Comments on "Revisiting the key results in MARAM/IWS/2019/PENG/P4 in light of the 2019 Panel recommendations (FISHERIES/2020/JUL/SWG-PEL/53REV) by Richard B. Sherley".

By

Mike Bergh
OLSPS Marine
Silvermine House
Steenberg Office Park
Tokai 7945

24 August 2020

1. A debate that has been circulating for a number of years is the merit of the use of individual bird data versus standardised aggregated data. This debate has not been resolved in FISHERIES/2020/JUL/SWG-PEL/***53REV***. This dichotomy is one of the important differences between the results reported in FISHERIES/2020/JUL/SWG-PEL/***53REV***, and those reported in FISHERIES/2020/JAN/SWG-***PEL/09*** (and FISHERIES/2019/NOV/SWG-***PEL/27rev***). Further deliberations on the management actions required for island closures need to see this debate resolved by methods that occur in fisheries management deliberations in South Africa. This involves different researchers establishing 'common ground', viz. using the same data and methods to produce the same analytical results. Thereafter, researchers should demonstrate how departure from such 'common ground' leads to differences in the analytical results, and to provide step by step justification of the merits of these departures. This is both a scientifically defensible approach and it gives those on the periphery of the technical deliberations confidence to base management recommendations on a particular analytical result. Allusion to general theoretical support for departures from 'common ground' may have some merit but it is necessary to back all this up by comprehensive numerical analyses. Management deliberations are impeded when different methods are applied without the 'common ground' approach, and especially when they also produce different results. There are further differences between the methods in FISHERIES/2020/JUL/SWG-PEL/***53REV*** and FISHERIES/2020/JAN/SWG-***PEL/09*** (see next point). But at this stage, without the use of 'common ground' to facilitate deliberations, it is very difficult to accept one or another approach as a sound basis for decision making.
2. The international stock assessment review panel convened in 2015 (*International Review Panel Report for the 2015 International Fisheries Stock Assessment Workshop 30 November – 4 December 2015, UCT. A Dunn M Haddon, A M Parma, A E Punt*) proposed a methodology to be incorporated into any method that is applied in estimating the island closure effect. The relevant excerpts are points 3 and 4 of page 3, point 15 on page 4, point A3 of page 5, Table 1 and Appendix A. These points and recommendations have not been addressed or followed in FISHERIES/2020/JUL/SWG-PEL/53REV
3. The 2016 International Review Panel Report (*International Fisheries Stock Assessment Workshop 28 November – 2 December 2016, UCT, A Dunn , M Haddon , A M Parma , A E Punt*) noted that "the use

of disaggregated data in an estimator would require that steps 1-4 of Table 2 of MARAM/IWS/DEC16/Peng_Clos/P1a be followed" – these four steps are reproduced below:

Table 2: Summary of the power analysis procedure.

Step	Description
1	**Generating data**
(a)	Apply the OM to the actual data to get the estimates for V_{OM} (where V_{OM} is λ for a catch OM and δ for a closure OM) in terms of point estimate and standard error.The OM is used to condition other parameters.
(b)	Apply the EM to the actual data to get the estimate for V_{EM} (where V_{EM} is λ for a catch EM and δ for a closure EM) in terms of point estimate and standard error.
(c)	Fixing V_{OM} at a range of values on the interval $[\mu^{EM}_{data} - 3(\sigma^{EM}_{data});$ Threshold$]^a$, condition the remaining parameters of the OM for each value of V_{OM}. and use these estimates to generate future simulated data using the OM. These simulated data are appended to actual historical data.
2	**P_{min} bias**[b]
(a)	Use the OM model to generate past and future data for V_{OM}=Threshold.
(b)	Apply EM to data set from 2(a) and calculate detection probability at V_{EM}=Threshold to evaluate the value of P_{min} for which this detection probability at V_{EM}=Threshold is 0.5.
3	**Detection probability**
(a)	Apply EM to the data set generated in 1(c) to calculate detection probability with the bias-adjusted P_{min} value from 2(b), i.e. calculate the number of times in the 1000 simulations that $P_i(\lambda_i < T)$ is greater than the adjusted P_{min}. The detection probability is plotted against the V_{EM} variable (λ if the EM is a catch model or δ if the EM is a closure model).
4	**GLM bias**
(a)	Evaluate the GLM bias as follows. (i) Generate the historical data with the EM-like-OM[c] and apply the EM model to these data to obtain the mean from $\bar{V}_{EM}$ from the V_{EM}'s estimated for each of the 1000 generated data sets. (ii) Apply the EM to the actual historical data to obtain μ^{EM}_{data} as for 1(b). (iii) The bias is defined as $B = \bar{V}_{EM} - \mu^{EM}_{data}$. (iv) The adjusted value used for the mean of the starting distribution in Step 5 is $\mu^{EM*}_{data} = \mu^{EM}_{data} - B$. (v) There will thus be one GLM bias for each EM, irrespective of the OM. This (and this approach) follows conceptually from the fact that in reality one knows only the EM, whereas there are innumerable OMs which could reflect the underlying reality.

These four steps have not been followed in FISHERIES/2020/JUL/SWG-PEL/***53REV***.

4. The methods referred to in point (2) and (3) above include methods that need to be applied to deal with estimation bias (see in particular point 4 in the extract above) as is extensively elaborated and enlarged upon in MARAM/IWS/DEC15/PengD/P1. Such estimation methods have not been incorporated into FISHERIES/2020/JUL/SWG-PEL/***53REV*** although they were incorporated into FISHERIES/2020/JAN/SWG-PEL/09. This shortcoming needs to be addressed.

5. It is surely necessary to test whether effects which are being introduced as random effects satisfy the requirements for assuming that they are random. Or can one simply assign certain effects that need to be accounted/allowed for as random, and others for which specific estimates are require as fixed?

6. The use of different random effects has a large impact on the standard error of certain of the closure effects reported in FISHERIES/2020/JUL/SWG-PEL/***53REV*** (see for example M1 of Table 1). The reasons for this result needs explanation since it leads to a conclusion that the island closure effect is not statistically significant.

7. The Kaplan Meier results suggests that at Robben Island chick survivorship is dependent on time, or chick age (see Figure A4.13 of FISHERIES/2020/JUL/SWG-PEL/***53REV***). The potential that this has biased the closure effect estimate because of the selection of chicks at different ages/times for estimating chick survivorship needs to be fully explored. This has not been done in FISHERIES/2020/JUL/SWG-PEL/***53REV***.

8. FISHERIES/2020/JUL/SWG-PEL/***53REV*** and FISHERIES/2020/JAN/SWG-***PEL/09*** focus on different penguin response variables and data – the data that were used are indicated in shaded blocks.

	FISHERIES/2020/JUL/SWG-PEL/53REV			
53Rev	**Dassen Island**	**Robben Island**	**St Croix**	**Bird Island**
Chick condition	■	■	■	■
Max distance	■	■	■	■
Chick survival	■	■		
Chick growth				
Forage length				
Forage duration				
Fledging success				

	FISHERIES/2020/JAN/SWG-PEL/09		FISHERIES/2019/NOV/SWG-PEL/27rev	
	Dassen Island	**Robben Island**	**St Croix**	**Bird Island**
Chick condition	■	■		
Max distance	■	■	■	■
Chick survival	■	■		
Chick growth	■	■		
Forage length	■	■	■	■
Forage duration	■	■	■	■
Fledging success	■	■		

It would be useful to obtain a justification for which data and results are, and/or should be, relevant for future management deliberations. The 2015 international panel report made some recommendations in this regard, specifically:

i. ***All six response variables should be assessed with respect to how reliably they are sampled and how informative they are regarding potential fishery effects on population growth rates.***

ii. ***A response variable should not be considered further if there is no (objective) way to determine a threshold for it.***

iii. ***If a particular response variable is sub-ordinate or directly correlated with another then there may be little to be gained by considering it further.***

Points (i) to (iii) need to be applied to all the variables listed above to produce an agreed set of variables for use in management deliberations to pre-empt and avoid different researchers making difference choices about which variables to include in analyses.

9. In relation to the tables of data used in FISHERIES/2020/JUL/SWG-PEL/***53REV***, the possibility that particular results have been favoured because they provide a particular result needs either to be dispelled, or a Bonferroni adjustment should be used in assessing the statistical significance of results.

10. The use of an overall closure effect, as presented in Figure 5 of FISHERIES/2020/JUL/SWG-PEL/***53REV***, needs more care. The international panel report of 2016 submitted some ideas about this, pointing out that some variable are relevant to fledgling success while other may be relevant to adult survival ("*…while chick condition and chick growth are likely correlated, chick condition/growth and fledgling success affect processes that are sequential in the life history of penguins, which means that a fishery effect on each of chick condition/growth and fledgling success in combination could lead to a biologically meaningful population effect. Moreover, increases in forage trip length due to fishery impacts may have negative consequences for adult survival*."). Integration would seem to require at least some demographically sensible basis for their combination and a framework for

establishing where in penguin life history their effects are active, whether on adult survival or on chick survival and fledgling success.

11. The combination of posterior distributions in the manner proposed in Figure 5 of FISHERIES/2020/JUL/SWG-PEL/***53REV*** needs to address the opposite result for 'Condition Index' at St Croix versus the result for 'Max Distance' at St Croix, before attempting to integrate these results. Invoking methods used in local fisheries management deliberations, two opposite results should not be 'averaged out'. Firstly, the reasons for this would be explored insofar possible. Failing resolution at this level, the full management implications of each would be considered and any compromise position would be developed at a late stage once the management implications were known. A similar approach is applicable here. In addition, it is inappropriate to add up very different response variables – the demographic implications of different variables may imply the use of different weightings.

12. Figure 5 of FISHERIES/2020/JUL/SWG-PEL/***53REV*** duplicates the posterior distribution for the Robben/Dassen Island chick survival based island closure effect. If, as is the case here, the closure effects at Robben Island and Dassen Island have been constrained to be equal, then there is only one closure effect estimate and only one posterior distribution, not two. Therefore, notwithstanding other comments here about approaches used to combine results from difference response variables, only the one posterior distribution should contribute to the overall posterior distribution in Figure 5. This will change the overall posterior distribution presented in Figure 5 (bottom left panel).

13. FISHERIES/2020/JAN/SWG-***PEL/09*** provides a different island closure impact on chick survivorship for Robben and Dassen Islands. Figure 5 of FISHERIES/2020/JUL/SWG-PEL/***53REV*** provides estimates under the constraint that the closure effect is equal across the two islands (at least in its Figure 5). There seems no reason for these estimates to be equal. The separate estimates for the closure effect at these islands should be presented.

14. The use of maximum forage distance as a relevant and appropriate response variable needs more justification. Intuitively, foraging energy expenditure is relevant to population level outcomes such as adult survival. On that basis energy expenditure does not equate to the maximum foraging distance, but rather to total distance travelled (i.e. forage length). Additionally, it is necessary to translate this result to a value which is demographically meaningful.