

## SUGGESTIONS RE FURTHER PENGUIN POWER ANALYSIS AND ESTIMATOR SIMULATION TESTS IN RESPONSE TO PANEL RECOMMENDATIONS

D.S. Butterworth

MARAM, Department of Mathematics and Applied Mathematics, University of Cape Town

### SECTION I

First the existing methodology for the power analyses associated with the Island Closure Feasibility Analyses, from which suggestions for further work to pick up on the Panel's recommendations are most readily developed, are duplicated below (taken from Appendix B of MARAM/IWS/DEC14/Peng/B4).

#### Power Analysis Methodology

This Appendix indicates how the general linear model (GLM) analyses of the main text which estimate the fishing effect parameters  $\lambda_i$  are extended to estimate the power of an Island Closure Experiment. Statistical power reflects the probability that an experiment will detect an effect if it exists.

#### Methods

##### *Fixed year effects model*

The GLM for a reproductive success parameter  $F$  is:

$$\ln(F_{y,i,s}) = \alpha_y + \gamma_s + \lambda_i \frac{C_{y,i,p}}{\bar{C}_{i,p}} + \varepsilon_{y,i,s} \quad (\text{B.1})$$

for year  $y$ , island  $i$ , and data series  $s$ , where

$\alpha_y$  is a year effect reflecting prevailing environmental conditions,

$\gamma_s$  is a series effect (subsuming an island effect),

$\lambda_i$  is a fishing effect,

$C_{y,i,p}$  is the catch taken in year  $y$  in the neighbourhood of island  $i$  of pelagic species  $p$ ,

$\bar{C}_{i,p}$  is the average catch taken over the years considered, and (excluding years for which fishing was prohibited), and

$\varepsilon_{y,i,s}$  is an error term.

Following Brandão and Butterworth (2007), future penguin response data are generated as follows:

$$\ln(F_{y,i,s}) = \hat{\alpha}_y + \hat{\gamma}_s + \hat{\lambda}_i \frac{C_{y,i,p}}{\bar{C}_{i,p}} + \varepsilon_{y,i,s} \quad (\text{B.2})$$

where

$\hat{\alpha}_y$  are generated by sampling with replacement from estimates for  $\alpha_y$ ,

$\hat{\gamma}_s$  are the best estimates of  $\gamma_s$ ,

$\hat{\lambda}_i$  are the best estimates of  $\lambda_i$ ,

$C_{y,i,p}$  are generated by sampling with replacement from the time-series of observed catches for series  $s$  for years in which the island concerned is “open” to fishing, and zero otherwise, and

$\varepsilon_{y,i,s}$  are generated from  $N(0, \sigma_\varepsilon^2)$ , where  $\sigma_\varepsilon^2$  is the variance of the residuals when the model is fit to the historic data.

The future data are appended to the historic time-series.

The GLM is fit to obtain estimates for  $\lambda_i$  and the associated  $t$ -probability using a fixed year effects model.

The process is repeated a large number of times (for results in this paper 950).

Experimental power is calculated as the number of  $\lambda_i$  estimates which are statistically significant (at the 5% level) divided by the number of simulations performed.

### *Random year effects*

Calculating power based entirely on a fixed year effects model does however give rise to some difficulties, the chief on which is that for the data sets available such models often have relative few degrees of freedom so that ML estimates of residual variance  $\sigma_\varepsilon^2$  may be substantially negatively biased. Accordingly the power analyses have made use of results from some random effects models – specifically the GLM model parameter estimates which are more precise because of the greater associated number of degrees of freedom, and the estimate for the residual variance which is unbiased because of the use of REML – for the generation of future response data. The GLM fitted to these future data remains a fixed year effects model, but in future work a mixed model for which the year factor is treated as a random effect could also be applied.

### *Effect size*

Effectively the approach outlined above is taking the effect size for the power analysis to be equal to the current best estimate of the fishing effect parameter  $\lambda_i$  under the random year effects model. This does however raise the problem that if that estimate is very small (perhaps so small as not to be meaningfully different from zero biologically), it is of no real interest to ascertain the exact value of the rather large number of years which would be needed to collect sufficient data to determine that the value had been distinguished from zero at the 5% significance level.

Instead therefore, for cases where the point estimate of  $\lambda_i$  is small, it has been replaced by a fixed value, of the same sign as the point estimate of  $\lambda_i$ , but of a magnitude which is (arguably) biologically meaningful. The actual fixed value chosen is 0.1. The justification for this choice comes from the following consideration of penguin population dynamics.

If penguin reproductive maturity is assumed to occur at age 4, the basic equation used by Robinson (2013) for the mature female component of the population (numbering  $N$  in year  $y$ ) may be written:

$$N_{y+1} = N_y S + H_{y-3} S^3 N_{y-3} \quad (\text{B.3})$$

where  $S$  is the mature female annual survival proportion and  $H$  is a measure related to the product of egg production and fledging success. In a situation where the population is changing at a steady rate:

$$\eta = N_{y+1}/N_y \quad (\text{B.4})$$

then

$$\eta^4 = \eta^3 S + HS^3 \quad (\text{B.5})$$

which if  $H$  changes by  $\Delta H$  leads to a corresponding change in penguin growth rate  $\Delta\eta$  given by:

$$\Delta\eta = \frac{S^3}{4\eta^3 - 3\eta^2 S} \Delta H \quad (\text{B.6})$$

Now results in Robinson (2013) suggest that for  $S=0.88$ , the Robben island penguin population abundance was approximately steady, so that substituting  $\eta=1$  in equation (B.5) yields  $H = 0.176$ , and hence from equation (B.6):

$$\Delta\eta/\eta = 0.088 \Delta H/H \quad (\text{B.7})$$

Now from differentiating equation (B.1), the relative change in the penguin response variable  $F$  arising from a suspension of fishing ( $C$  changes from  $\bar{C}$  to 0) will be given by:

$$\Delta F/F = -\lambda \quad (\text{B.8})$$

so that if one assumes as a first approximation that a relative change in  $F$  results in the same relative change in  $H$  (i.e.  $\Delta H/H = \Delta F/F$ ), it then follows that:

$$\Delta\eta/\eta = -0.088 \lambda \sim -0.1\lambda \quad (\text{B.9})$$

If then 1% is to be regarded as a meaningful change in the penguin population growth rate (to be achieved, conceivably, by a suspension of fishing in the neighbourhood of the colony concerned), it follows that the corresponding value for the magnitude of  $\lambda$  is about 0.1, which is why this value was chosen for what is in effect a default minimum effect size above.

### *Future closure sequences*

At the International Panel Review meeting in 2010 when the feasibility study was discussed (Parma *et al.*, 2010), the schedule of alternating closures, each of three years' duration, which was agreed was for Robben and then Dassen Island commencing in 2011. For St Croix, a three year closure period was to be completed by a further closure in 2011, which then was to be followed by three years of closure around Bird Island.

This schedule was implemented, with closures extending for 10 nm around the islands (taken to correspond to a single grid block – see Coetzee, 2014). Thus closures are assumed to impact only the catch within this area, which is reduced to zero. However for models fitted to catches over greater distances from the islands, such as 20 and 30 nm, it is assumed that closures have no impact, as any catch that would have been made within the 10 nm distance from the island seems most likely simply to be displaced to the area between 10 and 20 nm from the island.

Thus the results reported in the main text contrast future alternating closure approaches [denoted C/O for closed/open] with those with no closures at all (and hence typical catches continuing every year) [denoted O] only for models related to catches within a 10 nm distance from islands. Models for catches within greater distances are treated only as “no closures” scenarios [O].

### **References**

Brandão A and Butterworth DS. 2007. An initial analysis of the power of monitoring certain indices to determine the effect of fishing on penguin reproductive success from an experiment where pelagic fishing is prohibited in the neighbourhood of Robben Island, but continues around

Dassen Island. Unpublished report, Marine and Coastal Management, South Africa. Report No. EAFWG/OCT2007/STG/04.

Coetzee JC. 2014. The current set of available data for evaluation of the Island Closure Feasibility study. Document MARAM/IWC/DEC14/Peng/C1: 1-15.

Parma A, Punt AR and Stefansson G. 2010 International Review Panel report of the 2010 International Fisheries Stock Assessment Workshop, 29 November – 3 December. Document MARAM/IWS/DEC10/REP/1: 1-14.

Robinson WML. 2013. Modelling the impact of the South African small pelagic fishery on African penguin dynamics. PhD thesis, University of Cape Town. xiv + 207 pp.

## SECTION II

This section proceeds to make specific suggestions for implementing the Panel recommendations under consideration here.

### Power analysis given residual autocorrelation

FISHERIES/2015/MAY/SWG-PEL/18 reports coarse estimates and associated standard errors for the autocorrelation of the residuals in the GLM analyses of penguin response variables from the Island Closure Feasibility Study. For Dassen and Robben Islands these can be as positive as 0.71, and the associated standard errors can also be high particularly given the shortness of some of the series.

This indicates that it might be unwise to update estimates by relying on a series specific and hence imprecise estimates of this autocorrelation. Instead it is suggested that the methodology above be extended to consider autocorrelations of 0.2 and 0.5 in addition to the current (implicit) value of 0.

Thus rather than generating residuals for the power analysis as:

$\hat{\varepsilon}_{y,i,s}$  from  $N(0, \sigma_{\varepsilon}^2)$ , where  $\sigma_{\varepsilon}^2$  is the variance of the residuals when the model is fit to the historical data,

instead:

$$\varepsilon_{y,i,s} = \mu\varepsilon_{y-1,i,s} + \sqrt{1 - \mu^2}\eta_{y,i,s} \quad (\text{II.1})$$

where  $\eta$  is from  $N(0, \sigma_{\varepsilon}^2)$  and  $\mu$  is the autocorrelation.

A further point raised last year concerns sample size impacting the sampling component of the residual variance  $\sigma_{\varepsilon}^2$ , so that this variance could be year dependent. However, while that might impact the process for optimal estimation of the fishing effect parameter  $\lambda$ , it is reasonable to assume fixed sample sizes into the future for power computations, and this consideration is orthogonal to the bias issue addressed in the sub-section following.

### Simulation testing to quantify possible estimator bias

The simulation tests recommended by the Panel arise from the analyses of MARAM/IWS/DEC14/Peng/A10 which points to possible bias in estimates of the fishing effect parameter  $\lambda$  in the regression:

$$\ln(F_{y,i}) = \mu B_y + \beta_i + \lambda_i \frac{C_{y,i,p}}{C_{i,p}} + \varepsilon_{y,i} \quad (\text{II.2})$$

if the catch  $C_y$  is partly determined by, and hence correlated to, the forage fish biomass  $B_y$ . Note that this applies also for the equation (B.1) formulation above, as the year effect  $\alpha_y$  there is a surrogate for biomass  $B_y$ .

Simulation testing thus requires recasting the procedures of Section I to provide Operating Models that allow for the  $C_y$  values to be generated in a way that reflects some positive correlation with the  $\alpha_y$  (or  $B_y$ ).

Simulation tests need to be conditioned on the scenarios and data under consideration to address the question raised by the Panel. However, sampling values with replacement from the past series concerned becomes problematic in circumstances where the correlation needs to be reflected, so that the following is suggested.

- Sample the  $\hat{\alpha}_y$  (or  $B_y$ ) from a normal distribution with the mean and variance of the series in question, but truncating any values generated outside the range covered by the series.
- If considering  $B_y$ , add observation error when generating pseudo-data. Work with  $\ln B_y$ , to ensure positive values  $B_y$  are generated when adding normally distributed error with mean zero, standard deviation equal to the average CV of the estimates in the series concerned, and truncated at  $\pm 2$  standard deviations to avoid undue “outlier” influence.
- To avoid negative values of  $C_y$  when generating using a normal distribution, truncate below at 5% of the mean value of the catch series in question, and above at  $\pm 2$  standard deviations.
- Generate  $C_y$  values from  $C_{y,i,p} = \phi \hat{\tau}_y + \sqrt{1 - \phi^2} \chi_{y,i,p}$   
 where  $\phi$  is the correlation between  $C_y$  and  $B_y$ ,  
 $\chi_{y,i,p}$  is generated from the mean and variance of the  $C_y$  series in question, subject to truncating as above, and  
 $\hat{\tau}_y = \psi \hat{\alpha}_y$  (or  $\phi B_y$ ) where  $\psi$  is the ratio of the means of the historical  $C_y$  and  $\hat{\alpha}_y$  (or  $B_y$ ) values of the series concerned.
- The choice for the value(s) of  $\phi$  considered will be informed by the results of regressions in MARAM/IWS/DEC14/Peng/B9. Note that that paper (see Figures 8-11 thereof) reports values of the correlation coefficient ( $\phi$ ) between  $C_y$  and  $B_y$  of 0.15, 0.19, 0.41 and 0.45 for anchovy catches within 10 and 20 nm of Robben and Dassen Islands. Note also that given that the associated linear regressions suggest weaker dependence between catches and biomass than linear proportionality, the formulation suggested above will over-emphasise the impact of this effect. Sensitivity to this specification might be explored by considering also an adjusted relationship  $\hat{\tau}_y^* = \psi^* \sqrt{\hat{\alpha}_y}$  where  $\psi^*$  is the ratio of the means of the historical  $C_y$  and of the square root of the historical  $\hat{\alpha}_y$  (or  $B_y$ ) as a possibly improved reflection of the regression relationship.

- The length of the time series of data simulated will correspond to that available for the scenario for which series are under consideration.

It then remains to select the scenarios for which these estimates of bias in the values estimated for the  $\lambda$  parameter will be carried out. The following are associated suggestions/aspects needing consideration:

- Work first with equation B.1 before II.2, as existing results are for the former.
- Clarify whether unstandardized mean (i.e. nominal) values for the response variables  $F$  are to be used (as already available), or whether these are first to be GLM-standardised based on the raw data if now openly available.
- There are many possible choices across islands (though suggest to first focus on Dassen and Robben), as well as distance from island, prey species, and response variable. To get started it might be best to select two scenarios – one corresponding to a fairly lengthy historical series and the other a fairly short one – and further select three more to move onto in the second phase.
- Simulations will assume three-year periods of opening and closing a 10 nm area around islands as for the power analyses, and testing will include estimators that consider both closure and catch as the fishing effect.