INTERNATIONAL REVIEW PANEL REPORT FOR THE 2014 INTERNATIONAL FISHERIES STOCK ASSESSMENT WORKSHOP

1 – 5 December 2014, UCT

[A Dunn[1], J S Link[2], A E Punt[3], G Stefansson[4], and R S Waples[5]]

Introduction

The Panel recognised the very high quality of the research presented at the 2014 International Fisheries Stock Assessment Review Workshop. This included research on Southern African hake and sardine, as well as research associated with whether pelagic fishing near island breeding colonies has an impact on African penguins. The Panel thanked the workshop participants for their hard work preparing and presenting the workshop papers, for the extra analyses undertaken during the workshop, and for the informative input provided during discussions.

This report starts with observations from the Panel on some general issues for the species / programmes reviewed, and then focuses on the more detailed technical review and recommendations concerning each. The recommendations are annotated by their priorities (H, M, L and conclusions are indicated by asterisks).

Summary of general issues

Penguins

The Panel recognized that the issue surrounding penguins and small pelagic species is effectively one of multiple, competing objectives. As such, ultimately the key issue for the topic is an evaluation of trade-offs among different objectives. In this context, the Panel notes that both globally, and locally in South Africa, ecosystem considerations have emerged as an important management issue for fisheries. Nevertheless, there remains a conflict between the local scientific communities that primarily focus on ecosystem modelling and on fish stock assessment, which needs to be resolved. Both science and management advice would benefit from closer collaboration between these communities.

The Panel recognized that the two groups ('A' and 'B') had evaluated different sets of questions. Specifically, the 'B' group focused primarily on the estimation of residual variance using the results of the Island Closure Feasibility Study and the impacts of catches on penguin biological parameters, while the 'A' group focused on the impact of closures on penguin biological parameters. The Panel noted that greater clarity on the objectives of the work might have been able to reduce some of the disagreements on methodology and results that were apparent.

The Panel clarified that the fundamental question is whether small pelagic fisheries removal/catch of targeted fishes near islands can impact severely declining penguin populations. This question was phrased as: Can the Island Closure Feasibility Study elucidate whether an experimental closure programme could yield definitive conclusions regarding the impact on penguin populations, where fishing occurred in close vicinity to penguin breeding islands. The Panel notes that these differences in the question being addressed by both groups in some ways reflect the differences in disciplinary perspectives. These different perspectives could ultimately be a strength in approaching this issue if managed well.

The Panel notes that both groups report inconsistent findings regarding the impact of closures / reduced catches. Thus, some effects of closures have been detected, even if found

[1] NIWA, Wellington, New Zealand
[2] NOAA Fisheries, Silver Spring, USA
[3] University of Washington, Seattle, USA
[4] University of Iceland, Reykjavik, Iceland
[5] University of Washington, Seattle, USA

to be marginally significant, with some highly significant results. The challenge is these detected effects can be in either the "positive" or "negative" direction (see Table 1). This is most apparent across groups, but also even within a group with respect to penguin responses at a given island. These inconsistencies must be considered when interpreting the results from either set of analyses, as well as the observation that the two groups answered different questions. One plausible explanation is that there is at least one factor which drives penguin dynamics that has not been included in any of the models, but which is confounded with the closure periods. This can easily happen because there have only been few closures over a short time period. Alternatively, effects could be occurring, but are not fully detected due to insufficient data or a change in how catches were reported with respect to catches in the area of closure.

One important element of information that emerged from the Panel review was an evaluation of catches and penguin foraging distance/path length relative to area closed. The vast majority of penguins forage inside the 20km closure during the chick rearing phase. This was robust whether or not an island closure was in effect.

Effects that are statistically significant are not necessarily biologically important. Nevertheless, given the literature, it is not surprising to find some statistically significant results indicating a positive effect of reduced fishing on penguins (Furness and Tasker, 2000; Witherall *et al*., 2000; Daunt *et al*., 2008; Frederiksen *et al*., 2008; ICES, 2014); in contrast, the inference that increased fishing improves penguin population status seems unlikely given the literature. However, the key challenge remains to estimate the relative magnitude of any such effects and how they impact both the penguin population (compared to other pressures) and the fishery (compared to total landings).

Despite the differences, the Panel notes that there were areas of agreement between both groups. Examples include positive effects on penguins of reduced catches / closures at St. Croix, the general recognition of the importance of foraging path length and fledging success, that multiple factors can effect penguin dynamics, that catch is not a direct measure of biomass, and that cessation of fishing around the islands by itself is unlikely to be sufficient for the penguin population to recover. Of primary note and most importantly, both groups agreed that the feasibility study was successfully conducted (even if for different reasons). The Panel concurs.

Hake

The process implemented following the 2013 review involved biologists and modellers collaborating to synthesize multiple sources of information for hake off South Africa and Namibia to identify plausible stock structure hypotheses. It also identified the data that might be used to fit models based on those hypotheses. This process has been very successful. The hypotheses identified through this process, while not yet final, represent an effective synthesis of a diverse set of information. The Panel recommends reducing the number of hypotheses considered during the initial stages of modelling transboundary hake issues, owing to problems that some of the analyses identified during the review.

Model formulations have been developed to conduct assessments of the hake off South Africa and Namibia that take spatial structure into account. These formulations range from those representing spatial structure using fishery selectivity, to ones which model spatial structure explicitly. The initial fits of the latter class of model are currently poor, and much work likely remains before a model is found which provides an adequate fit to all of the data when they are represented spatially.

The Panel was impressed with the progress made on developing an assessment method that can take into account both cannibalism between the two hake species and inter-specific predation. This line of work holds as much if not even more promise than the movement model in terms of representing hake dynamics, because its results may appreciably change perceptions of stock trajectories for the two hake species - in particular that for *M. paradoxus*, which is subject to predation by *M. capensis* as well as cannibalism. The Panel looks forward to further progress of this work.

Sardine

The process implemented following the 2013 review involved biologists and modellers collaborating to synthesize multiple sources of information for sardine off South Africa to identify plausible stock structure hypotheses. The process also identified the data that might be used to fit models based on those hypotheses. This process has been very successful. The hypotheses identified through this process, while not yet final, represent an effective synthesis of a diverse set of information.

The work of the Panel focused on ranking the various model formulations and identifying the technical changes to these formulations so that maximum use can be made of available data, in particular the data that has recently become available from parasites.

General considerations

The Panel highlights the critical importance of reviews of work used for management purposes including a detailed review of the quality and accuracy of the basic data. For example, doing so for hake might have avoided the problems encountered with the conclusions drawn from the GeoPop model (see below).

The Panel was presented with a model with many random effects based on the Template Model Builder (TMB) software (Kristensen, 2014). Many of the models on which assessments of South African fish and invertebrate stocks are based involve parameters that are effectively random effects. Use of the most efficient methods for parameter estimation is therefore essential. Holding a TMB workshop in South Africa would both allow local scientists the opportunity to learn this package and also to evaluate whether they wish to use it in future work.

The Panel was provided with an enormous amount of background material and primary papers. This represents a very large investment of time by the analysts, and each paper warrants careful consideration. However, even some primary papers were not presented let alone fully discussed, given the time available to the Panel. The Panel therefore recommends:

- For each species being reviewed, a document should be produced that provides an overview of the fishery, its history of exploitation, and the types of data available, expanding, for example MARAM/IWS/DEC14/Hake/BG4. This "fishery description" document should be provided to the Panel well in advance of the review meeting, as it will help Panel members who are not familiar with South African fisheries and fisheries management techniques to become fully prepared for the review. Bentley *et al.* (2014) provides an example of this type of document for sea perch in New Zealand.
- Terms of Reference (or sets of 'key questions') should be developed for each species being reviewed. These should be provided to the Panel well in advance of the review to enable the Panel to focus their effects when reviewing the rest of the material on those aspects that are most pertinent to the questions of greatest interest. Document MARAM/IWS/DEC14/Sardine/P3 provides such an overview, but was one of the last documents to be circulated to the Panel.
- The documents provided should be restricted to those that are directly relevant to issues at hand.

African penguins

Background

Any evaluation of the impact of catches or closures on penguin biological parameters needs to be placed in the context of all of the impacts that might be hindering recovery of the island penguin populations. The Panel notes that projections conducted by both groups (although evaluating somewhat different questions) suggested that eliminating pelagic catches / implementing long-term closures around islands by itself cannot be expected to lead to recovery of penguin populations (MARAM/IWS/DEC14/Peng/A11, MARAM/IWS/DEC14/Peng/3a, Weller *et al.* 2014), although there are some predicted benefits. Further, the Panel notes that fishing is a factor that can be controlled and reductions in fisheries (via closures) have had positive effects in other places around the world in similar situations (Furness and Tasker, 2000; Witherall *et al.*, 2000; Daunt *et al.*, 2008; Frederiksen *et al.*, 2008; ICES, 2014). What also merits consideration is why the mainland Stony Point population of the same penguins exhibits different population trajectories than the island populations. Additionally, the Panel notes that the island closures will not only affect penguins and small pelagic fishes, but also potentially many other species, either directly or indirectly. The consequences of these effects merit consideration.

As such, and in the context of multiple objectives, the Panel recommends that the appropriate authorities (DEA and DAFF) work together to identify goals for both the pelagic fishery and penguin recovery, to develop and implement a comprehensive research program that aims to identify the core reasons for the reduction in penguin population numbers, and identify any potential mitigation measures. Fishery impacts on prey should be one important component of this program, but there will likely be other components, as identified, for example, by the penguin pressure model (Weller *et al.*, 2014).

The Panel was presented with an on-going debate on whether to use a simple factor of closed/open or catch-in-area as the explanatory variable. In a general regression setting, **if** the explanatory variable is an actual control variable, which can be pre-set to specific values, and **if** the response is a linear function of the control variable, then it is common statistical design to pre-set the control variable at several values, dispersed over a range of levels, not just 0 and 1. On the other hand, this explanatory variable is **not** a control variable in the present setting, since it depends on the behaviour of the fleet, which again depends on the behaviour of the resource as well as on economic aspects. In addition, because catches are correlated to (though not a proxy for) biomass to some extent, there is a serious risk that coefficients may change sign, in which case a regression of a penguin demographic variable on catch may reflect a relationship with biomass and not the true control variable which is whether the area is closed or not. Thus, although a regression using catches may have greater power than an ANOVA using a closure factor, there are several problems with the approach, and interpretation needs to take this into account. The use of a local biomass estimates may alleviate some of these issues, but then one will need to take into account the large(r) measurement uncertainty. These issues are serious, and can be avoided if a simple ANOVA-type model is used, but at the likely cost of a loss in power.

Responses to Terms of Reference

The initial Terms of Reference for this component of the review provided to the Panel (MARAM/IWS/DEC14/Peng/C2) were:

a) What scientific conclusions can be drawn from the present island closure and related analyses regarding the benefits or otherwise of closures to penguin recovery?
b) What further analyses of available information could help resolve uncertainties about the conclusions?

c) What would be the benefits to formulating scientific advice of continuing the closures taking into account limited further resources that DAFF could commit to further monitoring and analysis?

Immediately before the workshop, Dr. Kim Prochazka proposed a change to these Terms of Reference to the Panel. The Panel requested that the proposed changed Terms of Reference be further clarified to be more precise and hence allow it to respond appropriately. This new set of Terms of Reference (Figure 1) was a hierarchical tree, depending upon evaluation of sequential issues to better focus the Panel's review and ability to provide usable advice. This new set of Terms of Reference was also compatible with the original Terms of Reference, and was simply aimed to clarify and specify them. In this report, the Panel has considered both the initial Terms of Reference and the Terms of Reference in Figure 1.

1. Can the Island Closure Feasibility Study tell us whether an experimental closure programme could yield definitive conclusions regarding the impact of fishing close to penguin breeding islands on penguin populations?

The Panel requested each group to comment on whether they believed that the feasibility study had been successfully completed. Both groups replied in the affirmative. The Panel also **agrees** that the feasibility study has been completed successfully (but see 8 below). The Panel **concludes** that an experimental closure programme could yield definitive conclusions regarding the impact of fishing close to penguin breeding islands on penguin populations. Consequently, the Panel reports on questions 5 – 8 in Figure 1.

5. [if yes to question 1]Can it [the Island Closure Feasibility Study] be used to draw conclusions re responses of penguin populations?

Both groups have found that statistical tests for the impact of fishing on biological parameters of penguin populations led to significant results (Table 1). However, these statistically significant results were often in different directions. In some cases the results were conflicting across islands for the same response metric, even within each group's analysis (e.g. the impact of closures on chick condition and trip duration [group A]). The lack of consistency between and within groups may be because there are factors other than those considered to date that impact biological parameters for penguins. The Panel therefore **concludes** that it is premature to draw final conclusions.

6. If no [to question 5], what closure regime, data and analyses will assist in moving towards a conclusion?

The Panel **recommends** that the most effective way to address the impact of fishing near islands on penguin populations is to conduct a full scale experiment. Any experiment should be preceded by the development of an experimental design using an appropriate power analysis. The Panel **strongly recommends** that the current closure regime be continued if an experiment is conducted. The exact number of years needed to achieve a given outcome, in terms of the effect of fishing on penguin demographics, will depend on a power analysis accepted by DAFF under a process specified by their Director Research, which is not available at this time. If a full experiment is to commence, there will be a considerable gain in ensuring that it forms a natural continuation of the feasibility study. Unless a clear improvement to the current design is found, the Panel **recommends** that the design in the feasibility study (Table 2) continue to be used in the interim period until a design is finalised. The Panel provides additional recommendations below regarding the design of a full experiment.

7. If yes [to question 5], what conclusions can be drawn from the existing data and analyses regarding the penguin response, and what additional analyses can help in determining the penguin response?*
N/A.

8. If yes [to question 1], is it advisable to proceed to a full-scale experiment?
The Panel **concludes** that sufficient data are available to conduct a scientifically appropriate power analysis to evaluate how long an experiment might have to be conducted for. The Panel notes that while a full-scale experiment is likely to lead to a conclusive outcome in the medium term, the design of such an experiment would need to be fully evaluated using an appropriate power analysis to determine the level of impact that has high probability of being detected within a specific timeframe. Issues that need to be considered when conducting a power analysis, and ultimately an experiment include:

1. a range of effect sizes based on the impact on the dynamics (recovery) of the penguin population;
2. the observed residual patterns (some of the fits in MARAM/IWS/DEC14/Peng/B16 exhibit substantial serial correlation which needs to be reflected in the power calculations);
3. the foraging range of African penguins and the potential overlaps of that foraging range with fishing effort;
4. use of the specific location (latitude/longitude) of catches in analyses of impact, rather than 10 n.mile catch reporting blocks;
5. the methods, collection and recording of appropriate response variables and covariates be specified to reduce conflict with post-experiment analyses; and
6. consideration be given to how the experiment be designed such that data from the feasibility study can usefully be incorporated into any final analyses.

In relation to the initial Terms of Reference:

- *What scientific conclusions can be drawn from the present island closure and related analyses regarding the benefits or otherwise of closures to penguin recovery?*
  There is sufficient information to evaluate the feasibility of an experimental study to elucidate effects. The Panel **concludes** that detection of effects is feasible, and suspects they may be directional. However, available data do not allow firm conclusions regarding the effects on penguin populations to be drawn at present. The Panel **encourages** both groups to more broadly consider the literature from related situations to better constrain the biological plausibility of analyses and interpretation of results.
- *What further analyses of available information could help resolve uncertainties about the conclusions?*
  See response to new Terms of Reference 6 and 8 above.
- *What would be the benefits to formulating scientific advice of continuing the closures taking into account limited further resources that DAFF could commit to further monitoring and analysis?*
  Conducting an experiment will allow DAFF to make informed decisions regarding the impact of fishing near islands on penguin populations even given that there are multiple conflicting objectives. The Panel does not presume to develop future experimental designs, cognizant of the limited resources, but would prioritize development of a suitable design and a suitable process to both inform and make decisions related to this issue. Relevant research recommendations, prioritized taking

account of whether they are critical for decision making immediately and to some extent cost, are given below.

Other Considerations

The Panel's deliberations concern only the scientific aspects of the impacts of catches/closures. The Panel recognizes that the ultimate decision on whether to continue or terminate the island closures has a large policy component, including aspects related to the costs of the closures to the industry, and the benefits of closures to penguin populations. The Panel did not attempt to weigh these societal costs and benefits, and instead focused on whether the experiment could provide sufficient information for others to make informed decisions given these inherent trade-offs.

Nevertheless, although the Panel is cognizant of not proceeding beyond scientific merits of the work conducted, it also understands the need to directly address the trade-offs across differing mandates. The Panel understands there are mandates to sustainably manage fish and conserve threatened species. In that context, the Panel notes that decision making could be improved with the following additional scientific analyses:

- an economic evaluation of the catch of sardines and anchovies relative to the closures, and of economics of tourism to the islands.
- development of a multi-criteria decision analysis (MCDA; e.g. Linkov and Moberg 2012) support tool. The MCDA approach, or similar forms of risk analysis, would provide benefits beyond the purely statistical/experimental approach adopted to date, would certainly use the information from such experiments, and importantly would establish a transparent process by which decisions could be jointly explored.
- development of additional, data-driven, approaches (besides the largely modelling-emphasized efforts to date) that better: elucidate penguin population responses to island closures; explore responses to other, multiple threats impacting penguins; and examine a range of possible management interventions.

Research recommendations

A.1 (H) MARAM/IWS/DEC14/Peng/A10 provides an analysis indicating that the application of fixed effects GLM-type models (such as those in MARAM/IWS/DEC14/Peng/B12) to a system in which both local biomass and catch can impact penguin populations will lead to biased outcomes. The Panel notes that the model on which MARAM/IWS/DEC14/Peng/A10 is based does not match exactly the error structure on which the analyses of MARAM/IWS/DEC14/Peng/B12 are based. In addition, some of the analyses in MARAM/IWS/DEC14/Peng/B12 are based on a random effects and not a fixed effects structure. It might be possible to evaluate potential biases for models such as those in MARAM/IWS/DEC14/Peng/B12 using analytical methods. However, a simpler way to examine this issue would be through simulations; and the Panel recommends that simulations to evaluate bias in estimation methods be explored, which are conditional of the types of scenarios reflected in MARAM/IWS/DEC14/Peng/A10

A.2 (H) Various elaborate models have been applied to test the effect of fishing on penguin demographics. More elementary analyses directly aimed at evaluating the questions and statistical power should be applied because with this sole focus on highly parameterized statistical and analytical methods, the proverbial "missing the forest because of the trees" remains an important risk. Consider ANOVA-type models such as those provided to the Panel but make the fishing/closure effect parameter independent of island. More details (factor levels and new explanatory variables) may well be needed, but this is the type of model that the Panel would have liked to have seen as the initial test of the effect of fishing /

closure. Notably there is no issue with multiple testing since the rudimentary model only contains a single parameter describing the effect of the closure on the demographic parameter.

This simple ANOVA/ANCOVA model can then be expanded in a number of directions, towards MANOVA/MANCOVA models or other, more detailed, response-pressure multivariate models for a single response. Finally, simple summarizations of the results would be beneficial to avoid obfuscating main results.

A.3 (H) Explore whether existing parameter estimates can be used to determine an upper limit on the rate of rebuilding that is likely to result from island closures and should be of value in making a decision on whether to proceed to an experiment. Some initial analyses along these lines are available in Appendix B of MARAM/IWS/DEC14/Peng/B4.

A.4 (H) Develop and implement a comprehensive research program that aims to identify the core reasons for the reduction in penguin population numbers, and identify any potential mitigation measures.

A.5 (M) Consider the suggestions noted in the “Other Considerations” section above to better facilitate decision making.

A.6 (M) As future data are collected, models (such as GAMs or GAMMs) which allow for non-linear relationships between penguin biological parameters and covariates should be applied. Furthermore, using a similar parsimonious approach via non-parametric models might also elucidate key responses without the plethora of assumptions needed.

A.7 (M) Develop improved methods for obtaining precise estimates of the local biomass of small pelagic species. This may not be practical given limited resources.

A.8 (L) Explore, using MSE, a range of strategies for managing the impact of pelagic fishing near islands on which declining penguin populations are found. This could involve assessing the biomass near each island relative to penguin foraging needs (*sensu* thresholds noted in the literature) and reducing or eliminating fishing, depending on whether the needs of penguins will be satisfied. Implementation of this recommendation is likely to require the availability of methods to estimate local at-sea density of penguins near islands with much higher precision than is possible now.

Hake

Stock structure

B.1 (*) Genetic data provide evidence related to what might be termed “breeding stocks” – that is, these data provide insights into the degree of reproductive isolation over space and time. In general, neutral genetic markers will not be able to evaluate another class of scenario that might be of interest to management: breeding occurs randomly in a single location, but portions of the progeny either passively drift or actively migrate into two or more geographic areas, where they then grow and mature. If the areas differ in environmental features, this could produce differences in growth rate, parasite load, age at maturity, or other phenotypic characteristics that might be interpreted as evidence for multiple stocks. Such scenarios should be considered when evaluating management options, even though they would not include more than a single breeding population.

B.2 (*) The stock structure hypotheses that assumed a hard boundary between the stocks of *M. capensis* and *M. paradoxus* (C2d and P2b) within Namibia should be assigned low priority for implementation. This is because the basis for the hypotheses (Reimer, 1993) was largely qualitative – retrieving the original data and conducting appropriate statistical analysis would be necessary before these data could be used to derive stock structure hypotheses. The assumption of a northern stock of *M. paradoxus* off Namibia that is separate from a southern stock, which underlies hypothesis P2b, is problematic given the perceived lack of spawning of this species in Namibia. Further, the assumption of a hard boundary between two stocks of *M. capensis* is inconsistent with the mixing that can be inferred from the genetics studies.

B.3 (*) The discrepancy between the mtDNA results for 2005 and for 2012-13 for *M. paradoxus* has not been resolved. The hypothesis that the different results were due to different sampling locations in the earlier and later time periods is not consistent with the assumption of a single panmictic population. Under those conditions, it should not matter from where the samples are taken, as they all should be derived from the same random-mating population. However, such results can occur if, for example, animals from the same family or cohort are found and sampled together. This can lead to a "chaotic" pattern of statistically significant results that do not provide consistent results over time (Planes and Lenfant 2002; Iacchei *et al.* 2013).

B.4(*) Figure 2 uses results of the power analyses conducted by Henriques *et al*. (unpublished data) to depict levels of stock differentiation of *M. paradoxus* that can and cannot be excluded based on available genetic data. If cryptic stock structure (more than a single stock) exists, it must be characterized by combinations of migration rate (*m*) and effective population size (*Ne*) that produce *Fst* values too small to detect with available data. As illustrated in Figure 2, how confident one can be that the power analyses of genetic data can rule out more than one population of *M paradoxus* depends on how large *Ne* is. Two general approaches can be used to estimate recent or contemporary *Ne*. First, one can use genetic methods (reviewed by Luikart *et al*. 2010) to estimate *Ne* based on any of several indices. However, because all of these indices are sensitive to a signal proportional to 1/*Ne*, they are most effectively used to study relatively small populations, particularly those of conservation concern. Once *Ne* reaches 10^3 or 10^4, 1/*Ne* is so small that it becomes difficult to distinguish between values that are "large," "very large," and "very very large" (Waples and Do 2010). Therefore, although genetic methods have considerable power to detect relatively small *Ne*, they have difficulty distinguishing between values in the range 10^3 – 10^4 and higher. Second, one can use the ratio *Ne/N* estimated for other species, together with a species-specific estimate of *N* (number of mature adults), to derive an estimate of *Ne*. However, published estimates of the ratio *Ne/N* in marine species span such a wide range (about 10^{-1} to 10^{-6}; see review by Hauser and Carvalho 2008) that considerable uncertainty would still remain after application of this approach.

B.5 (*) Further review of the data on which the GeoPop model results used to develop stock structure hypotheses are based suggests that the boundary at the Olifants River for *M. paradoxus* does not appear in the raw survey data for South Africa nor the model output. This strongly suggests that stock structure hypotheses for *M. paradoxus* based on GeoPop output should not be taken forward at present, and any future attempt to develop stock structure hypotheses using GeoPop should utilize plots of the raw data, as well as the predicted distribution of abundance by age-class.

B.6 (*) The following stock structure hypotheses should form the basis for initial stages of future modelling work:

- *M. paradoxus*. The most likely hypothesis is that there is a single stock off Namibia and South Africa (hypothesis P1).
- *M. capensis*. The most likely hypothesis based on the genetics data is that there are two stocks off Namibia and South Africa (hypothesis C2c). There are northern and southern stocks and an area of mixing.

The stock structure hypotheses that should be included the second stage of future modelling work are:

- *M. paradoxus*. One or more multi-stock hypotheses based on the results of the GeoPop analyses (see recommendation B.9). The group developing stock hypotheses based on GeoPop should consider that one interpretation of the genetics data is a single breeding stock with sub-stocks that have different migration patterns.
- *M. capensis*. One or more multi-stock hypotheses based on the results of the GeoPop analyses (see recommendation B.9) and a three-stock hypothesis if the analyses of genetics data supports such a hypothesis.

B.7 (H) Genetic data for *M. capensis* provide convincing evidence for more than one stock. Results from the STRUCTURE analysis conducted by Henriques *et al.* (unpublished data) are most compatible with two stocks, which have an area of mixing that varies across years within the approximate range (28-33 lat). Fitting the STRUCTURE results to a scenario involving three populations does not provide convincing evidence for the existence of another stock. However, because it is well known that clustering programs such as STRUCTURE respond primarily to the strongest genetic signal and hence might miss a weaker signal of less-well differentiated stocks, two additional analyses should be conducted to further explore the potential for more than two stocks of *M. capensis*.

- Using the data for all three years and a STRUCTURE run with k = 2, select those individuals assigned with high confidence to the “northern” stock. Run STRUCTURE with k set to 2 and evaluate evidence for an additional stock that is genetically similar to the “northern” stock. Repeat the above using just the individuals assigned with high confidence to the “southern” stock. Removing the main signal of north *vs* south might reveal cryptic structure within either of those “stocks”.
- Using all the data for all three years, conduct a principal components analysis as described in Patterson *et al*. (2006), which provides a formal statistical test of the number of gene pools in a mixed sample.

B.8 (H) Use the DNA already extracted from the 2005 samples to generate microsatellite data comparable to those that are available for the 2012-2014 samples. This will provide a multi-generation perspective on the stability of the mixing pattern of putative stocks seen in the more recent samples.

B.9 (H) Establish a group under the auspices of ECOFISH to review the GeoPop model in greater depth. This review should consider whether the model output is consistent with the raw survey data (and any commercial catch-rate data, even if such data cannot be included in the GeoPop analyses). It should then use the model output to identify stock structure hypotheses, including potential migration routes for putative stocks.

B.10 (H) Although microsatellite data for *M. paradoxus* do not provide evidence for more than a single stock, results of a “factorial component” analyses conducted by Henriques *et al.* (unpublished data) show a few individuals as substantial outliers. Is it possible that those are mis-identified *M. capensis*? Repeating the same analyses with simulated data for a single, random-mating population would provide a useful context for interpreting this result.

B.11 (M) Efforts should be made to determine whether historical collections of scales or otoliths exist. Several published examples exist where scientists have been able to extract DNA from samples up to a century old, and this can provide a very valuable temporal dimension to information related to stock structure.

B.12 (M) Develop next-generation DNA sequencing markers to further evaluate evidence for stock structure in *M. capensis* and *M. paradoxus*. With 10^3 to 10^5 new markers, it should be possible to considerably increase power to detect weak population structure. Furthermore, it might be possible to identify markers in or closely linked to genes under selection, which would provide information about the extent of local adaptation.

B.13 (L) The ability to develop and compare stock structure hypothesis would be enhanced if spatial data on age and growth and maturity curves were available. In addition, the use of parasite studies and the application of otolith microchemistry approaches, morphometric analyses and meristic methods have the potential to inform the selection and weighting of stock structure hypotheses. However, the data required are currently not available. There is a consequently a need to evaluate (and implement) the sampling schemes that could inform future discussions regarding stock structure.

Population dynamics modelling

B.14 (*) The proposed schedule for model development in MARAM/IWS/DEC14/Hake/P10 is an appropriate way to move forward. However, taking account of mixing in a "fleets as areas" model will be difficult. Consequently, it may be best to move from Stage 1 directly to Stage 3.

B.15 (*) The fit for case 1 of the multispecies hake model that include cannibalism / inter-species predation is better than that for case 3, in particular in terms of the fit to the early west coast ICSEAF and more recent *M. paradoxus* GLM CPUE series. However, case 3 is better (though not perfectly) able to mimic the diet data. At this stage of the development process, do not exclude cases unless there is clear visual evidence for marked model mis-specification as additional data / model assumptions may change the residual patterns.

B.16 (H) The multispecies hake model that includes cannibalism / inter-species predation should be expanded in the following steps:

- Apply the model ignoring the spatial availability matrix to assess whether this feature of the model is needed to allow the model to mimic the observed diet compositions by age.
- Consider alternate formulations of stock-recruit models for hake that incorporate cannibalism, both directly as a covariate and indirectly in how spawning stock biomass is defined (e.g., Link *et al*., 2012).
- Explicitly account for spatial structure, either using a movement model or by treating predation on the west and south coasts as separate 'fleets' (base initial analyses on diet data for the west coast only).
- Disaggregate the model by sex to better fit, for example, the longline catch-at-age data. It should be possible to disaggregate the diet data by predator sex but not by prey sex.
- Include other predators (re-evaluate sources of hake mortality to identify which predators to add to the model).

B.17 (H) Plan, and then implement, a review of the sampling strategy for diet data given the results of the current model as well as other needs for diet data.

B.18 (H) Development of a spatial model for hake requires catch data split to species, and ideally to depth strata. Such data already exist for South Africa but were not available for Namibia. A group of scientists, and others with appropriate expertise, should be convened under the auspices of BCC to develop catch series by the spatial strata in the spatial models.

B.19 (H) The current fit of the hake explicit movement model (MARAM/IWS/DEC14/Hake/P5) to the survey data by stratum is very poor. An evaluation of whether reasonable fits to these data are possible in principle should be conducted by (a) fitting a fleets-as-areas model to data disaggregated to the same extent as are used in the spatial model, and (b) increasing the weight applied to the fits to the survey data in the spatial model.

B.20 (H) Some of the predicted movement directions in the movement model appear to be biologically unrealistic. A group of scientists should be established to provide guidelines for the appropriate qualitative structure of the movement matrices, such as that movement is towards deeper water with age.

B.21 (H) Identify the spatial strata that need to be implemented for the full set of movement models. This information is needed to allow the data used for fitting purposes (survey, catch, composition) to be extracted for use in these models.

B.22 (H) Explore alternate approaches to estimating or evaluating daily ration. These include C/B ratios, Q/B ratios, stomach/body mass data, caudal fin ratio, Essington's growth check method, evacuation rate approaches, and functional response models (NEFSC, 2007, 2011; Essington *et al.*, 2004).

B.23 (M) Consider alternative models for how predation rates change with predator age including the Essington *et al.* (2004) model and the suitability functions developed by ICES (e.g. Daan and Sissenwine, 1991; Magnusson, 1995; ICES 2010; 2011).

B.24 (M) Examine the role and range of "other prey" more thoroughly, perhaps as in a sensitivity analysis to better reconcile how much food hake need to eat with what specific prey they eat.

B.25 (M) Explore data-poor methods for estimating diet composition in the absence of substantial amounts of stomach contents data. Approaches (e.g. Link, 2004) exist that can be used to predict diet that are robust. This may help to interpolate diet estimates to those locations, times, etc. that have limited stomach sampling.

B.26 (M) Explore and contextually inform the topic of hake cannibalism using a literature survey, particularly taxa-oriented reviews and the multi-species modeling work conducted by ICES (e.g. Daan and Sissenwine, 1991; Magnusson, 1995; ICES 2010; 2011).

B.27 (L) Examine the residual patterns for the fits to the diet data by year and consider the inclusion of random effects if these patterns are indicative of major model mis-specification.

Sardine

C.1 (*) The alternative models (see MARAM/IWS/DEC14/Sardine/P2) should be implemented in the following priority order:

- Alternative D (Varied Adult Movement). Fits to the parasite data should be used to evaluate the plausibility of the sub-variants of this model variant.

- Alternative C (Varied Adult Distribution). Fits to the parasite data should be used to evaluate the plausibility of the sub-variants of this model variant.
- Alternative A (Effective Spawning Biomasses). An additional sub-variant in which the proportion of south SSB off the west coast exceeds 20% should be considered. Whether this sub-model is taken forward depends on whether auxiliary information suggests that it is plausible. See also recommendation C.4.
- Alternative B (Varied Recruit Distributions).

Alternative B is assigned lowest priority because the extent to which south stock recruits are assumed to be distributed west of Cape Infanta at the time of recruit survey is essentially arbitrary. Alternative A is given next lowest priority because it depends on the results of the linked biological-hydrodynamic model (MARAM/IWS/DEC14/Sardine/BG8), but those results are likely subject to considerable uncertainty owing to an inability to model the response of early life stages to environmental conditions.

C.2 (H) The parasite data have the potential to evaluate the relative plausibility of the alternatives listed in MARAM/IWS/DEC14/Sardine/P2. These data should be included routinely in the likelihood maximized when fitting the model. The approach outlined in MARAM/IWS/DEC14/Sardine/P3 will need to be extended to allow predictions of prevalence and abundance of the parasite to be made. In particular, the prevalence of the parasite by age and over time needs to be tracked for the "South from west stock" animals. More complex models will be needed to implement alternative 'C' variants. For example, it will be necessary to keep track of 'South visited West" animals.

C.3 (H) Initial focus for model development should relate to including prevalence of the parasite, but the abundance (intensity) of the parasite may provide additional information and should be considered as a qualitative reality check if possible.

C.4 (H) Alternative B should be implemented accounting for the implications of south stock recruits being found throughout the west coast, and hence being subject to the fisheries there, and possibly being infected by the parasite.

C.5 (H) The alternative models should be evaluated in terms of their ability to fit the available data and to mimic more generic expectations for the dynamics of sardine populations. In particular, individual future projected trajectories of population size should exhibit the boom-and-bust dynamics typical of small pelagic fishes.

C.6 (H) An attempt should be made to estimate the recruitment from winter spawning on the south coast. The two approaches in MARAM/IWS/DEC14/Sardine/P3 should be explored.

C.7 (H) The selectivity of the trawls used during the November survey is estimated to be domed-shaped. This was unexpected given the nature of an acoustic survey which should survey the entire biomass (uniform over all ages), and suggests that the trawls provide a biased view of the length-structure of the population. However, this will be inconsequential for assessment purposes if the trend in abundance is not sensitive to selectivity. The following ways should be explored to determine the extent to which this selectivity could impact the results of sardine assessments: (a) compute length-frequencies based on biomass-weighted length-frequencies stratified by survey stratum and broad size group, and (b) compute biomass estimates using length-frequencies where the abundance of the length-classes assumed not to be fully selected are scaled up by the estimates of selectivity by length-class in Figure 6 of MARAM/IWS/DEC14/Sardine/BG6.

C.8 (M) Show “worm plots” of movement probabilities to evaluate whether the behaviour of, in particular, the MoveAutoC option results in realistic behaviour.

C.9 (M) Age-0 animals should be included in the model predictions of the biomass available to the November survey.

C.10 (M) If computationally feasible, full account should be taken of the uncertainty associated with the parameters of the stock-recruitment relationship if this relationship is fitted outside of the process of fitting the population dynamics model.

C.11 (M) Consider an alternative model in which the effective spawning biomass is the spawning biomass off the south coast, under the assumption that all of the west coast recruitment originates from the south coast, as a sensitivity test.

C.12 (M) Further examine whether an environmental variable exists that links environmental variables and recruitment / movement in the operating model (not necessarily the candidate OMPs). Although evaluations of the potential for environmental variables to explain movement and recruitment have been conducted in the past, substantially more data appear to be now available.

C.13 (L) The Panel continues to support its earlier recommendation that an alternative model be developed that assumes that recruitment is dependent on total west+south biomass and in which the proportion recruiting to the south and west is estimated.

References

Bentley, N., Kendrick, T.H. and D.J. MacGibbon. 2014. Fishery characterisation and catch-per-unit-effort analyses for sea perch (*Helicolenus* spp.) in New Zealand, 1989–90 to 2009–10. New Zealand Fisheries Assessment Report 2014/27.

Daan, N. and M.P. Sissenwine eds. 1991. Multispecies models relevant to management of living resources. *ICES Marine Science Symposium*, **193**

Daunt, F., Wanless, S., Greenstreet, S. P. R., Jensen, H., Hamer, K.C. and M.P. Harris. 2008. The impact of the sandeel fishery closure on seabird food consumption, distribution and productivity in the northwestern North Sea. *Canadian Journal of Fisheries and Aquatic Sciences* 65: 362–381.

Essington, T.E. Kitchell, J.F. and C.J. Walters 2001. The von Bertalanffy growth function, bioenergetics, and the consumption rates of fish. *Canadian Journal of Aquatic and Fishery* 58**:** 2129–2138.

Furness, R.W. and M.L. Tasker. 2000. Seabird –fishery interactions: quantifying the sensitivity of seabirds to reductions in sandeel abundance, and identification of key areas for sensitive seabirds in the North Sea. *Marine Ecology Progress Series* 202: 253–264.

Frederiksen M., Jensen H., Daunt F., Mavor R.A., and S. Wanless. 2008. Differential effects of a local industrial sand lance fishery on seabird breeding performance. *Ecological Applications* 18:701-710.

ICES. 2010. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 4–8 October 2010, San Sebastian, Spain. ICES CM 2010/SSGSUE:05. 95 pp

ICES. 2011. Report of the Working Group on Multispecies Assessment Methods (WGSAM), 10-14 October 2011, Woods Hole, USA. ICES CM 2011/SSGSUE:10. 229 pp

ICES. 2014. Report of the Herring Assessment Working Group for the area south of 62^{o}N. ICES C.M. 2014/ACOM:06, 1406 p.

Hauser, L., and G.R. Carvalho. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* 9**:** 333–362.

Henriques, R., Matthee, C. and S. von der Heyden. This meeting. Assessment of the genetic population structure of the Cape Hakes (*Merluccius capensis* and *M. paradoxus*) around southern Africa. Draft document for BCC.

Iacchei, M., Ben-Horin, T., Selkoe, K.A., Bird, C E., García-Rodríguez, F.J. and R.J. Toonen. 2013. Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. *Molecular Ecology* 22: 3476–3494.

Kristensen, K 2014. TMB: General random effect model builder tool inspired by ADMB. R package version 1.1.

Link, J.S. 2004. A general model of selectivity for fish feeding: a rank proportion algorithm. *Transactions of the American Fisheries Society* 133: 655-673.

Link, J.S., Lucey, S.M. and J.H. Melgey. 2012. Examining cannibalism in relation to recruitment of silver hake *Merluccius bilinearis* in the U.S. northwest Atlantic. *Fisheries Research* 114: 31-41.

Linkov, I. and E. Moberg. 2012. Multi-Criteria Decision Analysis: Environmental Applications and Case Studies. CRC Press. Boca Raton, FL.

Luikart, G., Ryman, N., Tallmon, D.A., Schwartz, M.K. and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA–based approaches. *Conservation Genetics* 11**:** 355–373.

Magnusson, K.G. 1995. An overview of the multispecies VPA and applications. *Reviews in Fish Biology and Fisheries* 5: 195-212.

NEFSC 2007. Assessment Report (44th SARC/SAW). Section B.8. [TOR 6]. Northeast Fisheries Science Center Reference Document, 07-10. pp 332-344, 504-547.

NEFSC. 2011. 51st Northeast Regional Stock Assessment Workshop (51st SAW) Assessment Report. Northeast Fisheries Science Center Reference Document 11-02; 856 p.

Patterson, N, Price, A.L. and D. Reich. 2006. Population Structure and Eigenanalysis. *PloS Genetics* 2:e190

Planes, S. and P. Lenfant. 2002. Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Molecular Ecology* 11: 1515-1524.

Reimer, L.W. 1993. Parasites of *Merluccius capensis* and *M. paradoxus* from the coast of Namibia. *Applied Parasitology* 34: 143-150.

Waples, R.S., and C. Do. 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications* 3: 244-262.

Weller, F., Cecchini, L.-A., Shannon, L., Sherley, R.B., Crawford, R.J.M., Altwegg, R., Scott, L., Stewart, T. and A. Jarre. 2014. A system dynamics approach to modelling multiple drivers of the African penguin population on Robben Island, South Africa. *Ecological Modelling* 277: 38–56.

Witherell, D., Pautzke, C. and D. Fluharty. 2000. An ecosystem-based approach for Alaska groundfish fisheries. *ICES Journal of Marine Science* 57: 771–777.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Table 1. Penguin response to closure from consideration of data from the period from 2008 onwards by Group A, and to reduction of catches from consideration of data for all years from Group B. Results shown are self-reported by the two groups. Symbols refer to the effect on penguins, rather than the effect directly on the trait measured.

	Group A				**Group B**			
	Dassen	Robben	Bird	St Croix	Dassen	Robben	Bird	St Croix
Chick condition	-	+*	+	-*	-	0		
Chick growth	+	+	+	-	-	-*		
trip duration	-*	+	+*	+	-*	-	0	+
Foraging path length	-	+	+	+*	-	-	-	+*
Maximum foraging distance	+	+	+*	+				
Active nest proportion					-*	-*		
Fledging success		+*			-	+*		

0 = none or indeterminate; - = negative; + = positive; -* = significantly negative; +* = significantly positive

Table 2. A possible implementation of a full-scale experiment. The actual number of years needed will depend on the final outcome of a power analysis. Note that the St. Croix-Bird Island pair will be a complete design.

	Dassen	Robben	St Croix	Bird
2008	X			
2009	X		X	
2010			X	
2011		X	X	
2012		X		X
2013		X		X
2014	X			X
2015	X		X	
2016	X		X	
2017		X	X	
2018		X		X
2019		X		X
2020	X			X
2021	X		X	
2022	X		X	
2023		X	X	
2024		X		X
2025		X		X
2026	X			X
2027	X		X	
2028	X		X	
2029		X	X	
2030		X		X
2031		X		X
2032	X			X
Years(T)	25	25	25	25
Closures	12	12	12	12
(T from 2009)			24	24

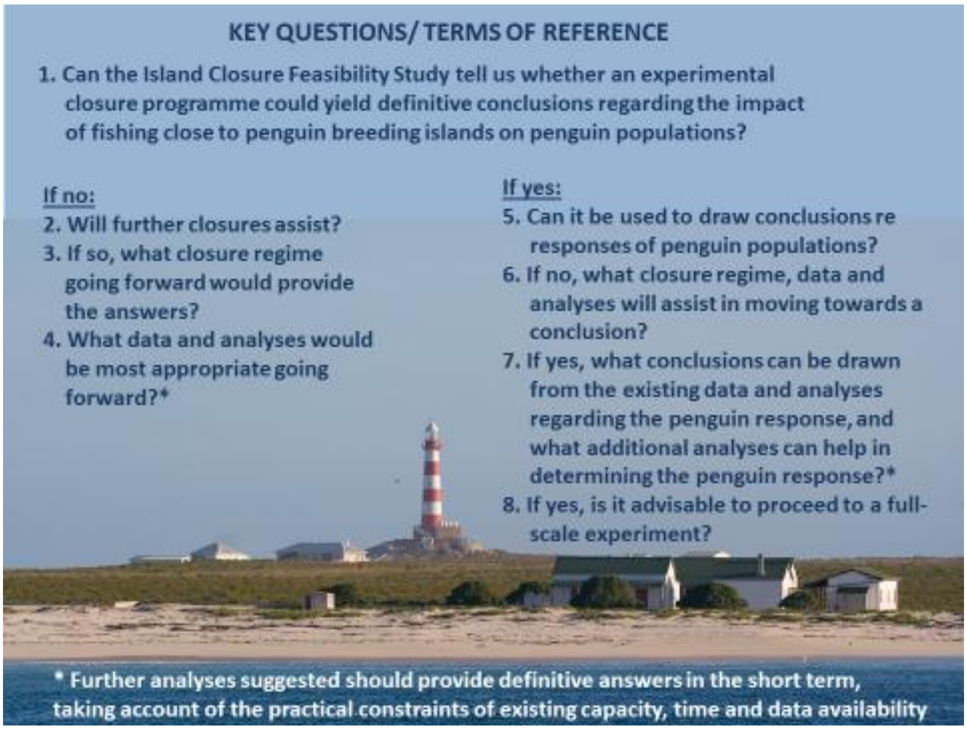

Figure 1. The final Terms of Reference for the penguin component of the review.

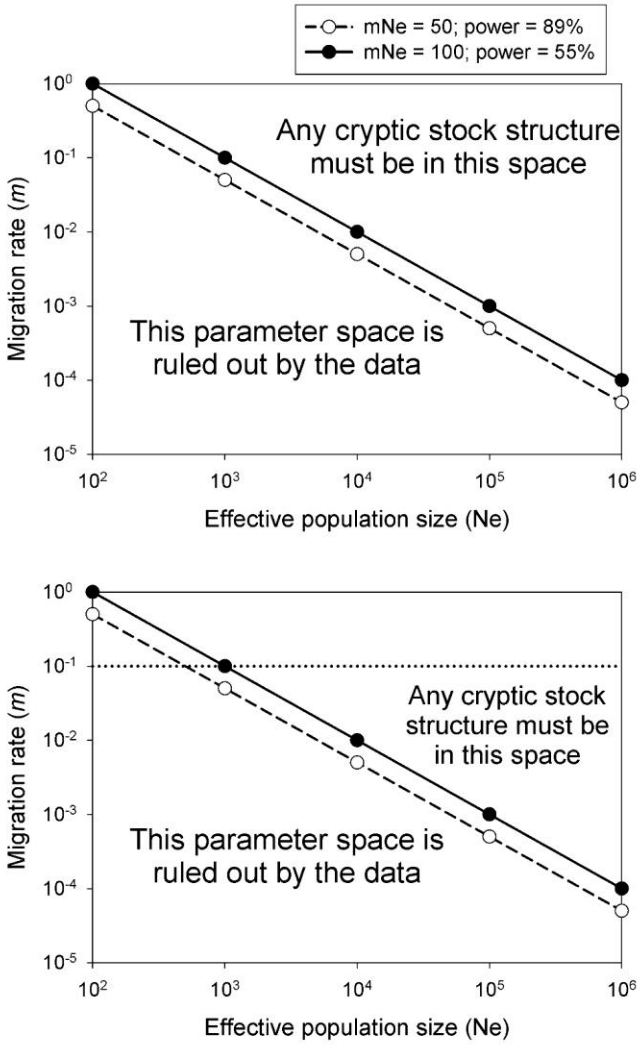

Figure 2. Graphical depiction of the parameter space that is and is not compatible with more than one stock of *M. paradoxus*. The lines plot combinations of effective population size (*Ne*) and migration rate (*m*) that are expected to produce *Fst* values of 0.005 (*mNe* = 50) and 0.0025 (*mNe* = 100). The values from the power analysis come from Henriques *et al.* (unpublished data). The parameter space below the lines can be ruled out as implausible with specified probabilities based on genetic data. The bottom figure shows that the parameter space consistent with multiple stocks is further constrained if one assumes that separate stocks must exchange migrants at a rate below a certain threshold (in this case *m* = 0.1 = 10% per generation). Two caveats about the above relationships between *m*, *Ne*, and power: 1) They are based on a widely-used but somewhat simplistic relationship between *mNe* and *Fst* [$E(Fst) \approx 1/(1+4mNe)$] developed by Wright (1931). The relationships shown above are probably qualitatively robust but caution should be used in quantitative applications. 2) Wright's relationship assumes that an equilibrium has been reached between the homogenizing effects of migration (*m*) and divergence due to genetic drift (indexed by *Ne*). Under an alternative scenario, *Fst* can be modeled as a value that increases over time in a system in which populations are completely isolated. A comparable figure could be developed based on the relationship $E(Fst) \approx t/(2Ne)$, where *t* is elapsed time in generations since the populations diverged. For example, *Fst* = 0.01, which produced 100% power according to Henriques *et al.* (unpublished data) , could be achieved if 2 populations of size *Ne* = 1000 each were isolated for 20 generations, if two populations of size 100,000 were isolated for 2000 generations, or any other combination of *t* and *Ne* that satisfied the above relationship. This means that very large populations might have to be isolated for large numbers of generations before a detectable signal of genetic differentiation develops.