REPORT ON

BERING SEA & ALEUTIAN ISLANDS

CRAB STOCK OVERFISHING DEFINITIONS

24-27 APRIL, 2006

SEATTLE, WASHINGTON

Prepared by

Patrick Cordue
Fisheries Consultant
New Zealand

for

University of Miami
Independent System for Peer Review

6 June 2006

EXECUTIVE SUMMARY

A CIE Review Panel considered a proposed overfishing definition for Bering Sea and Aleutian Islands crab stocks from April 24-27, 2006 at Alaska Fisheries Science Center, Seattle, WA. The existing definition had been found to be in need of revision and an interagency work group had been charged with developing a new definition. They had encountered difficulties in doing this and a two-day workshop had been held to discuss and resolve issues. The CIE review took place about 8 weeks after the workshop. In the interim, the work group had continued working on the overfishing definition framework. In particular, they had attempted to find suitable default parameter values and proxies needed to complete the overfishing definition.

The proposed overfishing definition is an improvement on the existing definition in that it provides some constraint on fishing mortality. The existing definition is flawed in concept, does not constrain fishing mortality, and needs to be replaced.

The proposed definition is:

- an improvement on the existing definition
- comprehensive (as a framework)
- borrowed from groundfish (so is already reviewed to some extent).

Weaknesses of the proposed definition:

- complicated
- it is still a work in progress
    - default values for parameters are not yet determined
    - sensible definition of biomass in the stock recruit relationship is not determined/specified
    - criteria for determining optimal default parameters are not determined/specified
- extensive simulations are needed to determine suitable default parameters
- potentially, it may unnecessarily constrain harvest strategies.

I make several recommendations. The most important of these concern two central issues: the definition of biomass in the stock recruit relationship, and the criteria for choosing between overfishing-definition MSY control rules.

The issue of the definition of "biomass" in the stock recruit relationship is peculiar to crabs because fishing mortality is only directed at males. In groundfish stocks it is not an issue because female spawning biomass is a good proxy for total fertilized egg production. For crabs it is a crucial issue for the proposed overfishing definition because the biomass proxy for total fertilized egg production is a primary determinant of F_{MSY} and F_{MSY} proxies. To date, the analysis of this issue has been inadequate. Immediate efforts

need to go into the derivation of appropriate functional forms. In the short term, if a default definition is needed, mature male biomass should be seriously considered.

Also, there is the issue of what constitutes a “good” overfishing definition, in general, and for Bering Sea Aleutian Islands crabs in particular. The answer to this question needs to be clearly stated. It is then relatively straightforward to define the analysis and simulations needed to test alternative overfishing definitions (and to determine default parameter values for the MSY control rules in the proposed tier system). The function of an MSY control rule in an overfishing definition must be acknowledged. The preliminary simulations aimed at determining default parameter values tested MSY control rules as rebuilding plans and harvest strategies. They are neither. MSY control rules must be evaluated in conjunction with harvest strategies (either existing harvest strategies, or a default harvest strategy).

The parameterization of the proposed MSY control rules implies a reduction in F at B_{MSY}. It does allow flat control rules (alpha = –infinity) but it precludes the suggested default overfishing definitions of Restrepo et al. 1998 (where the reduction in F occurs below B_{MSY}). I suggest that an extra parameter is added to the framework to allow MSY control rules of the form proposed by Restrepo et al. (1998). In the absence of this parameter, the proposed framework may unnecessarily restrict harvest strategies.

BACKGROUND

A three person CIE Review Panel considered a proposed overfishing definition (OFD) for Bering Sea and Aleutian Island (BSAI) crab stocks from April 24-27, 2006 at Alaska Fisheries Science Center, Seattle, WA. The North Pacific Fishery Management Council had determined that the existing OFD needed revision. A four member interagency work group had been charged with developing the new OFD. They had already participated in, and taken direction from an interagency workshop on crab OFDs which had met February 28-March 1, 2006. Simulation studies, aimed at determining default parameter values and proxies needed in the proposed OFD framework, were undertaken between the OFD workshop and the CIE review meeting.

This report presents my personal view with regard to the proposed OFD and the methods and techniques needed to determine appropriate default parameter values and proxies. I also comment on the stock assessment models and estimation methods in general. Finally, I suggest some research priorities. This report should be read in conjunction with those of my fellow reviewers Dr Mike Bell and Dr Nick Caputi. Although there was no attempt to reach a consensus on any of the issues it was apparent that the Review Panel shared many common views with regard to the proposed OFD and associated research.

REVIEW ACTIVITIES

Meeting Preparation

Prior to the meeting I read the main documents and consulted the background material made available on a website (Appendix 1). I also consulted material on the Web and conversed with colleagues with regard to crab biology.

Meeting Attendance

A brief narrative of the meeting is given below.

24 April

The meeting was convened at 8.30 am and began with a round of introductions. The meeting Chair, Dr Anne Hollowed, gave an introductory presentation on the purpose of the review and the "charge for the CIE". Dianna Stram reviewed the history of crab management and the existing OFD and the reasons for revision. Simply put, the existing OFD had been rushed through; it was conceptually flawed and provided no constraint on fishing mortality.

The four member Working Group then covered material relating to their statement of work, that of the two-day workshop, the proposed OFD structure (tier system and parameters) and two example stock assessments (snow crab and red king crab).

The Review Panel asked many questions during the presentations. We were aware that slow progress was being made in terms of the original agenda but thought that it was best to fully explore the issues during the presentations. We had already advised the Chair that we would not need to use the whole week. The scheduled "writing team" days were not needed as Panel members agreed that we could best do this after returning to our home locations.

25 & 26 April

The meeting resumed at 8.30 am with Dr Jim Ianelli in the Chair. We began with a presentation on the projection model structure (Dr Siddeek). This was followed by a presentation on approaches for estimating F_{MSY} and B_{MSY} proxies (Dr Turnock). The report on the interagency workshop (Anon. 2006) was reviewed briefly since we had already discussed most of the issues considered in it.

During the rest of the day and during the next day, preliminary simulation results were presented by the Working Group members. Attempts had been made to evaluate different alpha, beta, and gamma parameter values. Also, some proxies for F_{MSY} had been tested. However, in all cases the results were preliminary and no firm recommendations could properly be made with regard to proxies or default parameter values on the basis of the simulations.

27 April

The Review Panel convened at 9.30 am to identify, discuss, and clarify all relevant issues relating to the proposed OFD and to supporting research. We covered points a.-e. as per our Statement of Work (Appendix 3). Late in the day we had a question and answer session with Dr Hollowed, Dr Turnock, and Dr Rugolo.

Post Meeting Activities

Prior to and during my return journey to New Zealand I considered the two main problems that the Working Group were grappling with.

First, they had not fully defined the criteria for choosing between alternative tier-structure parameter values (in terms of being the best defaults). This, I believe, stemmed from the fact that the problem had not been fully specified. In order to determine the best defaults, one must define what it is for one MSY control rule to be better than another when they are used as part of an OFD.

Second, there had been inadequate analysis used to define “biomass” (B) in the stock recruit relationship (SRR). The Working Group had found that their results were very sensitive to the definition of B. They did not have an adequate definition and had no means of choosing between the alternatives they had proposed. I spent considerable time exploring alternatives for deriving appropriate functional forms – the aim being to illustrate how total fertilized egg production could be expressed as a function of population parameters (which could conceivably be measured or estimated).

The lead reviewer, Dr Bell, was to present our findings at two meetings which were scheduled earlier than the original deadline for production of our reports. On my return to New Zealand I produced an interim report for Dr Bell, in advance of his first meeting, which, while short on detail, differed little in the conclusions and recommendations of this report. I also undertook to produce my final report well in advance of Dr Bell’s second meeting (but some days after the new deadline specified in the revised SOW – see Appendix 3).

SUMMARY OF FINDINGS

The existing OFD is conceptually flawed and as a consequence places no constraints on fishing mortality. It clearly needs to be replaced, but care must be taken to ensure that its replacement does not overly constrain potential harvest strategies.

To my mind, there are two central issues to consider with regard to the proposed OFD.

First, there is the issue of what constitutes a "good OFD", in general, and for BSAI crabs in particular. If the answer to this question is clearly stated it is relatively straightforward to define the analysis and simulations needed to test alternative OFDs (and to determine default parameter values for the MSY control rules in the proposed tier structure). Related to this issue is the question of whether an OFD MSY control rule can be appropriately tested in isolation from a harvest strategy (HS). In reality, the MSY control rule imposes constraints on the HS which is used and so management strategy evaluation must test MSY-control-rule:harvest-strategy pairs.

The second central issue is the definition of B in the SRR. This issue is peculiar to crabs because fishing mortality is only directed at males. In groundfish stocks it is not an issue because female spawning biomass is a good proxy for total fertilized egg production (TFEP). This is a crucial issue for the proposed OFD because the biomass proxy for TFEP is a primary determinant of F_{MSY} and F_{MSY} proxies.

What constitutes a good OFD?

We should first consider the question, exactly what is an OFD? We should distinguish between an OFD for a particular stock and an "OFD framework" which specifies a family of OFDs. It is the latter which the review is concerned with and the "family" consists of OFDs for BSAI crab stocks. Central to an OFD is the concept of an MSY control rule, which defines F_{OFL} as a function of biomass and from which derives the overfished threshold (MSST). A "good OFD" (framework) can sensibly be defined as one which specifies "good MSY control rules".

The proposed OFD has a five level tier structure to accommodate stock assessments with different levels of reliability (Anon. 2006, page 8). The fifth tier is for stocks which are not formally assessed. In the first four tiers a linear parameterized MSY control rule is specified. F_{OFL} is constant above B_{MSY} and set equal to F_{MSY} or a proxy. Below B_{MSY} there is a linear reduction in F_{OFL} governed by two parameters alpha and beta. In tier 4, the F_{MSY} proxy is the product of the parameter gamma and M. The fishery is closed when estimated biomass (as a proportion of B_{MSY} or its proxy) is less than beta.

As it stands, the OFD appears incomplete until some default parameter values and proxy definitions are specified. In order to do this, criteria must be specified for determining when one MSY control rule is better than another. Given the criteria, alternative parameter values and proxies can be tested by doing model simulations over an appropriately broad range of population models (i.e., with different biological parameters and/or SRRs and/or model structures; the range being appropriate to the tier being tested).

The criteria for determining whether one MSY control rule is better than another were not discussed during the review meeting. From the preliminary simulations it appears that the implicit criteria relate to their performance as rebuilding strategies (since simulations were done from starting values less than BSST or at beta, with catches set at the OFL). The ranking of MSY control rules on the basis of their performance as rebuilding plans, or more generally as harvest strategies, is inappropriate given that is not their *function* (when specified in an OFD). The function of an OFD MSY control rule is to constrain (estimated) fishing mortality and to provide Status Determination (i.e., MFMT and MSST). It impacts on whatever harvest strategy is used for setting OY but it is not the harvest strategy (or the rebuilding plan).

I note that simulations using an MSY control rule as a harvest strategy are required to determine MSST (Restrepo et al. 1998). This is because the full definition of MSST is the maximum of two values: half B_{MSY} and "the minimum stock size at which rebuilding to the MSY level would be expected to occur within 10 years if the stock or stock complex were exploited at the maximum fishing mortality threshold". (During the review meeting no such simulations were discussed and it was (implicitly) assumed that MSST always equaled half B_{MSY}. In general, this should not be taken for granted.)

Restrepo et al. (1998) offer some advice on choosing an MSY control rule. Two factors are mentioned. First, the position of MSST may be of interest in that a council could "minimize the range of stock sizes within which special rebuilding plans would be required" if it opted "for an MSY control rule that afforded a good deal of 'built-in' rebuilding". The proposed OFD has such MSY control rules in that the linear decrease in F_{OFL} begins at B_{MSY} (which is even more conservative than the default MSY control rule suggested by Restrepo et al. 1998). Second, they suggest that the "tradeoff between magnitude of yield and constancy of yield" could be used. This involves testing the MSY control rules as harvest strategies. As already discussed this is inappropriate since that is not their function in an OFD setting.

In practice, an (OFD) MSY control rule is never used as the harvest strategy. Councils are required to "adopt a precautionary approach to the specification of OY" (Restrepo et al. 1998). Obviously, from a management strategy evaluation perspective, MSY control rules cannot be tested in isolation. They must be tested with an associated harvest strategy.

The choice of an MSY control rule is primarily a management decision. The tradeoff is between potential yield and risk. If an MSY control is too constraining on harvest strategies it may unnecessarily reduce long term yield. Conversely, if it is too liberal it may allow harvest strategies which are not precautionary. Given the current requirement for an (OFD) MSY control rule and a precautionary harvest strategy, constrained by the MSY control rule, it is necessary to test MSY-control-rule:harvest-strategy pairs.

Once this conclusion is reached it becomes a matter of detail on how to determine appropriate defaults to complete the specification of the proposed OFD framework (or to justify a choice of OFD for a particular stock assessment). Any existing harvest strategies are candidates to be tested. In their absence I suggest adopting some "default" harvest strategies in the simulations (e.g., those derived from the MSY control rule by applying 75% of the estimated OFL in each year).

When doing such simulations it is important to distinguish each of the individual components. There is the "operating model", which is the model of reality, within which everything is known exactly (e.g., B_0, F_y for each year y, $F_{MSY,}$ etc). There is also the "estimation frame" where quantities are estimated, such estimates being a function of the truth (from the operating model) and error (e.g., an estimate from a stock assessment). When evaluating an OFD MSY control rule there must also be a HS. The role of the MSY control rule is limited. It defines F_{OFL} for any given biomass *estimate* and it defines whether the stock is overfished or not (on the basis of the biomass *estimate*). On the other hand, the HS is used to set the TAC in each year that an assessment is conducted (within the simulation model). There is a requirement for a buffer between the OFL and the TAC. Hence, simulations using the MSY control rule as the harvest strategy (i.e., $F_y = F_{OFL}$ in every year y) are entirely inappropriate.

Definition of B in the SRR

The definition of B in the SRR is of crucial importance in obtaining a precautionary OFD. Since the directed fishing mortality is only on males, females suffer fishing mortality only as incidental bycatch (and subsequent handling/discard mortality). If the usual groundfish definition for B, of total female spawning/mature biomass, is used then F_{MSY} and proxies for F_{MSY} (such as $F_{50\%}$) are very large in an absolute sense. Crab biology is such that the role of males is crucial in the production of fertilized eggs and it is clear that the males must be brought into the definition of B.

In the long term, a suite of deterministic population dynamics models should be derived specifically for crab stocks, taking account of the important role played by males in the SRR. In the interim, it is probably best to derive an appropriate functional form for TFEP and simply assume that mean recruitment is a Beverton Holt or Ricker function of TFEP.

The review material contains several alternative proposed definitions for B. Total female mature biomass was, I assume, used for illustrative purposes only. Total male and female mature biomass was put forward as a candidate. This must be rejected because as male biomass approaches zero, TFEP approaches zero, but total mature biomass does not. There were at least two variations of female mature biomass scaled down by an "effective fertilization factor" (derived from an assumed "mating ratio"). The concept behind these definitions is that TEP is proportional to female biomass and that successful fertilization depends on the proportion of mature males in the mature population and the average number of females that each male can mate (the "mating ratio" which is assumed to be constant).

The concept of a "mating ratio" is sound in principle. In practice, it was found that F_{MSY} and F_{MSY} proxies were sensitive to the assumed mating ratio. So, even if one of the proposed formulations was accepted it still leaves the problem of determining an appropriate parameter range for the mating ratio.

During the review I questioned the validity of the assumption that TEP is proportional to mature female biomass. Dr Rugolo presented results from trawl survey and experimental data on the total number of eggs per female as a function of clutch fullness, shell condition, mating category, and carapace width. It is known that older females tend to have lower clutch fullness and that crabs with very old shell condition (4 & 5) tend to be barren. This is a problem for the proportionality assumption in that increasing biomass (with age) is inversely proportional to EP. Though, if the proportion of older females stays relatively constant it may not of itself be a major problem. However, it was also indicated that clutch fullness is strongly related to mating category, at least in snow crabs, with primiporous females typically having a0.75 clutch fullness and first time multiporous females typically having full clutches. Further, within clutch fullness category, the number of eggs appeared to be linearly related to the carapace width. While these data cast considerable doubt on the biomass proportionality assumption, their existence provides the very means by which to construct a sensible functional form for TEP and possibly to estimate a mating ratio.

I have undertaken some preliminary work on the derivation of a suitable equation for TFEP (Appendix 2). This work is illustrative and not definitive. An experienced mathematician should work with crab biologists to derive appropriate forms (at different levels of complexity) for TFEP. I also indicate how the trawl survey data (available since 1995) could be used to estimate unknown parameters, including a mating ratio, within the equation for TFEP (Appendix 2). In the absence of this sort of work (i.e., given time constraints), the best proxy for TFEP may be total mature *male* biomass (TMMB).

This suggestion was made by Dr Caputi at the review meeting and at the time, after discussion, was considered to be deficient in that it was inappropriate for stocks near their virgin level. It was considered that we needed a relationship which would be sensible over the full range of stock sizes. In a severely depleted stock, it is clear that sperm availability is the determining factor in fertilization success (since there are plenty of females). It is reasonable to argue that TMMB could be approximately proportional to TFEP when a stock (through removal of males) is depleted below some level. The effective mating ratio doesn't need to be known – the assumption is made that there are always enough females and that the mating ratio is constant. Of course, above some level of TMMB the proportionality assumption must fail. In Appendix 2 I have suggested an appropriate functional form to adjust for this effect. It adds an extra level of complexity to the SRR but will be more realistic than assuming full proportionality.

Other issues

Most other issues are minor in comparison to the two central issues already discussed. However, they are numerous and potentially time consuming in their detail. Below, I give some general comments on some of the issues.

Having two modeling groups is both a strength and a weakness. The exchange of ideas is valuable. The natural competition which arises can be stimulating and lead to improved methods and models. However, differences in modeling approaches can become entrenched; argument rather than discussion can be the outcome. While all members of the Working Group were cordial, helpful, and professional during the review meeting, there was clearly some tension between the two groups. In New Zealand, "contested stock assessment" is a common feature of our annual stock assessment cycle (Starr et al. 1998). We have recently agreed on some principles to help competing modeling groups work together:

- consider all components of the models and estimation procedures
- identify where differences exist between the two approaches
- where there is a "best" way to do something, agree to do it that way
- where there are two (or more) reasonable alternatives, implement both (all) options
- ideally, each modeling group should be able to reproduce the results of the other group (but, if totally different estimation procedures are being used, this is probably not an option).

Difference results do not present a problem if the reason for the differences is understood. In New Zealand the two competing groups use Bayesian estimation methods implemented with their own software packages. The use of the same estimation method is very helpful in terms of making comparisons. The two crab team groups use completely different estimation methods, neither of which is entirely satisfactory.

The weighted least squares method (Zheng 2004) does not allow the production of standardized residuals. It is not a "fully statistical" model: diagnostics cannot be properly evaluated. The maximum likelihood approach (Turnock & Rugolo 2005) at least allows the production of standardized residuals even if this has not been routinely done. However, the use of penalty functions is not ideal and where possible they should be replaced by properly formed priors. Indeed, both modeling groups need to move towards fully Bayesian assessment methods as soon as possible. There simply is no other generally accepted method for incorporating prior/ancillary information and statistically accounting for uncertainty. It is not perfect, but it is currently the "state of the art" and will be so for some time to come.

The assumption by both modeling groups that the trawl survey q is known exactly (on the basis of "under-bag" experiments) ignores the uncertainty due to unknown aerial availability (i.e., the proportion of the population within the survey area). The estimates are conservative but they are definitely biased. The assumption is neither necessary nor desirable and the trawl survey q should be estimated. The information from the under-bag experiments, and whatever else is "known" about q, can be incorporated in a prior (or, if necessary, a penalty function).

The estimation of natural mortality (M) is always problematic whether it is done inside or outside of a stock assessment model. This is true for a single M assumed constant over the whole history of a fishery. Attempts to estimate different M for different time periods in a stock assessment (Zheng 2004) are ill-advised unless there are ancillary data which can reasonably be argued to index M in some way (e.g., a biomass index for a known major predator).

Many of the problems facing the crab team are generic in nature, some are crab specific (definition of B) and some are even more general (testing of OFDs). Wherever possible, efforts should be made to establish collaborative projects to share the workload.

CONCLUSIONS

The conclusions are organized according to the headings provided in the SOW (Appendix 3).

a. Strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.

The proposed OFD is:

- an improvement on the existing OFD
- comprehensive (as a framework)
- borrowed from groundfish (so is already reviewed to some extent).

The existing OFD does not provide any sensible constraints on fishing mortality and in that regard it appears fatally flawed. The proposed OFD will at least provide constraints.

Weaknesses of the proposed OFD:

- complicated
- it is still a work in progress
    - default values for parameters are not yet determined
    - sensible definition of B not determined/specified
    - criteria for determining optimal default parameters not determined/specified
- extensive simulations are needed to determine suitable default parameters
- it may potentially unnecessarily constrain existing harvest strategies.

The Review Panel were shown the results of preliminary simulations aimed at determining suitable default values for alpha, beta, and gamma. One can envisage an extensive suite of simulations which could determine suitable default values, but this can only happen after:

- sensible definitions of B are derived (being proportional to total fertilized egg production)
- the criteria for optimal default parameter values are defined.

An important issue, relating to the optimality of default parameter values, is how to define a "good OFD". The current simulations test an OFD MSY control rule by using it as a HS (i.e., assuming that catch is always set at the OFL) and testing its performance when the stock is initially overfished. However, this ignores the fact that a council is required to act in a precautionary manner when setting TACs and, as such, the simulations are testing something which will never occur.

b. Recommendations for improvements to proposed overfishing definitions or alternative definitions.

The parameterization of the proposed MSY control rules implies a reduction in F at B_{MSY}. It does allow flat control rules (alpha = –infinity) but it precludes the suggested default overfishing definitions of Restrepo et al. 1998 (where the reduction in F occurs below B_{MSY}). I suggest that an extra parameter is added to the framework to allow MSY control rules of the form proposed by Restrepo et al. (1998). In the absence of this parameter, the proposed framework may unnecessarily restrict harvest strategies.

c. Review of model configurations, formulations and methods used to account for uncertainty.

The model configurations and formulations are generally appropriate, but there may have been some implementation error in some of the models (e.g., mating dynamics not consistent with expert opinion). There needs to be more effort made to ensure that both modeling groups correctly implement the agreed population dynamics. When alternative dynamics are considered possible they should also be implemented to allow sensitivity analyses to be performed.

d. Review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in the simulation models.

The determination of an appropriate SRR is one of the central issues of the review. All of the alternative definitions of B used in the preliminary simulations were inappropriate. They were either demonstrably inadequate (e.g., female mature biomass, total mature biomass) or inadequately justified (i.e., no analysis or derivation). Total fertilized egg production does not appear to be proportional to female mature biomass. Therefore, definitions of B should not be based on scaled female mature biomass (e.g., through an assumed mating ratio) .

Other life history parameters appeared to be appropriately estimated (except, in one model where M was estimated to change during different time periods – not appropriate without additional data – see recommendations).

e. Research priorities to improve understanding of essential population and fishery dynamics necessary to formulate best management practices.

Recommendations, with regard to all aspects of the review, are given in the section below.

RECOMMENDATIONS

It appears that the proposed overfishing framework can be considered "acceptable" (complete) without default parameter values, F_{MSY} proxies, or a definition for B. If this is the case, then the first assessment for each stock, under the new framework, will require a full suite of simulation results justifying the OFD used. As more stocks are assessed "default" definitions and parameter values will materialize as scientists borrow them from the previously accepted assessments. Such an evolution is far from ideal and the process will need to be managed carefully. It would be better to get agreement on as much as possible in the proposed OFD before it is "accepted". Certainly a default definition for B is desirable.

In any case, irrespective of the timing relative to the acceptance of the proposed OFD framework, I have the following recommendations.

- Derive sensible definitions of B:
  - being defensibly proportional to total fertilized egg production
  - consider primiporous and multiporous matings separately
  - get the cycle consistent with the best available expert opinion (i.e., which males can participate in which matings)
  - B is not proportional to mature female biomass (e.g., clutch size is not proportional to biomass)
  - use an analytical approach to derive suitable functional relationships (see Appendix 2)
  - estimate parameters of the relationship in the stock assessment models using available data on egg production by color class (see Appendix 2)
  - mature male biomass appears to be defensible (use as a default?)

- Agree on the criteria and method for testing (OFD) MSY control rules:
  - these methods could be applied to tiers 1-4 (e.g., not only to determine "good" alpha and beta values, but also to choose between different proxies, e.g., $F_{50\%}$ or $F_{60\%}$)
  - it must be decided what makes one MSY control rule better than another when they are part of an OFD (i.e., test their *function*, they are not rebuilding plans or harvest strategies)
  - test MSY control rules in conjunction with a HS (e.g., an existing HS or a "default" OY control rule which takes 75% of the OFL – see Restrepo et al. 1998)

- compare with a flat control rule ($F = F_{MSY}$, i.e. are alpha and/or beta even needed?)
- examine performance over a range of starting biomasses (not just overfished; you want to know how they perform “going down” as well as “going up”)
- incorporate observation error (i.e., true B and observed B can differ)
- incorporate stochastic recruitment
- examine trade-off statistics (e.g., what is forgone in yield to achieve higher biomass/lower probability of being declared overfished)
- use the full definition of MSST (i.e., not just 0.5 B_{MSY} – see Restrepo et al. 1998)
- include an extra parameter in the MSY control rules so as not to exclude the suggested default rules of Restrepo et al. (1998) (this parameter can have a default of 0 if desired)

The following two recommendations only apply if it is decided to use tier 1-2 simulations to derive default alpha and beta for tiers 1-4. It may not be the case that “good” alpha and beta values in tiers 1-2 will necessarily be any good when used in conjunction with F_{MSY} proxies. However, it may be a necessary assumption given time constraints.

- Agree on criteria for testing F_{MSY} proxies (stock specific, Tier 3):
    - using expert judgment choose a range of steepness/SRR relationships (after sensible definition of B)
    - use minimax or some other agreed principle to choose the best proxy

- Agree on criteria for testing gamma (group specific, Tier 4):
    - explicitly and precisely define gamma (in relation to selectivity and timing of the fishery)
    - use the same approach as for tier 3, but wider parameter space
    - obtain default gamma for each of several species/stock groups

- Consider what simulations, if any, could help for tier 5:
    - to define the period over which catches should be averaged (e.g., guiding principles on “not too much catch variability”; not a “declining trend in biomass” over the period)

- Stock assessment models
    - estimate the survey catchability q
    - start with parsimonious models
    - only introduce extra parameters if absolutely necessary
    - do not confound M with possible changes in catchability
    - estimating changes in M is only defensible if supported by auxiliary data on known predators/disease
    - calculate standardized residuals
    - iteratively re-weight indices so that residuals are consistent with variance assumptions

  - as soon as possible move to fully Bayesian assessments

- Trawl survey
  - if feasible, routinely retain a sample of female crabs with orange colored eggs to estimate the proportion of fertilized orange-colored egg-production (i.e. to estimate, at the time of the survey, what proportion of orange colored eggs are actually fertilized)
  - if feasible, routinely retain a sample of females (of the relevant species) to estimate “sperm load” (i.e., for those species which retain sperm).

REFERENCES

(see Appendix 1 for further references)

Mace, P.M.; Doonan, I.J. 1988: A generalised bioeconomic simulation model for fish population dynamics. New Zealand Fisheries Assessment Research Document 88/4. 51 p.

Starr, P.; Annala, J.H.; Hilborn, R. 1998: Contested stock assessment: two case studies. *Can. J. Fish. Aquat. Sci. 55*: 529–537.

APPENDIX 1: MATERIAL PROVIDED

The website from the interagency workshop was made available to the Review Panel. This contained documents and presentations, but also contained links to other related material. Below I list the material which I obtained from the website (or related links) and additional documents which were emailed to the Review Panel or provided as hardcopy, before or during the review meeting. I do not include several documents which were emailed to the Review Panel after the meeting (as I did not consult them).

Anon. 1998. Fishery Management Plan for Bering Sea/Aleutian Islands King and Tanner Crabs. Exec summary. July 18, 1998. 5 p.

Anon. 1999. Draft for Secretarial Review: Environmental Assessment for Amendment 7 to the Fishery Management Plan for the commercial king and tanner crab fisheries in the Bering Sea/Aleutian Islands. 53 p.

Anon. 2005. Magnuson-Stevens Act Provisions; National Standard Guidelines; Proposed Rule. Federal Register Vol 70, No. 119. 21 p.

Anon. 2006. Center of Independent Experts, Alaska Crab Overfishing Definitions, April 24-29, 2006. Alaska Fisheries Science Center, Seattle, WA. Apr 14th 2006 Draft Agenda. 2 p.

Anon. 2006. Current Overfishing Definitions in Crab FMP (FMP Section 6.0 revised from Amendment 7 1998). 2 p.

Anon. 2006. Workshop Report Crab Overfishing Definitions Inter-agency Workshop, February 28-March 1, 2006. Alaska Fisheries Science Center Seattle, WA. 21 p.

Anon. 2006. Alaska Crab Overfishing Definitions Workshop, February 28 – March 1, 2006. Alaska Fisheries Science Center, Seattle, WA. Feb 22, 2005. Draft Agenda . 2 p.

Anon. 2006. Draft report of the Scientific and Statistical Committee to the North Pacific Fishery Management Council, April 3-5, 2006 . 17 p.

Anon. 2006. Participant list for interagency workshop. 1 p.

Anon. 2006. Roadmap for Crab Workshop Documents. 1 p.

Anon. 2006. Tier system. 7 slides.

Maunder, M.N. 2003. Review of the stock assessment and harvest strategy for the eastern Bering Sea snow crab. CIE review report. 29 p.

Punt, A.E. 2003. The performance of a size-structured stock assessment method in the face of spatial heterogeneity in growth. *Fisheries Research 65*: 391–409.

Restrepo, V.R. et al. 1998: Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. 56 p.

Rugolo, L.J.: Simplifications in population models and fishery dynamics modeling approaches. 16 slides

Rugolo, L.J.: Statement of work: NPFMC BSAI King & Tanner Crab Working Group. 30 slides.

Rugolo, L.J.; Siddeek, M.S.M., Turnock, B.J.; Zheng, J. 2004. North Pacific Fisheries Management Council Bering Sea / Aleutian Islands King and Tanner Crab Working Group Progress Report to the Crab Plan Team 22 September 2004. 34 p.

Siddeek, M.S.M: Parameters input to SPR models. 8 slides.
Siddeek, M.S.M.: Preliminary results. 7 slides.
Siddeek, M.S.M.: Model structures. 3 slides + 2 spreadsheets.
Siddeek, M.S.M.: Approaches to estimate proxy BRP values. 9 slides.
Siddeek, M.S.M.; Zheng, J. 2006. Reference point estimation analysis for the Bering Sea and Aleutian Islands (king and tanner) crab revised Fisheries Management Plan. 68 p.
Stram, D.L.: Overview of crab management and background on current overfishing definitions. Alaska Crab Overfishing Definitions Workshop, AFSC, Seattle, WA, February 28-March 1, 2006. 10 slides.
Thompson, G.: National Standard 1 Guidelines: use of SPR reference points, and incorporating uncertainty. 14 slides.Turnock, B.J. Snow crab stock assessment. 30 slides
Turnock, B.J.: Proposed tier system. 9 slides.
Turnock, B.J.: Input values to SPR models estimated from stock assessment models. 15 slides.
Turnock, B.J.; Rugolo, L.J. 2005. Stock Assessment of eastern Bering Sea snow crab. 96 p.
Turnock, B.J.; Rugolo, L.J. 2006. Analysis of proposed overfishing tier system for BSAI king and tanner crab stocks. 28 p.
Zheng, J.: Population Dynamics and Stock Assessment of Red King Crab in Bristol Bay, Alaska. 29 slides
Zheng, J. 2004. Bristol Bay red king crab stock assessment in 2004. 72 p.

APPENDIX 2: DEFINITION OF B IN THE SRR

Below I present three suggestions for the definition of B in the SRR, each being a proxy for total fertilized egg production (TFEP). They range from simple to complex. For the most complex method I also illustrate how some of the unknown parameters in the functional form might be estimated.

The three suggested definitions for B in the SRR are:

- total mature male biomass
- a function of total mature male biomass
- a function of total female egg production and a “fertilization factor”.

The first suggestion was made by Dr Caputi in the review meeting. It is appealing in its simplicity and it makes sense for crab stocks when the male population is severely depleted. It is deficient for crab populations when the sex ratio is near its virgin level. However, we can derive a functional form to correct for its deficiency.

Suppose that,

$$TFEP = \min \{ P, q\ SP \}$$

where P = total egg production, SP = sperm production, and q is a constant which translates sperm production into number of eggs. In reality, q is not a constant; it depends on any and all factors which affect fertilization success (e.g., size and age structure, environmental effects on sperm potency, spawning migration patterns). In fish stocks we expect that $q\ SP >> P$ and thus accept any reasonable proxy for P as a proxy for TFEP for use in the SRR. However, in crab stocks where fishing mortality is directed only at males we must incorporate sperm production.

Consider a deterministic model for a crab population where fishing mortality (F) is primarily on males. Let,

B_F = male mature biomass at equilibrium under fishing mortality F
P_F = total egg production at equilibrium under fishing mortality F
$a_F = P_F / B_F$

and let $TFEP_F$ and SP_F denote TFEP and SP respectively at equilibrium under fishing mortality F.

We then have,

$$TFEP_F = \min \{ P_F, q\ SP_F \}$$

Now, suppose that sperm production is proportional to male biomass: $SP_F = s\ B_F$.

Hence, when q $SP_F \leq P_F$, we have

$$TFEP_F = q\ s\ B_F$$

and when q $SP_F \geq P_F$, we have

$$TFEP_F = P_F = a_F\ B_F.$$

Note, that when q $SP_F = P_F$, we have a_F = q s. Denote the F at this point as F_a and the associated B as $B(F_a)$.

Now, as F varies from 0 to infinity, B_F varies from its virgin level, B_0, to 0. From 0 to $B(F_a)$, $TFEP_F$ is a linear function of B_F (which passes through the origin). The form of $TFEP_F$ between $B(F_a)$ and B_0 depends on the nature of a_F. However, since males are preferentially exploited it follows that as F decreases that a_F also decreases. Hence, $TFEP_F$ is a linear function of B_F from 0 to $B(F_a)$, and then is convex from $B(F_a)$ to B_0.

This deduction allows us to use an approximate functional form which is independent of the details of any particular model. We will use an exponential function which is approximately linear over part of its range and then convex. Changing the notation somewhat, let,

$$TFEP(B) = b\ [\ 1 - \exp(\ -aB)\ ]$$

where B = male mature biomass, and a, b, are unknown parameters.

Let, $TFEP(B_0) = P_0$, then since,

$$TFEP(B_0) = b\ [\ 1 - \exp(\ -aB_0)\ ] = P_0$$

it follows that,

$$a = \frac{-\ln(1-\frac{P_0}{b})}{B_0}$$

and

$$\text{TFEP}(B) = b\left[1-\left(1-\frac{P_0}{b}\right)^{\frac{B}{B_0}}\right]$$

This can be better expressed as,

$$\text{TFEP}(B) = \frac{P_0}{1-\eta}\left[1-\eta^{\frac{B}{B_0}}\right]$$

where $0 < \eta < 1$.

This equation provides a simple generalization of total male mature biomass as the definition of "B". The range of η values to consider depends on how effective one believes the males can be at fertilizing eggs when at depleted biomass levels (or how many surplus males there were at virgin levels). Values of $\eta > 0.5$ provide fairly linear functions, with TFEP at 20% B_0 not much more than 20% P_0. For approximately 40% P_0 at 20% B_0 use $\eta = 0.1$ and for approximately 60% P_0 at 20% B_0 use $\eta = 0.01$.

This equation should be used in conjunction with an assumed SRR as a function of TFEP to produce a SRR as a function of mature male biomass. For example, if mean recruitment is given by R(TFEP), then use R(TFEP(B)) – i.e., to get mean recruitment as a function of B rather than TFEP.

For example, if a Beverton Holt SRR is assumed with steepness Δ (Mace & Doonan 1988), then

$$R = R_0 \left[\frac{4\Delta\left(1-\eta^{\frac{B}{B_0}}\right)}{(1-\Delta)(1-\eta)+(5\Delta-1)\left(1-\eta^{\frac{B}{B_0}}\right)} \right]$$

which behaves much like a Beverton Holt SRR unless both Δ and η are small (i.e., is similar to a Beverton Holt SRR with Δ set equal to the proportion of R_0 obtained at 20%B_0 from the full relationship). Information on η may be available from trawl survey data in which case it may be possible to estimate η within the stock assessment model, or even externally.

The third suggested definition for B is arrived at using a somewhat different approach. The idea is to start with an unspecified functional form and then, through a series of assumptions, and bearing in mind what data are available, arrive at a particular form.

Consider a particular mating (i.e., the primiporous or multiporous mating, or perhaps a combined mating for modeling convenience) in a particular year and suppose that sperm are not stored by the females ("sperm storage" across matings could perhaps be incorporated if data on stored sperm levels from the trawl survey were available).

Let $C = \{ c_i \mid i \varepsilon n \}$ be the set of all crabs involved in the mating (i.e., there are n crabs labeled 0,…,n –1). Each crab has various biological characteristics. E.g,: $s(c_i)$ = sex of the ith crab, $cw(c_i)$ = carapace width of the ith crab, $a(c_i)$ = age of ith crab, $f(c_i)$ = clutch fullness category of ith crab. Subsets of crabs can be specified. E.g., $s(C, male) = \{ c_i \mid s(c_i) = male \}$. Let TFEP = $\Gamma(C)$. That is, the total fertilized egg production (of this mating) potentially depends on every characteristic of every crab involved in the mating. True, but unhelpful. We will split TFEP into tow components: total egg production (P), and a fertilization factor (G):

$$\Gamma(C) = P[\ s(C, fem)\ ]\ G(C)$$

Now, we have assumed that egg production depends only on the females; fertilization is still dependent on all crabs (e.g., male and female length frequencies, as well as sex ratio).

From data presented during the review meeting and subsequent discussion it is clear that quite a lot is known about clutch fullness as a function of various categorical variables (e.g., primiporous females typically have 0.75 clutch fullness, 2nd clutches are typically full, shell condition 4 and 5 females are typically barren, older females have lower clutch fullness). For a given clutch fullness it appeared that egg production of an individual female was a linear function of carapace width. In any case, the data exist and can be analyzed to provide an appropriate functional form and parameterization for P. One approach would be to use a GLM to explain individual egg production as a function of variables which could reasonably be incorporated in the stock assessment model.

For example, for a combined mating, the data may be consistent with a single categorization within which a linear function of carapace width may be adequate for average individual egg production. The functional form could be as follows:

$$P[s(C, fem)] = \sum_{i \varepsilon cat} \sum_{j \varepsilon cat_i} a_i + b_i cw(c_j)$$

where cat = { primiporous, 2nd mating, shell condition > 3, other }, a_i and b_i are the linear coefficients for each category member, and cw denotes carapace width.

The above form is just an example which may or not be suitable. However, given the available data I am confident that a suitable form will be derived. It will be defensible and I doubt that female mature biomass will be seen to be an adequate proxy.

The fertilization factor is a more difficult challenge. However, there are also data available which may enable the estimation of relevant parameters if an appropriate form can be hypothesized. The simplest form for G(C) is a constant. However, this would make TFEP independent of males – clearly not appropriate. The fertilization factor, at a minimum, must use the relative number of males and females. Other elements, such as relative size distributions and the propensity for males to fight for “desirable females” could also be brought in (but not easily).

A candidate, already used as a component of some of the Work Group definitions of B is:

G(C) = min { n[s(C, male)] / n[s(C, fem)] r, 1 }

where r is an unknown “mating ratio” and n[] denotes cardinality of a set (i.e., the number of members).

There are data available from the trawl surveys which could be used to estimate r, preferably within the stock assessment model (so that trawl survey catchability and selectivity can be estimated simultaneously). I refer to the individual egg production data which includes a color classification. Clutches are orange to begin with and at the time of the survey clutches are either orange or another color. If they are non-orange then they are fertilized, but some proportion of orange clutches are also fertilized (and haven't changed color yet). To use these data in a stock assessment model, we need to be able to formulate predictions for orange and non-orange trawl-survey egg production. Below I give a sketch of how to do this.

We already have an expression for total egg production, P[s(C, fem)] which could be modified in the model to account for trawl survey selectivity and catchability to become an expression for "trawl-survey egg production". We shall denote this simply by P. In the following, assume that trawl-survey selectivity and catchability have been appropriately dealt with in all components of (predicted) egg production.

Let,

P_o = orange egg production
P_{of} = orange fertilized egg production
P_{oun} = orange unfertilized egg production
P_{non} = non-orange egg production
p_f = proportion of fertilized eggs
p_{fo} = proportion of fertilized eggs that are orange

We have,

$$P_o = P_{of} + P_{oun} = P\ p_f\ p_{fo}\ +\ P\ (1 - p_f)$$

and

$$P_{non} = P - P_o.$$

Which gives,

$$P_0 = P\ [\ p_f\ p_{fo}\ +\ 1 - p_f\ ]$$

$$P_{non} =\ P\ p_f\ [\ 1 - p_{fo}\ ].$$

We have two unknown parameters: p_f and p_{fo}.

However, p_f = n[s(C, male)] / n[s(C, fem)] r. So there is only one extra parameter: p_{fo}. Data on this could be available for each (future) survey if a random sample of females with orange clutches was retained and observed in the lab to see what proportion of orange clutches were fertilized (noting that p_{fo} = "proportion orange and fertilized"/p_f).

Alternatively, some assumptions about the distribution of mating times would be needed together with some knowledge of how long it takes fertilized eggs to change color.

For example, assume a normal distribution for the mating time X of a female: $X \sim \mathrm{N}(t_m, \sigma_m^2)$. Suppose that a females' clutch will change color after a time interval δ, and let $Y = X + \delta$. Suppose that the survey occurs at time t_s and let $q(t_s)$ = the proportion of (fertilized) eggs that are non-orange.

Then,

$$q(t_s) = \mathrm{Prob}\left(Y < t_s\right) = \mathrm{Prob}\left(Z < \frac{t_s - (t_m + \delta)}{\sigma_m} \right)$$

where Z is the standard normal random variable. Some educated guesses will help define a range for $q(t_s)$ and hence to $p_{fo} = 1 - q(t_s)$. Another approach is to look at the relative distribution of clutches within color classes to try to directly estimate the proportion of orange clutches which are fertilized (e.g., a disjunction between the proportion of orange clutches and the proportion of the next color class may indicate that very few orange clutches are fertilized – for a particular survey).

Trying to include competition between males is an interesting exercise. It can be approached by setting up a system of differential equations with coupling, decoupling, and fighting rates. It gets sufficiently complicated that it may not be a worthwhile exercise in itself. Perhaps it is better to do the full job and look to set up a system of differential equations for crab-specific population dynamics – a medium to long term project.

APPENDIX 3: STATEMENT OF WORK

The statement of work given below was received in early May after I returned from the Seattle meeting. It differs from the original statement of work in two respects. First, it clarified some issues which the Review Panel raised while we were in Seattle. Second, it contains a new date for submission of reports. I was not able to accommodate the shift of the deadline from 1 June 2006 to 12 May 2006. However, I did produce an interim report with the highlights of my findings and recommendations which I supplied to Dr Bell before his attendance at the May 17 meeting.

Consulting Agreement between the University of Miami and Reviewer

April 27, 2006

Background

The Alaska Fisheries Science Center (AFSC) requests review of proposed overfishing definitions and simulation models used to evaluate biological reference points for Bering Sea and Aleutian Islands King and Tanner crab stocks. The North Pacific Fishery Management Council (NPFMC) has determined that the existing overfishing definitions for Bering Sea and Aleutian Islands King and Tanner crab stocks need revision. The AFSC is seeking review of the population dynamics models developed for revising the overfishing definitions.

There are currently 22 Bering Sea and Aleutian Islands crab stocks under the Federal Bering Sea Aleutian Island Crab Fishery Management Plan (FMP) of which 7 are considered major stocks. Four of the seven major crab stocks have been declared overfished and rebuilding plans developed within the last 7 years. Of the remaining three stocks, only one has been relatively stable at a low level, another has maintained stable catch for several years, however, even for this stock it appears recruitment may be declining. While the remaining stock has increased, survey abundance estimates have low precision and the fishery is closed due to bycatch concerns. There is no consensus on the principal cause of declines in Bering Sea crab stocks.

Review Requirements

A panel of three consultants is requested for this review. In aggregate, the panel will need to be thoroughly familiar with various subject areas involved in the review: crab biology; analytical stock assessment, including population dynamics theory, length-based stock assessment models, rebuilding analyses, estimation of biological reference points and harvest strategy modeling for invertebrates; and AD Model Builder. The CIE consultants will travel to Seattle, Washington to meet with the Interagency Work Group charged with developing the new overfishing definitions. We request that one member of the Panel should be present at the May meeting of the NPFMC Crab Plan Team in

Anchorage, Alaska. We also request that one member of the Panel be present at the June meeting of the NPFMC Scientific and Statistical Committee meeting in Kodiak, Alaska. It would be preferable that the same individual attends both of these meetings, but this is not a requirement.

The report generated by each consultant should include:

a. A statement of the strengths and weaknesses of the proposed overfishing definitions, simulation models and analytical approaches.
b. Recommendations for improvements to proposed overfishing definitions or alternative definitions,
c. A review of the model configurations, formulations and methods used to account for uncertainty.
d. A review of input parameters (fishery, biological and life history parameters and spawner recruit relationships) used in simulation models.
e. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

AFSC will provide copies of the NPFMC Work Group statement of work, proposed overfishing definitions, preliminary results of simulations, discussion of input parameters, a copy of the code for the snow crab stock assessment, and the AD Model Builder and Fortran code used for reference point estimation. The panel will meet with scientists from the Alaska Fisheries Science Center and the Alaska Department of Fish and Game from April 24 to April 28, 2006, in Seattle, Washington (see attached agenda).

It is estimated that the duties of each reviewer will occupy a maximum of 14 days each: several days for preparation, five days for the workshop, several days for writing their reports, and two days for travel. In addition, a maximum of nine reviewer days will be allowed for attending the two council meetings, including preparation time, travel, and one day to attend each meeting. The total level of effort is 51 days of reviewer time.

Products

- One member of the panel will attend the May meeting of the Crab Plan Team on May 17, 2006 in Anchorage, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- One member of the Panel will attend the June meeting of the NPFMC Scientific and Statistical Committee meeting on June 5, 2006 in Kodiak, Alaska, to discuss the panel's findings regarding the strengths and weaknesses of proposed definitions and modeling approaches.
- No later than May 12, 2006, each panelist shall submit a written report of findings, analysis, and conclusions. See Annex 1 for details on the report outline. The reports should be sent via e-mail to Dr. David Die at ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani at mshivlani@rsmas.miami.edu.

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

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

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

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and the center and a copy of the statement of work.

4. Individuals shall be provided with an electronic version of a bibliography of background materials sent to all reviewers. Other material provided directly by the center must be added to the bibliography that can be returned as an appendix to the final report.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cimas/Report_Standard_Format.html