Response to Comments on : Robben Island penguin survival rates from Bayesian analysis by William Robinson and Doug Butterworth (Altwegg: MCM/2010/SWG-PEL/08)

Doug Butterworth and William Robinson – MARAM, UCT

16 February, 2010

We thank Res Altwegg for his speedy and helpful reply to our original submission (MCM/2010/SWG-PEL/07). For ease of cross-reference, portions of Res Altwegg's original text are reproduced in italics below, with our responses following.

Doug, Will and I are on the same page about the need for quantitative models to address this problem, and I commend them for taking this forward. However, at the last SWG meeting on 13 January I felt that Doug made claims about his and Will's modeling approach that I felt were not warranted. The first was that they are fully accounting for all sources of uncertainty. An important source of uncertainty in this kind of work is structural uncertainty; the uncertainty whether ones modeling approach captures the system adequately. My first comment at the meeting therefore was to point out that by using one approach only, this type of uncertainty is not easily quantified. A different approach might well give rather different results. We should at least consider one alternative approach to be reasonably confident, and with others I am working towards one.

There is a semantic confusion here. Any claim made would have referred to uncertainty within the structure of the model under consideration. There is wide recognition in MSE of the need to explore the consequences of structurally different models (as such differences are usually the greatest source of uncertainty), and there are no problems with alternative models being put on the table which are also reasonably able to reflect the available data, so that their implications can be evaluated. One of us made this point at the 13 January meeting.

My second comment at the meeting related to Doug's claims that the models were showing no appreciable effect of fishing on the penguin population. I pointed out that there were still problems with their model, and this is the criticism Will and Doug are addressing here. I am not convinced they have solved the problem, however.

Again it appears that responses made at the 13 January meeting were not fully grasped. The model referenced in the context of an absence of appreciable (negative) effect of fishing on penguins is not the model which MCM/2010/SWG-PEL/07 considers – that model was developed for a different purpose, *viz.* as a framework to allow for testing the power of future tagging to estimate effects on penguins of suspending fishing in the vicinity of breeding colonies. A structurally different model as set out in MCM/2008/DEC/SWG-PEL/27 is the one that has been used to evaluate the impact of lower sardine and anchovy TACs on penguin populations and suggested little effect. But that too was not the one that was being referenced in the arguments being made by one of us on January 13. Those arguments referred rather to the final conclusion mentioned in the abstract of EAFWG/OCT2007/STG/04, which is reproduced below:

> Historical data on fledging success and the breeders per adult moulter ratio for penguins at the Robben and Dassen island colonies are analysed using General Linear Modelling (GLM) to provide the basis for a simulation evaluation of the power of an experiment which closes a neighbourhood (about 20 km radius perhaps) around Robben Island to detect a statistically significant (5% level) impact of the (assumed multiplicative) effect of fishing on penguin reproductive output. Process error for each index is estimated to be about 20%. For an actual effect size of 30% (corresponding to a change in the population growth rate of about 3% p.a.), the experiment would have to continue for some 10 years for a better than 50:50 chance of such detection for the breeders per adult moulter index, and at least 20 years for fledging success. An initial GLM analysis to determine the direction and magnitude of the effect of pelagic catches on penguin reproductive output in terms of the two indices and colonies considered suggests that these catches enhanced this output in all four instances (a potential consequence of the hypothesis that fishing scatters large shoals, thus increasing the probability of natural predators finding their prey), though none of the estimates is statistically significant at the 5% level.

I did point out at the meeting (and during a telephone conversation with Doug in early December) that the model they presented returned adult survival estimates of 0.96 (in the previous document, and here it is 0.98), which was the (arbitrary) upper limit they set in their model. This is a problem in itself, and Will and Doug are now addressing this. Much more importantly, however, I fear it is a sign of more fundamental problems with the model. This is what I am really worried about and apparently I did not make myself clear. Clearly, for many of the years in question, the model did not estimate an adult survival rate, but rather just returned the set limit for this parameter. This to me is a sign of inadequate information in the data to estimate this parameter. Just looking at counts, it is difficult to tease out whether a population changes because of a change in survival or reproduction, or both. This may well lead to a ridge in the likelihood, parameters that trade off, and one of them ends up at the boundary of the allowable parameter space. This is what I'm worried has happened in Will's and Doug's approach (the same problem in various versions of their penguin model). I understand that the model uses the information on the proportion of young birds in the moulter counts to estimate reproduction. Maybe this information is not enough to reliably separate survival and reproduction. If it turns out that the likelihood is not flat, another possibility leading to the observed outcome would be that the proportion of young birds in the moulter counts is biased low as a measure of reproduction. To compensate, the model would then push survival to the boundary. If moulting young penguins were less likely to be counted than moulting adults, e.g. if they tended to moult away from the main sites or if young birds can be mistaken for adults, then the proportion of young in these counts could not directly be used to measure reproduction. If that were the case, the model would have to account for different detectabilities of the *two age classes. This would require extra information, as for example from the capture-mark-resighting data.*

Certainly MLEs of parameters on the boundaries of their feasible range are indicative that the information content of the data is not substantial. But this is a problem that occurs often in marine mammal population modelling, and the general approach to handling it is well-known and regularly applied – specifically it is the Bayesian approach applied in MCM/2010/SWG-PEL/07, which takes appropriate account of the uncertainties mentioned above in developing posterior distributions for parameters of interest. That notwithstanding, a sequence of MLEs on the same boundary, or posteriors heaped towards that boundary, may be indicative of some factor of importance having been omitted from the model. In this case, it is probable that immigration in the early years (from Dyer Island?) may be "biasing" survival rate posteriors high for those years.

Regardless of what causes the adult survival estimates to be pushed to the boundary in this model, clearly no strong conclusion can be drawn from it. Especially problematic are conclusions that rely on adult survival being estimated correctly. The conclusion that fishing does not affect penguins critically depends on the estimate of adult survival. If adult survival is high, it will be the most important parameter for population dynamics, and if it is in the range of 0.96 or 0.98 it almost completely determines dynamics. Clearly, with adult survival effectively set at these values for most of the time, not much can happen to the population. I'm therefore not surprised that the model suggests fishing has no effect on penguins.

There is again a misunderstanding here. This model is not the one that has been used to draw inferences about the effect of fishing on penguins. But if the results from this model were to be used in that way, the full joint posterior distribution would be used, **not** the MLEs, so that uncertainties ("estimation" errors) are properly taken into account, and the results in MCM/2010/SWG-PEL/07 show that median survival rates for the earlier years are nowhere near the upper bound imposed, but typically about 0.9.

The present document treats the symptom without solving the problem. Fitting the same model with Bayesian methods does not solve the underlying problem. I note that the adult survival rates are now a bit lower and suspect that this is the effect of the prior pulling the values down from the boundary. Will and Doug are using a uniform prior. It is not stated, but I assume it is uniform between 0 and 1. This prior is not uninformative on the logit scale on which survival is estimated. The left diagram below shows 1000 random numbers drawn from a uniform distribution. The diagram on the right shows the distribution of the same values after logit transformation. The fact that priors can be uninformative on one scale but informative on another is a well-known problem in Bayesian analyses. This is a minor point, though. Using Bayesian methods to fit the same model doesn't solve the more fundamental problems I described above.

The point about the interpretation difficulties surrounding what constitutes an "uninformative prior" is well taken. In the original document only one such distribution was used, as the purpose was primarily illustrative. We show below survival results for the original distribution (a) and alternative distributions (b) to (d) in Fig. 1-3. The prior distributions tested are as follows:

(a)	Uniform on mortality rate *M*:	U[0.02; 1]
(b)	Uniform on mortality rate *M*:	U[0.04; 1]
(c)	Uniform on survival rate *S*:	U[0; 0.98]
(d)	Uniform on survival rate *S*:	U[0; 0.96]
(e)	Uniform on ln *M*:	U[ln 0.02; ln 1]

The relationship between survival rate and morality rate is $S = e^{-M}$.

For all cases, results are very similar with the most obvious differences noted in the figure captions. The original (a) prior should have had a higher upper bound, but it is clear from the results for the *S* priors (c) and (d) that this would not have made much difference.

Aside from issues such as possibly incorporating immigration, which has been done in earlier models more geared to addressing the impact of fishing, we do not see that fundamental problems have been raised – to the contrary the Bayesian approach illustrated in MCM/2010/SWG-PEL/07 is fairly standard in similar marine population assessments.

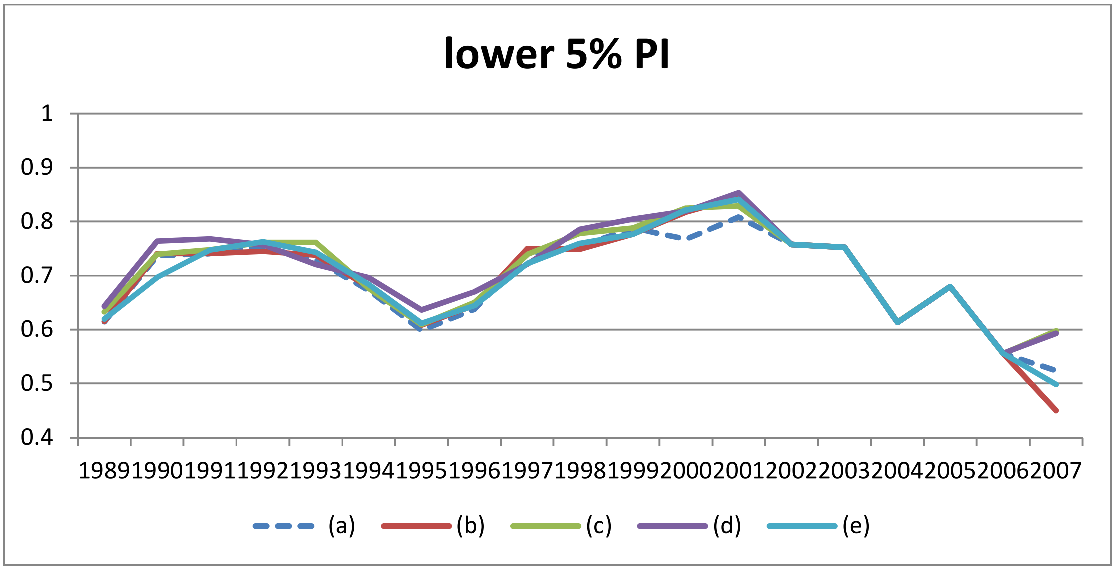

Fig. 1: At the lower bound of the 90% probability interval, the survival curve for each model sensitivity attained a slightly higher peak in the years 2000 and 2001 when compared to the original model (a), shown here with a dashed line.

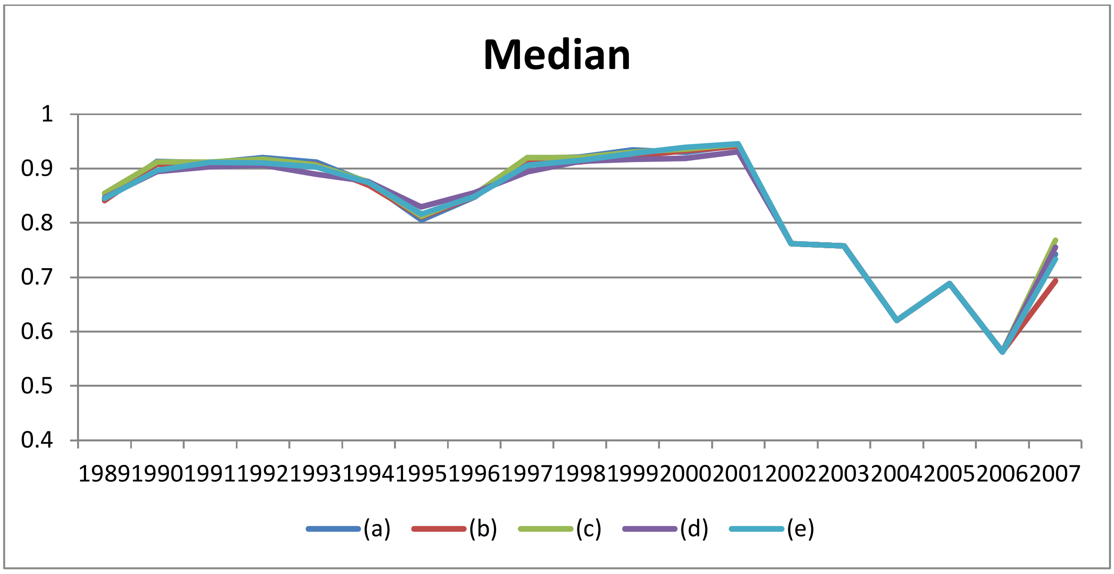

Fig 2: The median survival rates for all model sensitivities are practically indistinguishable from the original.

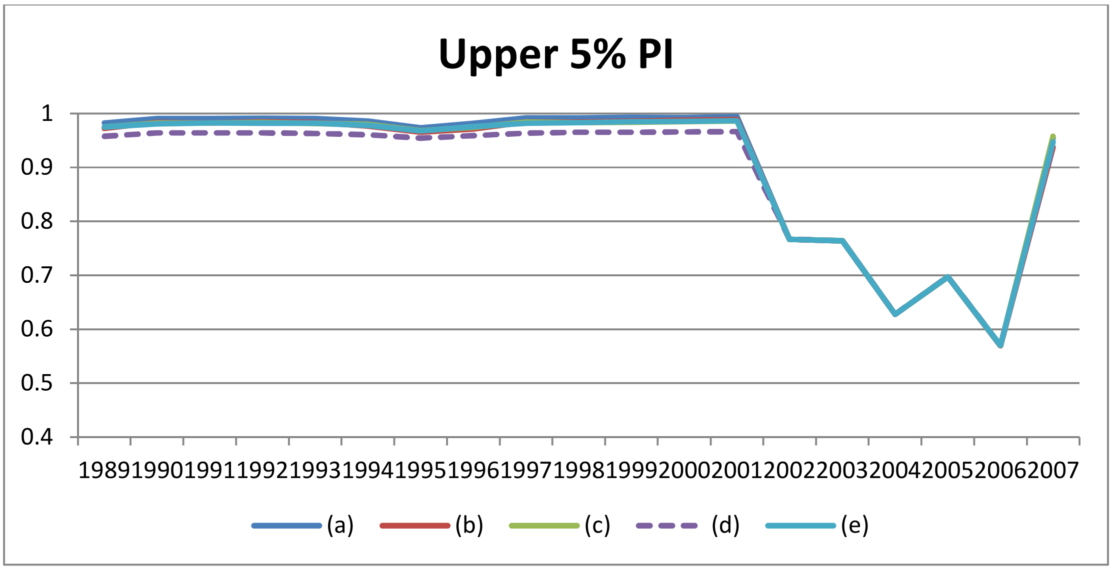

Fig. 3: At the upper bound of the 90% probability interval, the only noticeable difference from the original model, apart from the limit imposed by the edge of the prior, is that the survival curve for sensitivity (d) is slightly lower for the years up to 2001.