Question Q4: Summary document

This document combines pertinent comments and responses contained in various documents submitted to the Small pelagic Working Group during the course of 2020 that are associated with Question Q4.

<u>Given the large amount of additional material that this question would require the Panel to review in order to answer this question, and the delay in getting material to the Panel members, they have recommended that this question not form part of the ToR for this review. In the interests of transparency and for the benefit of others who have not seen these summarised arguments, this document will be retained for background purposes.</u>

QUESTION Q4: A 2020 SWG-PEL document has indicated that both the aggregated and disaggregated data approaches lead to broad agreement when asking "has the experiment conducted thus far shown that fishing around the island has a "biologically meaningful" negative impact on penguins" (FISHERIES/2020/SEP/SWG-PEL/95 Table 1, page 4). Are the panel able to comment on whether they would support this notion that the two sets of analyses point to the same conclusion about the closure experiment?

Overview of material included under 4 steps: 1 = assertion, 2 = response, 3 = response to response, 4 = further responses. Note interpretation of document purpose below is that of Janet Coetzee. Author DSB = Doug Butterworth, RS = Richard Sherley, MOB = Mike Bergh.

	Step	Author	Original source document and main reference documents	Pg.
Exchange 1: Sherley vs Butterworth	1	RS	**FISHERIES/2020/JUL/SWG-PEL/53rev (**Asserts that the two independent sets of analyses, FISHERIES/2020/JUL/SWG-PEL/53rev and FISHERIES/2020/JAN/SWG-PEL/09 agree that biologically meaningful effects of fishing around African penguin breeding colonies are apparent)	2
			FISHERIES/2020/SEP/SWG-PEL/95 (Compares results from the two sets of analyses (FISHERIES/2020/JUL/SWG-PEL/53rev and FISHERIES/2020/JAN/SWG-PEL/09)	
	2	DSB	**FISHERIES/2020/SEP/SWG-PEL/96rev** (Argues that Figure 1 negates the assertion above) **FISHERIES/2020/SEP/SWG-PEL/110** (Draws attention to various caveats and their associated complications and contradictions, which confound clear interpretations)	3
Exchange 2: Sherley vs Bergh	1	RS	Same as for Step 1 above	2
	2	MOB	**FISHERIES/2020/OCT/SWG-PEL/106** (Argues that there is not a clear and consistent negative impact due to fishing and that for cases where such negative effects have been shown to be biologically meaningful, the overall effect is weak) **FISHERIES/2020/SEP/SWG-PEL/99 (**Points out that there are still outstanding technical issues regarding the analyses presented in FISHERIES/2020/JUL/SWG-PEL/53REV and notes the existence of a mathematical proof against using those results for informing decisions) **FISHERIES/2020/OCT/SWG-PEL/107** (Points out remaining technical issues in response to FISHERIES/2020/SEP/SWG-PEL/87 and FISHERIES/2020/JUL/SWG-PEL/53rev) **FISHERIES/2020/OCT/SWG-PEL/113** (Comments on a proposal for future island closure, FISHERIES/2020/OCT/SWG-PEL/105rev, which was based on results reported in FISHERIES/2020/JUL/SWG-PEL/53rev, and states that the results of the island closure experiment are not sufficiently consistent nor of a scale that such closures are likely to benefit penguin population trends substantially.	6

EXCHANGE 1: Sherley/Butterworth

Step 1 – Sherley assertion:
Document FISHERIES/2020/JUL/SWG-PEL/53rev

Regardless of their differences, and the technical debates above, both the analyses by Sherley and colleagues and the analyses by Butterworth and colleagues point to the same conclusions about the presence of biologically meaningful effects of closures/fishing around African penguin breeding colonies:

Sherley [bottom of page 11, FISHERIES/2020/JUL/SWG-PEL/53REV] asserted that "we have now iterated to a place where two independent sets of analyses agree that biologically meaningful effects of fishing around African penguin breeding colonies are apparent and, importantly, that some of those effects are on variables (chick survival, fledging success) that contribute to the demographic process" because "the most recent update of the MARAM power analysis (Ross-Gillespie and Butterworth 2020) also finds "evidence in the current data of a biologically meaningful fishing effect" on chick condition at Robben Island, fledging success at Robben Island (chick survival measures a component of fledging success), and chick survival at Dassen Island. Those results from Ross-Gillespie and Butterworth (2020) concur with results presented here in section 3.1 and 3.3 [of FISHERIES/2020/JUL/SWG-PEL/53REV]. Plus, Ross-Gillespie and Butterworth (2020) also find that "a biologically meaningful fishing effect is likely to be detected" if the experiment continues for 2–5 years (using a dataset that ended in 2015) for chick survival at Robben Island".

This point was supported by an independent assessment of the inference to be drawn from the two sets of analysis [de Moor 2020, FISHERIES/2020/SEP/SWG-PEL/95] which stated:
"It therefore appears that:
- Both sets of analyses currently show that the closure of Robben Island to fishing will benefit penguins.
- Both sets of analyses currently show that the closure of St. Croix Island to fishing will benefit penguins.
- Both sets of analyses currently show that the closure of Dassen Island to fishing will benefit penguin chick survival, but the overall benefit is not clear given some results from Butterworth and Ross-Gillespie which suggest the closure of Dassen Island to fishing may be detrimental to chick growth, path length and trip duration?
- Sherley's results suggests this experiment could not inform on the impact of fishing to penguins on Bird Island, while a power analysis from Butterworth and Ross-Gillespie may help ascertain whether extending the experiment at this Island may produce more informative results.

Step 2 – Butterworth response:

Direct responses to the assertion made in Step 1 in relation to QUESTION Q4: Step 1

Paragraph 1

The comparison of results from the two approaches shown in Figure 1 below negates the assertion. Quite clearly there are substantial and meaningful differences in many of the pairs of estimates from the two approaches, some of which would certainly lead to different inferences being drawn as to the benefits or otherwise of island closures for penguins. For example, the precision of the closure effect estimates from chick survival data at Dassen and especially Robben island differ substantially between the two approaches. For St Croix, the chick condition response variable provides no clear result for the effect of closure under the aggregated data approach, but a strong indication of the absence of a positive impact under the individual data approach; this is in circumstances when the foraging related variables suggest a positive impact under both approaches. Since both these variables are not closely tied to demographic measures, and the foraging data point in the reverse direction for Robben and Dassen islands, an attempt at integration of this information would broadly suggest overall that the aggregated data results indicate a positive impact of closure at St Croix, but that the individual data results do not do so.

Paragraph 2

The conclusions cited rest on superficial examination of results for the experiment without caveats giving due attention to their associated complications and contradictions, which confound such clear interpretations as are offered. The comments quoted in *red italics* below, which expand on these aspects, are taken from FISHERIES/2020/SEP/SWG-PEL/96rev and FISHERIES/2020/SEP/SWG-PEL/110.

Foraging track-related indices give inconsistent indications for different colonies, and there are difficulties in linking changes in these response variables quantitatively to penguin population dynamics.

Regarding the statement in Q4 Step 1 that: "Both sets of analyses currently show that the closure of St. Croix Island to fishing will benefit penguins":

For the individual data analysis (see Figure 1 of PEL/96rev [Figure 1 below]*) the analysis of chick condition data provides a strong indication of a counter-beneficial effect of closure.*

Regarding the statement in Q4 Step 1 that: "Both sets of analyses currently show that the closure of Dassen Island to fishing will benefit penguin chick survival":

While in principle chick survival data have relatively high potential information content for the purposes of the closure experiment, present results need to be considered in the context of a number of caveats:

- *The substantial increase (which is indicated to be relatively precisely estimated) in the estimated survival at for Robben island (but not Dassen island) by the Kaplan-Meier method after some 50 days of exposure (PEL/53REV); this aspect of these results first needs to be explained, so as to be clear on whether or not its source results in confounding estimates of cumulative chick survival. The fact that exposure time does not relate directly to chick age*

may be contributing to such confounding.

- *The result for fledging success at Dassen island is (in simple terms) in the opposite direction and is nearly statistically significantly different at the 10% level from that for the chick survival; this is not as might be straightforwardly expected, as the latter is a component of the former. This may indicate that negative correlations between different stages of the fledging process need to be taken into account.*

- *A mathematical-statistical demonstration (see the Annex of PEL/82) indicates that individual data based estimates of closure effects (including those for chick survival) are unreliable; this demonstration has not been falsified, rendering consideration of such estimates scientifically unjustifiable at this time.*

- *The assumption made in baseline individual data-based estimation using chick survival data for Dassen and Robben Islands of a common value for the impact of a closure effect is unjustified, given general indications of inter-island differences in this effect (see PEL/82 and also PEL/84).*

Note that the "reverse" result for fledging success for Dassen is of greater importance, given its direct linkage to penguin population growth.

In that regard, Sherley (FISHERIES/2020/SEP/SWG-PEL/85) introduces the possibility of regime shifts in presenting arguments to explain the discrepancy between the chick survival and fledging success estimates of δ for Dassen island. But this approach is problematic if it is to be used a posteriori in this manner; especially because if such additions are to be allowed in a comparative exercise, the associated rules have to be agreed a priori – for example, such rules would need to cover acceptable criteria to justify the possibility of allowing for such assumptions, because in other years and for other variables such assumptions could change the conclusions to which their analyses would otherwise have led.

A further important point to clarify is the role of scientific inputs (on different topics) to the formulation of management advice. Suggestions made above for Robben and Dassen Islands are helpful in providing an example to illustrate this. The suggestions above recognize that the strength of evidence from the closure experiment to support the contention that closure will benefit penguins differs for these two islands. The proposals then made are ones in the direction of strong precaution with respect to the penguins, in treating this result as sufficient to confirm a beneficial effect and hence to justify closing Robben Island for an extended period despite the fact that ending close/open alternation will prejudice further improvement of estimation of the presence and magnitude of that effect at that island. Conversely, similar statements could be made, though in the reverse direction, about a proposal to open Dassen Island for an extended period and to continue close/open alternation at Robben Island, where this is motivated by strong precaution with respect to the industry and possible associated socio-economic impacts, and the contention that the present evidence for Dassen Island is so weak as to render the possibility of any real beneficial effect to penguins of closure there minimal. But such arguments are not alone a sufficient basis to advise a decision.

Basically, the ultimate decision on this matter by a decision maker will involve trade-offs, and the responsibility of the PWG is to summarise and quantify (to the extent possible) each of the various components of benefits and losses involved. In particular, this will need to involve, and on an island-to-island basis:

- *Quantification of the benefit (if any) of closure to penguins*

- *Linkage of that benefit to the impact on the growth rate of the colony concerned relative to its current population trend*

- *Similar quantification of the impact on the pelagic fisheries (and other sectors of the economy, such as tourism) in socio-economic terms (relating to employment and additional costs).*

Clearly, although complete objectivity in such an exercise is desirable, some subjectivity introduced in the form of expert opinion will be inevitable (and desirable in areas where the potential for more fully objective evaluations is limited). Differences in such expert opinions may well exist, and it will be important that the information accompanying recommendations is summarized in a way that attempts also to provide some indications of the associated weights of supporting evidence for these. But furthermore, although likely to be difficult, such differences must be distinguished from differences in personal risk preferences in such summaries, as ultimately the decision relating to risk preference rests with the decision maker.
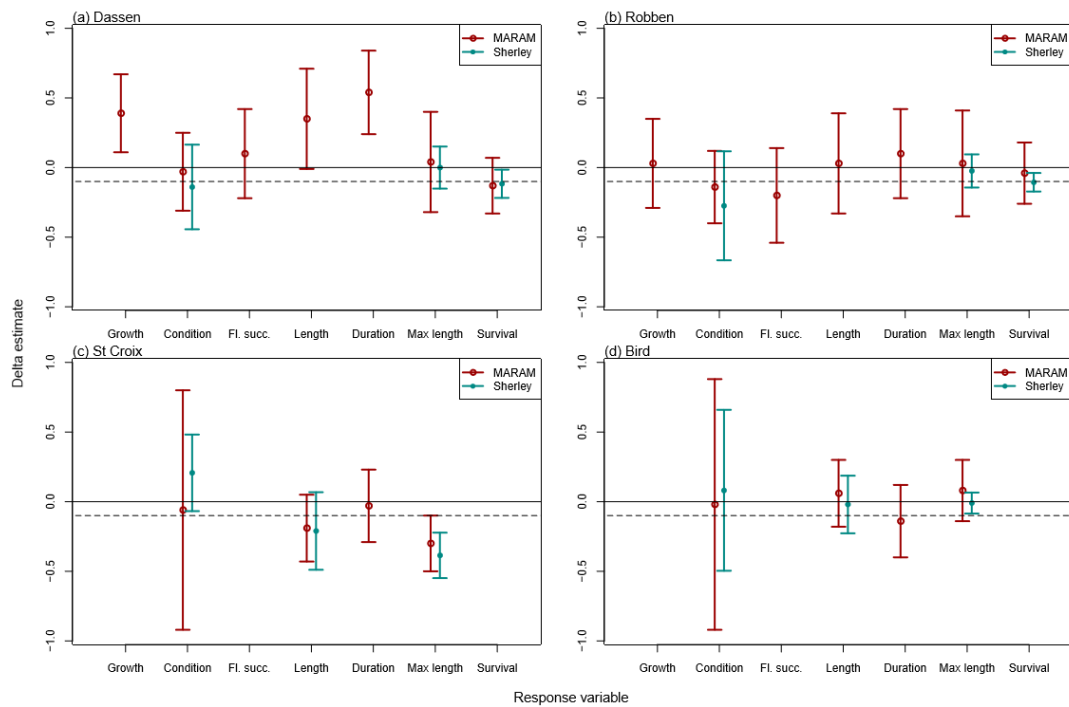

Figure 1: Zeh plots of the δ estimates and rough 95% confidence intervals are shown for the MARAM (aggregated data-based) and Sherley (individual data-based) models. The results for the MARAM models are taken from FISHERIES/2020/JAN/SWG-PEL/09rev for Robben and Dassen islands, from MARAM/IWS/2019/PENG/P2 for the foraging data for St Croix and Bird islands, and from FISHERIES/2019/NOV/SWG-PEL/33 for the chick condition data for St Croix and Bird islands. The values for the Sherley models have been derived from the last table of FISHERIES/2020/SEP/SWG-PEL/95 by use the following formula: $\delta = \ln(1 - p/100)$ where the p values are those reported in that last table as a simple approach to transform from normal to log-space to achieve improved comparability. The confidence intervals have been converted in a similar manner, and a rough standard error may be calculated as (max(CI)-min(CI))/4.. The Figure has been kindly provided by A. Ross-Gillespie.
[Figure taken from FISHERIES/2020/SEP/SWG-PEL/96rev]

EXCHANGE 2: Sherley/Bergh

Step 1 – Sherley assertion: Same as above (Page 2)

Step 2 – Bergh comment:

This question raises issues about the acceptability of the results in FISHERIES/2020/JUL/SWG-PEL/53REV, and touches on the reasons that it was recommended in FISHERIES/2020/SEP/SWG-PEL/99 that "**Given the problems that are pointed out above regarding the results reported in FISHERIES/2020/JUL/SWG-PEL/53REV, it is ill-advised to allow these to inform decisions that must be made this year**" and "**The results produced in FISHERIES/2020/JAN/SWG-PEL/09 and MARAM/IWS/2019/PENG/P2 form the basis for decisions that are made in 2020**". These reasons have links to points made and questions raised in FISHERIES/2020/OCT/SWG-PEL/106, FISHERIES/2020/OCT/SWG-PEL/107 and FISHERIES/2020/OCT/SWG-PEL/113. All these documents are more generally relevant to Q4. The bulk of these documents are therefore included below as excerpts relevant to Step 2 for Q4.

Document FISHERIES/2020/OCT/SWG-PEL/106

Proposals for Experimental Design Decisions and Island Closure Decisions in relation to the island closure experiment.
By
Mike Bergh
6 October 2020

"2.1 Fledgling Success and Chick Survival
2.1.1 Dassen and Robben Islands
Based on application of the analytical methods recommended by successive international stock assessment review panels dating back to 2014, the following results are available for the demographically meaningful response variables of fledgling success and chick survival (**FISHERIES/2020/JAN/SWG-PEL/09**):

- At Dassen Island the results suggest that fishing has a positive impact on fledging success but a negative impact on chick survival. These results are therefore **in conflict with** each other and this reality detracts from the fact that the chick survival results meet the 'biologically meaningful threshold' set out by the international panel.
- At Robben Island the effect of fishing is estimated to be negative for both fledgling success and chick survival, although only the fledgling success result meets the 'biologically meaningful threshold'.

The above comments notwithstanding, a significant issue has arisen for the chick survival estimates at Dassen Island and Robben Island. These are that the underlying data exhibits evidence of substantial dependence of the survival rate on the time since observations per chick were initiated, as shown in Fig A4.13 of FISHERIES/2020/JUL/SWG-PEL/53REV. Further work is imperative to dispel

concerns that this may have substantially biased estimates of the related closure effect and its variance.

2.1.2 St Croix and Bird Islands

- There are no estimates for chick survival and fledgling success at St Croix Island or Bird Island Algoa Bay, and hence no conclusions can be drawn that are directly relevant to penguin demographics for these response variables.

2.2 Chick growth, chick condition, foraging length, foraging duration, maximum foraging length.

Although these results cannot be interpreted in the same way as chick survival and fledgling success variables in the sense of being directly related to penguin population growth rate (with a 10% increase in chick survival translating roughly to a 1% increase in the population growth rate), the directions of the relevant closure effects are nonetheless relevant and important, as is considered here.

2.2.1 Dassen and Robben Islands

Other results that are available at Dassen and Robben Islands are for the following response variables

- Chick growth,
- Chick condition,
- Foraging length,
- Foraging duration
- Maximum foraging length.

The only result for which the existence of a biologically meaningful result has been demonstrated is for chick condition at Robben Island (under the assumption of linearity mentioned earlier). The point estimate indicates a negative impact on chick condition at Robben Island. For reasons set out earlier it is unclear how much of an impact this has on penguin population growth rates.

2.2.2 St Croix and Bird Islands

At St Croix and Bird Islands the available results are for

- Chick condition,
- Foraging length,
- Foraging duration,
- Maximum foraging length.

The analyses suggest the following:

MARAM/IWS/2019/PENG/P2: For Bird island, there is no evidence to suggest a significant (at the 5% level) impact of the fishery on the penguin foraging behaviour. For St Croix, the impact is estimated to be significant for maximum foraging length but not foraging length or foraging duration.

FISHERIES/2019/NOV/SWG-PEL/33, For chick condition at St Croix and Bird "none of the fishing effect impacts are estimated to be significantly different from zero, or from -0.1, and the standard errors for these estimates are large, indicating that any inferences drawn from this analysis would not have great reliability".

2.3 Overall

The overall impression from Figure 1 of **FISHERIES/2020/SEP/SWG-PEL/97rev** (which sets out the results for all response variables) is that these do not suggest a clear and consistent negative impact due to fishing. In general, confidence intervals are very broad and span both 0 and the biologically meaning threshold level. Point estimates on either side of zero and/or the biologically meaning threshold are obtained.

Overall the impression from the point estimates in part (a) of Figure 1 of **FISHERIES/2020/SEP/SWG-PEL/97rev** is that the closure effect at Dassen Island is not positive, and less than at Robben Island."

"As regards the potential for further experimental work to provide useful results in less than 10 years, **FISHERIES/2020/JAN/SWG-PEL/09** indicates that the potential exists at the following islands and for the indicated response variables:

2.3.1 Potential for further experimental work, Dassen Island and Robben Island

Robben Island: The main area where work is likely to have further benefits is at **Robben Island, for chick survival**.

Dassen Island: It is however important to resolve the conflict between the chick survival results and the fledgling success estimates at Dassen Island. Although **FISHERIES/2020/JAN/SWG-PEL/09** suggests that this will involve at least 10 years of further experimental work, this result is based on calculations that assume that there is a true underlying negative effect due to fishing. Since the results so far estimate an effect in the opposite direction, more than 10 years of work would be required to turn around the results achieved thus far and confirm the assumed reality. It may however be that the true fledgling success impact is counter to the design assumption, in which case some more work, but perhaps less than 10 years, is necessary to confirm this.

There is thus some justification to continue with monitoring work at Dassen Island for both chick survival and fledgling success. However, since there is no clear indication in results in Figure 1 of **FISHERIES/2020/SEP/SWG-PEL/97rev** that across the point estimates for all response variables there is a clear negative effect due to fishing, there is also a rationale to keep Dassen Island open from now on. Further deliberations are required as to whether Dassen Island should remain open, or alternate in an open/closed fashion.

2.3.2 Potential for further experimental work, St Croix and Bird Islands

There seems to be no further benefit to be derived in the next ten years from further work on chick condition (the confidence intervals are too wide).

For the foraging response variables there are apparent contradictory results within and between islands that might benefit from further experimental work, but this needs further discussion.

There is also some benefit to be derived from initiating fledgling success monitoring at these islands."

"The absence of island closure recommendations in this document in the main (aside from the preference expressed to keep Dassen Island open) is consistent with the overall weak effect ascribable to fishing, and hence the absence of the need for urgent action w.r.t. fishing."

Document FISHERIES/2020/SEP/SWG-PEL/99

Summary comments on the Penguin Island Closure Experiment
By
Mike Bergh
20 September 2020

"**A**greement on the next steps to take with regard to the Penguin Island Closure Experiment is being impeded by the following:

1. Different researchers have analysed a different set of response variables.

2. Results presented by different researchers (FISHERIES/2020/JUL/SWG-PEL/*53REV* and FISHERIES/2020/JAN/SWG-*PEL/09*) have been produced using different methods. This is in addition

to the fact that FISHERIES/2020/JUL/SWG-PEL/*53REV* analyses individual bird data while FISHERIES/2020/JAN/SWG-*PEL/09* analyses aggregated bird data.

3. Some researchers have used data that were not available to other researchers.

4. Results produced by different researchers differ in a number of important respects.

5. No common ground has been established between different researchers about using a default method that provides a basis for such common ground. Non-technical participants in the debates cannot therefore rely on this to judge the reliability of the results.

6. The analytical methodology used in FISHERIES/2020/JUL/SWG-PEL/*53REV* is not consistent with IWS panel recommendations dating back to 2015. Those produced in FISHERIES/2020/JAN/SWG-*PEL/09* are.

7. There are a number of outstanding technical issues with the methods and results in FISHERIES/2020/JUL/SWG-PEL/*53REV* that have not been answered.

These unresolved matters weigh heavily on the scientific deliberations which are now ongoing, and force participants to take a position on one or the other set of results, since both cannot be reliable. In addition, there is now a mathematical proof (see the annex of FISHERIES/2020/AUG/SWG-PEL/82) that the standard error of the island closure effect achieved using aggregated bird data cannot be improved upon by using data from individual bird data. In the absence of any submission that contradicts this proof, there is no reason to question the correctness of this proof. It follows that any results that provide estimates with standard errors that are smaller than the s.e. achieved using aggregated bird data must either be in error, or be negatively biased (presumably because the random effect used to adjust for pseudo-replication in the case of analyses using individual bird data is failing to account fully for this pseudo-replication). These results are therefore producing a misleading impression of the precision of estimates of the island closure effect.
Another consideration is that since decisions on the Penguin Island Closure Experiment must be made this year, it is likely that, given the complexities associated with the statistical analyses and the time it will take to resolve these, decisions will have to be made on the basis of results that have been tabled thus far. Given the problems that are pointed out above regarding the results reported in FISHERIES/2020/JUL/SWG-PEL/*53REV*, it is ill-advised to allow these to inform decisions that must be made this year.

It is proposed therefore that

1. A halt be called with regard to any new analytical results for 2020, or until management decisions can be finalised.

2. The results produced in FISHERIES/2020/JAN/SWG-*PEL/09 and* MARAM/IWS/2019/PENG/P2 form the basis for decisions that are made in 2020. It should be noted that for some response variables at some islands there is a meaningful island closure effect. But the deliberations cannot ignore the totality of the results nor that some of these response variables offer contradictory results.**"**

Document FISHERIES/2020/OCT/SWG-PEL/107

Comments on FISHERIES/2020/SEP/SWG-PEL/87

by

Mike Bergh
20 September 2020

"FISHERIES/2020/SEP/SWG-PEL/87 responds, using numbering R1 – R14, to the 14 points made in FISHERIES/2020/AUG/SWG-PEL/84 on FISHERIES/2020/JUL/SWG-PEL/53REV. The following further comments on points R1-R14 in FISHERIES/2020/SEP/SWG-PEL/87 use numbering R1-R14 as well (text in italics taken directly out of FISHERIES/2020/SEP/SWG-PEL/87):

R1: Common Ground (raised in Doc 84). A table comparing results produced using different methods does not provide common ground in the sense described in FISHERIES/2020/AUG/SWG-PEL/84. It would rather be achieved by obtaining agreement first between different researchers using the same data and the same analytical approach, followed by justification for, and demonstration of the implications of, departures from a common approach in a step by step manner.

R2 – R4: The original points 2-4 in FISHERIES/2020/AUG/SWG-PEL/84 were that the methodology in FISHERIES/2020/JUL/SWG-PEL/53REV is not consistent with the totality of IWS panel recommendations going back to 2015. FISHERIES/2020/SEP/SWG-PEL/87 has responded that these methods relate only to power analyses, and that the bias correction recommendations proposed by the panel are only applicable when the catch covariate is used. This is not correct, the methods set out by the panel are intended to correct for biases which arise even when the catch covariate is not used, as can be seen by the results presented on page 6 of FISHERIES/2020/SEP/SWG-PEL/95. Further, to the point about power analyses, the estimation methods recommended by the panel are a key component of the power analyses, and the results of those power analyses are also required for further management deliberations.

R5. Verification of Random Effects: FISHERIES/2020/AUG/SWG-PEL/84 raised the point "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?" The R5 response was:

"The diagnostic plots presented FISHERIES/2020/JUL/SWG-PEL/53REV make this check and that is why they were asked for by the 2019 panel: "This should include standard residual analysis as well as residual analyses that are tailored to the problem at hand (e.g., temporal, spatial or with in season plots of residuals)" (Die et al. 2019). The residual plots presented show no evidence of deviation from the assumptions underlying the use of the relevant mixed effects models."

What we need to see are time plots of the best estimates for the random effects to verify that these estimates show no systematic trends.

R6. Impact of the Use of Different Random Effects. Comment 6 in FISHERIES/2020/AUG/SWG-PEL/84 notes that 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, and argues that the reason for these differences needs an explanation. FISHERIES/2020/SEP/SWG-PEL/87 responds that

"As outlined in FISHERIES/2020/JUL/SWG-PEL/53REV, one possibility is that this is an issue with having Island in both the fixed and random components of the model. M1 in all cases in the maximal model (the most complex possible random effect structure); maximal models are "generally wasteful and costly in terms of statistical power for testing hypotheses" (Stroup 2012, pg. 185) and maximal

models – even when they converge – can result in overparameterization that leads to uninterpretable models (Bates et al. 2018). The maximal model may actually trade-off power for some conservatism beyond the nominal Type I error rate, even in cases where the maximal model matches the generating process exactly (Matuschek et al. 2017)"

This is speculative and should be backed up by numerical results which would probably need to be derived from data analyses as well as simulation studies. It would also need to address the annex of FISHERIES/2020/AUG/SWG-PEL/82 which demonstrates mathematically that "**as far as estimates of the precision of the closure effect in the island closure experiment is concerned, there is no advantage to be gained from analysing the individual data each year rather than an aggregate value such as their means**."

R7. Kaplan Meier, Possible Bias in Chick Survival Estimates. The R7 argument in FISHERIES/2020/SEP/SWG-PEL/87 suggests that the chick survival estimates may be biased. This is a reason to carry out further work to correct for this bias, by, for example and as suggested in FISHERIES/2020/SEP/SWG-PEL/87, excluding from analyses chicks that were not monitored from their hatching date.

R8 Choice of Response Variable for Analytical Purposes. We clearly need an agreed list of response variables which should be used in analyses, and hence whose results are admissible input into management deliberations. It is extremely confusing if different analysts not only use different analytical methods, but also use different datasets and different response variables.

R9. Bonferroni Adjustment. Some of the posterior distributions in Figure 5 of Rev53 show an exceptional result. However, if these exceptional results were derived by starting with a larger number of results, many unexceptional and a few exceptional, then if the non-exceptional ones are filtered out, the remaining 'exceptional' results are not really exceptional. This principle applies regardless of the methods used, whether Bayesian or otherwise. Had this filtering occurred, then some adjustment is needed for the measure of what is exceptional. **Did this filtering occur to any extent? There seem to be many possible measures to be derived from the foraging GPS data. Were all explored and were then only those that showed a significant results reported?**

R10. Value and Logic of an Overall Closure Effect. FISHERIES/2020/AUG/SWG-PEL/84 questioned the logic and value of the overall closure effect presented in Fig. 5 of *FISHERIES/2020/JUL/SWG-PEL/53REV*. FISHERIES/2020/SEP/SWG-PEL/87 responded that:

"*The Overall Closure Effect does not claim to integrate the impact of these variables on penguin demography.*" and *"The Overall Closure Effect basically says, "without making any additional assumptions, what is the probability, given all the uncertainty in the penguin responses, that, overall, this pre-agreed threshold of 10% has been exceeded".*

This comment raises a question about the usefulness of the combined closure effect in Fig. 5 of *FISHERIES/2020/JUL/SWG-PEL/53REV*. Further use of this approach requires a final discussion and mutually binding agreement about the suitability of the use of different response variables in analyses and how these could be combined, should that be necessary.

R11. Against Averaging Out Conflicting Results. The R11 response in FISHERIES/2020/SEP/SWG-PEL/87 has three parts:

- **Part I '1st para of R 11 of PEL/87'.** "*This comment seems to stem from a misunderstanding; the Overall Closure Effect in FISHERIES/2020/JUL/SWG-PEL/53REV does not claim to show demographic implications, it merely reports on whether the effect overall (taking account of all uncertainty in the model outputs) meets a pre-agreed threshold for management action (Cochrane 2016)".* Understood but this still then involves combining results which contradict each other, and it remains important to interrogate the implications that there is this contradiction.

- **Part II 'Secondly'.** Commonly accepted local practice on how to address conflicting results does not need to be backed up by a citation from the peer reviewed literature.

11

- **Part III 'Third'**. The dynamics and demands of local fisheries management are such that numerical results which might be revised versions of work submitted for publication but which are available before publications are finalised, must be used and final quantitative estimates used in management cannot be restricted to peer reviewed work. In general a pass in a peer reviewed process does not guarantee the infallibility of published results or arguments. There are errors in published work (as can be verified by a quick on-line search), and any scientific work is always open to question and can never represent the final word on any topic. In RFMO scientific committees publication of results is in no way sufficient for their acceptance – they will be independently interrogated and frequently not accepted.

R12 – Duplication of Posterior Distributions. This needs some discussion since it is not clear whether the response in FISHERIES/2020/SEP/SWG-PEL/87 is saying that only one of the posterior distributions shown is used in the overall distribution. Figure 5 shows two posterior distributions for chick survival, and states that the estimates at the two islands have been forced to be equal. Showing two posteriors under these circumstances is duplication. Only one posterior distribution should be shown. And only one of these posteriors should contribute to the overall posterior distribution. It would help to clarify whether for the overall posterior only one posterior for the Robben/Dassen Island complex is used when their closure parameter estimates are forced to be equal.

R13 – Chick survival estimates forced to be equal for Dassen and Robben Islands FISHERIES/2020/AUG/SWG-PEL/84 argues that there is no reason for these estimates to be equal and that separate estimates for the closure effect at these islands should be presented, as well as their associated standard errors. I note that these are now provided in FISHERIES/2020/SEP/SWG-PEL/87 under R12.

R14 – Question RE Maximum Forage Distance. FISHERIES/2020/AUG/SWG-PEL/84 argues that the use of **maximum** forage distance as a demographically relevant and appropriate response variable is counter-intuitive and needs more justification. FISHERIES/2020/SEP/SWG-PEL/87 responds that

*"**Maximum** foraging distance has a strong link to resource availability in seabirds and has recently been found to respond to directly estimated local prey abundance (measured by the small-scale surveys) around Robben Island (Campbell et al. 2019)".*

This does not link maximum foraging distance to penguin demographics. **"**

Document FISHERIES/2020/OCT/SWG-PEL/113

Comments on FISHERIES/2020/SEP/SWG-PEL/105REV, "Recommendations for island closures around African Penguin colonies".

by

Mike Bergh
19 October 2020

"2. For reasons described here, the results of analyses of the island closure experiment as reported in **FISHERIES/2020/JUL/SWG-PEL/53REV** and **MARAM/IWS/2019/PENG/P4** should not form part of the deliberations about island closures.

3. Island closure experimental results differ across the four islands included in the experiment and cannot therefore be extrapolated to Stony Point and Dyer Island."

"The results of the island closure experiment are not sufficiently consistent nor of a scale that such closures are likely to benefit penguin population trends substantially; there is therefore no reason for immediate urgency with regard to action to implement island closures".

"2.1.3 "2-3 times more evidence for positive effects of fishing closures on breeding penguins than no effects."
This statement apparently makes use of results in **FISHERIES/2020/JUL/SWG-PEL/53REV**, in particular a composite posterior distribution for which a technical question regarding the possible duplication of information was raised in a previous commentary (see **FISHERIES/2020/SEP/SWG-PEL/93**). This technical question has not been answered, other than to present its negative, which does not deal with the concern underlying the original question."

"The experimental results reported in **FISHERIES/2020/SEP/SWG-PEL/97REV** show that the island closure effect is not generally consistent across different islands and in some cases different response variables show contradictory results. The experimental results cannot therefore be reliably extrapolated to the Dyer Island and Stony Point penguin colonies. Also (see **FISHERIES/2020/SEP/SWG-PEL/97REV**), while there is some evidence of a positive island closure effect at Robben Island and the effect is positive for both chick survival and fledgling success, the positive effect on chick survival at Dassen Island is confused by a negative island closure effect for fledgling success, detracting from the chick survival results there.**"**

"The results at St Croix and Bird Islands do not involve any response variables that are directly interpretable in demographic terms and they are also not consistent in their direction.
The scale of the island closure effects, viewed across all results, are also not such that island closures could markedly alleviate or avert the rate of decline in penguin numbers, and do not justify the contention that urgent action to close islands to fishing is required. **"**

"2.2.1 *"The implications of inaction will not only have dire ecological impacts to penguins but will also influence the socio-economic benefits that this species provides."*
This statement implies that there is great urgency to implement island closures, failing which, dire consequences will follow. The results from the island closure experiment do not indicate that there is a critical role for island closures in the health of the penguin population. The islands in the experiment include about 70% of the breeding population. Only at Robben Island are there some reasonably firm and demographically relatable results which may suggest a positive island closure effect. These translate to a roughly 1% increase in the population growth rate at Robben Island, which in 2019 comprised 1216 breeding birds out of a South African population of 13312 breeding birds. It is of pertinence that the island closure experiment has involved closing Robben Island to fishing 50% of the time, so that only half of the possible benefit from closing Robben Island would be additionally realized by now closing Robben Island every year. This amounts to 0.5 % of the population size of the Robben Island breeding population of 1216 birds. Closure around Robben Island would therefore have a minimal positive effect on the overall South African penguin population for which, since 2008, the average rate of decline has been 6.5% per annum (see **FISHERIES/2020/APR/SWG-PEL/32**).
Another perspective on the island closure experiment is obtained by considering that the experiment has involved island closures since 2008 at four breeding islands comprising some 70% of the breeding population in South Africa, 50% of the time. If island closure had the potential to substantially retard penguin population declines, this would already have been obvious. But the penguin breeding population has continued to decline since 2008 at a concerning 6.5 % per annum. A conclusion from this is that the benefits of island closures for penguins is most likely not very large.**"**