CHAIR'S OVERVIEW OF PROGRESS IN THE SIMULATION TESTING OF PENGUIN CLOSURE EFFECT RESPONSE ESTIMATORS AS REPORTED IN MARAM/IWS/DEC15/PengD/P1-4

Kevern Cochrane, Chair, Penguin Island Closure Task Team[1]

In the light of research recommendations in the 2014 Panel report, at a meeting held in August 2015 (MARAM/IWS/DEC15/PengD/BG6) a Task Team was set up to address the first two of these recommendations, i.e.:

> A1. (H). That simulations to evaluate bias in estimation methods be explored, which are conditional of the types of scenarios reflected in MARAM/IWS/DEC14/Peng/A10
>
> A.2 (H) Various elaborate models have been applied to test the effect of fishing on penguin demographics. More elementary analyses directly aimed at evaluating the questions and statistical power should be applied…

Documents MARAM/IWS/DEC15/PengD/P1-4 summarise the work carried out under the auspices of the Task Team and the agreed outputs from the associated computations.

The purpose of this document is to provide a brief "walk through" of those four documents to embellish the rationale for the choices of analyses pursued, to assist external readers follow the logic of this process as seen by the Task Team.

Note that the purpose of the simulation testing exercise recommended by the Panel was to quantify the estimation bias that may be present for various reasons when applying the estimators that have been suggested to assess the impacts of pelagic catches and closures of areas around islands to fishing on penguin reproduction-related response variables. Such bias is pertinent both to the evaluation of estimates of these impacts from the data currently available, and to the power of a future closure experiment to determine such estimates with statistical significance.

Before starting however, note that the data available to the Task Team are as listed in MARAM/IWS/DEC15/PengD/BG1. These comprise annual central estimates (usually means) of six different penguin response variables that have been collected at Dassen and Robben islands. Although data at an individual observation level have recently been made available, the Task Team focused initial attention on the analysis of aggregated data. Document MARAM/IWS/DEC15/PengD/P4 described below offers suggestions on how the approaches considered might be extended to take account of aspects that are further informed by the individual data.

[1]The Task Team consisted of M.O. Bergh, D.S. Butterworth, K.L. Cochrane (chair), T.L. Morris, R.B. Sherley and H. Winker. A. Ross-Gillespie undertook, on behalf of the Team, all the analyses and tests, under the supervision of D.S. Butterworth.

MARAM/IWS/DEC15/PengD/P1. Specifications for operating models to evaluate bias in estimation methods in accordance with recommendation A.1 of the 2014 International Review Panel

This document details specifications for the simulation testing process agreed by the Task Team.

Two different approaches are used to develop the operating models for this estimation. The sub-regional biomass surrogate approach of equation (1), which relies on the assumption that the abundance of fish available at Dassen and Robben islands each year will be highly correlated because the two islands are relatively close, and in estimation treats the associated year-effect as a random effect (mainly to increase the number of degrees of freedom). The second is the regional biomass approach of equation (9) which explicitly uses the biomass estimate from a survey, though once of much spatial coverage than the near vicinity of the two islands concerned. These approaches as used for analyses presented at the 2014 workshop, but now been extended to be able to consider both closure catch as covariates, instead of one or the other.

Following concerns expressed at the 2014 that the estimators used were applied to annual average values for the response variables without consideration of small size, an adjustment was agreed to the residual error formulation to be able to take this into account (equation (3)).

Given concerns raised during the 2014 workshop (see MARAM/IWS/DEC14/Peng/A10) that correlation between catch taken close to islands and local abundance could introduce bias into the estimators being applied, particular care was taken to be able to simulate pseudo-data with such an underlying correlation. That process is described on pgs 3-4 of the document, while Table 1 and Figure 1 show empirical linear regression of catches close to islands and survey biomass estimates which were used as a basis to choose (species-specific) values of this correlation for use in the simulations.

Equation (8) specifies the process used to include autocorrelation in the residuals of the pseudo-data generated. This follows a recommendation from the Panel in 2014 that such autocorrelation was evidently present, and needed to be taken into account in analyses. The basis underlying the autocorrelation values chosen is provided in document MARAM/IWS/DEC15/PengD/BG2.

The balance of the document sets out some other technical aspects of the pseudo-data generation process, as well as specifying the response variable series selected for initial analyses under the approach developed, together with some output statistics to be computed for pseudo-data generated to be able to check whether the generation code was addressing its intent adequately.

MARAM/IWS/DEC15/PengD/P2. Consolidated analyses produced in implementation of the approaches described in document MARAM/IWS/DEC15/PengD/P1

Note the sections below report on work over a period of weeks, and appear in the order the work was conducted, so that later sections build on the results in earlier ones.

Section 1: *Some initial results for the penguin simulation conditioning process*

Results of conditioning operating models for the simulation testing on the basis set out in MARAM/IWS/DEC15/PengD/P1. Note that this was a limited initial exercise to check that the processes set out in PengD/P1, and particularly that for dealing with possible difficulties when the catch size effect (λ) and the closure effect (δ) parameters were both estimated together, were working sensibly. Note that sample size weighting was not implemented in these computations.

Section 2: *Penguin process vs observation error*

Section 2rev: *Penguin process vs observation error - update with all available se values taken into account*

These sections provide comparisons of the sizes of the observation error contributions to the total error in the GLM/GLMMs. These showed that in almost all instances, the observation error made only a small contribution to the total error. Given that difficulties had arisen in implementing equation (2) of MARAM/IWS/DEC15/PengD/P1 to take observation error values into explicit account (as R did not offer the variance relationship required), these results led to the decision to continue assuming a constant error variance independent of sample size. The rev version simply takes account of further data that became available after the original Section 2 had been compiled.

Section 3: *Generating pseudo data – evaluating correlation for all years in which catch and biomass are available*

This document was produced in response to the wide distribution of catch-biomass correlation values for generated pseudo-datasets, which led to questions about the analytical formulation and coding used to generate this correlation. This section shows these distributions to be the result of the small sample sizes involved; in contrast use of a set of catch and biomass values over a longer period (the full set of data) yields the requisite unbiased results with variances compatible with the expectations from a jacknife estimate.

Section 4: *Generating pseudo data – some initial results*

First, this section shows diagnostic plots developed to check that the pseudo-data generation procedures were behaving as desired for the chick growth response variable and a catch-only based simulation model. The important results relating to estimation bias are shown in Table 4.2. These suggest a trend in bias as correlation between catch and biomass is increased that is in the direction indicated by the analytical contributions of MARAM/IWS/DEC14/Peng/A10 to the December 2014 workshop. There was, however, an indication of bias (in different directions for the two islands) when that correlation was set to zero – a result that seemed insensitive to different approaches to generating the pseudo-data. Table 4.2 showed results for a number of variants of the data generation process to investigate whether these zero-correlation biases were perhaps a consequence of the procedure adopted, but yielded little evidence in support of that. This matter is pursued further in Section 7 below.

Section 5: *A problem detected with the biomass surrogate approach*

At this stage of the process it was noted that in some instances the random effects model of equation (1) of MARAM/IWS/DEC15/PengD/P1 had produced estimates where the variance of these effects was zero. The immediate recommendation was that these cases were proposed to be taken off the table for immediate further analyses. This matter is discussed further in MARAM/IWS/DEC15/PengD/P4.

Section 6: *Possible further penguin analyses*

This document provided straw-man proposals for further analysis priorities in the light of the limited time remaining for the Task Team before the workshop with the Panel. The decision made was to give initial priority to the analyses related to the random effects model (equation (1) of MARAM/IWS/DEC15/PengD/P1). Section 8 reports initial associated results.

Section 7: *Evaluating "small-sample-size" bias*

The "zero-correlation" bias reported in Section 4 was suggested at that time to possibly be due to a "small sample size effect". Analyses reported in this Section extended the period considered for the zero correlation scenario, and found that bias then tended to zero, consequently indicating that there is indeed a small-sample-size bias for the estimator of equation (1) of MARAM/IWS/DEC15/PengD/P1.

Section 8: *Further runs for simulating pseudo data as recommended in Section 6*

Earlier results related to bias estimation are extended somewhat in this Section, broadly along the lines set out in Section 6, to include consideration of a further response variable (forage trip duration), different regions around the islands to define catches taken, closure only instead of or in addition to catch only as the estimator co-variate, and the impact of autocorrelation in the time series of residuals. These results, discussed in more detail in the section, show that the impacts on bias of some effects, for example the catch-biomass correlation, can be large, whereas those of correlation is the residuals seems small.

Section 9: *Testing a simple estimator*

This section extended analyses to consider closure as well as catch-based estimates for the first time, and also includes an initial attempt to address the 2014 Panel recommendation A.2 (H) (see above) that more elementary analyses directly aimed at evaluating the questions and statistical power should be applied. Some difficulties of interpreting results, e.g. for a closure-based estimator applied to pseudo-data generated by a catch-based operating model, and are amongst those raised for further discussion in MARAM/IWS/DEC2015/PengD/P4.

MARAM/IWS/DEC15/PengD/P3. Penguin power analysis methodology - draft

This document describes how the methodology of MARAM/IWS/DEC15/PengD/P1 is extended to generate pseudo-data 20 years into the future to allow the estimation power of an island closure experiment to be evaluated. The effect size is taken as the point estimate for the response variable/estimator combination under consideration, except that it is set to a higher threshold value if that estimate is smaller than would be biologically meaningful. The document also explains the basis for arriving at the specification of that "biologically meaningful" criterion, which is related to the extent of the change in the penguin population growth rate which would follow. The results of four examples of applying the process are reported.

MARAM/IWS/DEC15/PengD/P4. Simulation testing of penguin closure effect response estimators – where next?

This document sets out the issues arising in taking existing computations forward (those reported in MARAM/IWS/DEC15/PengD/P2 on the basis of the framework set out in MARAM/IWS/DEC15/PengD/P1) to address bias estimation, ultimately for use in adjusting estimates of penguin closure response parameters. The document starts by listing the various factors/scenarios under consideration, and goes on to set out a number of issues that arise in both the simulation testing exercise itself and in the interpretation and use of the results.

It is envisaged that the workshop will step through this document point by point, also considering related aspects the Panel may raise, to obtain the Panel's advice on choices amongst some of the possible courses of action that could be pursued in finalising the overall analyses of this issue in the new year.