Responses to Matters Raised in Documents MARAM/IWS/DEC14/Peng/A1 and Peng/A2

D. S. Butterworth and W. M. L. Robinson

Marine Assessment and Resource Management Group, University of Cape Town

Summary

Responses are provided to a large number of the comments and assertions made in documents MARAM/IWS/DEC14/Peng/A1 and Peng/A2, primarily insofar as they relate to earlier analyses by the MARAM Group. Amongst the more important of these responses are the provision of a number of reasons why standard sample size weighting may not be appropriate for the types of analyses reported in MARAM/IWS/DEC14/Peng/B4 of the data from the island closure feasibility study (though this is being examined further for subsequent report to the Panel), and why the use of simpler models using fewer data may be preferable for such analyses. Peng/A2 is considered to fail to provide a clear explanation of the rationale for its assertion that weak catch *vs* biomass correlations at the local level can lead to spurious indications from GLM analyses that the impact of fishing on penguins is positive; at the very least, this assertion needs to be confirmed by an appropriate simulation study before it might merit further consideration. The advocacy in Peng/A2 for closure to replace catch as a covariate in the feasibility study analyses is questioned, particularly in the light of the associated implication that annual catches of nearly (and in one case exactly) zero have the same potential expected impact on penguins as the much larger catches that have been made in some other years.

Background

This document details those responses to comments made in documents MARAM/IWS/DEC14/Peng/A1 and Peng/A2, primarily insofar as they relate to analyses by the MARAM Group to the penguin issues under consideration by the International Review Panel. For the reader's ease of reference, the material to which a specific response is offered is reproduced, with the quote in question shown in standard red when from Peng/A1, and in *red italics* when from Peng/A2, with the key words on which the response is focused indicated by the use of yellow highlighting. Details of references in those quotations are listed only if necessary.

Data usage in MARAM analyses

Currently there is some uncertainty around the modelling approach from MARAM (hereafter the MARAM models) which casts some doubt around the reported outcomes from those models.

First, the models do not use the best available data. Instead they use unweighted, annual means for many response variables, which in many cases exclude any variance or error estimates. *If individual observations are not available, this relationship can be approximated by modelling the means, provided that the analysis is weighted by the inverse of the standard errors of the group means (square root of group size over group variance) or some other weighting that accounts for the different amount of information contained in individual means. In the case of the penguin data sets, the individual observations were available but not requested for use in Robinson (2013). A mean that is based on 100 observations should contribute more in an analysis than one based on a single observation. For example at Dassen Island during the closed years, 2008 and 2009, there are only 13 foraging track observations, while during open years there are 158. In the Robinson (2013) GLMs, the 13 observations have the same weight as the 158, because unweighted means were modelled.*

Data availability

At a meeting held on 30 July 2014, at which a number of the authors of Peng/A1 and Peng/A2 were either present or circulated with the meeting Aide Memoire in its adoption process, the data types and series to be used in the analyses for presentation to the Panel were agreed. For the penguin response variables for each island, these were the unweighted means. This reconfirmed decisions made explicitly or implicitly at earlier meetings on 9 November 2012, 20 February 2013 and 15 October 2013, that these would form the basis used for the analyses of these response variables to inform management decisions.

There is no record in the Aide Memoirs of those meetings of any suggestion that weighting needed to be taken into account, nor was any mention made of this during the 30 July 2014 meeting when data types to be used in this series of analyses were agreed (such intent would have required proposals made then to include further information than the unweighted means). Furthermore that meeting was informed that further data not already in the public domain could be made available only through use of the Public Access to Information Act (PAIA). This is a time-intensive process, and such further unweighted means that were obtained through this process became available too late to be analysed before the 27 October deadline for the first round of submissions to the Panel.

We recall this issue being raised only once in previous meetings (though perhaps our memories are not 100%), when we responded with the argument given under the third bullet below. Since there was no further response at that time, we understood this matter to have been "asked and answered", and so not one seen to require further attention.

Why this approach was followed

The points made above should not be seen as suggesting that there were not in any case sound reasons for preferring the form of analyses of the unweighted means undertaken anyway.

- This was to maintain consistency with the analysis approach discussed at the 2010 International Stock Assessment Review Workshop and endorsed by the International Panel then present.
- The method can be applied in a consistent way to the different response variables available.
- Frequently for marine data such as these, process error swamps sampling error for all but the smallest sample sizes, so that downweighting of data is inappropriate except for sample sizes which are very small. Thus for example, in an analysis of South African sardine length frequency data, de Moor and Butterworth (2009) found that downweighting for low sample size should apply only for samples sizes as small as 40 or less, with weighting independent of sample size at larger sizes.
- "More" does not necessarily result in "better". Taking individual observations into account requires the development of a possibly complex model to account for covariance structures, which then necessitates the estimation of further parameter values from the data. There is no guarantee that this will necessarily produce more reliable and acceptable results which manifest less bias and/or greater precision. For example in the ICCAT assessment of western north Atlantic Bluefin tuna, there has been a discussion extending over 20 years about ways to improve on the current approach of giving different abundance index series different weights, with as yet still no progress beyond use of the unweighted default.
- Selection of the approach used needs to take account of the simulation study that would need to follow, based on the same estimator, to calculate experimental power. An estimator that uses many data with a complex covariance structure will require generation of future pseudo-data with that same structure. Given that this may lead to a complex analysis requiring considerable time to complete, practical considerations may dictate a preference for the use of simpler models provided that these are defensible.

Ultimately the determining test in such circumstances (although its application may be questionable given its requirements for further allocation of limited analytic resources) is to try both options to see which provides better estimation performance (see MARAM/IWS/DEC14/Peng/B12, which reports on such a comparison).

Further analyses

Sample size data for the response variables have been provided very recently only. The impact of taking this information into account in the GLM analysis approach used in documents MARAM/IWS/DEC14/Peng/B4 and Peng/B12 will be reported in a subsequent document.

In addition, using single data points for each year has the effect of creating over-parameterised models (because, for example, the degrees of freedom for a string of 14 numbers is only 13), with in many cases too few degrees of freedom to allow robust estimates of model error parameters. *The models contain too many predictors for the number of outcomes available, and therefore would have little power to detect significant results. A general rule of thumb is to have at least 10 observations for each parameter being estimated, because it is then possible to also assess the uncertainty in the estimates. The more observations per parameter, the more precise the estimates and associated uncertainty will be. In contrast, when a model is fit that estimates e.g. 9 parameters from 11 data points (see Table 1 in Robinson and Butterworth 2014a), there can be no confidence in any of the parameter estimates, nor in any of the conclusions drawn from such a model.*

This is unfortunate, because much richer, raw datasets exist and have been available for use, but were not requested by MARAM. Suggestions from penguin biologists to MARAM that these datasets should be considered were routinely ignored. *To avoid the concerns explained above, the raw data should be used in analyses, so that the variation within and between groups, including repeated measures on the same penguin or nest, can be used for valid inference. Use of the raw data also makes it possible to account for potential observation-level confounding factors. For example, environmental conditions or prey biomass levels can be included at the spatial and temporal scale of the observations.*

Our responses provided above deal with the comments repeated in the quotes immediately above about data availability, that we purportedly "routinely ignored" certain suggestions, and the use of models that can accommodate less-aggregated data inputs than single annual aggregates to reflect response variables.

The example given of one instance of a clearly over-parameterised model is essentially correct (in that the numbers given are wrong, but the criticism remains valid), but that quote omits to mention that this was immediately corrected in an Addendum to the paper referenced, with this Addendum having been tabled eight months ago.

If the models used indeed have "little power to detect significant results", it would seem surprising (Table A3 of MARAM/IWS/DEC14/Peng/B12) that the random effects variant of the approach detects results significant at the 5% level in 30% of cases.

But in any case, the primary purpose of applying these models was to provide a basis for a power analysis, which self-evidently is necessary only in a situation where insufficient data are available at present to achieve significant results in many cases. This matter has already been addressed under Item 14 of MARAM/IWS/DEC14/Peng/B10.

Nevertheless even given a large number of individually non-significant results, meta-analysis can still provide very important insights. Thus, for example, though relatively few stock–recruitment plots for fish populations show statistically significant evidence of a drop in recruitment at low spawning biomass, this effect becomes evident when the plots are considered jointly in a meta-analysis, as

demonstrated by Myers *et al.* (1999). Inferences we have drawn from the totality of the results from the GLM analyses (see MARAM/IWS/DEC14/Peng/B2, pg 3) are in the spirit of that approach, though do not pretend to be a formal meta-analysis for the reasons stated there.

Inclusion of pelagic catches in the set of predictors for GLMs

Linear regression is used to estimate the linear impact of a set of predictors on a response variable. Although causality is implied by the final estimates from linear regressions, causality cannot be revealed or confirmed by such analyses. This comment is relevant to the use of pelagic catches as predictors in linear regressions of various penguin responses. There is thus a need for a conceptual re-evaluation of the merit and meaning of the inclusion of pelagic catches in the set of predictors for GLMs which use metrics for penguin breeding success, survival and foraging as response variables.

Certainly causality cannot be confirmed by such analyses, but as these are often the only basis available on which to proceed, they frequently provide a valuable means to move forward to decisions in fisheries.

We note the following:
The correlations between catch and biomass at the macro level are positive, but not always large enough to require that model Figure 2 should replace Figure 1. Given the relatively optimistic correlation threshold of 0.7 used in Robinson (2013), the smaller scale of these correlations (< 0.7) rules out the need to address problems of collinearity in the regression for a type Figure 1 regression: this is the sole significance of the VIF statistic discussion in Robinson (2013).

We understand this to reflect agreement by the authors of the quotation that collinearity is agreed by them not to provide a basis sufficient to discount results of the GLMs reported in in documents MARAM/IWS/DEC14/Peng/B4 and Peng/B12.

The positive correlation between catch and biomass does not in itself indicate the direction of causality between catch and biomass. Logically, however, since an increase in catch cannot cause an increase in biomass, the direction of causality MUST BE in the direction of biomass to catch. Positive correlations between catch and biomass at the macro level have not been confirmed at the local island level, but it is likely that the correlation between catch and biomass at the local level is positive, although possibly weak (i.e. below the VIF threshold where collinearity would complicate interpretation of linear regression results, and invalidate model diagnostics).

The comment about direction is self-evidently correct. Document MARAM/IWS/DEC14/Peng/B9 confirms the supposition that the highlighted correlation is indeed positive but very weak.

In the absence of a direct link between catch and penguin response, the mechanism given in Figure 4 is the most likely to be in operation, but in which the effect from biomass to catch dominates any negative impact of catch on biomass (because of the positive catch to biomass correlation at the macro level which is assumed to exist at the local level, emphasis: positive but not necessarily strong, could be a weak correlation). In this situation, and under an open island condition, a reduction in catch can only be achieved by reducing the pelagic biomass, catch cannot, under fishing as usual conditions, be reduced independently of pelagic biomass when the island status is open. Application of a Figure 1 regression under circumstance where Figure 4 is in operation would produce the impression that catch is positively related to penguin response. It would also create an impression that it is possible to vary catch independently of pelagic biomass leading to the conclusion that fishing is good for penguin responses.

The highlighted statement is quite incorrect. For the same biomass, there are many other factors which can result in a different catch. For anchovy on the South African west coast, for example, processing

limitations (*inter alia*) limited the catch increase possible under the roughly three-fold increase in biomass across the turn of the century (see also other arguments in MARAM/IWS/DEC14/Peng/B9). Another limiting factor that comes into play from time to time is restriction on anchovy fishing to avoid exceeding the Precautionary Upper Catch Limit for the bycatch of juvenile horse mackerel in this fishery. Thus the inferences drawn the authors of this quote from the statement highlighted are also incorrect.

 If catch has a direct effect on penguin response, then Figure 5 is the most likely representation of the interaction, but in which the effect from biomass to catch dominates any negative impact of catch on biomass (because of the positive catch to biomass correlation at the macro level which is assumed to exist at the local level). However, application of a Figure 1 regression when Figure 5 is the operative mechanism will produce a similar outcome as when Figure 4 is operative. Thus the Figure 1 regression cannot give insight into which of Figure 4 or Figure 5 is operative.

 These comments have a bearing on the interpretation of the positive impact of fishing on penguin response variables reported in Robinson (2013). It seems likely that these are the result of a positive correlation between biomass and catch at a local level – it is not the strength of the correlations which is relevant here (i.t.o. VIF statistics etc.), but their average sign which is most likely predominantly positive. In view of this, estimating the impact of a reduction in fishing is only meaningful and causally possible if there is a concomitant reduction in prey biomass, and surely this will be negative for penguins. Since the causality is from biomass to catch, the proposition to reduce catches is in fact a proposition to reduce prey biomass.

Again document MARAM/IWS/DEC14/Peng/B9 has confirmed the highlighted suppositions that the highlighted correlations are indeed positive but very weak at both the macro and local level. But no clear explanation has been provided as to why the latter precludes identification of the impact of fishing on penguins (as indexed by some response variable). If, say, recruiting biomass is the same in two years, but catch differs for reasons such as those given above, why should this preclude identification of the relationship, from the data available, between the impact of that catch on penguins through the extent of its reduction of that recruiting biomass? The "theory" underlying the application of the GLM approach to disentangle the otherwise confounded effects of catch and biomass on penguin response variables is set out in Appendix A of MARAM/IWS/DEC14/Peng/B4. This was tabled seven months ago in Butterworth and Robinson (2014)—the authors of this quote do not offer any rebuttal to that.

 It seems possible that the plethora of positive but non-significant catch to penguin response relationships from the GLMs are due to weak but positive local level catch to biomass correlations under Figure 4 conditions.

This assertion is based on arguments that are, at best, unclear and further are contradicted by the responses above. At the very least the authors of the quotation need to offer a simulation study to evidence that the effect they suggest can occur, before serious consideration merits being given to this assertion.

The situation changes when islands are experimentally closed to fishing, since the correlation between catch and biomass is eliminated. Experimental closures are therefore necessary to predict the penguin response to setting catches to zero around breeding islands

The rationale for the correlation mechanism advanced has not been demonstrated above. See immediately below for comments suggesting that only "perfect" closures allow inferences concerning the effects of fishing on penguins.

It may be controversial that the local catch to biomass correlation is positive but weak. If so then the conclusion that catch has a direct positive impact on penguins is at least equally controversial. Under circumstances of unresolved controversy of this nature there would be confusion about whether the

operative mechanism is Figure 1, 4 or 5. Under these circumstances it is safest to exclude catch from the GLMs and use closure status and prey biomass as the relevant predictors.

The opening statement is not controversial; it is evident from the data (see MARAM/IWS/DEC14/Peng/B9). No clear rationale has been advanced for the necessity to distinguish the "mechanisms" (themselves unclear) advanced in the Figures referenced. This contrasts with defensible rationale advanced for the GLM approach in comments above and in MARAM/IWS/DEC14/Peng/B4. The final sentence highlighted, which suggests that the covariate needs to be closure status rather than catch, raises many questions.

- First note the discussion of this possibility under Item 16 of MARAM/IWS/DEC14/Peng/B10. Since this approach of using closure status was the one initially suggested some seven years ago, but later rejected in favour of the current approach of using catch, it is unclear why this suggestion should suddenly be resurrected a few months before the results of analyses are needed to inform management recommendations.
- The approach requires the assumption that the impact of catches on penguins is the same irrespective of the level at which the actual catch in a particular year lies between the historical maximum and (other than zero) minimum, and differs from that if the catch is exactly zero. This assumption hardly seems plausible unless past annual catches have all been very similar. Nor does it seem plausible that a changed position of a single catch from, say, just inside to just outside the 10 nm boundary for the closed areas could have a substantial impact on penguin dynamics.
- Inspection of the actual catches made within the 10 nm areas closed (Figure 1) suggests that the assumption might be reasonable as regards anchovy catches off Dassen Island (if one overlooks the fact that there were some small catches made in some of the nominally closed years). But for sardine the assumption becomes more questionable, with the catch in the "closed" year of 2008 almost half that in the open years of 2012 and 2013. For Robben Island, the very small anchovy catch in the "open" year of 2010 suggests conditions much closer to those of the three closed years following, than the much larger catches of the preceding four years. For Bird Island, although open in 2011, the sardine catch for that year in that region was zero, and hardly larger for 2010. Wouldn't one then expect penguin responses in those years to be similar to those of the closed years of 2011 to 2013, rather than to the open years of 2007 to 2009 when catches were 10- to 30-fold greater than in 2010?

Clearly models based on the closure/open distinction alone are open to serious question for these reasons. There seems no reason not to continue to use catch, as endorsed by the International Panel in 2010. (See also further discussion of this point in MARAM/IWS/DEC14/Peng/B12.)

Vessel effects

A second significant concern is that the core MARAM models use catch instead of closure as a predictor. Thus these models fail to address the fundamental aim of isolating the closure effect.

We suggest closure status, not catch, should be used in closure analyses as there may be effects on penguins from fishing vessels, other than removal of biomass. It has been suggested that penguins avoid areas where fishing is taking place (Harding 2012) ……………These are two plausible but untested mechanisms that could cause penguins to disassociate from vessels. There may be other mechanisms, but our collective experience of observing penguins and other seabirds at sea strongly supports the hypothesis that penguins, unlike many seabirds such as albatrosses, gulls and gannets, are not attracted to vessels.

The fundamental aim of the island closure exercise is not isolation of the closure effect *per se* (this is merely a device to maximise catch contrasts to improve estimation performance); rather it is to estimate the impact of catches close to the islands under investigation on penguin reproductive

success. The problems associated with use of closure rather than catch as an explanatory variable are summarised immediately above.

Many plausible hypotheses can be advanced for mechanisms at work in vessel–penguin interactions—doubtless many of them play some role, acting in various directions with different magnitudes in different conditions, and likely also many of them are untestable. That is why ultimately only an empirical approach can address the fundamental aim (see also Appendix A of MARAM/IWS/DEC14/Peng/B4). The Clark mechanism mentioned in that Appendix, which could lead to a positive effect on penguins in circumstances of sustainable fishing, requires neither that the penguins be particularly close to, nor attracted to the fishing vessels themselves.

Power analysis

The power analysis (Robinson and Butterworth 2014a) is based entirely on the linear modelling approach criticised above. In particular, the concerns around over-parameterisation call into question the validity of the estimates of the residual standard error in each model (a crucial determinant of variability in the simulations) and the lack of consideration of the precisions of the means used to estimate the effect of fishing calls into question the precision of the fishing effect in each model. Furthermore, because in future simulations catches are set to 0 in the 20 and 30 nautical mile (nm) blocks around each island, they represent an unrealistic scenario which has the effect of overestimating the actual power. In reality, the closure is ca. 15 nm, with fishing continuing within the 20 and 30 nm blocks while the closure is in place (Pichegru et al. 2012).

The matter of not taking account of the precision of the mean values used for the response variables in the GLM analyses conducted in documents MARAM/IWS/DEC14/Peng/B4 and Peng/B12 is addressed in the opening response above, and will be reported upon further in a subsequent document.

Robinson and Butterworth (2014a) present results from a number of GLMs in which the degrees of freedom is small. This casts some doubt on the reliability of the residual error variance which is the basis of subsequent power analyses. In order to address this, two options were explored. The first is to use the upper 95% confidence interval of the residual variance in the power analyses. The other is to treat the year effects as random effects (Robinson and Butterworth 2014b). Although this approach is an improvement because it introduces additional variance into the power analyses, viz. the year factor variance, it nevertheless suffers from the limitations which arise because the penguin response value which is used in the input data is a simple year average. Thus, since the random effects model assigns the same variance to each year effect, the nuances of the number of observations per year and their within year seasonal distributions are not dealt with, which would result in different effective weights for annual averages from different years. Lastly, it is not clear whether the problems with small degrees of freedom are addressed by the use of year as a random effect. Sample size recommendations for multi-level mixed effects models is an unsettled issue in statistics (Bell et al. 2010).

Again, responses to the questioning of the use of a single statistic (the mean), without consideration of its sampling precision, to reflect information on the response variable each year in the analyses has been addressed in the opening responses above. Whether the random effects model estimates of residual standard error for use in the power analyses are biased may well be an open question. Table 1 shows examples where these estimates are compared with bias-adjusted MLE estimates from the corresponding fixed effects models. The latter are about 30% higher on average, but this comparison nevertheless does not suggest that the random effects estimates are completely unrealistic. Nevertheless the main point to consider is that the primary purpose of the feasibility study is to obtain estimates of these residual standard errors which are a pre-requisite for the power analysis computations. The authors of the quote above raise possible problems about what has been done, but offer no alternatives.

Issues of interpretation

Based on observed positive correlations between catch and some penguin fitness components, Robinson (2013) and Butterworth and Robinson (2014a) conclude that fishing around Dassen and Robben islands does not affect penguins and in some cases may even have a positive effect. The mechanism that is used to explain this is that fishing may break up the shoals of fish, resulting in more smaller shoals, making it more likely that a predator (such as penguins) will encounter a shoal. Centuries of ecological research show that prey densities positively affect predator densities (e.g. Power 1992). The Butterworth – Robinson mechanism must therefore be strong enough to override basic ecological principles. Such a strong assertion requires solid empirical evidence before it can be regarded as a plausible explanation. Observations that would be needed to support this mechanism are a) whether fishing does indeed break up shoals of fish, b) if a) is true a measure of how big these remaining shoals are and how long it takes before the fish aggregate again and c) whether there is a threshold shoal size which penguins are able to detect. Furthermore, it should be noted that there are no observations that splitting fish schools assists foraging by penguins, whereas some research has indicated that penguins sometimes feed cooperatively and may benefit from exploiting large schools of fish (see Additional information).

However, there is no observational or past evidence that catch benefits penguins; rather the converse has applied with regard to catch and many studies have indicated a positive relationship between estimates or proxies of forage fish biomass and numbers of African Penguins breeding and their demographic parameters. Studies conducted elsewhere have demonstrated both the importance of sufficient availability of prey to sustain seabird processes at colonies and a negative impact of local depletion of prey by fishing on foraging parameters of seabirds that compete with fisheries for anchovy. Therefore, it is not defensible at present to expect fishing to benefit penguins.

Prey densities do not always correlate positively with predator densities. For example, increased trawl catches of South African hake over time have reduced hake abundance, but led to increases in populations of scavenging seabirds.

Empirical evidence is indeed required, and is provided by the analyses of MARAM/IWS/DEC14/Peng/B4 as updated in Peng/B12.

Fishers frequently assert that fishing does scatter shoals, leading to catch rates falling until there is some break in fishing which allows the fish to re-aggregate. To investigate this possibility in a simple way for the South African anchovy fishery, average catch per set values for the directed anchovy catch for each day of the week for 2013 were computed (courtesy Naseera Moosa), anticipating that such an effect might be evident assuming that fishers take weekends off. The results are shown in Table 2. There is some indication of this effect given that catch per set drops as the week progresses, *except* that the value for Fridays spoils the pattern hypothesised! The Table also indicates that fishers work much too hard, resting only on Saturdays, and even then not enough that day of the week for this approach to provide the extent of data contrast sought. A more extensive analysis would seem necessary to try to identify more firmly if there is indeed such an effect in this fishery.

The absence of direct observations indicating that scattering shoals assists (or does not assist) penguins in foraging does not falsify either possibility. As mentioned above, many plausible mechanisms may not be observationally testable – what matters is empirically monitored outcomes.

Clearly over a wide range of fishing mortality, there must be a domain at the high end where fish abundance is reduced so much that the impact on predators must be negative, but the overall relationship need neither be linear nor even monotonic. Furthermore the trade-off between the impacts of different mechanisms (see Appendix A of MARAM/IWS/DEC14/Peng/B4) is likely to be situation

specific (e.g. the anchovy "river" effect impacting the west coast colonies at Dassen and Robben Islands – see MARAM/IWS/DEC14/Peng/B5 – would not be relevant to the Eastern Cape's Bird and St Croix Islands). Thus net impacts could be positive in one region and negative in another, and hence reference to an effect found elsewhere does not constitute definitive evidence for what is to be anticipated locally.

It is also pertinent to document prior observations regarding the impact of food availability on African Penguins. If catches of sardine or anchovy, the main prey of African Penguins, have a positive impact on penguins, it might be expected that numbers of penguins in each of three regions where they breed (southern Namibia, Western Cape, Eastern Cape) would increase following the commencement or expansion of commercial fishing of these species. Instead, in each region large decreases of penguins were observed (Shelton et al. 1984, Crawford et al. 2001, 2009).

Note the overall possible non-linearity and even perhaps non-monotonicity of the relationship between the impact of fishing and penguin response variables mentioned above, and also the comments about comparisons with Namibia under Item 12 of MARAM/IWS/DEC14/Peng/B10.

Alternatively, if African Penguins are benefitted by a ready availability of prey, numbers breeding and demographic parameters of penguins might be expected to be positively related to prey availability and vice versa. This has almost invariably been the case. In each of the three regions noted above, numbers breeding were positively correlated with abundance of sardine and anchovy (Crawford 2007, Crawford et al. 2008, 2011). Similar positive relationships between numbers breeding and prey abundance have been observed at the colony scale (Crawford et al. 2008, Sherley et al. 2013). By contrast, severe fragmentation of penguin colonies occurred under conditions of food scarcity (Cordes et al. 1999).

Relationships to penguin **numbers** are confounded with the population dynamics of the penguin group concerned. Relationships need to be demonstrated through (essentially) fitting to residuals after accounting for those dynamics, as is effectively what has been done for the Robben Island population in MARAM/IWS/DEC14/Peng/B3a and b. These points also need to be viewed in the context of the very small contribution (about 0.5%) that consumption by penguins makes to the overall loss of anchovy to natural mortality (see MARAM/IWS/DEC14/Peng/B3b, pg 25).

Decreases of penguins at colonies between Lüderitz in Namibia and Dassen Island in the Western Cape, South Africa from the 1960s to the 1980s followed collapses of sardine stocks off Namibia and South Africa. Increases at Mercury and Ichaboe islands north of Lüderitz after the 1960s probably resulted from an increased local abundance of pelagic goby (Crawford and Shelton 1978, 1981, Shelton et al. 1984, Crawford et al. 1985, 2001). An increase and later decrease in numbers of penguins at Dyer Island, South Africa coincided with an increase and then a fluctuating decrease of anchovy off South Africa (Crawford 1998). Large numbers of African Penguins abandoned breeding at Dyer Island in 1991 when anchovy in its vicinity was scarce (Crawford and Dyer 1995).

Again the comments about comparisons with Namibia under Item 12 of MARAM/IWS/DEC14/Peng/B10 apply. Furthermore care needs to be taken in claiming close linkages between trends in penguin and anchovy abundance trends; for Dyer Island the downward trend in penguin numbers from 1986 referenced by Crawford (1998) continued through 1991 and 1992 (see Figure 2 of Ludynia *et al.*, 2014) despite a marked increase in anchovy abundance in those years (see MARAM/IWS/DEC14/Peng/C1, Table 11).

In broader terms though, are the circumstances that pertained in the examples above pertinent to the issues under current consideration? If they were, then wouldn't the matters currently at issue have been considered resolved, so why then institute the current closure feasibility study?

The formation of new colonies at Stony Point, Boulders on South Africa's mainland and Robben Island and increases at these colonies and at Dassen and Vondeling islands took place during a

period of recovery of South Africa's sardine in the 1980s and 1990s (Crawford et al. 2001, Underhill et al. 2006). After 2004, large decreases in numbers of penguins breeding off western South Africa were associated with an eastward shift of adult anchovy and sardine and a collapse of South Africa's sardine (Crawford et al. 2011). As the prey of penguins shifted south and east, there were decreases in penguin numbers at colonies located off northwest South Africa (the northernmost colony at Lambert's Bay became extinct). Numbers also decreased at the more southerly Dyer Island, while numbers at nearby colonies such as Boulders Beach and Stony Point, stabilised or increased (Crawford et al. 2011). More than 220 000 tonnes of sardines were caught within about 30 km of Dyer Island during 2002- 2004 (JJ van der Westhuizen, unpublished information) and the exploitation rate of sardines west of Cape Agulhas rose to 30 - 45% in 2002 and 2005 - 2007 (Coetzee et al. 2008). This high catch of sardines had no apparent benefit for penguins at Dyer Island, but rather appears to have had an adverse influence. When colony size at Dyer Island was > 3 500 pairs, numbers breeding there were negatively related to sardine catches made within 20 nautical miles of the island (Ludynia et al. 2014).

The four responses immediately preceding this apply also to the above.

Several demographic parameters of African Penguins have been positively related to estimates or proxies of the abundance of sardine and/or anchovy at various spatial scales, including adult survival (Robinson 2013, Sherley et al. 2014), breeding participation (Crawford et al. 1999, Durant et al. 2010) and breeding success (La Cock 1986, Adams et al. 1992, Crawford and Dyer 1995, Crawford et al. 1999, Crawford et al. 2006, Cury et al. 2011, Sherley et al. 2013, Pichegru et al. 2014). There was likely delayed onset of breeding by African Penguins during a period of food scarcity (Crawford and Dyer 1995). The fledging period for chicks has been negatively correlated with prey abundance (Sherley et al. 2013).

It would have been interesting to have had these argued relationships set out in detail to allow readier thorough review. The reference to "proxies" of fish abundance does however raise concerns, particularly given earlier arguments (including in at least some of the papers referenced immediately above) that catch may be used a proxy for such abundances, as these arguments have since been demonstrated to be spurious (see MARAM/IWS/DEC14/Peng/B4, B9 and B10). Once again though the question arises: were these arguments definitive, wouldn't the issue have been considered resolved, so why then the current closure feasibility study?

Given (a) ambiguity about how catch in the vicinity of African Penguin colonies may influence penguins, (b) the absence of any empirical proof for a benefit of fishing on penguins, and (c) the above substantial body of evidence that a ready availability of food benefits African Penguins, it is not defensible at present to expect fishing to benefit penguins. Furthermore, studies conducted elsewhere have demonstrated both the importance of sufficient availability of prey to sustain seabird processes at colonies (e.g. Hunt et al. 1986, Gaston et al. 2007, Cury et al. 2011) and a negative impact of local depletion of prey by fishing on foraging parameters of seabirds that compete with fisheries for anchovy (Bertrand et al. 2012).

These points have essentially been addressed already above. Thus, for example, empirical evidence is provided by the analyses of MARAM/IWS/DEC14/Peng/B4 as updated in Peng/B12. Furthermore the trade-off between the impacts of different mechanisms (see Appendix A of MARAM/IWS/DEC14/Peng/B4) is likely to be situation-specific, so that net impacts could be positive in one region and negative in another, and hence reference to an effect found elsewhere does not constitute definitive evidence for what is to be anticipated locally.

Additional Information: Foraging behaviour and strategy of African Penguins

Adult African Penguins tend to forage in groups (Frost et al. 1976, Wilson and Wilson 1990). …………Since swimming is slower and more energetically expensive than flying (Pinshaw et al. 1977,

Schmidt-Nielsen 1999), penguins require predictable food resources close to their colonies during breeding (Sherley et al. 2013). The sine qua non for African Penguins to breed successfully at colonies and after that to survive to moult will be a sufficient density of prey in the neighbourhood of colonies.

Again, as above, over a wide range of fishing mortality, there must be a domain at the high end where fish abundance is reduced so much that the impact on predators must be negative, but the overall relationship need neither be linear nor even monotonic. However what matters for informing imminent management decisions is the relationships that apply under prevailing conditions, which is the basis for the proposals put forward in MARAM/IWS/DEC14/Peng/B2.

The African Penguin Biodiversity Management Plan (BMP-AP) was gazetted in 2013 (DEA 2013) andaims to halt the decline in African Penguins within two years of implementation. The plan outlines various actions to be taken to improve the conservation status of the African Penguin and it is anticipated that all of these in combination will need to play a part in achieving the aims of the BMPAP. Scientific results from available analyses thus far, suggest that there is some benefit to penguins of purse-seine fishing closures around islands, even with the limited duration of the study and remaining uncertainty in the process model. This lays the foundation for several candidate strategies to be suggested for further evaluation.

The highlighted statement is impossible to evaluate in the absence of provision of any details.

Weller et al. (2014) developed a systems dynamic (process-based) model of the pressures acting on African Penguin demographics, and parameterised this for Robben Island. This model is strategic and not intended for prediction purposes. The model showed a weak benefit to the penguin colony of short-term (three-year) closure to purse-seining around Robben Island, which would be strengthened with longer and more continuous periods of closure. The approach highlighted the importance of managing the combined negative pressures acting on the population, including oiling. The model was parameterised based on the best available biological estimates for the locality (Annex 3) and is being adapted for application to the Dyer Island African Penguin colony, for which sparser data are available. In light of analyses currently underway and which are suggesting existence of an asymptotic relationship between food availability and penguin survival (R. Sherley pers. comm.), the model will be refined accordingly once the relationship has been finalized.

The fundamental problems associated with this approach, and consequently the questionable nature of the conclusions which it draws, are addressed in MARAM/IWS/DEC14/Peng/B7 and Peng/B14.

Another component of this work is the small scale pelagic fish abundance surveys around the islands. As shown in Figure 1 of Annex 2, the link between closure and local fish biomass is key to understanding the impacts of closures on African Penguins. On the West Coast, these small scale surveys were conducted by the DAFF and on the East Coast by Dr Lorien Pichegru and Alistair McInnes (University of Cape Town). Due to weather and logistical constraints, surveys on the west coast have been limited in frequency during the penguin breeding season. Nevertheless penguin tracking was achieved in conjunction with several of the small scale surveys. The small scale surveys show high variability in pelagic fish biomass, such that variability and the gaps between surveys would only allow for imprecise biomass estimates. Survey effort was not equal over the years of interest, for example in 2013 only one small scale survey was conducted around Robben Island. For these reasons the small scale data were not included in the analyses in Annex 2. The surveys on the East Coast were performed more regularly (once per month during the penguin breeding season) and often in conjunction with the tracking of breeding African Penguins. These data are currently being analysed as part of two PhD theses and show promise for use in linking the finer scale movements of penguins to local fish abundance (A. McInnes and K. Robinson pers. comm.) Efforts should be made to continue conducting these surveys on both coasts, at a frequency which will allow insights into finer scale changes in fish abundance and thereby elucidate the links between fishing, available biomass and penguin parameters. If necessary, additional resources should be sought to do this.

Analysis of similar surveys around Dassen and Robben Islands showed their information content to be disappointingly low (MARAM/IWS/DEC14/Peng/B6). Hopefully these surveys in the Eastern Cape turn out to give better results, but the information provided above is insufficient to allow this prospect to be evaluated.

In this context and in the absence of consensus or converging results showing that island closure is *not* beneficial to penguins, a precautionary approach to managing fishing around these islands is justified. However, without alternation of closures, simultaneous closures of all islands may not provide sufficient variability to show significant results, as there would be no reference point for general ecosystem conditions. In line with an Ecosystem Approach to Fisheries, there are three dimensions that need to be evaluated: the ecological dimension (maximum protection to penguins), the human dimension (potential social and economic impacts), and the dimension of ability to achieve (which includes minimizing existing scientific uncertainty). An in-depth discussion is required regarding the trade-offs between these three dimensions namely (a) contributing to the conservation aims of the BMP-AP, (b) the scientific objective of generating statistically significant results and (c) potential social and economic impacts of closures. In the light of the above, a few strategies should be considered:

- Permanent precautionary closure of all four islands simultaneously (maximising conservation benefit to penguins);
- Continuing three year alternations of closure between island pairs in order to improve the sample size for detection of any significant effects (maximising scientific knowledge gain);
- Consecutive years of experimental closure (5 years) of all four islands, using each island as its own baseline (compromise).

The underlying logic here is flawed. Closure of islands is taken to maximise conservation benefit to penguins. But that is assuming as true the hypothesis whose truth or otherwise the whole closure initiative was set up to determine. Given that the majority of the GLM results for Robben and Dassen Islands indicate the effect of fishing on penguins to be positive, how then can a position that island closures reflect a precautionary approach to penguin conservation follow logically?

References

Butterworth, DS and Robinson, WML. 2014. Response to: "Response to comments on the Robben Island penguin pressure model". DAFF Branch Fisheries document: FISHERIES/2014/APR/SWG-PEL/ICTT/25: 14pp.

de Moor, CL and Butterworth, DS. 2009. How should the sardine length frequencies be weighted? Exploring alternative methods. Marine and Coastal Management document: MCM/2009/SWG-PEL/42. 23pp.

Ludynia, K, Waller, LJ, Sherley, RB, Abadi, F, Galada, Y, Geldenhuys, D, Crawford, RJM, Shannon, LJ and Jarre, A. 2014. Processes influencing the population dynamics and conservation of African penguins on Dyer Island, South Africa. African Journal of Marine Science, 36 (2): 253-267.

Myers, RA, Bowen, KG and Barrowman, NJ. 1999. The maximum reproductive rate of fish at low population sizes. Canadian Journal of Fisheries and Aquatic Sciences, 56: 2404–2419.

Table 1: A comparison of residual standard error σ_ε estimates for each penguin response series is given for the random year effect and fixed year effect models for the case of total catch within 30 nmi for the Western Cape colonies and sardine catch within 30 nmi for the Eastern Cape colonies considered. The estimates for the random year effects model are REML, while those for the fixed year effects model are MLE inflated by the standard linear model $\sqrt{n/(n-p)}$ bias correction factor.

(a) Dassen and Robben islands

Penguin response	σ_ε (random)	σ_ε (fixed)
Chick condition	0.212	0.251
Active nest proportion	0.411	0.577
Fledgling success	0.083	0.321
Foraging path length	0.293	0.255
Foraging trip duration	0.207	0.309

(b) St Croix and Bird islands

Penguin response	σ_ε (random)	σ_ε (fixed)
Foraging path length	0.106	0.176
Foraging trip duration	0.128	0.196

Table 2: Daily average catch-per-set values for directed (>50%) South African anchovy catches during 2013.

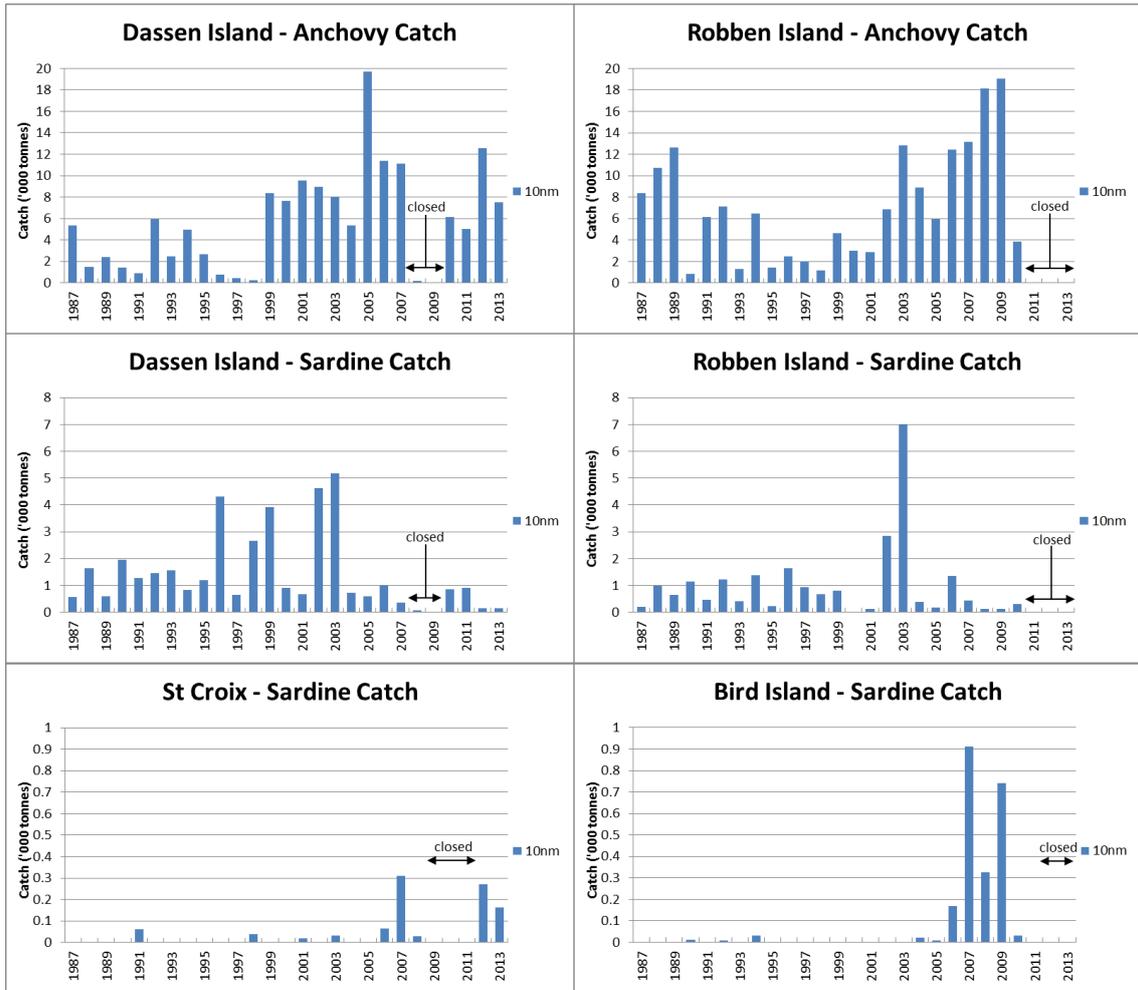| Day of the week	Average catch-per-set (tons)	Number of sets
Monday	32.3	316
Tuesday	32.4	349
Wednesday	25.0	383
Thursday	28.2	483
Friday	34.8	493
Saturday	26.0	223
Sunday	28.4	341

Figure 1: Annual catches of anchovy and sardine within 10 nm of the islands under consideration, with periods during which such regions were closed to pelagic fishing also shown. Note that anchovy catches in these regions around St Croix and Bird Islands were negligible.