Center for Independent Experts (CIE) independent peer reviewer's report of an age-based, integrated stock assessment for Antarctic krill (*Euphausia superba*) with projected catches to 2035

Robin Thomson
608 Nelson Road, Mount Nelson, Tasmania 7007, Australia
Robin.Thomson@csiro.au

Executive Summary

The use of an Integrated Analysis model for assessment of Antarctic krill in FAO Subarea 48.1 is appropriate and would constitute an improvement over the modelling strategies currently in use to manage krill. The rich, but patchy and varied, data sources available for the region support the implementation of an age-structured model conditioned on all currently available survey and commercial fisheries data. In its present form, the IA model presented to this review has considerable difficulty converging on a "best fit" parameter set. This appears to be the result of over parameterization. The likelihood components used in the model's objective function are weighted using pre-specified values. Recommendations made in this report aim to improve convergence, primarily though identifying parameters that might be poorly estimated and should therefore be fixed at plausible values and subsequently simulation tested. Survey length frequency data reveal that variability in krill abundance is much larger than previously thought, such that existing CCAMLR management rules would prevent any fishing in the area. This is an unintended consequence of existing CCAMLR regulations. The proposal that future catch limits should be set by comparing projected biomass under fishing with projections in the absence of fishing is supported by this review. Once convergence problems have been overcome, the IA modelling framework reviewed here has the potential to substantially improve the scientific basis for future krill management.

Contents

1. Background

Antarctic krill (*Euphausia superba*, hereafter "krill") is a cornerstone prey species in the Antarctic food web. The fishery for krill is regulated by the Commission for the Conservation of Antarctic Living Marine Resources (CCAMLR), which recognizes the importance of taking the needs of krill predators into account by ensuring higher levels of escapement (abundance left in the water after fishing has taken place) than often occurs under classical single-species fisheries management regimes. CCAMLR were arguably the first fisheries management body to operationalize the principle of ecosystem management through paragraph 3 of CCAMLR Article II, which seeks to conduct harvesting and associated activities in a way that does not severely deplete any population to levels at which recruitment is impacted; maintains ecological relationships and restores depleted populations; and prevents or minimizes the risk of changes that cannot be reversed within two or three decades. There are three management rules for krill (Constable et al 2000):

(1) the probability that krill spawning biomass drops below 20% of its pre-exploitation median level over a 20-year harvesting period should not exceed 10%;
(2) the median krill spawning biomass over a 20-year period should not fall below 75% of its pre-exploitation median level;
(3) implement the lower catch limit of (1) and (2) above.

Fishing for krill began in the 1970s but even by 1984, when the CCAMLR convention was signed, catches are thought to have been very small in relation to the large available biomass of krill. However, expansion of the fishery was anticipated so that, given the pivotal role of krill in the Antarctic ecosystem, it was considered imperative to implement conservative management measures despite considerable information gaps concerning the size, productivity and certain aspects of the biology of krill. Consequently, a simulation model that uses a small number of pre-set parameter values, rather than a more conventional stock assessment model tuned to fisheries and survey data, was used to set a precautionary catch limit. This model, which has come to be called the Generalized Yield Model (GYM, Constable et al. 2000) describes krill catch limits as a proportion (γ) of virgin (unfished) krill biomass. The implementation of management rules (1) to (3) above, in conjunction with the GYM, yields a value of γ that is then applied to an estimate of unfished biomass, typically from a survey.

The implicit assumption is that krill biomass varies around an unchanging median virgin biomass, that krill fishing depresses the population to a lower level, and that if fishing were stopped, the population would return to this level. Changes to the environment (such as a permanent reduction in sea ice extent), and to the ecosystem (such as recovery of baleen whale stocks) are not taken into account in this formulation.

The GYM was developed at a time when data were relatively sparse. However, the availability of relevant data has improved considerably over the last two decades. Time series of biomass indices from surveys, as well as length frequencies from these surveys, and from the commercial krill fishery are now available for FAO subarea 48.1 (Kinzey 2015c). Although many unknowns still exist, krill biology is better understood than it was when the GYM was first proposed (Butterworth et al. 1994). The time has come for the application of a fisheries stock assessment model of the kind commonly used in modern fisheries management

(Maunder & Punt 2013). Such a model has now been developed and is the subject of this review (Kinzey et al., 2016).

2. Description of review activities

This document describes the activities that occurred during a three-day review meeting held at the offices of the Antarctic Ecosystem Research Division (AERD) of the South West Fisheries Science Center (SWFSC) in La Jolla, San Diego over three days: Monday 5th April to Wednesday 7th April 2016, with additional informal discussions on 8 April 2016.

The first day consisted of presentations, considerable discussion, clarifications and generation of ideas. Doug Kinzey was able to implement some of the recommendations made regarding the assessment model on the first day, the results of which he presented on the second morning. This led to further discussion, additional recommendations, as well as clarification and elaboration of those already made. Much of the second and third days consisted of a group coding exercise during which the assessment model was modified in accordance with some of the recommendations. Further debugging was discussed on 8 April. Unfortunately, little time remained for exploration of the model's ability to estimate parameters and debugging (by Doug Kinzey) is ongoing. The third day also included a discussion around the review Terms of Reference, with particular focus on the CCAMLR management regulations.

At the end of the meeting, the krill assessment model had been modified so that (pending further debugging):

(1) the Beverton Holt stock recruitment relationship is used in a more conventional way;
(2) recruitment deviations are constrained to sum to 1;
(3) recruitment deviations are not estimated prior to 1978 (the earliest length composition data is for 1981);
(4) spawning potential is estimated using a fecundity-at-length instead of a weight-at-length relationship;
(5) the six survey abundance time series (summer and winter acoustics; two types of nets in each of summer and winter) can be arbitrarily aggregated for the purpose of estimating selectivity functions (aggregation of the 4 net time series into 2 - one for each of summer and winter - was explored during the meeting).

2.1 Presentations and subsequent discussion

The meetings began with George Watters providing a description of the CCAMLR decision making framework and scientific and fishing activity in CCAMLR Area 48. This was followed by a presentation by Christian Reiss on krill biology and the acoustic survey process, of which net tows are a part. Finally, Doug Kinzey presented the Integrated Analysis stock assessment model for krill. Discussions occurred during and after the presentations.

George Watters highlighted that the present catch limit for Area 48 of 5.61 million tonnes was derived using a virgin biomass level (B_0) for that area from a survey conducted 16 years ago (in 2000) and that, with a life span of typically less than 10 years, all of the krill surveyed have since died. A trigger limit of 620,000 tonnes has been set for the area as an interim measure, based on historical catches (Kinzey et al. 2013). An agreed catch limit of 155,000 tonnes for Subarea 48.1 alone expires after 20 November 2016, after which management would revert to the original trigger limit of 620,000t across Area 48. The commercial fishery preferentially targets white, not green, krill and consequently concentrates on a narrower set of areas than the more wide-ranging AMLR surveys run by AERD. It also uses a courser net. Fishing historically concentrated in Subareas 48.2 and 48.3, but is now primarily in 48.1. In response to ecosystem models showing that the scale at which predators operate is smaller than that of the existing Subareas, CCAMLR divided the Subareas into small scale management units (SSMUs). The fishery and predators seem to concentrate on the same areas of high krill aggregation and abundance. CCAMLR has agreed to (but not yet operationalized) feedback management, which comprises:

- management at the SSMU level
- allowing catch limits to change on a time scale equivalent to changes in krill abundance
- monitoring of early season krill abundance
- reduction of fishing impact when predator offspring are young and energy demands are high

Christian Reiss presented background information on krill biology, fishing in Area 48, and krill surveys in that area. Reduction in sea ice extent as a function of global warming is opening up new ice-free areas and allowing a fishing season that extends further into the autumn and winter. Surveys have been conducted by US AMLR (included in the krill modelling) and also by the Longer Term Ecological Resource Network (LTER). The LTER surveys overlap with the southern part of the AMLR range and extend further south. Surveys conducted by Germany and Peru are also used in the krill model. The surveys used two, and, more recently, three sound frequencies to better resolve krill. Length frequencies of krill encountered by the acoustic surveys are collected using either RMT8 or IKMT nets, and from these net tow swept area style estimates of krill abundance can be calculated. Additional abundance estimates are derived from acoustic gear, which is capable of detecting krill to depths of between 10 and 250m. Acoustic data from night time operations are discounted, because at night krill are more likely to rise above 10m. The net tows operate between 170m and the water surface. It is believed that the bulk of krill are present shallower than 100m during summer in the area surveyed, although during winter the krill fishery sometimes operates as deep as 100m or a little deeper. Krill shape, swimming orientation, and material properties are accounted for using a physical model adopted by CCAMLR to convert acoustic signals to krill biomass.

Krill are believed to mature at approximately 34mm (roughly age 2). Krill recruitment is thought to be tightly linked to sea ice extent. Krill length is typically somewhat smaller in the waters at the southern end of the survey range, which could lead to some bias in the collected length composition information in years when the full survey area is not covered.

Doug Kinzey presented the stock assessment model for krill. He highlighted the poor correlation (30% or less) between abundance estimates from the acoustics and net tows. The acoustic data has much greater coverage, and is therefore to be preferred. It is theorized that larger krill and gravid females are to be found feeding on the sea bottom, out of the range of surveys. This is believed to remove between 2 and 20% of the population from the acoustic depth range. The model is no longer estimating catchability parameters (q), instead allowing the selectivity curve, over the range of lengths seen in the nets, to have a maximum of less than 1. The curves nevertheless have an asymptote of 1. For the krill fishery, q is also effectively absorbed into the model parameter for the median fishing mortality rate. The reported weight of krill caught is estimated by the fishing companies by converting the amount of product obtained, to green weight. Each fishing company uses its own conversion factor, and these are not typically reported to CCAMLR. There could be an unknown amount of cryptic mortality resulting from krill sustaining fatal damage as they pass through fishing nets.

The model-estimated vulnerable biomass for one of the winter surveys did not pass through the cloud of observed points, and the estimated selectivity for that survey was 1 for all ages. This appears to be an error, perhaps resulting from the considerable convergence problems encountered by the model. Simply changing the selectivity to 0.5 for all ages, and then recalculating the negative log likelihood without changing other parameter values, is expected to reduce the overall negative log likelihood value.

The estimation procedure finished with relatively large gradients (values close to zero indicate convergence). A large number of randomized phase sequences were used to achieve even those gradients. A discussion ensued around the likelihood that the model is over-parameterized, and recommendations were made to remove parameters that might be contributing to the problem.

A number of recommendations were made, some of which Doug Kinzey was able to implement overnight, presenting new model runs on the second morning of the meeting. The subsequent discussion led to new recommendations, and some modifications and clarifications of the originals. The final set of recommendations is given below.

Length composition data from the surveys show episodic recruitment and clear modal progression. During the 2012-2014 period, length composition data are available from both summer and winter surveys. The modal pattern, with somatic growth, is evident across this time series. This suggests that both summer and winter surveys index the same population of krill, and should both be included in the model. However, the presence of ice in winter complicates the estimation of krill abundance so that separate scaling parameters (q) should be maintained for summer and winter surveys. A larger CV might be required for weighting the likelihood component relating to the winter abundance surveys to account for greater uncertainty in those abundance estimates relative to the summer estimates.

2.2 Management decision rules

George Watters indicated that there is much greater variability in krill abundance evident in Area 48 than that previously assumed when the GYM was used to establish the CCAMLR catch limit for the area. The inherent variability, exclusive of any fishing effects, alone is sufficient to violate the CCAMLR decision rule that prohibits abundance from falling below 10% of the median pristine value over a 20-year period. The same problem was encountered for icefish, and this was solved using a 2-year projection from a survey estimate of abundance. Krill surveys do not index absolute abundance as closely as icefish surveys do, therefore this method cannot be extended to krill.

George Watters stated that several papers have indicated that an escapement limit of 75% is likely to be sufficient to maintain predator populations. The current regulation concerns 75% of pristine, rather than 75% of a future projection in the absence of fishing. A number of factors are likely to result in differing future krill abundance, e.g. the recovery of baleen whale stocks, and sea ice retreat. Ideally, desirable escapement levels, by SSMU, would be chosen by estimating the needs of the predators in each area. Attempts are underway to calculate these levels, but are not yet at a point where the results can be used to set krill escapement limits.

3. Terms of Reference

3.1. Summary of findings

The terms of reference provided to the independent reviewers are given in Appendix 1, and are repeated below. Specific recommendations made during the meeting are listed below (lettered A to R).

1. Evaluation of the ability of the integrated model for Antarctic krill, combined with the available data, to provide the parameter estimates required to assess the current status and productivity of Antarctic krill in FAO Subarea 48.1.

The model presented at the start of the meeting had serious trouble converging on a set of parameter values that minimize the negative log likelihood. This is evident in high maximum gradient components, much higher than the ADMB default criterion of 1e-4, indicating that estimation procedure stops when it reaches the maximum number of allowed iterations, not when it satisfies the convergence criteria. Profile likelihoods show a very uneven likelihood surface, indicating that local minima will be a serious problem. These problems are most likely the result of over-parameterization, particularly the inclusion of parameters for which the data are uninformative, and parameters that are confounded. The model needs to be stripped back to a minimum number of estimable parameters. After that, stepwise addition of parameters will reveal which can, and cannot, be estimated. It was recommended that the following set of parameters could be causing convergence problems:

A) Median fishing mortality with annual deviations does not need to be estimated

Instead, the weight of the catch can be assumed to be known without error (but see further recommendations under TOR 5 below). The level of fishing mortality that would give rise to the observed catch, given model estimated krill abundance, is calculated within the model without the fishing mortality rates being model parameters. R code illustrating how Newton's method, run three times, can be used to calculate F was supplied to Doug Kinzey by Robin Thomson after the end of the review meeting.

B) Natural mortality rate (M) might not be estimable

While natural mortality rate is typically a difficult parameter to estimate, it might be possible to estimate it for krill given the clear modal progression in the length frequencies, combined with selectivity that seems to be typically flat for most ages. However, the parameters of the selectivity curve are also estimated and these will be confounded with M. The model's ability to estimate both will be dependent on how informative the length composition data are, particularly with respect to selectivity. M should not be part of the minimum set of parameters, but can be considered in the step-wise addition of estimated parameters.

C) A separate selectivity curve for each of summer and winter, and each of the three gear types (a total of six curves), as well as one for fishing selectivity, is probably too many

The model is unlikely to be capable of estimating 7 different selectivity curves, especially given that some of the time series are of relatively short duration. It was evident from the first presentation of model outputs that one of the selectivity series was not being correctly estimated (the estimated vulnerable biomass did not pass through the cloud of observations), which might result from convergence problems. Aggregation should be considered. During the meeting, aggregation of the IKMR and RMT-8 net gears was considered, reducing the number of survey selectivities estimated from 6 to 4. Aggregation across season, into just an MT-8 and an IKMT survey should be considered as an alternative. The model fits to the survey data from each of the original 6 surveys types, and the commercial fishery, should continue to be depicted graphically, even when aggregation has been implemented. During the review meeting, code was added to the model that allowed flexible selection of aggregation.

D) sigmaR – the variability in recruitment might not be estimable

This parameter is often fixed in fishery models, but adjusted later during the model tuning phase. Tuning is discussed in more detail below. Again, it might be possible to estimate sigmaR and this possibility should be explored, with careful attention to convergence and collinearity with other model parameters.

E) Don't estimate recruitment deviations prior to 1978

The earliest length composition information is available for 1981, therefore the data are not informative for recruitment deviations prior to approximately 1978 (assuming that the length composition shows clear modes for the first 2 to 3 age classes).

F) the "steepness" parameter of the stock recruit relationship should not be estimable

The Beverton-Holt stock recruit relationship allows recruitment to remain high (essentially varying around an unchanging median level) over a wide range of abundances. Only at very low abundance does median recruitment tend to drop. Detecting the level at which this drop occurs requires that stock abundance drop to a low level, and remain at that level (or visit it frequently enough) to detect the new, reduced, recruitment even in the face of the (usually large) variability around this level. Krill recruitment variability is unusually high, and it is thought that the level at which it would drop is very low. Furthermore, the krill fishery in Area 48 seems to be operating at a relatively low level relative to abundance, so that sustained low biomass has not been observed. Estimation of steepness is likely to be impossible given the available data.

G) the variance in length-at-age (a single value is assumed for all age groups) might not be estimable

Conventionally, some kind of relationship (usually linear) between median length-at-age and variance in length at age is assumed. However, a single value for all ages is a reasonable starting point when little is known. Although a great deal of length information is available for the krill stock, age data is not. Hence the model infers age from modal progression in the length data, coupled with estimated growth (length-at-age). Although it is probably feasible to estimate the asymptotic maximum length and growth rate parameters of the von Bertalanffy growth curve (length at zero age is fixed at 0, a reasonable assumption), variances are typically more difficult to estimate. This parameter should not be part of the minimum parameter set, and the model's ability to estimate this parameter, and any key confounding with other parameters, should be explored.

H) R0, the median recruitment prior to the start of fishing, is a parameter that scales population abundance, and this might be poorly estimated

Like steepness, median recruitment can be difficult to estimate when recruitment variability is high, and fishing has little impact on abundance (so that the size of the catch cannot be used to scale the size of the population by noting the drop in catch rates that result after sustained removals). If this parameter does prove difficult to estimate, it might be possible to guide its estimation by imposing prior distributions on at least one of the catchability (q) parameters that scale the relative abundance estimates from surveys to absolute abundance.

A deliberate effort has been made to capture the full range of uncertainty in the krill assessment model by incorporating all of the parameters listed above. However, when parameters are poorly informed by data, convergence problems render the model unusable. Full uncertainty must instead be captured through sensitivity tests, in which inestimable parameters are assigned values at the extremes of their plausible ranges.

Additional recommendations concerning the model were made:

I) Estimate (or fix, or use priors) catchability parameters (q) and ensure that every selectivity curve takes a value of 1 for at least one age group.

Although q parameters have been used in earlier versions of this model, the framework presented to the review meeting does not use q. Instead, it allows the selectivity curves to take values lower than 1 over the range of ages included in the model. However, the standard form of the logistic relationship was used, in which the asymptote is fixed at 1. Therefore, if lower values of selectivity are required to achieve better fits to the abundance data, then the shape of the logistic curve has to be distorted. For example, an estimated curve that was effectively flat at 0.5 for all ages was presented. This brings the survey length composition data, which provides information on the shape of the selectivity curve, into conflict with the abundance data. It would be better to separate the parameters that shape the selectivity curve from the q's, which scale abundance. The effect of the missing q parameter for the krill fishery would be absorbed into both the selectivity and the median fishing mortality parameter (F is actually q*F) which would, typically, cause estimated fishing mortality to be much lower than it actually is (q is usually $\ll 1$ so $q*F \ll F$); which would distort the model estimate of the expected length composition.

J) Allow annual recruitments to be a function of spawning biomass in the previous year, with variability around a median value from the Beverton-Holt stock-recruit relationship

Recruitment was implemented in an unconventional way, the effect of which is difficult to predict. It seemed to favour higher estimated values of steepness by estimating recruitment deviations (which are constrained to be small) around a median level, which is in turn constrained to match the Beverton-Holt – ultimately favouring a flatter Beverton-Holt. A more conventional implementation would be easier for other researchers (such as the CCAMLR community) to understand, and has a lower likelihood of having undesirable properties. The recruitment deviations need to be constrained to sum to zero so that they are truly deviations around a median. (The same would be true for the fishing mortality deviations, if these were to be retained in the model).

K) Attention needs to be given to the relative weights given to data sources (model tuning)

The likelihood component for the length composition data is orders of magnitude larger than that for other data sources. This could give undue influence to those data and has the potential to exacerbate convergence problems. In addition, conventional fisheries assessment practice involves a tuning step during which the CVs or effective sample sizes ("weights") that are used to weight likelihood components are compared with those calculated from the residuals from the model fit. Typically, weights are altered during several iterations until input and output weights converge. An additional, somewhat arbitrary, adjustment is also often made to keep likelihood components at a similar scale. The sigmaR parameter (if not estimated) is compared to the standard error of the model outputs, and is also tuned at this stage. See presentations given to the 2015 CAPAM workshop (http://www.capamresearch.org/workshops/data-weighting/presentations).

L) The ADMB code should be simplified

The ADMB computer code has been developed over a number of years. Alternative model structures (e.g. one that allowed migrations between areas) that have been explored in the past are still implemented, but not used, in the code. This is a common phenomenon that results in computer code that is understandable to its developer but less so to others. AERD are hoping to use this model to set catch limits in CCAMLR Area 48.1, which would mean lodging the code with the CCAMLR Secretariat for use by scientists from other nations. It would therefore be advantageous to simplify the code as much as possible, removing sections that are no longer used and reducing, where possible, the number of steps and named variables that are used. This will also greatly reduce the likelihood of undetected coding errors.

M) The Gompertz growth model should be considered as an alternative to the von Bertalanffy

Simon de Lestang indicated that the Gompertz model has been found to better represent crustacean growth patterns than the von Bertalanffy. However, the models primarily deviate only during the early, slow growing stage of crustacean life. The estimated selectivity curves ensure that recruitment to the model occurs at a relatively large size (20-25mm) so that the von Bertalanffy model is probably appropriate.

N) A length-based model might be more appropriate

Simon de Lestang raised the possibility that a length-based model might be more appropriate for a species for which age could not be determined. The krill model relies on the presence of modal progression in the length frequency data for estimation of a growth curve, which is used to convert lengths to ages. While it would be interesting to explore the differences between a length based model and an age based model, it is my opinion that the data are sufficient to support the age based model.

O) The length data should be standardized for area of collection

It has been noted that krill collected towards the south of the AMLR survey area are generally smaller than those collected in the north. Simon de Lestang recommended that krill length should be standardized before inclusion in the model. However, it is not clear how this could be achieved, and the effect might not be strong enough to justify the associated workload.

2. Evaluation of the strengths and weaknesses in the use of the integrated modeling approach to assess whether harvest recommendations meet the CCAMLR decision rules.

The integrated (analysis) modelling approach (IA) allows the incorporation of different sources of data into a single analysis. The approach easily allows for gaps in data time series - unlike methods such as VPA where such gaps have to be estimated externally to the model. Krill survey data, in the form of biomass estimates and length composition data, are available for many years, but not over the same time periods. Alternative survey estimates collected using different gear type are sometimes available for the same year. Length composition data are also available from the krill fishery. The IA modelling framework lends itself to such data, incorporating all available data along with time series gaps, and multiple estimates for a given year, without necessitating the invention of missing data or the selection of a "best" single data value for each year. The IA approach is now widely used in the management of fisheries (Maunder & Punt 2013).

The Generalized Yield Model, currently used by CCAMLR to set catch limits, requires a small number of biological parameter values which are calculated (or guestimated) outside of the model. It is an appropriate model to use when very little information is available. However, a great deal of information is now available from Area 48.

The IA model estimates the size of the krill stock, relative to its median unfished size, over a period of years. It incorporates krill fishing. As such, the model is in a good position to assess whether the CCAMLR decision rules have been met, in so far as krill population variability and krill escapement is concerned. In the absence of survey estimates of absolute abundance, the model may struggle to produce precise estimates of krill population size. However, this has been solved in the past by assuming that survey estimates are absolute (q=1 for surveys). The IA model offers the opportunity to improve on this assumption by incorporating into model results, the uncertainty associated with the estimate of abundance.

3. *Evaluation of the spatial scale over which the model estimates may be applied.*

The IA model reviewed here applies to Area 48.1. Surveys are available for a wider area (Kinzey et al. 2011) so there is a potential to apply the model to all of Area 48. However, attempts to do so, which included estimation of krill movement between four regions, proved problematic (Kinzey et al. 2011). The data for Subarea 48.1 appear to be internally consistent and it would be best to complete the proposed modelling work for that Subarea before trying to extend the scope of the model. The model does not have sufficient data to allow disaggregation into smaller regions (such as SSMUs). Additional survey data for the Subarea exist, and it is recommended that:

P) LTER survey data be obtained and incorporated into the krill assessment, if possible.

4. *Evaluation of strengths and weaknesses of a proposed alternative to the current CCAMLR decision rules, which compare spawning biomass under projected levels of fishing to initial unfished spawning biomass. The alternative compares spawning biomasses during the period of projected fishing to spawning biomasses with no fishing during the projection period.*

The CCAMLR decision rules were developed in conjunction with the GYM, in its original form (Butterworth et al. 1994). Variability as great as that detected in Subarea 48.1 was not considered, but the intent was to allow krill fishing to occur, despite the presence of natural biomass fluctuation. The proposal that the effects of fishing be measured against projected biomass in the absence of fishing seems a reasonable solution to the problem that the existing management rules would not allow any fishing at all. The proposal seems consistent with the intent of CCAMLR Article II (Constable et al. 2000).

Two recommendations are made with respect to future spawning biomass. First, that:

Q) future projections should use recruitment deviations, not past numbers of recruits.

The model presented to the review uses actual numbers of recruits estimated in the past to project into the future. This results in relatively narrow confidence bounds on future numbers at age. It is better to draw from past recruitment deviations, around the Beverton-Holt relationship, for the future rather than estimated numbers of animals. This more accurately reflects uncertainty in future recruitment. Krill show marked periodicity in recruitment with roughly 5 years between peak recruitments. Ideally, a time series approach (such as AR(1)) could be used to model this periodicity and that, with added noise, could be used to draw future recruitment deviations. In the meantime, an easier approach would be to discount the two most recent recruitments (which will be less reliably estimated because those cohorts have had less time to be observed), block the previous years of recruitment deviations into groups of 5 consecutive years, and randomly draw (with replacement) from those groups for each set of 5 future years. Recruitments that occurred more recently should be drawn with greater frequency. It was recommended that the following set of weights be used when drawing 5-year time blocks, from most recent to most distant: 0.4, 0.3, 0.2, 0.1.

The second recommendation is that:

R) spawning potential should be measured as fecundity at age (integrated across length at age) rather than as weight at age. Only krill aged two or older should be considered mature.

A fecundity relationship, taken from the literature, was incorporated into the model during the 3-day meeting. It is recommended that a fecundity relationship be used in future, and that a fecundity relationship for the survey area could be calculated using the data collected by AMLR.

Simon de Lestang, who has expertise in crustacean biology, indicated that one-year old krill are unlikely to be mature, even if they are large. The model currently allows 47% of one-year old krill to be mature, based on an assumed length at age distribution and a length at maturity. Because the model estimates a growth curve, the proportion of animals of age 2 and older that are mature should be allowed to vary as the parameters of the growth curve vary (i.e. the proportion mature should be calculated within the IA model, based on the estimated distribution of length at age).

5. Evaluation of the suitability of the integrated assessment in comparison to the GYM approach to determining precautionary catches.

This has been covered under TOR 2 above.

(i) Recommendations for further improvements to the assessment model, including frequency of surveys and types of data collection.

It is essential that any model that is used to recommend future catch levels should satisfy basic convergence criteria, such as a maximum gradient component below 1e-3 (preferably below 1e-4). Recommendations A to J listed under TOR 1 aim to achieve convergence and are therefore essential. Recommendation K, regarding tuning, is highly recommended as an important part of modern stock assessment. Recommendation L, regarding code simplification, is also highly recommended particularly for the detection of 'bugs' in the code. Recommendations M to O are not essential, and N in particular could be time consuming (but very interesting).

The AERD krill team hope to include length composition information from predator stomachs into the assessment, to refine estimates of predator consumption needs. This is an ambitious idea, which would require good time series of stomach contents data and additional information in predator foraging strategies. If possible, this work would be well worth pursuing.

Extension of the model to a larger area could be considered once the existing model is in good working order (i.e. converging). A natural first step would be to incorporate LTER data (recommendation P) followed by AMLR survey data for other Subareas of Area 48. This would require the re-introduction of the immigration and emigration processes that were modelled by Kinzey et al. (2011), and has a high risk of over-parameterizing the model and once again causing convergence difficulties. Better underlying understanding of the movements and recruitment processes of krill in the region are needed so that strong theoretical understanding can be included in the model to support estimation.

Once the model is converging and a base case set of parameter values has been chosen, the model can be used to investigate the effect of alternative future survey designs. For example, are both summer and winter surveys needed; should surveys be conducted annually or biennially? Through simulation modelling, the effect of alternative survey design on the model's ability to accurately and precisely estimate key quantities of interest can be determined.

The use of different net types and acoustic frequencies has complicated interpretation of the data. It would be preferable to use a single configuration in future, and to calibrate the results from equipment used in the past by conducting parallel investigations.

6. Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

The review panel was small, consisting of two independent reviews, and four AERD employees of whom two were present only intermittently. Discussions were friendly and mutually co-operative with AERD employees helping the reviewers better understand Antarctic krill biology, CCAMLR's management system, and the GYM and AI krill models. Both reviewers and AERD staff felt comfortable to advance new ideas that might solve existing problems. AERD staff, and in particular Doug Kinzey, are to be commended for their openness to discussing their work and their willingness to explore alternative approaches. I would like to thank the AERD team for making the review such an interesting and positive experience.

A good understanding of the life history of the stock being modelled is essential for an assessment scientist, and the reviewers asked many questions concerning the biology and life history of krill as part of their review. Krill have an unusual life history, in the context of marine resource stock assessment, so that the traditional assessment framework might need at least some small alterations to best represent krill. Although there were several question and answer sessions during the review concerning krill life history, whole books have been written on the subject and much is still to be learned about the species. It is probable that over future decades, as more information becomes available and clearer hypotheses are formulated, particularly regarding the stock structure, connectivity, movement, and recruitment processes, the stock assessment framework will be adapted.

3.2. Conclusions and Recommendations

Antarctic krill in CCAMLR Area 48, including Subarea 48.1, is managed using a catch limit and a trigger limit whose choice was guided using the GYM simulation model (Constable et al. 2000). An estimate of the absolute biomass of krill in the region from surveys conducted in 2000 was used, in conjunction with GYM output, to set the catch limit. The GYM was an innovative solution to the problem of setting a catch limit for a stock about which little was known and for which fisheries data were absent. In the decades since the implementation of the GYM, at least in Subarea 48.1, impressive time series of abundance indices from surveys, as well as length composition information from both surveys and a commercial fishery, have become available. A fishery has been operating in the Subarea for some time, providing information on the potential of a fishery to impact the stock. One of the greatest challenges to fishery management is the setting of catch limits for stocks that have not been fished before, or at levels too low to test the stock's ability to support catches. This was the problem to which the GYM was a clever solution. Now, with a history of fishing in Subarea 48.1 along with appropriate data collections and survey information providing feedback on the stock's ability to support that fishing, it is more appropriate to apply a fisheries stock assessment model. The current state of the art in this category is the Integrated Assessment Model (Maunder & Punt 2013). The IA model particularly lends itself to systems, of which krill in Subarea 48.1 is an example, in which data is available from multiple sources and in which there are gaps in time series. Not only is it appropriate to apply an IA model to krill in Subarea 48.1, in light of currently available data it would be inappropriate not to.

Simulation testing has been routinely conducted on the IA krill assessment model, using the model itself to generate data, with added noise, to which the model is then applied to see whether it can recover the "true" parameter values. The model developers are congratulated for taking the time to perform these important diagnostics. This should continue to be part of their model development work.

The IA model, in the form presented to the review, had considerable problems converging (i.e. identifying a unique set of parameter values that minimized the objective function value). This is most likely due to over-parameterization – the inclusion of parameters for which the data contained little or no information, and of parameters that were correlated with one another. The bulk of the recommendations from this review concern the suggested removal or aggregation of estimated parameters. It is recommended that poorly estimated parameters have their values fixed at plausible levels, and that sensitivity tests to varying those values be used to capture uncertainty in model results. Once the model is able to reliably converge on a set of "best fit" values, it would be appropriate to present the model to CCAMLR for use in setting catch limits.

The existing CCAMLR decision rules for krill preclude any fishing on a stock that has highly variable recruitment of the order observed for krill. The decision rules seek to prevent the occurrence of low levels of krill biomass, levels that occur naturally even in the absence of a fishery. This situation arose for icefish, leading CCAMLR to alter its decision rules for that species. It would be appropriate to alter the decision rules for krill too, to allow fishing while honouring the intent of CCAMLR's Article II which seeks to allow for the needs of krill predators. The ideal would be to estimate the foraging needs of all major predators in each SSMU and to set catch limits accordingly. However, in the absence of clear estimates (Plaganyi & Butterworth 2012), it would seem sensible to use an appropriately parameterized IA model for subarea 48.1 to set catch limits for the area that would be apportioned amongst SSMUs in accordance with the best available information on the needs of predators. It would be appropriate, and true to the spirit of Article X, to set the catch limit so krill biomass remains at or above 75% of the projected biomass in the absence of fishing. This allows for predictable fluctuation in krill abundance resulting from cyclic recruitment. If krill abundance were to be shown to suffer a long term decline due to factors other than the fishery, such as sea ice retreat, then it might be necessary to think beyond Article X, which provides guidance only for a system that is stable in the long term.

It must be noted that the IA model relies upon time series of data collected from surveys and from the krill fishery. While, in principle, it can be used in the same way that the GYM has been used, to set a catch limit that is then fixed for an indeterminate time into the future, it would be irresponsible not to revisit the catch limit at regular intervals, updating the model with data from continued collections. The long term accuracy and reliability of catch limits set by such an IA model rely on the current frequency and timing of the surveys and on accurate data collection of fishery length composition and tonnage caught. The influence on the accuracy of model estimates of any proposed alteration to the survey schedule should be simulation tested using the IA model. The model is reliant on survey estimates of biomass. If surveys were stopped, or became too infrequent, it would be necessary to consider using catch rate information from the fishery to index abundance. This method is notoriously inaccurate for patchily distributed stocks such as krill and would therefore be undesirable.

Good estimates of the tonnage of krill caught by the commercial fishery are essential for management of the resource. Existing estimates were calculated by the fishing operators by applying green weight conversion factors to the weight of the processed product. Only the resulting green weight estimate is supplied to CCAMLR, not the product weight and conversion factor. Operators reportedly do not all use the same conversion factors. It is important to standardize these calculations, that accurate conversion factors are calculated, and that information on how the estimated catches were calculated are supplied to CCAMLR.

Global warming is changing the Antarctic ecosystem. It will ultimately become necessary to revisit current management regulations and the aims of management as set out in the CCAMLR Convention, which assume that if fishing were stopped the system would return to an unchanging long term “virgin” biomass. If this is no longer true, then new management goals are needed.

Disclaimer

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

Bibliography

See Appendix 1.

Appendix 1: Bibliography of materials used for the review

Documents specified in the Terms of Reference:

Kinzey, D., G. Watters, and C. Reiss. (2015a) An integrated stock assessment with alternative decision rules for Antarctic krill harvests around the Antarctic Peninsula. 34 pp.

Kinzey, D., G. Watters, and C. Reiss. (2015b) Selectivity and two biomass measures in an age-based assessment of Antarctic krill (*Euphausia superba*). *Fisheries Research* 168:72-84.

Kinzey, D., G. Watters, and C. Reiss. (2014) Integrated models for Antarctic krill (*Euphausia superba*) using survey data from 1981–2014 in Subarea 48.1. CCAMLR WG-SAM-14/32. 30 pp.

Kinzey, D., G. Watters, and C. Reiss. (2013) Effects of recruitment variability and natural mortality on Generalised Yield Model projections and the CCAMLR Decision Rules for Antarctic krill. *CCAMLR Science* 20:81-96.

Kinzey, D., G. Watters, and C. Reiss. (2011) Modeling Antarctic krill: scale, movement and age-structure. CCAMLR WG-EMM-11/43 Rev.1. 37 pp.

Constable, A.J., W.K. de la Mare, D.J. Agnew, I. Everson, and D. Miller. (2000) Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). ICES Journal of Marine Science. 57: 778-791.

Additional materials supplied to reviewers:

Kinzey, D., G. Watters, and C. Reiss. (2015c) Estimating future krill catches that meet the CCAMLR and alternative decision rules for FAO Subarea 48.1 using an integrated assessment model. CCAMLR WG-EMM-15/51. 21 pp

Kinzey, D., G. Watters, and C. Reiss. (2013) Model diagnostics for the Antarctic krill assessment for Subarea 48.1. CCAMLR WG-FSA-15/60. 15 pp.

Kinzey, D., G. Watters, and C. Reiss. (2016) An integrated stock assessment with alternative decision rules for Antarctic krill harvests around the Antarctic Peninsula. 34 pp. *A document produced for this review.*

Model files were for a selected version of the model: the ADMB template file *krill.tpl*, data input file *krill.dat*, executable *krill.exe*, output parameter file *krill.pin*, report file *krill.rep*, and MCMC results file. Model files, input data and R code were also provided for the versions of the model documented in the Fisheries Research publication Kinzey et al (2015b), and the working group documents WG-EMM-15/51 and WG-SAM-14/20.

Additional documents referenced in this report:

Butterworth, D.S., Gluckman, G.R., Thomson, R.B., Chalis, S., Hiramatsu, K., Agnew, D. (1994) Further computations of the consequences of setting the annual krill catch limit to a fixed fraction of the estimate of krill biomass from a survey. *CCAMLR Science.* 1: 81-106.

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

Plaganyi, E.E. & Butterworth, D.S. (2012) The Scotia Sea krill fishery and its possible impacts on dependent *predators*: modeling localized depletion of prey. Ecological Applications. 22(3). 748-761.

Appendix 2: Statement of Work

National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

An age-based, integrated stock assessment for Antarctic krill (Euphausia superba) with projected catches to 2035

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards. (http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf).
Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The SWFSC Antarctic Ecosystem Research Division (AERD) requests an independent review of the integrated stock assessment it has developed for Antarctic krill. The Antarctic krill fishery is managed by the international treaty organization, the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR), of which the U.S. is a member. The fishery is currently expanding both in total catch and in the number of nations and vessels participating. Current catch limits of 5.61 million tons for the Scotia Sea, where all current krill fishing is conducted, were established using the Generalized Yield Model (GYM), a model developed in the 1990s. 155 thousand tons of this total catch limit has been apportioned to FAO Subarea 48.1, a subregion of the Scotia Sea.

The GYM is a simulation model rather than a statistical model and relies heavily on data from a single multination survey conducted in 2000. The integrated model developed by AERD is a statistical model that uses annual survey (1981-2014) and fishery data (1976-2014) from FAO Subarea 48.1, where available, and will continue to use new data as it becomes available. AERD in conjunction with the U.S. Department of State intends to propose to CCAMLR that the model developed by AERD be used to establish catch limits for the krill fishery rather than the GYM. Having the model framework scientifically vetted outside of AERD or CCAMLR will be an important step in this process. The Terms of Reference (TORs) of the peer review and the tentative agenda of the meeting are below.

Requirements
NMFS requires two reviewers to conduct an impartial and independent peer review in accordance with the SOW, OMB Guidelines, and the TORs below. The reviewers shall have working knowledge and recent experience in the application of fisheries stock assessment processes and results, including population dynamics, separable age-structured models, harvest strategies, survey methodology, ecosystem-based fishery management, and the AD Model Builder programming language. They should also have experience conducting stock assessments for fisheries management.

Tasks for reviewers

- Review the following background materials and reports prior to the review meeting:

  Kinzey, D., G. Watters, and C. Reiss. 2015. An age-based, integrated stock assessment for Antarctic krill (Euphausia superba) with projected catches to 2035. 30 pp (estimated page number, document to be developed)

  Kinzey, D., G. Watters, and C. Reiss. 2015. Selectivity and two biomass measures in an age-based assessment of Antarctic krill (Euphausia superba). Fisheries Research 168:72-84.

  Kinzey, D., G. Watters, and C. Reiss. 2014. Integrated models for Antarctic krill (Euphausia superba) using survey data from 1981–2014 in Subarea 48.1. CCCAMLR WG-SAM-14/32. 30 pp

  Kinzey, D., G. Watters, and C. Reiss. 2013. Effects of recruitment variability and natural mortality on Generalised Yield Model projections and the CCAMLR Decision Rules for Antarctic krill. CCAMLR Science 20:81-96.

  Kinzey, D., G. Watters, and C. Reiss. 2011. Modeling Antarctic krill: scale, movement and age-structure. CCAMLR WG-EMM-11/43 Rev.1. 37 pp

  Constable, A.J., W.K. de la Mare, D.J. Agnew, I. Everson,and D. Miller. 2000. Managing fisheries to conserve the Antarctic marine ecosystem: practical implementation of the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR). ICES Journal of Marine Science. 57: 778-791.

- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- After the review meeting, reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus
- Each reviewer may assist the Chair of the meeting with contributions to the summary report, if required by the TORs
- Deliver their reports to the Government according to the specified milestone dates

Foreign National Security Clearance
When 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 reviewers who are non-US citizens. For this reason, the 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/ and
http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-national-registration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance
The place of performance shall be at the contractor's facilities, and at the Southwest Fisheries Science Center in La Jolla, California.

Period of Performance
The period of performance shall be from the time of award through May 31, 2016. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Within two weeks of award	Contractor selects and confirms reviewers

No later than March 14, 2016	Contractor provides the pre-review documents to the reviewers
April 4-6, 2016	Panel review meeting
April 21, 2016	Contractor receives draft reports
May 15, 2016	Contractor submits final reports to the Government

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards: (1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
All travel expenses shall be reimbursable in accordance with Federal Travel Regulations (http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract. Travel is not to exceed $16,000.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

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

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

   c. Reviewers should elaborate on any points raised in the summary report that they believe 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 report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

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

Terms of Reference for the Peer Review

An age-based, integrated stock assessment for Antarctic krill (Euphausia superba) with projected catches to 2035

1. Evaluation of the ability of the integrated model for Antarctic krill, combined with the available data, to provide the parameter estimates required to assess the current status and productivity of Antarctic krill in FAO Subarea 48.1.
2. Evaluation of the strengths and weaknesses in the use of the integrated modeling approach to assess whether harvest recommendations meet the CCAMLR decision rules.
3. Evaluation of the spatial scale over which the model estimates may be applied.
4. Evaluation of strengths and weaknesses of a proposed alternative to the current CCAMLR decision rules, which compare spawning biomass under projected levels of fishing catch to initial unfished spawning biomass. The alternative compares spawning biomasses during the period of projected fishing to spawning biomasses with no fishing during the projection period.
5. Evaluation of the suitability of the integrated assessment in comparison to the GYM approach to determining precautionary catches.
   (i) Recommendations for further improvements to the assessment model, **i**ncluding frequency of surveys and types of data collection.
6. Brief description on panel review proceedings highlighting pertinent discussions, issues, effectiveness, and recommendations.

Tentative Agenda

An age-based, integrated stock assessment for Antarctic krill (Euphausia superba) with projected catches to 2035

TBD

Southwest Fisheries Science Center

8901 La Jolla Shores Drive

La Jolla, CA 92037-7000

March 28-30, 2016 9AM - 5PM

Point of contact: Front Desk

Appendix 3:Panel membership

Attendees at the SWFSC Antarctic Krill assessment review meeting:

Doug Kinzey, NMFS
Christian Reiss, NMFS (day 1)
Jeremy Rusin, NMFS (part of days 1 and 2)
George Watters, NMFS
Simon de Lestang, Independent reviewer
Robin Thomson, Independent reviewer

Appendix 4: Meeting report

No formal meeting report will be produced, but a highly summarized “vignette” will be produced by the AERD team.