Collated comments from some members of the Small Pelagic Scientific Working group in response to FISHERIES/2011/SWG-PEL/29.

Contributions from: Doug Butterworth, Carryn de Moor, William Robinson and Janet Coetzee.

Background

In February 2010 a method for assessing the implementation efficacy of an Ecosystems Approach to Fishing in the sardine fishery was presented to the SWG-PEL.

At that meeting, concern was expressed that the proposed tool (and its results) might be viewed as a basis for decisions by managers, when in fact it should be rather regarded as a “flag waver” (the analogy of the IUCN Red List was used to clarify this comment) that highlights issues of concern that would then be subjected to more detailed investigation. A request was also made at the same meeting for a document listing the various indicators, the rationale underlying their selection, as well as the thresholds that are employed for each to be circulated to allow for more informed suggestions and comments.

It was suggested that the human well-being component should be represented relative to similar income brackets in other sectors of the economy to generate a more realistic perception. As the output stands at the moment, it could provide a basis for political targeting of the fishing industry as “mistreating” workers.

A component of this tool, i.e., the ecological wellbeing component is currently under further development as part of an MSc project. A presentation of the planned indicators and their thresholds to the SWG-PEL on the 17th May 2011 (FISHERIES/2011/SWG-PEL/29) resulted in mixed views about the utility of such a tool and the relevance of certain of the indicators and their thresholds. As discussion time was limited, the Chair agreed to collate comments from working group members and observers for consideration by the student.

Comments

Although the SWG-PEL members were explicitly asked only to comment on the indicators and thresholds, some of the comments received do point to the objectives as well. This document contains all comments received by the chair within 2 weeks of the meeting:

[Note: For convenience reference is given via the numbers of the Figures shown in Appendix 1 of FISHERIES/2011/SWG-PEL/29]

Fig. 1:

<u>Doug:</u>
Sound objective, and conceptually the indicator is fine. However the current implementation makes no sense – it implies that vessels with observers dumped MORE fish than vessels without observers over 2005-2006. Until this analysis has been refined and this anomaly hopefully resolved, the only immediate basis for an index might be frequency of dumping as recorded in observer reports, though even that would need GLM standardisation to account for the non-random nature of

observer placements. It is impossible to judge the appropriateness of the levels suggested without re-expressing the catch/hour difference on a relative difference scale so that meaning in terms of proportion of fish possibly dumped would be evident.

Carryn**:**
Please give the details upon which this is based. What does a negative effect mean?

Janet:
I have checked with Sobahle and he informs me that these are his GLM observer effect outputs which are currently under revision. A negative value could potentially arise when vessels with observers follow a more cautious fishing strategy, by for example only fishing in areas where other vessels have had a low percentage of juvenile sardine bycatch. Currently these observer effects are relative to the observer effect in 2001 (a year for which he has a lot of observer data) – I would suggest review of this reference to a year in which we would have expected dumping to be low (no observer effect) such as a year of very low sardine recruitment. In terms of choosing thresholds, I would prefer to use outputs from the assessment models to inform on increased level of risk to the fishery when unreported catches exceeds certain values, similar to the analyses Carryn conducted in 2010 when we were faced with reports of dumping. In any event, I suggest we first review Sobahle's updated analyses at the next meeting on the 18th June before agreeing to thresholds, or how to select them.

Fig. 2:

Doug:
Sound objective, but the obvious and standard index internationally is fishing mortality *F*. The estimator used here (for *F/Z*) is well known to be flawed in *inter alia* its equilibrium assumptions. The obvious basis from which to obtain *F* values is the assessment which adjusts for those flaws. Thresholds should be based on standard internationally accepted reference points such as F_{msy} or (given estimation difficulties) a widely used proxy such as *F*30% or *F*40% (the fishing mortality that reduces the pristine spawner biomass per recruit to the percentage indicated). This could be calculated easily from assessment results. Given that this is a forage fish species, the higher percentage might be more appropriate.

Carryn:
Von Bertalanffy parameters should be based upon those developed by Mtengwane (more years of data and possibly more accurate in some of the years). Secondly, the time series showing an exploitation rate of almost 80% in 2007 seems questionable – that the industry was responsible for 80% mortality of fish in 2006? The method suggested by Fairweather et al 2006 appears to be a proxy and is compared to results from assessment models and EwE output. What does such a time series look like using these other available data? This method is very simple and will produce results of variable bias. There are more standard methods used to estimate *F*.

Janet:
Agree to recheck this time-series – possible mis-match in years chosen for catch vs biomass and would support the use of an indicator derived from assessment outputs. I understand the need for taking total mortality into account in the context of these being forage fish, but don't quite get my head around the usefulness of using such an indicator. For example, in 2007 the catch of sardine decreased by 35% compared to 2006, yet the total exploitation went up by 40 %. One needs to be able to judge the impact of the sardine fishery (by itself) surely if one wants to access its impact on the ecosystem??

Fig. 3:

<u>Doug:</u>
Sound objective, but the choice of indicator seems questionable as the bycatch of juvenile sardine in the directed sardine fishery is minimal, and effectively irrelevant compared to the bycatch in the anchovy fishery. Thresholds for that are already set in the OMP, but they don't translate to the monotonic form of relationship seemingly required for this approach, and would not sensibly do so because the acceptable level at any one time would be a function of both the current sardine and anchovy abundances.

<u>Carryn</u>**:**
Once again details are not provided. I suspect annual catch data has been used with an annual (May/June) cut-off length? Given the rapid growth experienced by recruits, this would not be accurate. Catch data need to be considered monthly (at least), with monthly varying cut-off lengths (eg see PEL/23). This is easily combined back to an annual percentage.

<u>Janet</u>**:**
The use of the recruit cut-off length would only be appropriate for catches taken up to May/June (the cut-off is the maximum size of a recruit). For some years, it is very difficult to determine the cut-off length, as recruitment occurs over an extended period (granted they are used in determining recruit biomass as well). For example in 2002 and 2003, this length is 16 cm (caudal length) and a nice size for canning – Industry would possibly even have been targeting that size.

Fig. 4:

<u>Doug:</u>
The objective is questionable. This is relevant only if there is more than one sardine stock, with catches being taken being spatially disproportionate to relative abundances of the stocks. Thus if this is to be considered, the methodology needs to factor in the relative plausibility of single to multiple sardine stock hypotheses. Threshold choices here would first require the development of a multi-stock model subjected to simulation testing to determine what level of spatial disproportionality led to biomass levels dropping below the thresholds associated with Fig. 6. There would be potential utility in the case of a single stock if some empirical linkage could be developed between the distribution of spawners and subsequent recruitment; that hypothesis would need to be tested against existing survey and assessment data before it might be taken further.

<u>Carryn</u>**:**
The 4th and 5th indicators place a 100% weight on the assumption of 2 (or more) sardine stocks. This is still under investigation.
4th indicator threshold: If you are measuring the proportion of sardine catch west of Cape Agulhas and you want to compare that to the proportion of total sardine biomass west of Cape Agulhas, the latter varies annually. You cannot have fixed thresholds of 10/20/30% which do not vary by year. If your program is hardwired such that you HAVE to have fixed thresholds, then you will need to rework your indicator such that the percentage in the catch is compared to the percentage in the population and thresholds of that ratio are considered.

<u>Janet</u>
If the sardine population consists of a single stock that does not show spatial tendencies in preference of spawning location and if spawning products from one location are as likely to be successfully transported to nursery grounds as those of any other location and therefore

recruitment success from spawning in any one location is no different from recruitment success when spawned in a different location, then I would agree that we do not need to consider spatially disproportionate fishing.

Evidence is starting to mount that this is not the case (although we probably still do not have enough data as conclusive proof) – some detail: Between 2004 and 2009 we had very poor recruitment despite widespread spawning by sardine on the shelf edge (mostly east of Cape Agulhas). SARP data suggests very few sardine egg and larvae in the jet current since 2002 – consistent with IBM modelling studies which suggested reduced transport success for eggs spawned on the south and east coast. First time many eggs noted on West coast again was Nov 2009, resultant good recruitment in 2010. Importantly though we need to finalize counting of egg samples from November surveys and analysing SARP data from last few years so that we can properly test the hypothesis that states: Above average sardine recruitment is more likely if sufficient eggs are spawned on the WAB (or something to that effect). Such data can then also feed into the sardine assessment model to provide us with the answers we need to set thresholds/manage the fishery.

At the current moment, we are looking at capturing some of the above scenario in the two-stock hypotheses and therefore don't yet know what the impact would be if the population was not comprised of one stock or if the population was comprised of one or two functionally discrete adult assemblages.

In that context (we don't yet know), it is difficult to judge indicators or thresholds. In terms of figure 4, it is unclear what data are plotted. Percentage of the total catch taken west of Cape Agulhas (catch west of CA/total catch) - if so, it does not look right, at least 90% of the sardine directed catch would have been caught west of Cape Agulhas up until about 2002 – they only started increasing fishing effort farther east after that. Nevertheless, it would be important to look at the catch taken west of Cape Agulhas as a proportion of the biomass available in that area (under the assumption that adult fish are not moving around too much). There is no problem in taking 100% of the catch west of Cape Agulhas if that 100% is only 10% of the population west of Cape Agulhas, but if that 100% is 90% of the biomass west of Cape Agulhas I would think there would be more reason for concern – even without conclusive proof that we don't have a single stock.

In choice of indicators and thresholds, I would also prefer to see assessment model outputs of the two stock scenario for various levels of fishing on each stock to inform on threshold choice. Until that is done, this threshold will remain a thumb suck – and it is no secret that I have been advocating spatial management for a while –to deal with this (in the context of limited time/data, one would ideally like to switch this on and off for now or apply some weighting linked to its plausibility).

Fig. 5:
<u>Doug:</u>
The same comments as under Fig. 4 apply.

<u>Carryn</u>:
It is not clear exactly what cut-off lengths are being used in this indicator. I suspect they are the cut-off lengths of recruits derived for the annual May recruit surveys? I hope I am wrong. However, if I am right, these cut-off lengths do not define "large" sardine, but rather define recruits of the year in May/June – and, as mentioned above, do not apply for the full year. In addition, large recruit catches are "bad" according to the 3rd indicator (Figure 3), but here large non-recruit catches (to the west) are also "bad". So you are penalising catches below and above the same cut-off length, leaving no possibility for an optimal size of sardine catch inbetween. I think a different cut-off length defining "large" sardine would be more defensible.

Janet**:**
I agree this is not feasible, i.e., nothing may be caught west of Cape Agulhas using this method. If one had to use a measure for large sardine, surely it would be linked to the reproductive potential of large vs small fish. Would something like the size at 50% / 75% maturity be something one could use as a starting point? Also surely the ratio would be relative to the abundance of “large” sardine in the area west of Cape Agulhas in the previous year? [Same argument as in 4 above]I assume the hypothesis here is that we need enough spawning sardine on the WAB to ensure successful sardine recruitment?

Fig. 6:
Doug:
Sound objective but inappropriately worded: delete the “in the presence of fishing”, so that this effectively becomes a biomass limit reference point as standardly considered internationally. The wording in the legend is confused. These values are presumably the reference case sardine assessment results, though it might be simpler to convert them using the estimated survey biomass bias factor so as to correspond to the survey results scale. This in turn would allow thresholds to be chosen also in relation to the current OMP exceptional circumstances thresholds which relate to the more readily comprehended November survey results.

Carryn**:**
Clarity needs to be provided as to where the values came from for this plot. I provided values for 1984-2006. I have not yet provided values for 2007-2009. The values provided for 1984-2006 can be compared to the risk threshold calculated for OMP-08, from whence the 459, 61**6**, 832 000t come from. An updated time series of 1984-2010 (yet to be finalised) will result in an updated pdf for Risk to be used for OMP-12.

Fig. 7:

Doug:
The objective is questionable here as it does not relate simply to a “good/bad” state of the environment. Sardine weight can be low either because of poor food production, or because the population is large thus reducing food availability per individual through density-dependence brought about through feeding competition. Accordingly evaluation on the basis of a monotonic relationship as the threshold system seeks is problematic, clearly rendering threshold selection problematic as well. Perhaps one might rather consider this as OK inside some range, but not OK outside, eg by transforming the index value *y* to mod(*y-ymean*)? The index suggested needs to be specified more clearly. Are these survey or commercial samples, and what GLM-standardisation procedure is being used to account for possible non-random sampling in space and time?

Janet**:**
Agree that there may be interactions, but am unclear on the calculation used to derive this index or the source of the data (commercial or survey) so cannot really comment**.** An observation possibly worth considering is that although the relative weight has decreased from 1987 to the present (I think in the absence of error bars), one would think that the sardine population (and ecosystem) benefitted from the large sardine population increase we saw during the late nineties to around 2003, yet this indicator suggests the opposite. Also it may be worth considering how fishing practices may bias this timeseries (spatial component).

Figs 8-12:

<u>Doug:</u>
The underlying intent here is perfectly reasonable, but problematic in the context of an EAF that relates to the ***sardine*** fishery ***only***. These birds' diets do not consist of sardine alone, so that unless there is evidence that their population status is nevertheless almost entirely dependent on sardine availability, these considerations belong in a separate analysis which considers the status of all forage species in some cumulative way.

A problem here is the definition of "acceptable levels" for defining thresholds. The numbers put forward appear near arbitrary. Some in principle basis first needs to be specified and justified, followed by consistent application across species and areas.

Fig. 8: The threshold choices (difficult anyway to understand given that the index is not defined) seem strange. The time series scarcely exceeds the lowest of these thresholds put forward, yet it includes the late 1990s where there is no question that the Robben and Dassen penguin populations were doing very well as they increased rapidly.

Fig. 11: Something seems wrong with the 2004 data point – if related to population status, the species dynamics are such that the implied 2004-2005 increase is likely demographically impossible.

<u>William:</u>
As I read it, in Underhill and Crawford (2005) the only factor included in the calculation of the index is the number of breeding penguins, and the baseline value is set at a 1950s population estimate. I gather from your talk that your index includes breeding pairs, moult counts as well as breeding success. First up, I think that trying to combine these series into a single index is dubious as they likely depend on things, but nevertheless it would be interesting to know how the different factors are weighted. Note that there is a high correlation between breeding pairs and moult counts... these data are certainly not independent.

The time series in Figure 9 is wrong. Note that the largest Eastern Cape colony is St Croix, which was not counted between 1990 and 1998, apart from 1993. This is why most of your data points in the 1990s are about 20000 lower than they should be.

Figure 8, the index values never go above 0.5, even in the "good" years up to the recent population peak in 2004. I guess that you are still using a relatively high 1950s population estimate as a baseline. If so, can you motivate why the thresholds of 15000 and 25000 were chosen for the Eastern Cape breeding pairs, when the number of breeding pairs in this area in the 1950s is only about 6000 pairs?

<u>Carryn</u>**:**
The derivation of the thresholds for penguins / seabirds is not clear. During your talk you mentioned why you chose some of the earlier thresholds. It would be good to have such descriptions for all the thresholds together with appropriate references, rather than just expert opinion.

<u>Janet</u>
Objective: "African Penguin population on Eastern Islands in good nutritional condition"- apart from problems mentioned during the meeting about interactions with other factors, the indicator chosen does not relate to the objective. Surely it is possible to have a large breeding population (possibly short-lived) in bad nutritional condition? Also in this context, the Indicator being used is not supported by recent diet studies of penguins in Algoa Bay (Pichegru in prep/subm). These suggest that of the small pelagic fish found in the diet of breeding penguins in Algoa Bay, 97% was anchovy. So using a penguin health indicator (whatever it may be) to evaluate EAF in the sardine fishery is not logical....unless otherwise can be shown.

Same general comment as Doug's to figure 8 applies to figure 10 and 12 as well. No positive response is observed in these indices, despite increases in sardine biomass in the late nineties/early 2000s; consequently the use of these to inform on the sardine fishery is questionable. Conversely, swift terns have continued to increase despite recent declines in sardine ....same argument.

More general comments

Doug:

While appreciating that the document was intended as an illustrative rather than comprehensive presentation, it is not possible to evaluate it as requested unless full specification of the suggested indices, including mathematical detail where relevant, is provided.

While it is reasonable to seek focus in responses on suggested threshold values, it is not possible to respond without putting such responses in context. As indicated above, logical inconsistencies arise for some of the indicators advanced.

The document refers to further analyses of results using "NetWeaver". The technical details of such an approach need (in due course) to be provided to allow judgment of whether such an analysis is appropriate.

The document (pg 3) refers to a thorough sensitivity analysis in relation to indicator threshold values, but does not give details, This would be complex as multiple non-independent indices are involved. A factorial design might be appropriate, followed by GLM analysis – Statistical Sciences staff may be able to give technical advice on this.

As indicated above, the predator components do not properly belong in this sardine-specific document, as they relate to overall forage fish abundance and hence are not affected by fishing on sardine only. The rest of the indicators essentially relate to resource status and level of fishing pressure. While the approach advocated might have some applicability for a data-poor fishery, its relevance (particularly given its advised objective of a summary report to managers) is questionable in a data –rich case. There is currently growing acceptance worldwide of a simple standard summary report format for managers for the latter situation – the *F vs B* "Kobe plot" as now adopted for US fisheries and all tuna RFMOs. Admittedly because of high recruitment fluctuations in a short lived species, the *B* axis thresholds might better relate to limit than target reference points in this instance, but otherwise this seems the obvious way to go for the purpose indicated.

In a previous presentation to the PWG, this approach was advanced as a flag-waver – as with the IUCN redlist to provide a simple first pass indication of possible problems (which if indicated lead to more detailed analyses), and as such would be perfectly acceptable in principle. For sardine, the flag was already effectively waved long ago, and for decades detailed analyses have been undertaken. However to use this as a vehicle to report resource status to managers, when effectively coarse and problematic indicators and near-arbitrary thresholds are used ahead of rigorous quantitative assessment can scarcely be advocated as the requisite best available science (unless supported by some further analysis showing that the assessment provides estimates than are less robust than do the crude indicators, whose consideration is in any case part of the assessment process). Suggestions that analyses such as those put forward in SWG PEL/29 might form the basis for resource status advice to Management have never been deliberated in the PWG, and should certainly first be so if this is a serious proposal, given that (as presented, though admittedly in somewhat initial form) they appear to advocate replacement of best available (albeit imperfect) science by considerably inferior analyses.

Carryn**:**

It appears that many of the thresholds are simply (expert) guesses? I'm glad to see, given the text on page 3, that sensitivity to all thresholds will be tested. I trust this will be done in different ways, eg varying upper threshold only, lower threshold only, median threshold only, upper and lower together (above / below), upper, lower and median together (above/below by smaller and bigger amounts) etc. And to account for the fact that some thresholds are not independent of others.