

RESPONSE TO MCM/2010/PEL/48:

Suggestions for Improving the Penguin Modeling Framework Developed by William Robinson, Éva Plagányi, Doug Butterworth and Carryn de Moor and expression of several concerns by outside, interested parties (Lynne Shannon, Res Altwegg, Rob Crawford, Astrid Jarre, Jessica Kemper, Jean-Paul Roux, Les Underhill, Lauren Waller, Ross Wanless)

Doug Butterworth, William Robinson, Éva Plagányi and Carryn de Moor

We thank the authors of PEL/48 for their comments, which have led us to formulate what we consider will be some important improvements to existing work. For readers' ease, this response is presented by inserting replies/comments into the original text of PEL/48. To distinguish these inserts from the original, the inserts are shown in *italics*. Note that some of these request the provision of priors to reflect biologists perceptions of the likely range of values for certain parameters: these are denoted, for example, by **P1** in these inserts, and are summarised at the end of this document, together with a list of further work planned in draft priority order for discussion at the next PWG meeting.

1) As general background, it should be noted that the brief for document MCM/2010/SWG-PEL/35 was to combine into a single convenient reference a number of documents that contributed to the discussion and decision on the issue of possibly amending the pelagic OMP to take explicit account of impacts of removals by the fishery on penguin dynamics that was taken at the December 2008 meeting of the PWG. No attempt was made in that document (PEL/35) to revise the calculations available at that time in the light of further work which has been ongoing – some of that work and associated future plans have been reported in the latter part of MCM/2010/SWG-PEL/36. Furthermore, it should be noted that those calculations had to be finalized rapidly to meet that December 2008 deadline, taking account of comments being made at an international review workshop taking place concurrently, so that certain elements were not reported upon in as full a manner as might otherwise have been the case.

This document is a compilation of comments, concerns and suggestions from a group of interested parties in response to the suite of penguin-pelagic fish interaction modeling undertaken by MARAM, as summarized in MCM/2010/SWG_PEL/35. Comments below are as of now, but further issues may arise during model revision and will necessitate further deliberations, thought and response. Throughout the process of model revision and development, further discussion with penguin specialists is encouraged.

Adult survival

The survival rate in the model comes out almost constantly at 0.96.

2) Concerns about what was originally, for reasons of time, a MLE estimation process, have been overtaken by an extension to Bayesian estimation – in particular the end product of the process depends on an integral over the distribution for survival rate estimates so requires a Bayesian approach. MLEs can fall on constraint boundaries as a result of observation error, but may also be indicative of model mis-specification (see further discussion below).

This is very unrealistic, and corresponds to no measured value from South Africa or Namibia, nor any other penguin population (except perhaps Emperor and King penguins) in the world! The model should

take available measured survival rates for African Penguins into account. The adult survival rate of African Penguins is well-known from Mark-Recapture studies over many years to be around 80%–85% per year.

3) Isolated estimates of survival rate in marine birds, fish and mammals, frequently obtained from tag-recapture data, are usually (quite substantially) downward biased because of confounding with the effects of, for example, emigration and tag-loss, and consequently are often not used as inputs to dynamics models in the manner that is suggested above (which was tried, failed, and consequently replaced by approaches of the type considered here at CCAMLR, for example). Although these estimates can be useful in developing priors to some parameters (if from data not used in fitting the dynamics model), best practice in modelling dynamics is where pertinent to consider using the original data themselves in the likelihood for the model fit, rather than the results from separate analyses, inter alia to ensure respect of self-consistency.

To accept a model which states, “for realism” the upper bound on survival is 96%, and then accepts results which use this boundary value year after year (Figure 6) is unacceptable. Independent estimates of adult survival of African penguins for both Dassen and Robben islands are available from capture-mark-recapture (CMC) analysis (Crawford *et al.* 1999, Wolfaardt *et al.* 2008, Altwegg 2009) but have not been used. At both Robben and Dassen islands, model estimates of survival are high in many years (see Fig. 4 – at the bound of allowable estimates) and considerably higher than CMC estimates. By not using CMC estimates, a lot of information is lost.

4) Subsequent work has incorporated tag-recapture data in the likelihood for fits to models of the dynamics of penguins at Robben Island, and this will in due course be considered for the penguin-fish-interaction model. It should be noted that this is a non-trivial exercise, because the existing software for the standard MARK tag-recapture analysis package cannot be readily integrated with efficient auto-differentiation software packages needed for population models with additional parameters and contributions to the log-likelihood; this necessitates reprogramming elements of the MARK multinomial likelihood approach from the bottom up.

One of the main problems is that the models cannot reproduce the observed increase of the African penguin population, especially during the 1990s on Robben Island. The model uses counts of penguin moulters, and essentially relies on the proportion of juveniles in these counts as a measure of reproduction. In an attempt to fit the population trajectory, the model pushes adult survival rates towards the limit, which the modelers set at 0.98 (or 0.96 in some versions). As a result, the model cannot be used to examine temporal variation in survival. It is worrying that this is what the modelers proceed to use it for when they correlate survival with fish abundance. Initial thought on our part was that the proportion of juveniles does not give enough information to reliably separate reproduction from survival. However, the Bayesian approach showed that the data do push the survival rates up, rather than that they drifted towards the boundary.

5) Although the characterisation above of how “the modelers proceed” is not entirely accurate, perhaps the more important aspect of this is that reflected in future plans in document PEL/36 which emphasise the importance to be given in ongoing work to testing the robustness of conclusions from such an approach to the impact of fishing on penguins to a wide variety of model assumptions. A particular option highlighted there is to take account of the possibility of immigration during the 1990s (conceivably from the Dyer Island colony) to the two colonies modelled, which would reduce posterior

median survival rate estimates, quite possibly to levels with which some scientists feel more “comfortable”. Suggestions of priors to use for the period and extent of such possible immigration of presumably mainly juvenile penguins from Dyer Island would be appreciated (P1).

Subsequently, we considered that the proportion of juvenile moulters may underestimate reproduction if juveniles are occasionally counted as adults, or some of them not counted at all (e.g. if juveniles and adults moult in different places). The modeler modified the model to allow for juveniles to be undercounted at a constant rate. This did not solve their problem; the survival rates are still biased high. Perhaps juveniles are undercounted at a variable rate, for example if more juveniles are counted as adults in years when the breeding season is relatively early. However, the modeling approach cannot handle such a time-variable detection effect.

*6) At present we have used a uniform prior over [0.5; 0.9] to relative undercounting of juvenile moulters (MCM/2010/SWG-PEL/27). As stated in PEL/36, we would welcome comments by penguin biologists familiar with the moult count exercise on the appropriateness of this choice (P2), because the original advice received from penguin biologists some 2–3 years previously was that the moult count process was definitely equally efficient at counting juveniles and adults. The models can also handle the time-variable possibilities suggested very straightforwardly through introducing a further random-effects parameter, and we would similarly welcome comments on an appropriate prior choice for the standard deviation of that effect (essentially the likely size of such variability) (P3). We caution however that this introduction of such an effect may make little difference to results because it can be compensated by changes in other parameters, and it would be expecting rather much of the data to be able to distinguish such effects and update the associated priors. The matter of whether to allow for possible relative undercounting of juveniles in the moult count is an important issue. Contrary to the statement above it **does** solve the “problem” of MLE estimates for the adult survival rate on constraint boundaries. PEL/?? shows that although the addition of the prior of PEL/27 does not result in much change in the Bayesian posterior median estimates of adult survival rates, most of the MLEs previously on the constraint boundary are moved well off that and are generally fairly close to the corresponding Bayesian posterior median estimates.*

We suggested that the modelers revisit the assumptions of a constant age at first reproduction, and the upper limit of 2 fledglings per year. William apparently played around with this and it did not change things much.

7) We are happy to include variation in age at first reproduction amongst our sensitivity tests, and would appreciate suggestions on an appropriate prior (P4). The same caution as in 6) above applies though. Many earlier studies for other species using essentially the same model type have shown that such many such effects are confounded with others, and all that really matters is appropriate consideration of the net composite effect as reflected by the composite parameter S' (with its random effects component) in the penguin population model. Increasing the upper limit for fledging success beyond 2 could result in lowering the posterior mean for adult survival (though the extent to which such compensation might occur is limited by the need for consistency with the juvenile/adult moult ratio in the moult counts. We had however already the impression that an upper limit of 2 was stretching things, when taking account of the comments by Crawford reproduced on pg 27 of PEL/35 that: “Averaging Anton’s values for those re-breeding within 2.5 months (25% of successful; 36% of failed at incubation; 21% of failed at brood) gives 27% of birds having a second clutch, which as Anton points out is the same as observed in an earlier study at Robben Island (27% p. 143, Crawford et al. 1999). Therefore to get an estimate of

chicks fledged per pair per year, I would multiply Anton's chicks produced per breeding attempt by 1.27.” Further comments on an appropriate choice of a prior for this upper limit would be appreciated (P5).

The models do not seem to pay much attention to the observation process. The model that uses moult counts only assumes that counts are perfect. As a result, any variability in the counts induced by the necessarily imperfect observations, is forced on the survival estimates.

8) This seems to evidence a major misunderstanding of the model. Molt (and similarly breeder and juvenile proportion) counts are NOT treated as perfect. This is clear from equations (16) to (18) which specify the contributions of these data to the log-likelihood, and the plots in Figs 1 and 2 certainly do not indicate exactitude between model and data.

At the May 2010 meeting, Doug argued that variability in the observation process could have produced survival estimates that are biased high. It is very difficult to count every last penguin during breeding, and even more so during moult, and so it seems possible that the population is undercounted by a bit more in one year than in another. Looking at the counts, it would now appear that more penguins were added to the population than seems biologically possible. However, to get survival rates on the boundary year after year as the modelers did with their model, would require a trend in counting error so that an ever larger fraction of the penguins are being detected as the population increases. This seems highly unlikely, but it may be worth examining whether there could have been an increase in counting effort over the years.

Breeding proportion and age at first breeding

The model fails to take account of real penguin biology. In particular, the model fails to incorporate the biological realities that: (1) not all adult penguins breed every year, and that the proportion breeding may be dependent on food abundance; (2) not all adult-plumaged penguins are breeders:

It is uncertain how the model allows for breeding participation to vary. Yet, this parameter is likely to be important in allowing a rapid response to good feeding conditions (e.g. Crawford 2007, Crawford *et al.* 2008). A multi-state model with defined transition probabilities would account for some of this uncertainty, and would, at a minimum, allow the assessment of the uncertainty on the model outputs with more clarity.

9) The model most certainly takes account of the biological realities (1) and (2) mentioned through the random effects parameter μ in equation (13) which was deliberately introduced for this purpose (originally by international reviewer Mark Maunder), and the Bayesian estimation approach will properly allow for the uncertainty associated with the estimation of this variability given the limitations of the data.

Age at first breeding: it is possible for birds to breed from age 3 as stated on p. 27 (cf. equation 4). Age at first breeding is hard to estimate in the field and is likely to be related to food availability. The authors assume that all adults aged 4+ are potential breeders (pg 3): “**Number of potential breeders each year is the sum of adults aged 4 and over**”, yet citing Crawford *et al.*'s (1999, see pg 27) estimate of 100% of breeders are age 5+. Also see Whittington's estimates of age at first breeding, based on re-sightings of banded individuals for Robben Island are: ~55% of individuals first breed when aged 4; mean age of first age at breeding 5.02 years. For Dassen Island ~60% first breed at age 4; mean age is 4.97 years.

Nevertheless, the point we make here is that age at first breeding is variable and models should account for this. The transition from “pre-breeder” to breeder is likely to be influenced by food availability. If food is abundant, then they might begin breeding at an earlier age, if food is scarce, they might delay breeding. This is also the case with birds of breeding age, which may or may not breed, depending on food availability. Has a sensitivity analysis on age at first breeding and the age structure of breeders been done?

10) See response 7) above.

Also note the following findings from Kemper’s PhD thesis: “From the age structure [based on observed annual adult survival rates and estimates of penguins in juvenile and adult plumage], it is estimated that between 60% and 66% of all individuals in adult plumage were 5 years or older [i.e. assumed to be potential breeders], with a mean percentage of 62.1%” . . . “Annual peak active nest numbers consistently underestimated the number of breeding pairs (i.e. half the potential breeding population) by between 21.4% and 46.8%, and on average by 32.2%. On average, the number of individuals aged five years or older was 3.0 times greater than the peak annual active nest count.” Although these data come from the study of the Namibian population, it is highly likely that the situation is similar with the South African populations, albeit with slightly different percentages. This discrepancy implies (a) that nest counts are a poor proxy for breeding pairs, because not all pairs breed at peak and/or (b) only a portion of birds of breeding age may breed; this proportion is likely to vary with food availability.

11) Bias in counts compared with true numbers is taken into account through the q proportionality constant parameters in the model – see equations (16) and (17). The current assumptions in their regard are as follows (pg 5 of PEL/35):

It is assumed that at both Robben Island and Dassen Island the nest count is thorough, including nearly all breeding pairs. Hence a high proportionality constant for breeders (q^B) is chosen, namely 0.95. Likewise, the count of Robben Island coastal moulters is considered a reasonably accurate measure of the population because very few birds have been observed moulting inland (Crawford and Boonstra 1994), so that q^M_{Rob} was set equal to 0.9. This is not the case at Dassen Island, so the constant q^M_{Das} is estimated.

These assumptions were based on input from penguin biologists some 2–3 years ago, including in particular one of the authors of PEL/48. We would appreciate updated priors for these parameters (and also their possible changes over time – certainly also very recent moult counts at Dassen Island in particular are affected by penguins first hauling out farther south to moult) (P6). However a value of q^B for Robben as low as suggested by Kemper would seem difficult to reconcile with the moult counts at Robben Island as it would imply many more breeders than adult moulters.

Breeding success

The rapid increase of numbers of penguins breeding in the late 1990s and early 2000s can be plausibly modeled through accounting for improved breeding success and increased participation in breeding, without the need to force survival to very high levels (Crawford 2007).

Pairs are able to fledge more than the 1.4 chicks per year, which is indicated as a maximum on p. 3 because pairs may rear two broods annually (e.g. Crawford *et al.* 1999). However, mean measured average breeding success is lower than this maximum value. Time series of average values for Dassen and Robben islands have been tabulated for the Closures Task Team (July 2010).

12) See response 7) above. Note also that problems with the analyses of Crawford (2007) were explained in detail in MCM/2007/MAY/SWG-PEL/11.

It appears that the parameters S^j merge breeding success and immature survival of African penguins. Whereas measures have been made of breeding success, and will increasingly account for other forcing factors such as heat and rain, immature survival is harder to estimate. By merging breeding success and immature survival, empirical information is lost. (Note also that juvenile survival is linked to food availability, as discussed under the "Food" sub-section later in this document). Merging of breeding success and immature survival also disallows comparison of model output (Fig. 4) with measured estimates. Use in the model of values estimated by other means will lessen the number of parameters required to be estimated from a limited number of observations.

13) Information is not lost in this case, because estimates of breeding success are biased down, and the upper limit applied is not in conflict with observations. As the effects of reproductive success and juvenile survival are completely confounded in the absence of any direct measures of juvenile survival in isolation, there is no saving on the number of parameters which need to be estimated.

Carrying capacity

The assumption that the penguin colony at Robben Island was in equilibrium in 1986 (p. 3) is incorrect. African penguins recolonized Robben Island in 1983 and the colony subsequently grew rapidly with a substantial part of the growth attributable to immigration (Crawford *et al.* 1995, 1999).

The aim of the model is to measure the effects of different fish management strategies (and hence fish biomass) on penguins. As K is probably a (fairly) direct function of fish biomass at present, using a predefined value for K and a fixed density dependence relationship (which seems in this case to have a large influence) can possibly mask the signal that we are seeking as the modeled population will be mainly driven by these effects rather than changes in food abundance. This model structure has the effect of rendering the modeled population insensitive to fish abundance. This is a major flaw of the model.

The carrying capacity in terrestrially breeding top predators is limited by several factors, i.e. suitable breeding space and food availability (within the breeding range, at least during the breeding season). When food is abundant, a population may be limited by breeding space; when breeding space is abundant, the population may be limited by food abundance. Given historical breeding population sizes in the Western Cape compared to current levels, breeding space carrying capacity is not relevant. The assumptions behind the definition of carrying capacity should therefore be (mainly) a function of food availability, and is therefore not a fixed value, i.e. carrying capacity should be higher when there is more fish. We do not understand how carrying capacity was defined for the model, and why current (adjusted) population estimates were multiplied by 5 for Dassen Island and by 2 for Robben Island, particularly when they have overlapping feeding ranges.

The carrying capacity of Robben Island for penguins was determined by Crawford *et al.* (2007) as being of the order of 8500 pairs, which using information in Crawford and Boonstra (1994) translates to about 27000 adult-plumaged birds. This is not dissimilar to the 24436 birds indicated on p. 3, but provides some empirical support. A present carrying capacity of 186000 adult-plumaged birds for Dassen Island seems too high and it is not estimated as indicated on p. 3. The maximum number estimated to moult around the island's shoreline was about 26000 birds in 2000/01, which if divided by 0.35 (p. 3, arbitrary) provides an estimate of about 74000 adult-plumaged birds. About 25000 pairs bred at Dassen Island in 2004, the recent maximum. If information in Crawford and Boonstra (1994) is used, this equates to about 80000 adult-plumaged birds.

Although carrying capacity may influence breeding success and immature survival (equation 5), this is not yet empirically demonstrated. However, recruitment of immature birds to islands appears related to carrying capacity (Crawford *et al.* 2007). Hence recruitment of first breeders to colonies probably is too. Recruitment of first breeders to colonies probably also is influenced by local food availability (e.g. Crawford *et al.* 2001), which can be expected to be influenced *inter alia* by extraction of fish in the vicinity of islands.

It is unclear how much is actually known about density dependence; however, the effect of density dependence on reproductive success appears to be considerable in the model (see formula 5, page 3). It seems that the carrying capacity / density dependence assumptions have a big effect on reproductive success.

Note that in colonial animals such as African Penguins, density dependence is not linear, but may follow a non monotonous trajectory, for example with reproductive success being low at very low breeding densities, high at intermediate densities and low again at very high densities.

14) There is considerable international debate about the functional form for density dependence, and in general sensitivity of results to alternative forms merits examination, in circumstances where a linear form is as reasonable an assumption as any for a reference case which should foremost respect parsimony. Nevertheless, the inclusion of density dependence in the model of PEL/35 was a hangover from earlier penguin models developed with different objectives. Although there may still be what effectively are space-limitation restrictions on penguins because of the joint effects of guano-harvesting and seals, to address the penguin-fish interaction problem, which is the purpose of this particular model, inclusion of density dependence is not essential. Therefore further analyses intend to omit the term concerned from the reference case, though variants on that assumption may be considered in sensitivity tests. As regards the equilibrium assumption used for starting the populations modelled in 1986, further estimable parameters could be introduced to reflect initial dis-equilibrium, but generally these cause only brief transient effects and make little difference to the results of particular interest.

Food

Simply adding sardine biomass to anchovy biomass (equation 14), without taking into account relative food value to penguins, is an oversimplification as the model assumes that the food value to penguins of 1 kg anchovy = 1 kg sardine, whereas the calorific value of anchovy is less than that of sardine. Also, capture effort for multiple anchovy is likely to be substantially higher than capture effort for a single sardine, further imbalancing the energetic value of different prey species to African Penguins.

15) We would welcome taking this point into account in the model, and request the provision of a prior reflecting biologists' assessment of the overall relative value of the same mass of sardine and of anchovy to a penguin (P7). Nevertheless note that it was addressed as a sensitivity (see PEL/35 pg 41, test 1) and made little difference to results (see pgs 43 and 44).

African penguins have a limited foraging range when breeding. Therefore, large extractions of fish in the immediate vicinity of colonies may have substantial impacts on penguins, including on breeding success, immature and adult survival, recruitment of first breeders to colonies and hence breeding participation. The model does not account for the impact of localized extractions of fish on penguins.

*16) That is not the central focus of this particular model. The matter has though been examined empirically, and the most recent and thorough analysis (MCM/2010/SWG-PEL/Island Closure Task Team/19) came to the conclusion that "Likely the most that could be said with some confidence is that the results of the analyses do **not** support the hypothesis that suspending fishing around Robben and Dassen Islands would enhance penguin reproductive success there.", suggesting that this aspect is not an immediate priority for future extensions of the interaction model.*

Not all demographic parameters for penguins in stratum B need be influenced only by biomass in stratum B – e.g. breeding participation, which is likely in part to depend on feeding conditions during post-moult fattening and may take place at some distance from breeding localities. For example, it is likely that at least some penguins from west of Cape Point have moulted to the east of Cape Point (e.g. Wolvaardt *et al.* 2008, Crawford *et al.* 2010). This also has implications for fitting numbers of birds observed moulting at islands to model estimates of numbers moulting (e.g. Fig. 1).

17) We would welcome provision of a prior from biologists on the likely relative importance of sardine and anchovy in stratum C to those in stratum B to penguin dynamics (if felt necessary separately for reproductive success cum juvenile survival and for adult survival) (P8). See also response 11) above.

The section headed "Annual variation in adult survival and reproductive success" only includes equations for survival. There is no further mention of reproductive success after the first sentence of the section. Is this an oversight, how was variable reproductive success dealt with?

18) There is no oversight. The parameter S^j includes both effects and their variation over time is accounted for through the random effects parameter μ . See also response 9) above.

Food availability is likely to affect juvenile survival (and therefore recruitment into the breeding population) more significantly than the survival of (more experienced) adults and even breeding success. Annual survival is consistently lower for juvenile penguins than for adult penguins. In Namibia, juvenile

recruitment appears to be highly sensitive to food availability. Therefore, the role of juvenile survival and recruitment in penguin population dynamics needs to be given more prominence in the model. Juvenile survival and recruitment is not only influenced by food in the direct vicinity around an island and the food biomass and area considered in the model is therefore not adequate to model the fate of penguin populations. We acknowledge that due to their mobility, population parameters for juvenile penguins are difficult to measure, but survival estimates exist (e.g. Whittington 2002, Kemper's PhD thesis 2006). Thus, if reproductive success, defined as encompassing breeding effort, fledgling success AND 1st year survival is modeled, as well as the population of non-breeding/moulting penguins, then the extent of the area where food availability significantly influences the penguins' population dynamics should be far bigger than "Spawner Stratum B" (MAP A2). Also note that the impact of juvenile population dynamics may be more marked in the Western Cape than in Namibia, because the juvenile phase is, on average, six months longer in the Cape (Kemper and Roux 2006).

19) The model defers to the data to estimate the relative prominence appropriately given to reproductive success cum juvenile success and adult survival (the S^j and S parameters). Re choice of area, see response 17) above.

In relation to the river model, it is questioned why breeding success is related to the fish in area B either the November before the breeding season or after the November after it, because it bears no relation to the fish (recruits) in the river during the breeding season (winter). Further,, the model needs to account for the high uncertainty around abundance estimates of fish.

20) There seems to be some misunderstanding here. First, the introduction to PEL/48 says that it relates to PEL/35, but a simplified implementation of the "River model" was only recently presented to the Island Closure Task Team. Secondly, the River model considers the anchovy recruitment biomass present on the west coast during the April-September period. This is calculated from results for the overall anchovy assessment model which estimates recruit numbers in November of the previous year, and then projects these forward under the mortality levels that apply to the period of interest (MCM/2010/SWG-PEL/Island Closure Task Team/10). This work has been extended to take within season variability into account, with no qualitative difference to the earlier conclusions arising as a result (MCM/2010/SWG-PEL/Island Closure Task Team/20).

Immigration

Numbers of birds breeding at islands will be influenced in the long term by emigration and immigration by recruits, not by established breeders which are site-faithful (e.g. Crawford *et al.* 1999) but this seems not to be accounted for in the model. Modeling of regional populations will be less dependent on such emigration and immigration.

The modelers apparently looked at the possibility of including immigration from Dyer Island to Robben and Dassen Islands, and we understand that taking this into account did in part solve their problem. However, we understand they abandoned this again to make the model simpler? There may not be sufficient banding data to obtain an independent estimate of these movement rates.

21) There are trade-offs involved here, because data available for the region as a whole is not generally as comprehensive as those for Dassen and Robben – in particular estimates of population trends (which are fundamental inputs for reliable modelling results) are poor overall compared with the two islands

modelled in the interaction paper. In the course of time the model may be extended to involve more colonies, but for the moment the priority is to address the issue along the lines also referenced in response 5) above. Earlier advice to us from penguin biologists downplayed any likely major role for emigration and immigration.

The authors subtly imply that the increase of the penguin population in the Western Cape is due to 2000 pairs moving from Namibia to the Western Cape. This statement is based on a sample of 14 banded birds which were observed to have migrated to South Africa from Namibia and is therefore highly speculative. What proportion of the number of banded penguins in Namibia do these 14 immigrants constitute?

22) That comment must be read in context. Its intent was not to suggest that immigration from Namibia had made a major contribution to the increase, but rather that immigration from the Eastern Cape certainly had not.

Moulters

The projections for Dassen Island will be heavily dependent on assumptions made concerning several crucial parameters, such as trends in the proportion of adults that moulted around the island's coastline and full breeding participation. For example, the ratio of potential to active nests at Dassen Island has increased markedly since 2000 (Crawford *et al.* 2010).

Some birds now moult inland at Robben Island (see p. 5).

23) In respect of both paragraphs above see response 11) above.

Immature penguins moult for the first time after being at sea for some 12 months or longer. Therefore, the number of immature birds moulting should be related to the number surviving to the end of the year after that in which they fledge, i.e. equation 9 should relate N^{juv} to $N_{y-1,1,j}$ and additionally multiply the left hand side by $S_{y,i}$, assuming adult survival after the end of the year in which chicks fledge.

24) Please see Figure A.3 from PEL/35. The year notation may be a little confusing, because the model year y runs from April of calendar year y to March of calendar year $y+1$ (see bottom of pg 2 of PEL/35). There is a typo in equation (9) of PEL/35 where S^j should be S . Calculations have been performed correctly, using S here instead of S^j , consistent with equation (2) of PEL/35. See also response 30) below.

It should be noted that population estimates based on moult counts should not be seen as absolute values, but rather as indices, which are still valuable for evaluating trends.

25) They are treated as relative counts in the model – see response 11) above.

Statistical concerns

There is concern that the model fails to use standard statistical model selection procedures. How many parameters are fitted in relation to the number of data points is not documented, nor are the standard estimates of the parameters, nor whether they are formally statistically significant.

26) See response 1) for context. As stated in document PEL/35 CI estimates were approximate being based on Hessians, and were provided with an illustrative intent only. We have not focused on perfecting these specific estimates, because the work has progressed from that analysis and eventual results will naturally be more rigorously presented.

The analyses referred to as GLMs (which are not true GLMs) are questionable in their usefulness as fitting a model with more parameters than residual degrees of freedom is not recommended practice. In addition, any positive correlation is more likely to be a breakdown of the assumption that colonies are synchronous, rather than that fisheries benefit penguins. It would be equally valid to conclude that penguins help the fishing industry, perhaps by scaring the fish into their nets.

What are the degrees of freedom of the t-tests performed? Were one-sided or two-sided tests performed? Two-sided tests should have been selected to argue significance in both directions.

27) It is unclear as to what the last two paragraphs are referring. The introduction to PEL/48 says it relates to PEL/35, but some of these comments appear to relate to quite different analyses addressing another issue in (MCM/2010/SWG-PEL/Island Closure Task Team/19). Some of the models in this last document do have relatively few degrees of freedom, but that is a frequent occurrence in this field given data paucity, and the best inferences possible still need to be drawn in these circumstances. Where necessary in related computations, REML approaches have been used to take due account of MLE bias. See also the conclusion of this document reproduced in response 16) above, which is not that reflected two paragraphs above. As far as the model of the paper under review, PEL/35, is concerned, the single key model selection was based on AIC in circumstances where the addition of a few fully free parameters was under consideration. The number of parameters given there is misleading, because most are not free but random effects "parameter", and properly results should be integrated over the distributions for those parameters so that with variance prespecified they do not constitute real addition of further parameters. However software to effect such integration was not available at the time, and has only recently become so. But in any case, with a move towards fully Bayesian estimation, the intention is to use the DIC model selection criterion where necessary.

The model fits (Figs 1 and 2) show substantial serial correlation of residuals. It would be useful to compare alternative models. The problem is intrinsically a time-series problem, and autocorrelation between successive values of the observations needs to be taken into account. At the absolute minimum, all models ought to be checked to determine whether there is autocorrelation between successive residuals. This is clearly the case in, for example, Figure 2, where the successive residuals between 1995 and 2002 are all negative (and the residual autocorrelation would be positive). These are also the years in which adult survival is forced to the unrealistic boundary survival rate. In order to fit the data, the proposed model would require survival rates larger than 1.00.

28) At later stages in modelling, the likelihood will be adjusted to take account of autocorrelation effects which are common in such marine data, because this is necessary for reliable estimation of precision. There are however different interpretations of the specific more systematic pattern referenced here. The reverse of the residual pattern for Dassen Island referenced in Figure 2 is evident in Figure 1. The model is trying to fit two dataserie (moult and breeder counts) which are in conflict, and this points to a possible need to consider a temporal component for the estimated q^M parameter for Dassen Island (see also response 11) above).

The reason for this string of negative residuals is simple. The model fails to take account the fundamentals of simple penguin biology. The concept of a regular single breeding attempt per year, with a maximal output of 1.3717 chicks per pair per year, is laughable – firstly, because no one believes this number to four decimal places, and secondly because penguins do not breed like clockwork, and this assumption fails to take account of the reality that penguins may breed successfully more than once a year in years of good food availability. Furthermore, the model assumes that the initiation of breeding is independent of food abundance. It is extremely likely that, when food abundance is good, birds commence breeding at earlier ages.

29) See response 7) and a number of others above. Attractive though this explanation might appear, it does not reconcile with information previously provided to us by penguin biologists, including one of the authors of PEL/48.

The models at equations (12) and (13) need a proper justification. They constitute a logit transform, and could be written far more simply. There is no justification offered of this particular functional form, nor of the parameterization. There are other alternatives that could be considered. The use of the terms “constants” and “random effects” in the sentence below equation (13) shows a misunderstanding of standard statistical terminology. Even within the current formulation, why has food abundance only in stratum B been taken account, when the penguin biologists are saying that food in stratum C is important for penguins in preparation for moult and for recovery after moult? On top of all this, the pelagic abundance index in equations (12) and (13) is indexed to the wrong year to be biologically meaningful. It needs to point to the previous year, or at least consider this as one of a family of biologically plausible models to be evaluated, and selected between.

30) Again see response 1), and also response 17). The logit transformation was used at the time as one simple means to ensure demographic bounds were respected. As explained in PEL/36, a priority for future work is checking robustness to such assumptions. Regarding the linkage to pelagic abundance, please note the year convention used and the fact that biomass estimates refer to November of the year specified – see response 24) above – hence the association is effected as the commentary above suggests. However we do now consider that the year labelling convention used in the model would better be changed for a variety of reasons to a calendar year. These are:

- *Less room for misunderstanding.*
- *Tag-recapture data may be incorporated at a later stage, and we understand these to be based on a calendar year.*
- *At present both moult and breeder counts are compared with the same model estimate of abundance, whereas the peaks for these two counts are about four months apart (December and April respectively). In a revision we plan to compare observations to model numbers corresponding to these specific months.*

Recommendations

Important note

The conclusions drawn concerning the impact of lessened catch on the penguin population are probably biased through considering only the two major colonies that have been worst affected by the eastward shift in prey resources.

31) *The argument for necessary bias is unclear. Data are best for these two colonies, plus the eastward shift provides more data contrast which enhances the possibility to detect the effect hypothesized.*

1. We recommend that the points of concern raised in this document be seriously accounted for in the next versions of the penguin-pelagic fish interaction modelling process undertaken by MARAM. Time series of input data need updating as per the values supplied to the Closures Task Team in July 2010. In addition to incorporating the new data, the model needs to be reformulated to account for the concerns expressed above.

32) *This is broadly in line with proposals we had already made in PEL/36. The question of exactly what data to best use will be addressed in another document currently under preparation.*

2. Overall, the presentation of the model results is too brief to enable assessment of the model according to normal statistical practice. Model results are not presented in a way that justifies the model in relation to alternatives that were considered (model selection requires statistical treatment whereas here we are presented with a single model and asked to accept this at face value). Thus, it is recommended that statistical models produced for consideration by the Pelagic SWG be subjected to a comprehensive in-depth investigation by a task force of appropriate qualified statisticians with expertise in model stochasticity and relevant expertise in modelling and model selection (for example, Professor Walter Zucchini, University of Gottingen, Germany, and Professor Ian MacDonald, University of Cape Town). This task should not be added to the already heavy burden carried by the review panel that oversees the stock assessment models.

33) *See response 1). Again the question of investigating robustness that underlies these comments is central to future work already planned, as has been advised to the PWG on numerous occasions, including in PEL/36. Various deadlines, including that for the revision of the pelagic fishery OMP during 2011, render the review planned for the end November international workshop the last opportunity for such an exercise, and the panel members are all more than competent in statistical model selection procedures and their application in marine resource population dynamics models. Given that Prof MacDonald is local, the opportunity will be taken to invite him to the pertinent sessions of this workshop as his input would certainly be appreciated.*

SUMMARY OF QUANTITIES FOR WHICH BIOLOGISTS ARE REQUESTED TO PROVIDE PRIORS

Note: For practical purposes, these can be as simple as advice of two limits between which a uniform prior will be assumed.

P1: Period and extent of possible emigration from Dyer Island to Robben Island.

P2: Extent of relative undercounting of juvenile moulters compared with adult moulters.

P3: Standard deviation of time-variable effect relating to detection of juvenile moulters.

P4: Variable age at first reproduction

P5: Maximum number of fledglings per pair per year

P6: Proportions of moulters and breeding pairs counted at Robben and Dassen Islands, and their possible time dependence

P7: Relative value of equal mass of sardine and anchovy to a penguin

P8: Relative importance of sardine and anchovy in stratum C to those in stratum B to penguins.

SUMMARY OF ISSUES FOR FUTURE ATTENTION, IN DRAFT ORDER OF PRIORITY

Note that the ordering here takes account of degree of technical difficulty in implementing as well as likely impact on results.

- 1) Change of year convention – see response 30) above (confirmation from penguin biologists on the appropriateness of the months specified to categorise different counts would be appreciated)
- 2) Detectability of juvenile relative to adult moulters
- 3) Simulation testing of the estimation procedure to provide estimates of the impact of pelagic TACs on penguin population growth rates
- 4) Robustness to alternative functional form for reproduction-cum-juvenile survival and adult survival relationships to prey abundance in the stratum of interest
- 5) Robustness to different error models for the relationships in 4)
- 6) Adding immigration (from Dyer Island)
- 7) Effect of according different benefits for penguins to the same mass of anchovy and of sardine consumed
- 8) Incorporation of tag-recapture data
- 9) Prior on the upper bound for fledging success
- 10) Different forms and extents of density dependence
- 11) Variation in time in the detectability of juveniles relative to adults in the moult count
- 12) Variation in time in the age at first reproduction
- 13) Extension of model to include Western Cape colonies other than Robben and Dassen Islands.

(New) References

Butterworth DS, Plagányi É E. 2007. Why plots of results for the simple penguin model of SWG/EAF/SEABIRDS/13APR/01 are misleading, and some broader consequential implications. Document MCM/2007/MAY/SWG-PEL/11.

Butterworth DS, de Moor CL. 2010a. A very simple implementation of the “River Model” to estimate the impact of fishing on the amount of anchovy available to West Coast penguin colonies. Document MCM/2010/SWG-PEL/Island Closure Task Team/10.

Butterworth DS, de Moor CL. 2010b. An extension to the simple implementation of the “River Model” to estimate the impact of fishing on the amount of anchovy available to West Coast penguin colonies which takes account of within season variability in recruitment. MCM/2010/SWG-PEL/Island Closure Task Team/20.

Butterworth DS, Robinson W. 2010a. Further responses to comments in Altwegg MCM/2010/SWG-PEL/23 regarding penguin analyses. Document MCM/2010/SWG-PEL/27.

Butterworth DS, Robinson W. 2010b. Summary of progress and plans for penguin-related analyses, with reference to comments made on previous submissions. Document MCM/2010/SWG-PEL/36.

Robinson W, Butterworth DS. 2010. GLMs relating penguin demographics to pelagic catches close to islands and to pelagic abundance. Document MCM/2010/SWG-PEL/Island Closure Task Team/19.

Robinson W, Plagányi ÉE, Butterworth DS, de Moor CL. 2010. Summary of penguin–pelagic fish interaction modelling during 2008. Document MCM/2010/SWG-PEL/35.

Robinson W, Butterworth DS. 2010 Penguin survival estimates from tagging data using a multinomial likelihood. Document MCM/2010/SWG-PEL/??

Shannon L, Altwegg R, Crawford R, Jarre A, Kemper J, Roux J, Underhill L, Waller L, Wanless R. 2010. Suggestions for improving the penguin modeling framework developed by William Robinson, Éva Plagányi, Doug Butterworth and Carryn de Moor and expression of several concerns by outside, interested parties. Document MCM/2010/SWG-PEL/48.

(Original) References

- Altwegg R, Barham BJ, Barham PJ, Crawford RJM, Dyer BM, Geldenhuys D, Leshoro TM, Makhado AB, Parsons N, Pichegru L, Sherley R, Underhill LG, Upfold L, Visagie J, Waller L. 2009. Report on monitoring of African penguins at colonies around which purse-seine fishing was disallowed or permitted in 2008, with early information for 2009. MCM/2009/SWG-PEL/16: 1–49.
- Crawford RJM. 2007. A simple model of the African penguin population. SWG/EAF/SEABIRDS/JAN07/xx: 3 pp.
- Crawford RJM, Altwegg R, Barham BJ, Barham PJ, Durant JM, Dyer BM, Geldenhuys D, Makhado AB, Pichegru L, PG Ryan PG, Underhill LG, Upfold L, Visagie J, Waller LJ Whittington PA. 2010. Collapse of South Africa's penguins in the early 21st century: a consideration of food availability. *MCM/2010/SWG-PEL/Island Closure Task Team/xx*: 1–37.
- Crawford RJM, Boonstra HGvD. 1994. Counts of moulting and breeding jackass penguins – a comparison at Robben Island, 1988–1993. *Marine Ornithology* **22**: 213–219.
- Crawford RJM, Boonstra HGvD, Dyer BM, Upfold L. 1995. Recolonization of Robben Island by African penguins, 1983–1992. In: Dann P, Norman I, Reilly P (eds) *The Penguins: Ecology and Management*. Chipping Norton, Australia: Surrey Beatty. pp 333–363.
- Crawford RJM, David JHM, Shannon LJ, Kemper J, Klages NTW, Roux J-P, Underhill LG, Ward VL, Williams AJ, Wolfaardt AC. 2001. African Penguins as predators and prey – coping (or not) with change. *South African Journal of Marine Science* **23**: 435–447.
- Crawford RJM, Shannon LJ, Whittington PA. 1999. Population dynamics of the African penguin *Spheniscus demersus* at Robben Island, South Africa. *Marine Ornithology* **27**: 139–147.
- Crawford RJM, Underhill LG, Coetzee JC, Fairweather T, Shannon LJ, Wolfaardt AC 2008. Influences of the abundance and distribution of prey on African penguins *Spheniscus demersus* off western South Africa. *African Journal of Marine Science* **30**: 167–175.
- Crawford RJM, Underhill LG, Upfold L, Dyer BM. 2007. An altered carrying capacity of the Benguela upwelling ecosystem for African penguins (*Spheniscus demersus*). *ICES Journal of Marine Science* **64**: 570–576.
- Wolfaardt AC, Underhill LG, Altwegg R, Visagie J. 2008. Restoration of oiled African penguins a decade after the *Apollo Sea* spill. *African Journal of Marine Science* **30**: 421–426.