Some observations on comparisons of fitting to the annually aggregated and the individual data, this time using JAGS and for the cases considered in FISHERIES/2020/JUL/SWG-PEL/53REV.

Richard B. Sherley[1,2]

1. Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Cornwall, TR10 9FE, United Kingdom.
2. FitzPatrick Institute of African Ornithology, DST-NRF Centre of Excellence, University of Cape Town, Rondebosch 7701, South Africa.

Email: r.sherley@exeter.ac.uk

Although similar results have been offered to the SWG-PEL on at least two occasions in the past (Sherley 2016, Sherley and Winker 2019), it seems that it may be useful to offer some additional observations on comparisons of fitting comparable models to the annual means and the observation-level data for some of the island closures datasets. Here, given time constraints, I have focussed on the datasets analysed in Sherley (2020a).

Table 1 below offers empirical results based on fitting comparable models to some of the annually aggregated datasets and individual datasets for the Island Closures experiment. They show that comparable models fit to the aggregated and individual data do not result in radically different precision estimates (as the panel put it, they have "*similar standard errors*" [Die et al. 2019]. And, contrary to the suggestion by Butterworth (2020), the maximal models (each M1 case in Table 1) from (Sherley 2020a) are indeed providing more precise estimates than the fits to the aggregated data in each case. They are not always larger, but they are (as would be expected from a model that has more statistical power) always more precise.

Crucially, however, Table 1 also shows that **there is a > 95% probability that the closure effect is genuinely positive for the penguins for chick condition at Robben Island, chick survival at both Robben and Dassen Island and for Maximum foraging distance at St. Croix whether the aggregated data or individual data are used**. Moreover, these probabilities barely change between the models fit to the aggregated data or the individual data. Ignoring these results – particularly when taken in the context of similar (albeit not the same) results for these penguin metrics in Ross-Gillespie and Butterworth (2020) – seriously risks a Type II error and would certainly not be in keeping with the application of the precautionary principle as required by the Marine Living Resources Act 18 of 1998 (MLRA), and the National Environmental Management Act 107 of 1998 (NEMA).

First, these results provide some context to Butterworth's (2020) purely theoretical example in the Annex of FISHERIES/2020/AUG/SWG-PEL/82 which is meant to demonstrate that "*even if the random effects approach to making use of individual data can fully account of their non-independence, and hence prevent this from negatively biasing estimates of the standard error of the island closure effect, the resultant estimates could not have better precision than those provided by corresponding models based on annually aggregated values of measurements of the response variables".*

Second, they underline why the 2019 IWS panel said: "*Given the nature of the experiment, use of individual data is to be preferred. However, this is only the case if an appropriate random effects structure is chosen*" and "*results presented to the Workshop suggest that estimates of closure parameters using models fitted to aggregated and individual data had similar standard errors*" (Die et al. 2019).

Third, they provided more support to the concluding statements in Sherley (2020) that meaningful, unbiased, positive effects of the fishing closures are apparent in all three of the following cases: a) When the observation level (disaggregated) data were modelled using random effects that were selected a priori by the analyst to adequately reflect the sampling structure of the data (Sherley et al. 2018, 2019); b) When the observation level (disaggregated) data were modelled using random effects that were selected using model selection (Sherley 2020a), as advised by the 2019 Panel (Die et al. 2019); and c) When the aggregated data (annual means) are modelled (Sherley and Winker 2019, Ross-Gillespie and Butterworth 2020, this document).

Fourth, they underline the conclusion that, no matter how we analyse these data, meaningful effects of fishing around African penguin breeding colonies are apparent (see also Table 1 in Sherley 2020b) and, importantly, that some of those effects are on variables (chick survival, fledging success) that contribute to the demographic process.

And fifth, it is worth noting that the evidence for meaningful or significant positive effects of the closure outweighs the evidence for meaningful or significant negative effects (Sherley 2020b). Sherley (2020c) finds 5 effects that show a positive benefit of the closures for penguins with > 95% probability (2 at Robben, 1 at Dassen, 2 at St. Croix) and only 1 effect (at St. Croix) that shows a negative effect of the closures for penguins with > 90% probability (4 times the evidence for a positive effect vs a negative effect). Moreover, the Overall Effect Size in Sherley (2020a,c) suggests 2–3 times the evidence for a positive effect versus a negative effect of the closures on penguins. This document finds those effects are consistent whether using the aggregated or individual data. And Sherley (2020b) demonstrates that this inference is not changed if the interaction Island × Closure interaction (Island specific effects) is retained in the chick survival model.

Meanwhile, Ross-Gillespie and Butterworth (2020) find 1 metric (from 7 considered) showing "evidence in the current data of a biologically meaningful fishing effect" in the direction that shows that the closures benefit the penguins at Dassen Island and 3 showing a meaningful effect in the opposite direction. For Robben Island, Ross-Gillespie and Butterworth (2020) find 2 metrics (from 7 considered) showing "evidence in the current data of a biologically meaningful fishing effect" in the direction that shows that the closures benefit the penguins", 1 that requires 2 to 5 more years before it would have sufficient power, and 0 showing a biologically meaningful effect in the opposite direction (4 trend in the opposite direction, but would require 10 more years of the experiment, thus are poorly estimated). Additionally, FISHERIES/2019/NOV/SWG-PEL/27rev, considering 3 metrics, found a significant (at the 5% level) negative impact on maximum distance of the fishery at St Croix and no significant effects either way for trip duration and path length. For Bird Island, FISHERIES/2019/NOV/SWG-PEL/27rev found no significant effects either way.

In conclusion, looking across all of the most recent analyses, we find a total of 3 (statistically significant/credibly different from zero/biologically meaningful) metrics showing a negative effect of the closures for penguins and 7 (statistically significant/credibly different from zero/ biologically meaningful) metrics showing a positive effect of the closures for penguins (split across Robben, Dassen and St. Croix), with 3 others showing strong support for a positive benefit (see Table 1 in Sherley 2020b). All the other effects are too poorly estimated at present to allow us to separate signal from noise. Of course, these 8 effects are not all independent (there is some double counting). But this does show that the most cases where effect sizes are well estimated, these effects point to a benefit of the fisheries closures for penguins (see Table 1 in Sherley 2020b).

Given the above, I suggest we cease having a technical debate on whether (generalised) LMMs can adequately balance Type I error rates and statistical power when random effects are selected by model selection approaches (since exactly this point has been established by simulation studies elsewhere; Matuschek et al. 2017). Instead, we should focus on the key conclusion that two independent sets of analyses agree that biologically meaningful effects of fishing around African penguin breeding colonies are apparent and focus on making recommendations that can aid management.

References:

Butterworth DS. 2020b. A response to Sherley: FISHERIES/2020/JUL/SWG-PEL/53REV. Department of Environment, Forestry and Fisheries Report: FISHERIES/2020/AUG/SWG-PEL/82. Pp. 1–38.

Die DJ, Punt AE, Tiedemann R, Waples R and Wilberg MJ. 2019. International Review Panel Report for the 2019 International Fisheries Stock Assessment Workshop, 2–5 December 2019, UCT. Department of Environment, Forestry and Fisheries Report: MARAM/IWS/2019/General/5. Pp. 1–18.

Matuschek H, Kliegl R, Vasishth S, Baayen H and Bates D. 2017. Balancing Type I error and power in linear mixed models. Journal of Memory and Language 94: 305–315.

Ross-Gillespie A and Butterworth DS. 2020. Updated implementation of the Algorithm recommended by the Panel for the 2016 International Stock Assessment Workshop for assessing whether or not to continue with the penguin island closure experiment. Department of Environment, Forestry and Fisheries Report: FISHERIES/2020/JAN/SWG-PEL/09. Pp. 1–17.

Sherley RB. 2016. Additional analysis suggested in response to differences in variance estimates between Sherley (2016) and Ross-Gillespie & Butterworth (2016). Department of Environment, Forestry and Fisheries Report: MARAM/IWS/DEC16/Peng Clos/P4. Pp. 1–4.

Sherley RB and Winker H. 2019. Some observations on comparisons of fitting to the annual means and the observation-level data for the cases in MARAM/IWS/DEC19/Peng/P4 that support a positive effect of the island closures experiment on African penguins. Department of Environment, Forestry and Fisheries Report: MARAM/IWS/2019/PENG/WP3. Pp. 1–5.

Sherley RB. 2020a. Revisiting the key results in MARAM/IWS/2019/PENG/P4 in light of the 2019 Panel recommendations. Department of Environment, Forestry and Fisheries Report: FISHERIES/2020/JUL/SWG-PEL/53REV. Pp. 1–27.

Sherley RB. 2020b. A reply to Bergh: FISHERIES/2020/AUG/SWG-PEL/84. Department of Environment, Forestry and Fisheries Report: FISHERIES/2020/SEP/SWG-PEL/87. Pp. 1–18.

Sherley RB. 2020c. Model selection results for the remaining penguin metrics that can currently be analysed using an individual data approach. Department of Environment, Forestry and Fisheries Report: FISHERIES/2020/SEP/SWG-PEL/89. Pp. 1–7.

Sherley RB, Barham BJ, Barham PJ, Campbell KJ, Crawford RJM, Grigg J, Horswill C, McInnes A, Morris TL, Pichegru L, Steinfurth A, Weller F, Winker H and Votier SC. 2018. Bayesian inference reveals positive but subtle effects of experimental fishery closures on marine predator demographics. Proceedings of the Royal Society B: Biological Sciences 285: 20172443.

Sherley RB, Barham BJ, Barham PJ, Campbell KJ, Crawford RJM, de Blocq A, Grigg J, Le Guen C, Hagen C, Ludynia K, Makhado AB, McInnes A, Meyer A, Morris T, Pichegru L, Steinfurth A, Upfold L, van Onselen M, Visagie J, Weller F and Winker H. 2019. A Bayesian approach to understand the overall effect of purse-seine fishing closures around African penguin colonies. Department of Environment, Forestry and Fisheries Report: MARAM/IWS/2019/PENG/P4. Pp. 1–25.

Table 1: Comparisons of (Generalised) Linear-mixed effects models implemented in JAGS and fit to either the annually aggregated datasets or individual datasets. Where annual aggregated data are used, the model structure uses Year as a random effect [as per the structures in model 2 in Sherley and Winker (2019) and the structure used in Ross-Gillespie and Butterworth (2020)]. Where individual data results are reported, they are based on the best fitting model and/or the model with the maximal random effects (M1) in Sherley (2020). If only M1 is given, this was the best-fitting model in Sherley (2020a).

Dataset	Region	Data type	Island	Closure effect mean	Closure effect SD	95% Credible Interval	Probability of a positive effect for penguins
Chick Condition	Western Cape	Aggregated	DASSEN	0.018	0.046	−0.073–0.109	66.4%
		Individual (M1)[1]	DASSEN	0.033	0.033	−0.031–0.098	84.6%
		Aggregated	ROBBEN	0.081	0.046	−0.010–0.172	**96.0%**
		Individual (M1)	ROBBEN	0.066	0.038	−0.009–0.142	**95.6%**
Chick Survival	Western Cape	Aggregated	No interaction	0.270	0.092	0.094–0.461	**99.7%**
		Individual (M1)[2]	No interaction	0.380	0.087	0.211–0.550	**100%**
Max. distance	Eastern Cape	Aggregated*	ST CROIX	−0.345	0.099	−0.539–−0.150	**99.9%**
		Individual (M6)	ST CROIX	−0.322	0.057	−0.434–−0.212	**100%**
		Individual (M1)	ST CROIX	−0.322	0.056	−0.433–−0.213	**100%**
		Aggregated[3]	BIRD	0.057	0.103	−0.145–0.263	27.7%
		Individual (M6)[4]	BIRD	−0.008	0.039	−0.084–0.071	58.2%
		Individual (M1)[5]	BIRD	−0.009	0.039	−0.086–0.069	58.8%

Notes: 1. Random effect structure was: Island/Year/Month; form for model equation = eqn. 1 in Sherley et al. (2019). 2. Random effect structure was: Island/Year/NestID; form for model equation = eqn. A2 in Sherley (2020). 3. Brood mass is omitted from this model and was used in M6 and M1 in Sherley (2020). 4. Random effect structure was: BirdID; form for model equation = eqn. 2 in Sherley et al. (2019). 5. Random effect structure was: Island/Year/BirdID; form for model equation = eqn. 2 in Sherley et al. (2019).