formal stock assessments capable of integrating a wide range of input data has been carried out on age-structured models (Maunder and Punt, 2013). Crab species, along with many other invertebrate species, are difficult or impossible to age routinely and cheaply, so instead of age-structured models, models with size-based dynamics have been developed (Sullivan et al., 1990; Zheng and Siddeek, 1995; Punt et al., 2013). Unfortunately, to date, no-one has developed what could be termed a generic size-based model that could be applied to the Alaskan crab stocks.

To address this need a project was initiated to develop such a generic code base for size-based integrated stock assessment models (Maunder, 2012). This aimed to generate a set of re-usable code modules relating to size-based dynamics that would be suited to handling the variations present in the typical size distributions, the need to distinguish gender and maturity status, some spatial details of a fishery, and the different shell conditions (old and new shelled animals) that occur in crab fisheries; specifically in the Alaskan crab fisheries. To meet the need, the notion of GMACs (Generalized Modeling for Alaskan Crab stocks), was developed and put into action (Whitten et al., 2014). This is not a minor undertaking but progress has now started to accelerate, and this current review is partly aimed at determining whether the current development is proceeding according the most effective path.

The Terms of Reference

The GMACs team requested that the original six terms of reference, detailed in the Terms of Reference document from the CIE, be replaced by the following five; given that these were primarily a re-structuring of the original six, involved no substantive changes to the ToR, and were the preferred headings for commentary to the GMACs team they were accepted.

Original Terms of Reference

1. Evaluation of modeling framework as a flexible assessment modeling tool.
2. Evaluation of functional forms and estimation frameworks, used in GMACS.
3. Evaluation, of the diagnostic products of GMACS. and the ability to compile assessment documents
4. Evaluation of utility of the modeling framework as a community supported modeling approach and practicality for managing into the future.
5. Evaluation, of the flexibility of the model to address the stock specific characteristics of assessed species including: spatial management (e.g., Tanner crab), data poor (Tier 4 assessments), and stocks with unique data sets (e.g., snow crab and Bristol Bay red king crab cooperative survey data).
6. Recommendations for further improvements and comment on the general applicability to the fishery management questions (risk assessment and MSE options).

Alternate Version of the Terms of Reference

1. *Evaluation of functional forms, estimation approaches, and diagnostics used in GMACS including uncertainty characterization and satisfying required assessment elements.*
2. *Evaluate application of GMACS to the BBRKC stock. Specifically, comment on what important features are missing relative to the current assessment approach?*
3. *Evaluation of GMACS as a flexible assessment modeling tool. i.e., potential application to other crab stocks e.g., Tanner crab, data poor stocks (Tier 4), and stocks with unique data sets (e.g., snow crab and Bristol Bay red king crab cooperative survey data).*
4. *Evaluation of utility of the modeling framework as a community supported modeling approach and practicality for managing into the future.*
5. *Recommendations for further improvements and comment on the general applicability to the fishery management questions (risk assessment and MSE options).*

Review Activities

Access was given to a website containing the documents listed in the first part of Appendix 1 on 16th June 2015:

http://www.afsc.noaa.gov/REFM/stocks/Plan_Team/crab/draft_assessments.htm

This contained links to other web sites, documents, and software repositories; some were difficult to download on site at the Alaskan Fisheries Centre, but download was straightforward elsewhere.

Other documents listed in Appendix 1 were available on the websites, provided once the review began or were already in my personal collection of size-based modelling papers and documents.

Review documents were downloaded within the first two days of going through the website, which enabled preliminary reading of the current Bristol Bay King Crab assessment (Zheng and Siddeek, 2015) and the GMACs version (GMACS Development Team, 2015b), plus the other GMACs documents. Only those parts of the code-base relating to primary population dynamics (growth, mortality, and recruitment) were examined.

On Monday 29th June there were PowerPoint presentations by Jim Ianelli, Robert Foy, and Jie Zheng:

Ianelli, J., Martell, S., and D. Webber: GMACs Overview
Foy, R.: BSAI crab stocks Management and Assessment
Zheng, J. and M.S.M. Siddeek: Overview of Bristol Bay Red King Crab Model

In addition, Jim Ianelli presented a demonstration of the versioning software Git and the online repository GitHub as it is being used in the development of the GMACs software.

Each of these presentations were accompanied by questions and responses.

On Tuesday 30th June there were three presentations, one relating to the GMACs equations, one relating to a comparison between the current model of Bristol Bay King Crab Fishery and the latest GMACs version of the same fishery, and the final presentation was on a second simulation framework (Stockhausen, 2015) aimed at providing an independent and external testing framework for the new software.

On Wednesday 1st July discussions continued along with further presentations, one on data poor applications of GMACs (Martell and Ianelli, 2015) and another, which was an extension of a previous talk where the equations used in GMACs were further articulated.

DISCLAIMER

The information in this review has been provided by way of review or materials provided only. The author makes no representation, express or implied, as to the accuracy of the original information and accepts no liability whatsoever for either its use or any reliance placed on it.

Summary of Findings

1. Evaluation of functional forms, estimation approaches, and diagnostics used in GMACS including uncertainty characterization and satisfying required assessment elements.

It must be emphasized that while the model is now functional and can be used, it is still heavily under development. This means that apparent omissions are mostly due to their still to be implemented. The direction of future implementations will also be discussed under Terms of Reference 3 – 5.

Functional Forms

The equations used to describe the population dynamics have similarities with early descriptions of size based models (Sullivan et al., 1990; Sullivan, 1992) but mostly reflect the first model of Bristol Bay King Crab (Zheng et al., 1995) and the current stock assessment (Zheng and Siddeek, 2015). These general equations are appropriate for most Alaskan crab fisheries, with modifications (some are already noted in the developing code), for example, relating to the occurrence or not of a terminal molt, needed to cover all species fished.

The growth dynamics are a major influence in size-based models (Punt et al., *In press*) and the description of growth used in GMACs accounts for both the probability of molting and, if molting has occurred, the subsequent growth increment; this also allows for the distinction between old and new carapaces (a function of whether molting has occurred). This approach to the description of growth more closely reflects the underlying biological processes, and can be expected to be more plausible and accurate. It is also more data demanding that other simpler approaches (Chang *et al*., 2012) and so may not suit some of the more data-poor crab species in the Alaskan fishery, which may need alternative growth descriptions, but is well suited to the Bristol Bay King crab.

The current implementation of GMACs uses a continuous-*F* approach that operates with instantaneous fishing mortality rates across a full year. As a preliminary step to getting the assessment software functional this is a good first approximation, but to adhere to the generic technical specification (Maunder, 2012), and to allow for the idiosyncrasies of each fishery it will be necessary to convert the equations describing the dynamics into a step-wise seasonal interpretation of events, with the duration of each 'step' being approximated by the proportion of natural mortality imposed during each such 'step'. This is required, for example, in the Bristol Bay king crab assessment to allow the male spawning stock biomass to be estimated as it stands on February 15th, which specific value is currently used to guide management actions. In addition, given that the fisheries for this and other species tends to operate only over a proportion of each year, then the sequence of the dynamics (e.g. growth happens mostly after or before commercial fishing) can only reasonably be captured by including a step-wise or seasonal structure to the within-year population dynamics.

The recruitment dynamics reflect the development of stock assessments in the high latitude fisheries, which have recognized that large changes in productivity can occur for periods of time in major fisheries. An outcome is that formal stock recruitment relationships are not used directly in the crab assessment model of dynamics, but rather different periods are assumed to have different mean recruitment levels with recruitment deviates being fitted within each of those periods.

There is a potential interaction between the predicted annual recruitment, and the fact that the size structure assumed in the model starts at 65mm and extending in 5mm size classes to 165mm for males and 145mm for females. If the recruitment entering the fishery each year is ever modelled directly as deviates from a stock recruitment relationship (rather than an estimated mean value for a given period) so as to introduce into the dynamics some density dependency between spawning stock size and recruitment, then notice would need to be taken of the time delay between the larval production from a given spawning biomass and the animals from such a cohort entering the model at 65mm. This issue is ameliorated in the current model by using a Gamma function to describe the proportional distribution of new recruits across the first five size classes; this is a reasonable sub-model of how new entrants to the model dynamics (the new recruits) grow into the model's size-structure.

In the code it appears that a number of variables, such as the number of size classes for males and females, and their initial and maximum sizes, are currently fixed in the model. When GMACs is extended to other species it would be reasonable to set such constants dynamically to avoid having to have a different model for each species.

Other details could be considered once the major structural assumptions (continuous vs seasonal) are dealt with. One that stands out is the size at maturity for females.

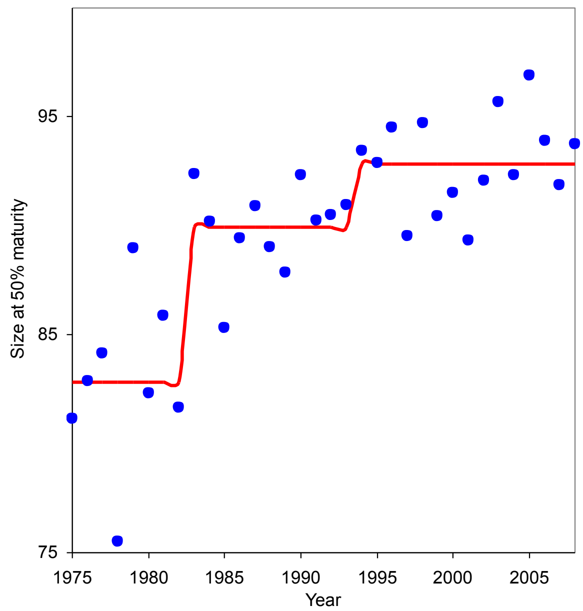

Figure 1. The size at maturity for female King crabs with three levels assumed for 1968 – 1982, 1983 – 1993, and 1994 – 2014. Copied directly from Zheng and Siddeek, 2015, PowerPoint).

In Whitten et al. (2014) maturity is also described as applying in blocks of time-steps. An alternative possibility is that the size at 50% maturity has been changing in a continuous manner from the 1970s to the 2000s (a curve could easily be fitted through the available data) so that possibly a two parameter mode could be used to replace a three (and possibly more) parameter model of the 50% maturity size.

While some such recommendations for future developments can be made, it also needs to be remembered that writing software is a heuristic and iterative process. Many such recommendations are possibly already in train but the state of development, which has focused on getting the main structural components in place and functional, has not led to their implementation yet.

Estimation Approaches

The model dynamics currently follow the numbers-at-length with details of gender, maturity, shell state, taking into account growth, natural mortality, and fishing related mortality. Changes will be required to describe the move to seasonal within-year dynamics.

The listing of the likelihoods used to fit the model to data appear completely appropriate; it is assumed that the equations given in Whitten et al. (2014) reflect the model structure except where they are replaced by versions presented in the latest documents (GMACS Development Team 2015a). There are likelihoods calculated for the catch (or discards) as biomass, the length-frequency of the catch, the index of relative abundance (in the case of the Bristol Bay king crab this is a survey index), and the fully selected effort data (this latter was yet to be fully implemented). In addition, there are penalty terms used to limit the range of the deviations in recruitment from the predicted average recruitment, to limit the between-size-class variation in the estimated initial size-structure, a penalty on the extent of inter-annual variation in fishing mortality for each fleet from its reference level, and a penalty on the sex ratio of recruits (GMACs Development Team, 2015a; Whitten et al, 2014; Zheng and Siddeek, 2015, PowerPoint). Others will be required as different aspects of the biology are included (the different periods of natural mortality rates, for example).

The fitting process iterates through generating a vector of trial model parameters and then compares the model predicted values for each of the fitted variables with those observed from the fishery, calculating the likelihood for each predicted variable and then combining the separate contributions into one likelihood. The iterations are conducted by the non-linear solver (ADMB) such that the combined likelihood is maximized (the negative log-likelihood is minimized) and when the likelihood is optimized the model is said to be fitted to the data. The model is termed an integrated analysis because it combines likelihoods from so many sources. This has the advantage that it allows for internal consistency in the model assumptions and the uncertainty associated with the multiple data sources can be appropriately propagated through to the final model outputs (Maunder and Punt, 2013; Punt et al., 2013); however, there is a large caveat on such statements. Each of the likelihoods and penalties has its own lambda value (λ), which is used to ascribe a relative weight to each of the likelihood components when they are combined into the single integrated likelihood, e.g. the λ_3 and λ_6 in the following taken directly from Whitten et al. (2015):

$$L_3 = \frac{1}{2(\lambda_3)^2} \sum_s \sum_t \left(1nI_t^s - 1n\hat{I}_t^s \right)^2 / \left(\sigma_t^s \right)^2 \tag{D.2}$$

$$P_1 = \frac{1}{2(\lambda_6)^2} \sum_t \varepsilon_y^2 \tag{E.1}$$

It is clear that the final 'weighted-likelihood' will be highly dependent upon what values are ascribed to the lambdas. How the lambdas are ascribed in either the current assessment or the GMACs software is not yet clear. This is not to claim that how their values are imposed is invalid or incorrect, merely that it is not yet documented. But it does mean that comparing the outputs of GMACs with the current model is further complicated through needing to know the values attributed to the lambdas. One innovation in GMACs is the projected use of the Dirichlet probability density function to model the size composition data. This seems to be a useful approach as it permits the prospective lambda value for the size-composition data to be estimated in the model. How this might interact with the other lambda values used in the model remains unclear.

This issue with the different weighting factors (Lambda values) is not unique to the Bristol Bay king crab assessment, and is common to all integrated assessments. Indeed, in the latest Bristol Bay king crab assessment the authors provide a response to a comment from the Scientific and Statistical Committee which read: *The SSC looks forward to the additional work planned by the author: implementing a random walk for natural mortality, investigation of recruitment dynamics, and investigation of survey weighting."* (Zheng and Siddeek, 2015, p 5). So the issue is well known and the development of GMACs is an opportunity to highlight and clarify its importance.

Use of AD Model builder means that the model should be capable of great expansion should this be required, and also that it can routinely provide asymptotic standard errors around the model parameters. It will also be possible to use Monte Carlo Markov Chain methods routinely to provide a more plausible characterization of the variability around parameter estimates and model outputs. At the current stage of development it is not surprising that a request to run an MCMC failed to proceed past about 12,000 steps. Despite the failure, which may have happened for a large array of reasons, the potential is there and it is used routinely in many other stock assessments. There are many advantages to using AD Model builder and the relatively simple availability of MCMC routines is one of them.

Diagnostics

Management Strategy Evaluation

The rsimGMACs package (Stockhausen, 2105) is an R based package designed to enable management strategy evaluation of the application of the GMACs model to data simulated from a population dynamics model of crab populations that is not necessarily identical to the GMACs software. It is still under development although already appears to be relatively flexible and capable of providing simulated data from a wide range of potential scenarios. The inclusion of MSE software independent of the GMACs software is a forward looking strategy that will enable the GMACs software to be well tested, along with alternative harvest strategies that might be considered (Zheng et al., 1997). Internal to the GMACs software there is also the option to simulate data from the model dynamics itself. This too is important as a fundamental test of stock assessment methods is that they are able to recover the hypothetical system status from data simulated from themselves. If they cannot this would be evidence that there are some fundamental model structural issues.

Model Diagnostics

Included in the GMACs software suit is another R package (gmr) which is used for plotting out the model results. This fulfills two functions, one of standardizing output plots for ease of communication, and completeness while providing the means of developing new plots and output formats should they be required. Secondly, this package provides for the automatic production of the very numerous diagnostic plots possible concerning model fits to the multiple data sources used in integrated analyses. This is analogous to the plots possible from the R package R4SS, which is used with Stock Synthesis, although, of course, customized to suit size-based modeling. While still under development there appears to be a full range of plots with better able to be added relatively easily. The plots can make good use of color, and provide for alternative views of the same results. The provision of such diagnostic plots and tables enables a closer approach to determining the representativeness of the sampling that provided the data that was used in the modeling. It also enables the quality if the model fit to the data is to be examined in great detail. This approach constitutes an enhancement to openness and communication, especially as the use of the package is simple and could be conducted by anyone wishing to do so.

2. Evaluate application of GMACS to the BBRKC stock. Specifically, comment on what important features are missing relative to the current assessment approach?

GMACs is still currently under development and its continuous-*F* strategy means that comparisons with the current Bristol Bay King crab model will undoubtedly find differences. Any differences will also be influences by the potentially different weightings (lambda values) applied to the likelihood values from different data sources. Strictly a formal comparison could only be made once the GMACs model is further along in its development, nevertheless, some comparisons are possible (GMACs Development Team, 2015b).

The current Bristol Bay red crab assessment (Zheng and Siddeek (2015), its initial GMACs equivalent, and the Excel spreadsheet comparing the structural assumptions of GMACs and the current assessment provide numerous areas that can be compared. A perfect match between the current custom model and the GMACs outputs is aspired to, but given even slight differences in assumptions, then while the outcomes can be expected to be very similar there is no reason to expect a perfect match.

There are similarities between the two assessment outcomes, with the current model appearing to fit the male survey biomass time trajectory somewhat better than the GMACs model although the differences in the male fit are not extreme. The GMACs two-sex model appears to follow the female biomass estimates rather better, at least up until the early 2000s (Figure 1 in GMACs Development Team 2015b).

Details of the implementation of the different periods of natural mortality are obviously important and affect the recruitment patterns, which because there is confounding with natural mortality, not surprisingly differ between the models. A major difference is found with the male spawning stock biomass trajectory and while the trends in spawning stock biomass are important, the differences in the trends between the two models is also very great.

The draft nature of the initial comparison means that this comparison is limited in how detailed a comparison can be made. It would be necessary to tabulate the different likelihoods and weightings used in generating the separate optimum model fits. Importantly, when the OFL from the two models (the BBRKC model and GMACs two-sex model) are compared the GMACs estimate is only half that of the BBRKC model, from a comparison of the table on page 22 of GMACs Development Team 2015b and tables in the executive summary of Zheng and Siddeek, (2015).

The structural differences between the two models are currently too great for such comparisons to hold much value; once the GMACs model is further developed the outputs from the two models should become more similar. But the potential of GMACs to provide a formal assessment is clear.

3. Evaluation of GMACS as a flexible assessment modeling tool. i.e., potential application to other crab stocks e.g., Tanner crab, data poor stocks (Tier 4), and stocks with unique data sets (e.g., snow crab and Bristol Bay red king crab cooperative survey data).

Currently the development of the GMACs software is focused on the Bristol Bay king crab fishery. As a strategy for developing the main structure and working modules in the software, this is sensible. The current intention appears to be that the next major changes will be to implement additions to the software to allow either snow crab or Tanner crab, or both to be assessed. Given the flexible and modular nature of the code structure it should be reasonably straightforward for the development team to implement the required changes to generalize the software to other species, though it may not be a rapid development. The output routines for tabulating and plotting the results should all remain the same, although there may be a need to implement some flexibility in the choice of seasonal structure to the events such as growth and fishing mortality, as well as selectivity curves for the various fisheries.

The review team were provided with an Excel sheet (Crab SA model feature.xls) that listed and compared the assessments for six species groups (Bristol Bay king crab, Norton Sound red king crab, Tanner crab, Snow crab, Saint Matthew blue king crab, and the Pribilof Islands red king crab) against the GMACs software specification. The idea was to see what data was shared and common across species, how different aspects of the dynamics were implemented. While, this spreadsheet remains a draft, it could certainly form the basis for more precise model specification to be used to guide the core development team as to the potential for extensions to the software in their next steps.

Some of the current issues that may need attention include:

- the size-structure used in the model; it currently appears to be very specific (16 classes for females and 20 for males). An obvious change to make the software more generic would be to make the allocation of size-classes to be dynamic rather than fixed.
- Any other constants hard-wired into the current model implementation should be reviewed to determine whether they should have the option of being allocated dynamically at the model initiation.
- The need to implement a terminal moult for females has already been noted in the code.
- The use for data-poor species would also need to be made optional, though such changes would presumably be of lower priority than the others listed.
- Similarly, the application of retrospective analyses should be automated as an option so it is available more easily.

Core Development Team

The code has now reached a stage where further steps forward may require guidance from prospective users. The Crab Planning Team would appear to be the obvious candidate for providing such guidance in that they are across the different fisheries and can bring their expertise as to what exactly would be required to go into the model structure and what they would expect to come out if it. An increase in the formal interaction between the GMACs development team and at least the Crab Planning Team is something that should be organized relatively soon so as to further facilitate the optimal progression of the software. The development of a concordance of opinion as to the best next steps and exactly what would be the optimal pathway forward is essential for the successful development of software that will be used widely.

Issues that will need immediate attention will be to agree on the seasonal structure to be implemented within each year inside GMACs (that a seasonal structure is needed appears clear), then an agenda should be developed for the next major steps, among the options suggested, in the development pathway. This will need close and formal communication between the development team and the Crab Planning Team.

4. *Evaluation of utility of the modeling framework as a community supported modeling approach and practicality for managing into the future.*

The use of Git and GitHub for software version control appears to have been implemented very effectively allowing for much more rapid collaboration and code development from multiple authors and coders (Ianelli et al., 2015). A great advantage of this software development environment is the relative ease with which the various components can be documented in a collaborative fashion. Writing software, particularly such a complex framework as GMACs, is time consuming and difficult. Classical means of documentation invariably slows development considerably and extends the time required markedly; which is why many individually developed pieces of software are rarely thoroughly documented. The current BBRKC assessment model is an exception with a number of formal publications included in its description, although without seeing its code it is not known whether the documentation of the code base is sufficient to allow someone new to come in and maintain the model as it currently stands. The GMACs development team have set up the GitHub system so that, once given access, almost anyone with some knowledge of size-based assessment models could pick up the fundamental structure and operation of GMACs in a relatively short time. As such, GMACs is in a very good position for future development and maintenance as the implementation is open and clear. For example, once access to the GitHub repository was granted, it was very simple to download the current version of the software and examine it in as much detail as desired. It would also have been very simple to have made suggested changes or incremental developments to particular routines, modules, or their documentation and re-submitted them back to the repository. This was a very powerful illustration of the utility of such software versioning tools for collaborative software development.

GitHub

The use of GitHub and the wiki documenting the GMACs software also proved to be a useful tool for allowing rapid review of wide array of documents and other details. The modelling framework being used to develop the GMACs software succeeds well in allowing this to be a ‘community supported modelling approach’, and makes managing the software development and maintenance into the future a very practical proposition.

In any collaborative arrangement where different people can work on the same piece of code or text, there needs to be some final arbiter (which may be a group of people) who makes the final decisions on what changes to implement and then which versions to adopt formally. The software is still in the process of being developed with rapid and relatively large changes continuing to be put in place. However, its development path has now reached a stage where design decisions need to be made so that implementations can be finalized, or at least crystalized (a major example is the contrast between the continuous-*F* and the within-year seasonal dynamics approaches). Currently, the GMACs development team appear to have an informal arrangement for such decisions. If GMACs is to become institutionalized and applied to a wide array of species, then this notion of a core development team also needs to become more formal. The idea is that this software not be dependent upon any particular individual, even though it takes individuals to develop it in the first place. The comments above about the cooperation between the development team and the Crab Planning Team (and potentially others) also apply here. The development team cannot be expected to know the idiosyncratic details of every Alaskan crab fishery to which this software may be applied, so this core development team should include representation from users who could assist with the requirements for the software (rather than the coding details of how to implement those requirements).

Open Source

Mention was made of the code being made open source. This is obviously a good idea with respect to the executables produced by the original code and a copy of the code itself could be made available. However, while it is likely to be beneficial to have the option of receiving suggested changes from anywhere and anyone, it is assumed that access to the core GitHub source would still be controlled to enable the core development team to manage the acceptance of changes suggested by external users.

Cooperation with Industry

The GMACs project is an Industry initiative and the project appears to be an excellent example of cooperation between Industry and the NMFS. Such cooperation will continue to be essential in ensuring that the code base is as useful as possible into the future and that its value to the Alaskan crab fishery is maintained and grows.

5. Recommendations for further improvements and comment on the general applicability to the fishery management questions (risk assessment and MSE options).

A number of possible changes have already been discussed in response to ToR 1 and 3, although from the discussion that ensued at the review it was clear that the development team were already aware of most of them.

MSE Options

The inclusion of the R-based MSE simulation framework (Stockhausen, 2015) is an innovation to be applauded. It is most common that harvest strategies that use a particular stock assessment method have been tested using management strategy evaluation in a relatively ad hoc manner, with the MSE simulation framework being developed in an ad hoc manner akin to how individual stock assessment used to be developed. In this case, a simulation framework devised in R, as a contrasting programming language to the C++ used in GMACs, should assist in both finding bugs and providing for the simulation of multiple test data sets. Working along, this reviewer often programs assessment models using both AD Model Builder and R separately. The R version is not used to fit the models, but rather as the model is intended to be a repeat of the formulation of the AD Model builder model they should generate the same answers when given the same parameter set. Having an independently written version of the same program always provides a further test of the programming.

The real function of the R-package is to conduct MSE testing of alternative scenarios to determine whether the GMACs software produced biased or inaccurate outcomes in the face of poor data or weak assumptions. The internal data simulator within the GMACs modules can play a similar part, although in this case using exactly the same dynamics as the assessment model. If GMACs cannot recover the state of the fishery, as defined by a given parameter set, this would highlight a serious flaw.

Risk Assessment

Risk assessments are understood to mean applying a harvest strategy with no feedback from year to year, so as to identify the implications of taking, for example, a constant catch or applying a constant fishing mortality. Generally, this is done stochastically so that the model is projected forward under uncertain inputs and conditions. The projection trajectories can be analyzed to determine the relative likelihood (probability) of different events occurring if the constant management being projected forward is maintained. Implementation of this is planned but is yet to be fully implemented.

Overall, the fashion in which the software is being developed and its open availability means that this will enable any group to explore the assessment and its inputs and outputs. This is a major step in openness and complete communication, which should be of great value to the industry and managers of the resources.

Conclusions/Recommendation

Given the current state of development of the code base, there are limits to the conclusions that can be drawn in relation to some of the terms of reference. The framework within which the code is being developed is certainly fully mature and is being used to excellent effect so that the collaborative development of code and documentation is now proceeding very well. The structural framework of the modeled dynamics is already clear and well set-up but some of the fundamental aspects of the model (continuous-*F* versus step-wise within-year dynamics) need final design decisions and implementation before complete comparisons can be made between a GMACs assessment and the current BBRKC model assessment.

With the current state of development the great potential of the GMACs software is apparent but the current functionality belies its future capacity. The GMACs software already has advantages over the current BBRKC assessment model in its readily accessible documentation (although this is still incomplete, but again the potential is obvious), in its use of an easily used R package to display model output results in a repeatable, potentially flexible, and clear manner, and in its ability to continue development and maintenance into the future independently of particular individuals. However, until the developments required to produce the necessary management outputs that are currently output from the BBRKC model, it does not yet form a replacement for that software.

It is recommended that:

- development of the GMACs software suite continue as rapidly as possible so that its potential can be fully realized (in particular implementation of the within-year dynamics proceed and the on-going generalization of the code to open it up to other species be progressed with priority).
- other development possibilities include considering alternative growth descriptions, examining alternative interpretations of how the size at maturity has changed through time, and generalizing many currently fixed constants, such as the size structure used, so that it becomes user selected.
- a formal committee is formed or a schedule of formal meetings is arranged to facilitate the interaction between the GMACs development team and, at least, the Crab Panning Team (plus any additional individuals with useful expertise), so that design details can be agreed upon and there is a rise in confidence that the new assessment model will produce what is required by industry and the managers.

- A core development team be identified to act as final arbiters over the accepted design to be implemented. This should include the main software authors but also some of the potential users (perhaps from the Crab Planning Team and or the Scientific and Statistical Committee charged with the care of the Alaskan crab fisheries).

1. ***Evaluation of functional forms, estimation approaches, and diagnostics used in GMACS including uncertainty characterization and satisfying required assessment elements.***

The current functions forms, estimation approaches, and diagnostics are all appropriate given the current state of development of GMACs. Having said that, it is recognized that major changes are still required to move forward. One of the biggest changes is to convert the current continuous-*F* arrangement of fishing mortality and other biological processes into more explicit within-year seasonal dynamics with the duration of each period being defined by the proportion of natural mortality expressed in each within-year time-step. It is recommended that this conversion to a seasonally explicit structure be the next large step undertaken. This will enable required assessment elements (for example, the male mature biomass of Bristol Bay red king crabs predicted to occur on February 15th each year; required for management decisions) to be more readily estimated and the sequential stock dynamics to be more convincingly modeled.

2. ***Evaluate application of GMACS to the BBRKC stock. Specifically, comment on what important features are missing relative to the current assessment approach?***

GMACs is still currently under development and its continuous-*F* strategy means that comparisons with the current Bristol Bay King crab model, which uses a more explicit seasonal within-year dynamics, will undoubtedly find differences. Strictly, a formal comparison could only be made once the GMACs model is further along in its development; nevertheless, some comparisons are possible (GMACs Development Team, 2015b). The comparisons that were possible indicate the great promise of the new assessment software, but this term of reference was difficult to complete given the current state of development of the model. Currently, while it can produce the required diagnostic plots and many of the required outputs the model needs numerous changes before it can be considered a replacement for the current BBRKC assessment model.

To make formal comparisons between the two models, more details concerning the likelihoods for each fitted data source, the λ values used in each case, the list of parameters being fitted in each case, and the full range of management related outputs need to be tabulated. Without these it is not possible to determine whether comparisons of like for like are being made. The potential of the new software is apparent as even with all the differences between the implementation of the model structures some of the outputs are very similar. However, some outputs appear to be greatly different and before the new model can be accepted the reasons for those differences, if they persist with the projected code changes, will need to be explained.

3. ***Evaluation of GMACS as a flexible assessment modeling tool. i.e., potential application to other crab stocks e.g., Tanner crab, data poor stocks (Tier 4), and stocks with unique data sets (e.g., snow crab and Bristol Bay red king crab cooperative survey data).***

Currently the GMACs model is aimed at assessing the Bristol Bay red king crab stock. With the addition of some further development (already noted by the developers) it will be applicable to the snow crab and Tanner crab fisheries. The potential is there for it to be applicable, or to be made applicable, to all king crab species in the Alaskan fishery.

The matrix of species vs properties (crab SA Feaures.xls) that was included in the review materials should be developed further and used to assist in guiding the application and potential for extensions to the software. Aside from the inclusion of the within-year seasonal dynamics and the implementation of a terminal molt for females, the priority and order in which further extensions should be made to the software should be guided by the formal communication between the GMACs development team and the Crab Panning Team. Without knowledge of the dominant issues across the Alaskan crab fisheries, then the further articulation of the GMACs software may proceed without reference to the priorities for wider fishery.

4. ***Evaluation of utility of the modeling framework as a community supported modeling approach and practicality for managing into the future.***

The use of Git and GitHub for software version control and development appears to have been implemented very effectively allowing for much more rapid collaboration and code, and documentation development from multiple authors. A great advantage of this software development environment is the relative ease with which the various components can be documented. Writing software, particularly such a complex framework as GMACs, is time consuming and difficult. Classical means of documentations typically slowed development considerably and extended the time required markedly; which is why many individually developed pieces of software are rarely thoroughly documented. The GMACs development team have set up the GitHub system so that almost anyone with some knowledge of size-based assessment models could pick up GMACs in a relatively short time. As such GMACS is in a very good position for future development and maintenance. The modelling framework being used to develop the GMACs software succeeds well in allowing this to be a ‘community supported modelling approach’, and makes managing the software development and maintenance into the future a very practical proposition.

The development path for the GMACs software suit has now reached a stage where design decisions need to be made so that implementations can be finalized (a major example is the contrast between the continuous-*F* and the within-year seasonal dynamics approaches). It is recommended that a Core Development Team be identified as the final arbiter on the final decisions about what changes to implement and then which versions to adopt formally. If GMACs is to become institutionalized and applied to a wide array of species then this notion of a core development team also needs to be formalized. The idea is that this software not be dependent upon any particular individual, even though it takes individuals to develop it in the first place.

5. ***Recommendations for further improvements and comment on the general applicability to the fishery management questions (risk assessment and MSE options).***

The design of GMACs already includes the provision of an external R-package designed to allow the generation of simulated data to be used in MSE testing of the GMACs software and any harvest strategies that use the software for an assessment. This will prove especially useful for any data-poor applications of the assessment. In addition to the external R-package, GMACs includes the ability to simulate data from its own dynamics, which provides for testing the model itself for internal consistency and stability. Given its structure and flexibility GMACs should be capable of generating outputs relating to any of the current fishery management questions that are likely to arise.

The inclusion of an MSE simulation framework as part of the overall design is an innovation. Most MSE testing is conducted using an ad hoc simulation framework devised after stock assessments have been conducted (Punt et al., *In Press*b). This inclusion from the outset could be a further advantage of using standardized stock assessment packages, although GMACs is the only one, to date, that includes such testing software as part of its design.

Recommendations about further improvements have been made in the other Terms of Reference. Once the within-year seasonal dynamics have been implemented other, possibly complementary changes will follow.

Appendix 1: Bibliography of Materials Provided

GMACS Development Team (2015a; Draft) A generalized size-structured assessment model for Crustaceans. 12 p. (GLBAM.pdf)

GMACS Development Team (2015b) Gmacs Example Stock Assessment. 23p. (bbrkc.pdf)

Maunder, M. (2012) Generic size based stock assessment model for Alaskan crabs stocks. Bering Sea Fisheries Research Foundation. 34p. (MaunderCrabSpec.pdf)

Martell, S., Ianelli, J. and D. Fournier (2015) GMACs: A generalized size-structured model and stock reduction analysis. (gmacsSRA.pdf Conference poster).

Whitten, A.R., Punt, A.E., and J.N. Ianelli (2014) Gmacs: Generalized Modeling for Alaskan crab stocks. 13p. (Whitten et al 2014 - Gmacs Model Description.pdf)

Whitten, A.R. and J.N. Ianelli (2014) GMACS: A generalized size-structured stock assessment framework: ADMB code. 47p (gmacscode.pdf)

Zheng, J., and M.S.M Siddeek (2015) Bristol Bay Red King Crab Stock Assessment in Spring 2015. 130p. (rkc-safe15s.pdf)

PowerPoint Presentations

Foy, R. (2015) BSAI crab stocks Management and Assessment. 53 slides

GMACS Development Team (2015b) Gmacs A generalized size-structured stock assessment model (gmacs.pdf 49 slides)

Ianelli, J., Martell, S. and D. Webber (2015) GMACS Overview. 45 slides

Martell, S. and J. Ianelli (2015) Some data poor applications. 7 slides

Stockhausen, B. (2015) rsimGmacs. 12 slides

Zheng, J. and M.S.M Siddeek (2015) Overview of Bristol Bay Red King Crab Model. 45 slides

No stated authorship:

Access to https://github.com/seacode/gmacs/wiki (the software manual to GMACS)

Access to http://seacode.github.io/gmacs/ (containing the code developed to date)

A Guide to the Preparation of Bering Sea and Aleutian Islands Crab SAFE Report Chapters (DRAFT) 13p. (guideline for crab assessments.pdf)

Crab SA model features.xls.xlsx (an excel sheet listing properties by species)

GKC Input Data and Likelihood 3p. (GKC Input Data and Likelihood.pdf)

Norton Sound Red King Crab modeling issues. Pp 38 – 70. (Norton Sound Red King Crab modeling issues.pdf)

Other Material Examined

Bull, B., Francis, R.I.C.C., Dunn, A., McKenzie, A., Gilbert, D.J. and M.H. Smith (2005) CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.07-2005/08/21. NIWA Technical Report 127.

Chang, Y-J., Sun, C-L., Chen, Y. and S-Z. Yeh (2012) Modelling the growth of crustacean species. *Reviews in Fish Biology and Fisheries* **22:** 157-187 DOI: 10.1007/s11160-011-9228-4

Maunder, M.N. and A.E. Punt (2013) A review of integrated analysis in fisheries stock assessment. *Fisheries Research* **142:** 61-74

Methot, R.D. Jr. and C.R. Wetzel (2013) Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* **B142**: 86 – 99.

Punt, A.E., Huang, T. and M.N. Maunder (2013) Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusk species. *ICXES Journal of Marine Science* **70**: 16 – 33

Punt, A.E., Haddon, M. and R. McGarvey (*In Press*) Estimating growth within size-structured fishery stock assessments: What is the state of the art and what does the future look like? *Fisheries Research* dx.doi.org/10.1016/j.fishres.2014.11.007

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A.A. and M. Haddon (*In Press*b) Management strategy evaluation: best practices. *Fish and Fisheries* DOI: 10.1111/faf.12104

Sullivan, P.J., Han-Lin Lai, and V.F. Gallucci (1990) A catch-at-length analysis that incorporates a stochastic model of growth. *Canadian Journal of Fisheries and Aquatic Sciences*, **47**: 184-198.

Sullivan, P.J. (1992) A Kalman filter approach to catch-at-length analysis. *Biometrics*, **48**: 237-257.

Zheng, J., Murphy, M.C. and G.H. Kruse.(1995) A length-based population model and stock recruitment relationships for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. *Canadian Journal of Fisheries and Aquatic Science* **52**:1229 – 1246

Zheng, J., Murphy, M.C. and G.H. Kruse. (1997) Analysis of the harvest strategies for red king crab, *Paralithodes camtschaticus*, in Bristol Bay, Alaska. *Canadian Journal of Fisheries and Aquatic Science*. **54**:1121 – 1134.

Appendix 2: Statement of Work

External Independent Peer Review by the Center for Independent Experts

Review of GMACS Modeling Framework

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 Alaska Fisheries Science Center (AFSC) is responsible for assessments of 5 Bering Sea crab stocks. Collectively these crab stocks support valuable commercial fisheries. Recently, scientists at the AFSC and the University of Washington have developed a **Generalized Model for Alaska of Crab Stocks (GMACS).** GMACS is a generalized and flexible size-structured modeling framework. The first version has been designed to develop stock assessment models for the stocks of red king crab in Bristol Bay and Norton Sound, Alaska. The framework makes use of most of the available data sources for both male and female crabs, including survey and fishery indices of abundance and fishery and survey size-compositions. A workshop held in January 2015 contributed to the implementation of GMACS to data for red king crab stocks in order to test its efficacy and determine priorities for development for application to other Alaskan crab stocks. A goal of the project is to provide a standardized size structured assessment that is more transparent than the current approach in which assessment scientists have written species-specific software. A CIE review of the GMACS modeling framework and issues related to further development and application for stock assessment is requested.

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. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein. CIE reviewers shall have the expertise, background, and experience to complete an independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have expertise and work experience in analytical stock assessment, including population dynamics, age/length based stock assessment models, data-poor stocks, survey design, and population structure and spatial management.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled during the week of July 29th, 2015 at the Alaska Fisheries Science Center in Seattle, Washington.

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 (name, affiliation, and contact details) to the Contract Officer Representative (COR), 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 information concerning other 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 COR 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/compliance_access_control_procedures/noaa-foreign-national-registration-system.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.

AFSC will provide copies of the statement of work, stock assessment documents, prior CIE review documents, and other background materials to include both primary and grey literature.

This list of pre-review documents may be updated up to two weeks before the peer review. Any delays in submission of pre-review documents for the CIE peer review will result in delays with the CIE peer review process, including a SoW modification to the schedule of milestones and deliverables. Furthermore, the 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.

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 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 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 at the **Alaska Fisheries Science Center** in **Seattle, Washington** during **June 29th - July 1st 2015** as called for in the SoW, and conduct an independent peer review in accordance with the ToRs (Annex 2);
3) In Seattle, Washington during June 29th - July 1st 2015 as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than **July 15, 2015**, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.com, and Dr. David Die, CIE Regional Coordinator, 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.

One month prior	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
Two weeks prior	NMFS Project Contact sends the CIE Reviewers the pre-review documents
June 29th - July 1st 2015	Each reviewer participates and conducts an independent peer review during the panel review meeting
July 15, 2015	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
July 29, 2015	CIE submits CIE independent peer review reports to the COR
August 7, 2015	The COR 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 made through the COR who submits the modification for approval to the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the CIE within 10 working days after receipt of all required information of the decision on substitutions. The COR can approve changes to the milestone dates, list of pre-review documents, and Terms of Reference (ToR) of the SoW as long as the role and ability of the CIE reviewers to complete the SoW deliverable in accordance with the ToRs and deliverable schedule are not adversely impacted. The SoW and ToRs cannot 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. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (the 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) each CIE report shall have 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 notification of 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 approved CIE reports to the NMFS Project Contact and regional Center Director.

Support Personnel:

Allen Shimada
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Allen.Shimada@noaa.gov Phone: 301-427-8174

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
NTVI Communications, Inc.
10600 SW 131st Court, Miami, FL 33186 mshivlani@ntvifederal.com
Phone: 305-968-7136

Key Personnel:

James Ianelli, Project Contact
NOAA National Marine Fisheries Service, Alaska Fisheries Science Center 7600 Sand Point Way, NE, Bldg 4, Seattle, WA 98115 paul.spencer@noaa.gov Phone: 206-526-6510

Anne Hollowed, Status of Stocks and Multispecies Assessment Program Manager NOAA National Marine Fisheries Service, Alaska Fisheries Science Center
7600 Sand Point Way NE, Seattle, WA 99115
PAnne.Hollowed@noaa.gov Phone: 206-526-4223

Steven Ignell, AFSC Deputy Science and Research Director
NOAA National Marine Fisheries Service, Alaska Fisheries Science Center 7600 Sand Point Way, NE, Bldg 4, Seattle, WA 98115

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.

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, 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 detailed summary of findings, 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 proceedings and findings of the meeting, 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 as separate appendices as follows:

   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

Review of GMACS Modeling Framework

CIE reviewers shall address the following Terms of Reference during the peer review and in the CIE reports.

CIE reviewers shall address the following Terms of Reference during the peer review and in the CIE reports.

a. Evaluation of modeling framework as a flexible assessment modeling tool.
b. Evaluation of functional forms and estimation frameworks, used in GMACS.
c. Evaluation, of the diagnostic products of GMACS. and the ability to compile assessment documents
d. Evaluation of utility of the modeling framework as a community supported modeling approach and practicality for managing into the future.
e. Evaluation, of the flexibility of the model to address the stock specific characteristics of assessed species including: spatial management (e.g., Tanner crab), data poor (Tier 4 assessments), and stocks with unique data sets (e.g., snow crab and Bristol Bay red king crab cooperative survey data).
f. Recommendations for further improvements and comment on the general applicability to the fishery management questions (risk assessment and MSE options).

Annex 3: Tentative Agenda

Review of GMACS Modeling Framework

Alaska Fisheries Science Center, Seattle, WA

June 29 – July 1, 2015

Contact for security and check-in: Anne Hollowed, Martin Dorn and William Stockhausen Contacts for additional documents: James Ianelli, Darcy Webber

Monday, June 29:

9:00 AM – 10:30 AM: **Introduction**

Topics: *Introductions, agenda, Overview, GMACS modeling framework. (Ianelli, Webber)*

10:30 AM – Break

10:45 AM – Current Bristol Bay red king crab (BBRKC) models used (*Zheng)* 12:00 PM – Lunch

1:00 PM -3:00 PM: **Input data**

Topics: *Model structure, likelihood formulations, data weighting, output diagnostics*

3:00 PM – Break

3:15 PM – **Discussions**

5:00 PM – Adjourn for day

Tuesday, June 30:

9:00 AM – 10:30 AM: **Assessment model**

Topics: *Community model development environment*

10:30 AM Break 10:45 AM **Discussions** 12:00 PM Lunch

1:00 PM -3:00 PM: **Demonstration runs BBRKC and NSRKC**

Topics: *Applications to Bristol Bay Red King Crab and Norton Sound Red King Crab; Issues and concerns for implementation for other crab stocks*

3:00 PM – Break

3:15 PM – Discussions 5:00 PM – Adjourn for day

Wednesday, July 1:

9:00 AM – 10:30 AM: Open Discussions Topics: *Open discussion with analysts as needed*

10:30 AM – Break

10:45 AM – Discussions 12:00 PM – Lunch

1:00 PM -3:00 PM: **Alternative model runs, further discussion as needed**

Topics: *TBA*

3:00 PM – Break

3:15 PM – Further discussions and summarize 5:00 PM – Adjourn meeting

Appendix 3: Panel Membership

Michael Bell – United Kingdom
Nick Caputi – Australia
Malcolm Haddon – Australia