SIMULATION TESTING OF PENGUIN CLOSURE EFFECT RESPONSE ESTIMATORS – WHERE NEXT?

Penguin Island Closure Task Team[1]

The purpose of the simulation testing exercise recommended by the Panel at the 2014 workshop is 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. In other words, there is a need to know estimation bias both now, and as it will change given further data in the future (the computations reported in MARAM/IWS/DEC15/PengD/P2 show that this bias depends on the scenario in question (response variable, distance from the island within which catches are considered, etc.), and also changes as further data become available.

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 such bias estimation, particularly with a view to obtaining the Panel's advice on choices amongst some of the possible courses of action that could be pursued. These choices relate, *inter alia*, to methods of estimation, methods for generating data for simulation-based evaluation of bias, and interpretation of the results particularly in circumstances where conclusions have to take account of the large number of plausible scenarios which may introduce bias into the estimation. It is intended to adopt such choices for the methodology to be used to take this work to completion in the new year.

There is no intention to limit such issues to the ones listed below when this document comes under discussion at the workshop. The Task Team did not itself have sufficient time to fully discuss the contents of this document, particularly its last few sections. Thus Panel members should feel free to introduce further separate issues related to this overall objective of bias estimation into the discussion should they see fit.

1) Options for scenario specification

More details on the options listed below can be found in MARAM/IWS/DEC15/PengD/P1 and MARAM/IWS/DEC15/PengD/P2. The figures in parenthesis after each sub-head denote the associated number of options.

[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.

Response variables (6)

Data are available for six response variables. These data are listed in MARAM/IWS/DEC15/PengD/BG1.

Islands (2)

Two (nearby) islands with penguin colonies are under consideration: Dassen and Robben (though note that most of the estimators considered evaluate them jointly).

Conditioning and Estimation models (2 x 3))

There are two fundamental approaches considered:

i) The sub-regional biomass surrogate approach (a GLMM model)
ii) The regional biomass approach (a GLM approach)

(see respectively equations 1 and 9 of MARAM/IWS/DEC15/PengD/P1.

Within each model type, there are three variants:

a) A catch only based variant
b) A closure only based variant
c) A variant incorporating both catch and closure effects.

Associated questions include:

I) Whether this provides a sufficiently wide range of conditioning models?
II) Whether further estimators (not necessarily of the same type as the conditioning models) also merit consideration?

Allowance for sample size in estimator (2)

This concerns whether or not equation (2) of MARAM/IWS/DEC15/PengD/P1, which makes allowance for sample size in the formulation for the residual variance, is implemented.

This raises a difficulty for the GLMM biomass surrogate approach of equation (1), as R does not include such a variant, so that coding from scratch would be required. A related question is whether years with only a very low sample size (e.g. 3) should simply be omitted? Note the results in section 2 of MARAM/IWS/DEC15/PengD/P2 showing the relative contributions of observation error variance to total variance which are pertinent to these questions. Note also, however, that observation error variance estimates are not available for all response variables under consideration.

Fish species considered (2)

Anchovy and sardine.

At the simplest level, this implies considering only the one or the other. Naturally penguin responses may reflect a reaction to some combination of the two, and in the past, following discussions in the Pelagic Working Group and with the Panel, one approach to such combination in which models were based on the sum of the biomasses and of the catches of both have also been considered. There is room for more discussion on this matter.

Areas considered around islands to define catches (4)

Block based distances of 10, 20 and 30 nm (C10, C20 and C30), and the area formally closed in the experiment (taken to be a circle or radius 18 km – $C_{closure}$).

Allocation of catches given closure to areas outside (2)

Options i) and ii) are detailed following equation (7) of MARAM/IWS/DEC15/PengD/P1.

Biomass series considered (3)

(See, e.g., Figure 1 of MARAM/IWS/DEC15/PengD/P1.)

i) Anchovy – May-June survey
ii) Sardine – May-June survey
iii) Sardine – Oct-Dec survey of previous year

Catch-biomass correlation (3 – differ by species)

See MARAM/IWS/DEC15/PengD/P1 above equation (6), and particularly also Figure 1, for the bases for these choices.

i) Anchovy – 0, 0.2 and 0.4
ii) Sardine – 0, 0.4 and 0.6

Autocorrelation in residuals (3)

Values of 0, 0.2 and 0.5 (see following equation (8) of MARAM/IWS/DEC15/PengD/P1).

The basis for these choices originates in the analyses reported in MARAM/IWS/DEC15/PengD/BG2.

Biomass and catch autocorrelation (2)

This is yet to be incorporated into computations, but would introduce a further factor affecting estimation bias and precision. The values, whose computation for catches excluded years where the island concerned was closed to fishing, are:

	Anchovy	**Sardine**
Biomass (May-Jun)	0.56	0.48
Catch Dassen 10nm	0.43	0.25
Catch Dassen 20nm	0.25	0.39
Catch Robben 10nm	0.54	0.69
Catch Robben 20nm	0.20	0.44

Note that when catch-biomass correlation is present, only the biomass values could be usedb when generating pseudo-data, though if that correlation is set to zero, the values for catch as well as biomass would both become pertinent.

2) Data to be used

Results reported in MARAM/IWS/DEC15/PengD/P2 have not all used the most up-to-date data on unstandardized annual means (and se's where available).

Final analyses will:

a) Use up-to-date data.
b) Rather than necessarily use nominal means for the value for each year, where pertinent these will be GLM-standardised for a month (or half-month) co-variate, as the data indicate to be appropriate.

3) Conditioning issues

Is the procedure set out immediately preceding equation (2) of MARAM/IWS/DEC15/PengD/P1 for dealing with questionable estimates of the catch and closure effect parameters when estimated jointly appropriate (see the results of applying this procedure, which are reported section 1 of MARAM/IWS/DEC15/PengD/P2)?

4) Issues related to generating pseudo-data for simulation testing

Table 2 of section 4 of MARAM/IWS/DEC15/PengD/P2 sets out five different approaches for generating pseudo-data. Results are reported in that Table and shown in subsequent Figures.

Approach (A) of that Table, corresponding to the specifications in MARAM/IWS/DEC15/PengD/P1, has thus far been used as the default for further calculations. Is that adequate/sufficient?

5) Procedure for adjusting initial estimates for bias

If: *a* is the original estimate of the catch/closure parameter used to condition the simulations for the scenario under consideration, and

b is the mean of the estimates obtained from the pseudo data, then

$c = a + (a-b)$ is currently being taken as the bias-adjusted estimate.

Is this approach (which assumes translation invariance) appropriate, or alternatively should *a* be adjusted on a trial-and-error basis towards a new conditioning value a^* - this process would continue until the corresponding $b^* = a$, with a^* then being taken as the bias-adjusted estimate?

For some scenario options (e.g. catch-biomass correlation), bias changes as the period for which further data become available increases. For a power factor analysis, it is impractical to re-evaluate bias for every extra year for which future data become available. Comments on the following procedure to address this issue are therefore requested:

i) Evaluate bias also for a 20 year extension to the series. If the bias-adjusted estimates after 20 years still fail to meet the significance requirement, no further consideration is necessary.
ii) However, if after 20 years this criterion is met, evaluate bias for the period after which the 80% power criterion is achieved. Iterate on such a process to ascertain the period after which (time-variable) bias-corrected estimates meet the significance requirement.

6) Aggregating results

There are potentially a very large number of scenarios to be considered, but certainly a full cross of all the options listed in section 1) above would be quite impractical. Advice is therefore sought on how best to "aggregate" over these options, considering issues that include the following:

a) There are three basic possibilities for dealing with multiple option factors as in the situation here: weighted averaging, considering some options only for single factor sensitivities, and considering robustness over a large number of combinations. Which of the factors in section 1) should be dealt with in which of these ways?
b) Should a goodness-of-fit criterion, such as AIC_c, be considered; further might an AIC weighting approach be used?
c) Should "expert judgement" based relative plausibility weighting across different scenarios be used, and what ground rules should apply to such a process?
d) Should scenarios be down-weighted/excluded from consideration if they result in situations for which there is consensus on a categorisation of unrealistic, e.g. a negative estimate of ψ which would imply that a larger abundance of fish results in worse penguin reproductive performance?
e) For scenarios for which the GLMM estimate for the biomass surrogate model of equation (1) of MARAM/IWS/DEC15/PengD/P1 leads to a variance estimate of zero

for the random effect (see section 5 of MARAM/IWS/DEC15/PengD/P2), should that scenario be down-weighted/excluded from consideration? An alternative is to fix a variance for that random effect typical of those (non zero) values for other models for that response variable in the conditioning estimation; note, however, that that option is not available in R so would require software development.

7) Miscellaneous

a) Section 9 of MARAM/IWS/DEC15/PengD/P2 provides comparative results for a very simple estimator applied to pseudo-data. This was intended to address the 2014 Panel recommendation that the performance of a simple estimator be considered. Is this specific estimator sufficient for such a test, or should other slightly more complex estimators also be considered in such an exercise?

b) Section 9 of MARAM/IWS/DEC15/PengD/P2 also provides results for instances of crossing original conditioning estimators with different estimators applied to the pseudo-data generated (e.g. generate data based on a catch-only scenario, but apply a closure-only estimator to the pseudo-data generated). Are results from such exercises meaningful – note that equation (1) of MARAM/IWS/DEC15/PengD/P1 is set up such that the λ and δ estimates each measure the impact of an average catch changed to zero catch, which arguably renders them comparable when the C10 and $C_{closure}$ options are used, but the situation is less clear for the C20 and C30 options where catches in these regions are not reduced to zero in years when the nearer neighbourhood of the island is closed to fishing.