QUESTIONS ARISING FROM WELLER ET AL. ARTICLES

D S Butterworth

Note:

a) The primary purpose of this document is, by posing specific questions, to facilitate focus in a verbal discussion of the two Weller et al. articles concerned with some of their authors in a meeting of the Pelagic Working Group, given that the contents of those articles address issues on which the Group is tasked to provide scientific recommendations.
b) What follows is not an exhaustive list of queries on the Weller et al. articles, but seeks rather to focus on the more important of these queries.
c) For readers' convenience this document shows extracts from those documents in *italics*, followed by comments and questions in standard text.
d) In a number of places below, the abbreviated references WEL, ROB and BUT are used. These refer to:

   WEL: Weller F, Cecchini L-A, Shannon LJ, Sherley RB, Crawford RJM, Altwegg R, Scott L, Stewart T and Jarre A. 2014. A system dynamics approach to modelling multiple drivers of the African penguin population on Robben Island, South Africa. Ecological Modelling, 277: 38–56.

   ROB: Robinson W, Butterworth DS and Plagányi ÉE. 2015. Quantifying the projected impact of the South African sardine fishery on the Robben Island penguin colony. ICES Journal of Marine Science *ICES J Mar Sci*. 72(6): 1822-1833doi: 10.1093/icesjms/fsv035.

   BUT: Butterworth, D.S., Plagányi, É.E., Robinson, W.M.L., Moosa, N. and de Moor. C.L. 2015. Penguin modelling approach queried. Ecological Modelling 316, 78-80. http://dx.doi.org/10.1016/j.ecolmodel.2015.08.001

e) Figures from other papers referenced in the text below and also duplicated in this document for the ease of readers may be found at the end.

<u>Penguins' perilous conservation status calls for complementary approach based on sound ecological principles: reply to Butterworth et al. (2015) – Ecological Modelling 337 (2016) 1-3</u>

1) *the extended model in Weller et al. (2016) used fitted non-linear relationships between observed adult and immature survival and hydroacoustic estimates of prey biomass to drive dynamics in these age classes [The Weller et al. (2016) reference here refers to the second of the two Weller et al. articles considered here – see below]*

   This advance is a clear improvement on the approach used in WEL.

2) *we suggest that the ROB model "indicated a lack of impact of changes in anchovy (Engraulis encrasicolus) abundance on annual reproductive success" chiefly because it fails in this last regard by modelling reproductive success from hatching to 8 months of*

age as a single parameter responsive only to the biomass of anchovy recruits estimated in May. ….. juvenile survival, which they suggest may be scarcely or not observable …. survival during the first year of life at sea (i.e. from fledging to returning to moult)can be estimated directly using capture-mark recapture methods. This has been done most recently by Sherley et al. (2014) … productivity and first-year survival can be the result of very different ecological processes. … the approach (used in ROB) of considering the influence of anchovy recruit biomass as the only driver of survival from hatching to full independence in January of the following year is not plausible from a biological standpoint. BUT are correct that "anchovy dominate the penguin prey around Robben Island during the period of the year when chicks are reared, and provide the chicks main food source". However they neglect to mention that chicks are only reared for ~2-3 months post-hatching – during which time they depend on being fed (by their parents) from prey available in relatively close proximity to their colony. After this they leave the breeding colonies (there is no post-fledging parental care) and search for food over a far larger area … indeed analyses in Weller et al. (2016) (TRACE appendix) indicate that post-fledgling (like adult)survival at Robben Island may well be best explained by the availability of adult sardine west of Cape Agulhas …

Sherley et al. (2014) state that estimates of juvenile survival are available for fewer than 2% of seabird species, and further that they were "poorly estimated in most years" for Robben Island and the other colonies which they analysed – this is why the words "scarcely or not observable" were used. Clearly many factors may affect first year survival, and if it is broken down into pre- and post-fledging periods, the major factors for each may differ. But that does not in any way invalidate the approach used in ROB. Say:

$$S_{\text{first year}} = S_{\text{pre-fledging}}(\mathbf{p}) * S_{\text{post-fledging}}(\mathbf{q})$$

where **p** and **q** are the vectors of the (values of the) factors on which those components of the first year survival depend. Now if **p** includes anchovy abundance (a), then using information on $S_{\text{first year}}$ and a to estimate dependence between the two is not invalidated by a not being a component of vector **q** (though the reason for practice elsewhere – see BUT – is to consider (equivalents of) $S_{\text{first year}}$ rather than $S_{\text{pre-fledging}}$ only is because some components of vector **p** may also appear in vector **q** because of indirect knock-on effects – e.g. better fed pre-fledging chicks may be thereby better equipped to survive post-fledging, or alternatively such good conditions may attract more predators to the vicinity and reduce $S_{\text{post-fledging}}$). Thus the fact that Figure 8 of ROB shows the estimated relationship of net reproductive success over the first year (incorporating $S_{\text{first year}}$) as a function of anchovy recruit abundance (alone) does not thereby invalidate the relationship estimated, nor imply that there are not other variables impacting reproductive success (their impacts are reflected by the residuals about the relationship estimated in Figure 8 of ROB, which is duplicated below).

Q2.1: The analyses of ROB are important to the management of the pelagic fishery as they form the basis for estimates of the different effects of alternative levels of sardine catch on penguins, so were it the case that the ROB model is flawed as implied above, this would require the PWG to make a major change to its methods of analyses. However, given the explanation above, which indicates that any flaw lies rather in the logic of the criticism levelled at ROB, is it accepted that that criticism was incorrect and that the ROB model is not in question for that reason; or if not, why not?

3) *Despite the methodological criticisms levelled by BUT, the ROB and WEL models yield essentially the same answer to the question of fisheries closures around Robben Island. … while management actions that improve prey availability close to the colony (e.g. through fishing restrictions) may yield benefits in terms of population increases, these are likely to be swamped by the food situation in the wider Benguela ecosystem as a whole. …. Certainly the converging results for the Robben island colony have increased our trust in the model by Weller et al. (2016) to the level where we believe that it can usefully support decision-making*

   One of BUT's key points was that the analyses in ROB (Figure 8) shows no impact of anchovy fishing around Robben Island, in contrast to WEL that indicated there was such an impact, and that this key difference needed investigation. Note also comments reported in 4) and particularly in 9) below, the latter indicating that the impact of fishing close to Robben Island is substantial.

   Q3.1: How can comments above that ROB and WEL models yield essentially the same results, and that these results are converging, be reconciled with these last-mentioned facts; is it not rather the case that the original key difference between the two approaches remains and that there has been no convergence, and hence that this difference remains in need of discussion to shed light upon the relative reliability of the two different conclusions reached?

4) *the extended model in Weller et al. (2016) … arrived at unaltered conclusions regarding the effectiveness of suspending fishing for small pelagics around Robben Island; a result supported by the increase in chick survival observed during the recent three year closure to purse-seining at this colony (Sherley et al., 2015).*

   Comments in 3) above refer, but it is also to be noted that the model upon which this Sherley et al. (2015) conclusion is based indicates also that chick survival increases as sardine biomass diminishes.

   Q4.1: As a model must be accepted in its entirety in terms of its management implications, is your view either that: a) management should increase fishing on sardine to reduce their biomass and hence benefit penguins, or b) that this points to unreliability of this model?

5) *The WEL model is not designed to make precise numerical predictions about what population increase will follow a given management action, but rather to determine which action(s) should yield the greatest benefit given a specific environmental parameter space. …we reject their implied notion that models at the strategic level can automatically be discounted if they disagree with more limited approaches, e.g. because they do not provide comparable facilities for error estimation.*

   Comparison of the impacts of different actions using any model necessarily implies an ability to estimate the magnitude of those actions and also some measure of the error associated with each estimate. That is not to say that the methods of estimation for both the size and the associated error may not be coarse and involve some subjective elements. But it remains essential, particularly if the factor involved plays a major role in the conclusion drawn, that the method used be set out clearly and rigorously, and in a manner that allows for replication, so that it is potentially subject to critical review. This is the standard bar for international fisheries science meetings tasked with the provision of scientific advice for management. Eva Plaganyi, an acknowledged international expert in this field, comments that she is "not aware of any instance where a broad strategic model based on expert opinion, which fails badly to replicate observed abundance data, is accorded a weight anywhere close to that to a rigorous tactical model that is consistent with those data, in the provision of formal scientific advice for management at a national or international level". (See also 10) below.)

   Q5.1: Would you concede that the provision of no more than a single value for each of a number of key parameters backed by no more than the statement of "Expert opinion" fails to meet this necessary bar?

   Q5.2: Would you agree that results from a strategic model that fails to meet this bar cannot be preferred to those from a tactical model that provides satisfactory fits to data in the provision of scientific advice for management.

   Q5.3: Can you provide any counter-example to the comment by Eva Plaganyi?

6) *BUT claim that use of Bayesian methods for statistical analysis is sufficient to handle the (in some parts sparsely parametrized) system modelled in WEL. This is, however, not completely valid. Bayesian techniques, even when incorporating substantial subjective elements, apply to parameter estimation for a given model structure. Such approaches do not address situations in which the model structure is inadequate or dynamically changing in response to endogenous (e.g. long term consequences of short term decisions, especially when delayed) or exogenous (e.g. climate change) drivers. … Data-driven model structures may be useful for precise short-term predictions, but are weak in capturing structural changes (even when using Bayesian estimation). System dynamic models are useful in understanding effects of (delayed) feedbacks in complex system, which may fundamentally affect system response.*

Bayesian techniques can also be applied to compare parameters estimated in different model structures by use of the Bayes factor approach.

Q6.1: If effects are expected to change over time, why can this not be handled by making allowance for time dependence in elements of the model underlying the Bayesian estimation?

Q6.2: Why are such models any less able or useful to understand delayed feedback effects which can certainly occur in tactical models?

Q6.3: What aspects of the Weller models could not be incorporated in a Bayesian framework, and along lines providing tactical advice?

7) *If, however, results differ substantially between approaches, the underlying data and assumptions must be further examined. We maintain that for forecasts in a multiple-pressure, weakly parametrized system, the inclusion of all available data with appropriate caveats and ranges of variation is more sound than completely excluding such data and thus assuming a null effect without adequate analysis of the underlying model assumptions.*

Certainly all information should be utilised to the extent practical and defensible, though that is not to exclude models which take account only of major factors under the (motivated) assumptions that others are small or random (e.g. MICE models – see 8) following). Thus, for example, bias-variance trade-off considerations have to be taken into account. Internationally in providing tactical fisheries-related management advice arising from marine multi-species/ecosystem models, the generally accepted best practice is to use MICE models with a stress on fitting the data concerned satisfactorily, rather than to attempt broader "ecosystem" models which attempt to expand to the full set of data available.

Q7.1: Given the response in 2) above, is it accepted that ROB did not constitute a model that assumed the "null effect" mentioned here; and if not, why not?

Q7.2: Would you concur with the best practice as stated above, and thereby accept the view you put forward asserting the necessary use of all data is incorrect; and if not, why not?

System dynamics modelling of the Endangered African penguin populations on Dyer and Robben islands, South Africa – Ecological Modelling 327 (2016) 44-56

8) *In particular system dynamics models allow expert opinion to be incorporated to qualitatively define relationships for which underlying data are lacking. This makes them particularly valuable … but the data to inform approaches like MICE (Plaganyi et al., 2014) are not available. …… Approaches such as MICE (Plaganyi et al., 2014) cannot*

currently be implemented for Dyer Island because of the scarcity of data for a data hungry method.

ROB (applied to the Robben Island penguin population) is a typical example of a MICE model. Similar data are available for the penguins at Dyer Island (Ludynia et al., 2014).

Q8.1: Is the statement that such models cannot be applied to the colonies in question incorrect?

Q8.2: Are the Weller models not effectively MICE models that have not (yet) been taken through to the stage of sound parameter estimation through fitting to the available data in the standard process used for fishery assessments?

Q8.3: Why not fit the "Weller" model to these data in this way; would that not improve your results as well as provide a useful test of the robustness or otherwise of the results in ROB?

9) *Overall these findings do not seem sufficient to suggest a departure from the previous conclusion that both zones exert a roughly similar effect on penguin population development. …. Management efforts aiming for improved prey fish availability should thus be aimed at both the area close to colonies and the wider forage area (see also Sherley et al., 2015).*

While otherwise different results do nevertheless seem to show a common theme as regards sardine abundance in the wider west coast area being of importance to penguin dynamics, the totality of evidence for such an impact (if any) to abundance/fishing in the neighbourhood of Robben Island is much weaker.

Q9.1: Why then do you seemingly advocate equal weight be given to recommendations for fishing restrictions in both regions?

10) *While predation at Robben Island appears to be at low levels (in the case of terrestrial predation, due to ongoing feral cat control)…. [Subsequently the associated penguin adult survival proportion is allocated a value of 0.9 on the basis of "Expert opinion".]*

In principle predation by feral cats could be the primary reason for the negative penguin trend at Robben Island, so that the basis for estimating the magnitude and trends of the effect need to be set out in some detail (even if the Expert opinion offered on this point is indeed correct). (See also 5) above.)

Q10.1: Would you concur that there is a need for the basis behind the 0.9 value offered here needs to be set down in much more detail if readers are to be placed in a situation where they can have at least some opportunity to evaluate such a claim, and that the credibility of your results would be substantially strengthened through such a process?

11) *Under these assumptions model output replicates recorded population dynamics with acceptable accuracy (Figure 10, upper shaded series). [This Figure 10 is reproduced here immediately following the text.]*

Figure 10 suggests that the modelled current penguin abundance is at about its average over the period considered, whereas the actual data indicate that it is at an all-time low. In contrast, Figure 7 from ROB also reproduced and situated below Figure 10 is able to reproduce the abundance data almost exactly. In any fisheries assessment process, a lack of fit between model and data as substantial as in Figure 10 would lead to rejection of the model. (Note also that the Weller et al. population trajectory for penguins at Dyer Island also shows clear systematic deviations from the abundance data.)

Q11.1: Why is the Figure 10 claimed to reflect "acceptable accuracy"? What credibility can be given to recommendations associated with the model underlying Figure 10 when its output is so incompatible with the basic abundance data?

Q11.2: Should not the maxim of Francis (2011), that the primary emphasis in marine population modelling should be that the model provide a good representation of the available abundance indices, apply also to modelling Robben Island (and other) penguin populations?

12) *Basic model assumptions ….*

- *The penguin population is at demographic equilibrium at time zero … a necessary and generally recommended modelling practice to establish a model base state (Renshaw, 1993).*
- *All adult age classes are assumed to share a common survival rate …*

Using these estimates [for Dyer Island; from Ludynia et al., 2014] did not produce a viable equilibrium model population. The low adult and immature survival rates led to extinction within 5-10 years. It was found that raising both estimates to the equilibrium levels used in the Robben Island model (adult survival: 0.88, immature survival: 0.51) resulted in a stable configuration in combination with the unaltered egg and chick survival estimates. This setup is a reasonable assumption …

The Renshaw advice is dated. In current fisheries assessment modelling the importance of setting an initial age structure in line with the circumstances of the resource and fishery at the time is well realised.

Q12.1: Why replace values of annual survival rates provided by estimates based on data with some equilibrium value supplemented by a few ad hoc adjustments? Is this not a key reason underlying the poor match between model and data that is evident in Figure 10 below?

References

Francis RICC. 2011. Data weighting in statistical fisheries stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68: 1124-1138.

Ludynia K, Waller LJ, Sherley RB, Abadi F, Galada Y, Geldenhuys D, Crawford RJM, Shannon LJ, Jarre A. 2014. Processes influencingthe population dynamicsand conservation of African penguins on Dyer Island, South Africa. African Journal of Marine Science 36(2): 253-267.

Plaganyi EE, Punt AE, Hillary R, Morello EB, Theobaud O, Hutton T, Pillans RD, Thorson JT, Fulton EA, Smith ADM, Smith F, Bayliss P, Haywood M, Lyne V, Rothlisberg PC. 2014. Multispecies fisheries management and conservation: tactical applications using models of intermediate complexity. Fish and Fisheries 15: 1-22.

Renshaw E. 1993. Modelling biological populations in space and time. Cambridge, Cambridge University Press.

Sherley RB, Abadi F, Ludynia K, Barham BJ, Clark AE, Altwegg R. 2014. Age-specific survival and movement among major African Penguin *Spheniscus demersus* colonies. Ibis 156: 716-728.

Sherley RB, Winker H, Altwegg R, van der Lingen CD, Votier SC, Crawford RJM. 2015. Bottom-up effects of a no-take zone on endangered penguin demographics. Biology Letters 11. 20150237.

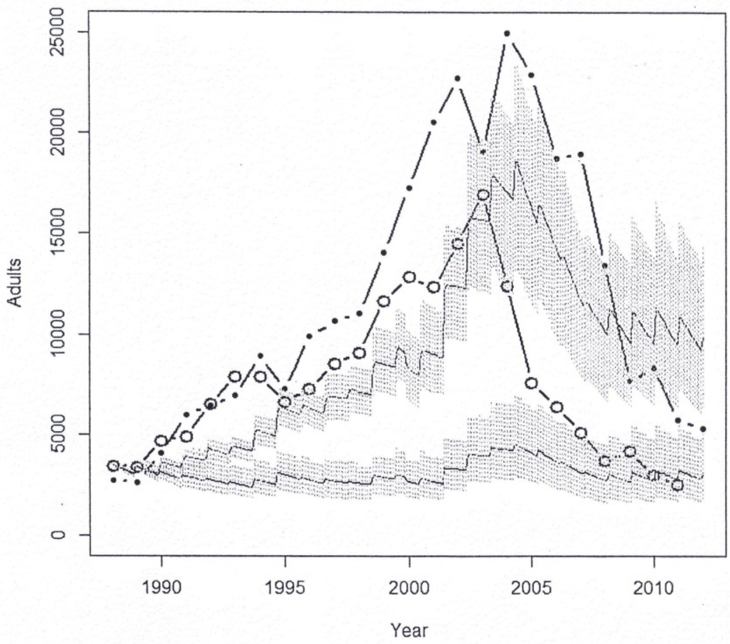

Figure 10. Recorded population development (number of adult African penguins) at the Robben Island colony 1988–2012, compared to model output using recorded food biomass for that period (5000 replicates). Empty circles: population based on moult count records; black dots: population exrapolated from nest counts. Lower shaded series: modelled series using default pressure parameters. Upper shaded series: modelled series using default pressure parameters and adding immigration, changes in shark predation, and changes in age at first breeding. Shaded ranges indicate 95% confidence intervals.

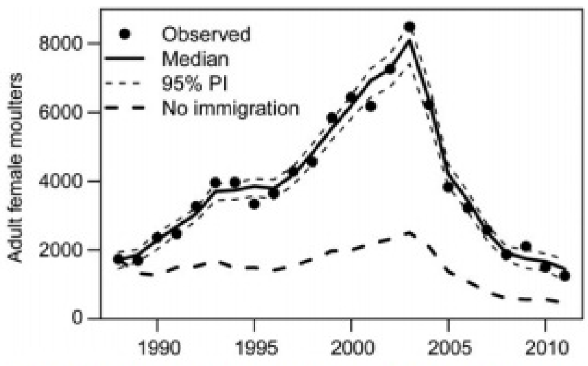

Figure 7. Time-series of female moulting penguins at Robben Island from observations (black circles). The medians and 95% probability intervals of the Bayesian posterior distributions for the model-predicted moult counts are joined by solid thick and light dashed lines, respectively. The thick dashed line indicates the median trajectory for the same demographic parameter estimates had the immigration estimated by the model not occurred.

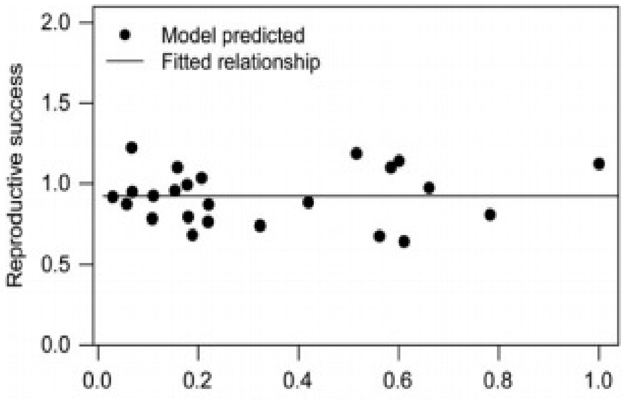

Figure 8. Estimated (posterior mode) and predicted values of penguin reproductive success plotted against an index of anchovy recruitment (prey abundance).