Agreement on the Conservation of Albatrosses and Petrels

Fourth Meeting of Seabird Bycatch Working Group

Guayaquil, 22 – 24 August 2011

ACAP At-Sea Priorities Framework – Identification of an appropriate scoring and weighting regime and other final steps

Michael C. Double, Juan Pablo Seco Pon, Spencer Clubb, Ian Hay and Barry Baker.

ACAP At-Sea Priorities Framework – Identification of an appropriate scoring and weighting regime and other final steps

Michael C. Double, Juan Pablo Seco Pon, Spencer Clubb, Ian Hay, Barry Baker.

Purpose

This document briefly summarises the development and current form of a framework for the identification of ACAP's at-sea priorities for action. It also describes a simple exercise that explored the performance of the framework in its current form relative to expert opinion. This approach will assist in identifying how the scoring and weighting regimes within the framework can be adapted to ensure that the high priority conservation actions identified by the prioritisation framework are credible, defensible and, as near as possible, are congruent with expert opinion. The final section describes a potential work plan that will complete this project by: 1) identifying the appropriate weighting and scoring regimes for the framework; 2) deciding upon the most appropriate format to present the results from the framework; and 3) providing recommendations for the future use of the framework by ACAP and its Parties.

Background

AC6 Doc 15 sets out a history of the development of both the 'on-land' and 'at-sea' frameworks that aim to identify priorities for conservation actions. Progress to date on the 'at-sea' framework by the AC and its associated Priorities Working Group (PWG) can be summarised as follows:

- The need for a process to identify conservation priorities was first identified at AC3. A conceptual framework was developed at AC4 and a more detailed framework and set of preliminary results was presented to AC5 (see below for a description of the current framework).
- While accepted in principle at AC5 there was a clear need to ensure that the larger and more complex at-sea prioritisation framework contained the best available information and that it would produce credible, defensible outcomes
- A significant peer review process has now been completed for the majority of the records within the at-sea framework.

The at-sea priorities framework – a reminder

Currently the at-sea prioritisation framework contains 1278 records. Each records is essentially an assessment of the threat posed by a single fishery to a single population of a ACAP listed species. This assessment is composed of the following factors:

- Population size (5 levels)
- Percentage of total species' population size (3 levels)
- Population trend (6 levels)
- Overlap with fishery (4 levels)
- Effort in fishery (4 levels)
- Risk of fishing method (5 levels)
- Use of effective mitigation in fishery (3 levels)

- Likelihood of successful mitigation if best-practice implemented (3 levels)

Each level within a factor is linked to a numeric score based on the scoring regime for the framework. The numeric scores are then used in the following equation to produce a total final score for each record:

A = (Population size + Percentage of total popln + Population trend) x (Vulnerability weighting/3)

B = (Fishery overlap + Fishery effort + Risk of fishing method) x (Threat weighting/3)

C = Use of effective mitigation score x Effective mitigation weighting

D = Likelihood of success score x Likelihood success weighting

Final score = A + B + C + D

Four weightings are included in this formula associated with groupings of factors that seek to capture:

1. the species vulnerability to fisheries bycatch;
2. the threat posed by the fishery to the population;
3. the current implementation of effective mitigation in the fishery; and
4. the likelihood of successful mitigation if best practice methods were implemented.

The significant progress to date has focused first on the identification of the appropriate factors and levels within each factor that will adequately describe the threat posed by a fishery to a population; second on the listing of all appropriate species/fishery records; and third on allocating for each record an agreed level for each factor.

Identification of an appropriate scoring and weighting regime for the framework

The next step in the development of the frame work was the identification of a scoring and a weighting regime that will produce intuitive and defensible final scores.

This process began by allocating scores to the levels within each factor. As these levels are mostly ordinal (e.g. steep decline, decline, stable increasing), the scoring regime reflected this order – but when the factor was scored as ‘unknown’ a decision was made as to how best to score this level (see Table 1 for examples).

Similarly, through discussion, the PWG allocated weightings to each factor or groups of factors (see equation above) based essentially on intuition. However, this is not easy to do because it is not simple to assess the interactions between the scores used for each factor and the weighting regime used to generate the final scores. As such, it may be difficult for experts to assess whether the scoring and weighting regimes, and even the equation itself, are behaving in the expected manner.

To assess the performance of our ‘first’ (subsequently named ‘original’; see Table 1) scoring and weighting regimes we investigated whether they would rank records in a similar order to those that have been ranked manually by recognised experts in this field.

To do this 20 records were selected randomly. The names of species, islands and fisheries were removed from each record to leave only the scored, descriptive categories – the

individual fisheries and species could not be identified. The same 20 records were then sent to two experts to rank independently on the basis of 'priority for action'.

The relative rankings of the two experts correlated well with each other (correlation coefficient = 0.82; Figure 1). This implies that based on the information described within each records these two experts at least could agree on which of the 20 selected fisheries should be prioritised for conservation actions. Unfortunately, however, the expert rankings correlated less well with the rankings derived from our 'original' scoring and weighting regime. The correlation coefficients were 0.32 and 0.39 for the two experts (Figure 2). This result was somewhat disappointing but not entirely unexpected given the reasons stated above. This demonstrates that while expert opinion should be influential in assessing the performance of the framework it could be overly influenced by one or two factors, an approach which may not on further reflection be entirely defensible.

The next question we posed was whether there are any scoring and weighting regimes that could generate rankings that correlate more closely with the experts' opinions. We investigated this by generating 60,000 scoring and weighting regimes, applying each to the records in the framework, ranking the records by their final scores and then calculating the correlation with the rankings of each of the two experts. The simulated weighting and scoring regimes were generated through a randomisation process but it maintained the ordinal structure of each factor. For example, in the factor 'population trend', a steep decline would always score higher or equal to decline, but the absolute score may differ between each scoring regime.

This process showed that modifications to the 'original' scoring and weighting regime could generate rankings that closely correlated with those of the two experts (see Figure 3). The three 'top' scoring regimes (highest combined correlation coefficients) are presented in Table 1; these regimes generated correlation coefficients with the expert's rankings from 0.82 to 0.94.

That the framework can deliver rankings that correlate well with expert opinion is reassuring and so these 'top' regimes can provide a basis on which to assess which factors, and levels within factors, are the most influential to the ranking decisions made by the two experts. From a superficial assessment of the top simulated regimes it appears that the experts ranked highly those records that included populations in steep decline, decline or unknown and very small populations. Unsurprisingly, therefore, each of the top regimes had a high score for the 'vulnerability' weighting. The two experts appear to have placed less weight on the threat posed by a fishery and whether effective mitigation was in place.

None of the 'top' simulated scoring and weighting regimes should be adopted as they are presented here as they may not be logically consistent within each factor or are not broadly intuitive or defensible – further thought and investigation is required, including potentially requesting experts to revisit their ranking decisions or to rank new records. For example, after the application of a generally well performing scoring and weighting regime to the framework, experts still identified clear anomalies. Some fisheries with strong use and implementation of effective mitigation and a proven record of significantly reducing bycatch were ranked very highly as priorities for further conservation action. This anomaly may not necessarily reflect a problem with the scoring regime, but may in fact indicate that the experts did not consistently rank such fisheries as the lowest priority for action. Greater understanding of such details will lead to refinement of specific criteria to assess the performance of scoring regimes (see recommendations below).

Findings and final steps

In summary, there has been a considerable amount of productive investment into developing and populating a framework to transparently identify and justify ACAP's at-sea priorities for conservation action. Although the bulk of the work is now complete a few important steps remain.

The simple analysis presented here has raised doubt over whether the regime performs in the manner expected. We believe it would to be premature to use the 'original' scoring and weighting regime. We suggest the identification of a suitable scoring and weighting regime can be resolved relatively simply and quickly in an intersessional meeting of the PWG that:

1. specifies simple criteria for assessing the performance of a scoring regime e.g.
    - scoring and weightings should be logically consistent and defensible;
    - fisheries that use strong effective mitigation should be ranked low as a priority for action;
    - rankings should correlate well with expert opinion;
2. identifies a single regime that performs well against the performance criteria; and
3. applies the regime to the c.1200 fishery/population records and determines highest priorities for management action.

In addition to the identification of a suitable scoring and weighting regime (although not a direct derivation of this particular study) we also recognise that to complete the entire prioritisation exercise it is remains necessary to:

Determine how the outputs of the prioritisation framework can best be presented by the Agreement.

Up to this point it has been assumed that the prioritised records will be grouped into broad priority categories, such as 'Highest Priority' and 'High Priority' and so on, but such categories, and the basis for such a process, have yet to be finalised. Once the categorisation method has been agreed it will be clearer how to assess the performance of the scoring regimes and define the performance criteria. It will be important not to focus too closely on the specific scores and their ranking in a 'league table', but whether the records fall into appropriate broad priority categories.

Provide recommendation on how the framework is managed, maintained and used by ACAP and its Parties.

The priorities for action identified through this process should influence the working of ACAP and its Parties, including decisions about resource allocation, the contents of the Advisory Committee (AC) work program and reporting processes. The PWG should make recommendations to ACAP and its Parties on how the results of the prioritisation process may be integrated into the work of the Agreement. Finally, the PWG should provide recommendations to the AC and Secretariat for the ongoing management and review of the framework.

Recommendations

The SBWG is asked to:

1. **note** that the prioritisation framework for at-sea threats is near completion but that an appropriate scoring and weighting regime has yet to be determined;
2. **agree** that useful criteria to identify an appropriate regime are:
    - that results correlate well with expert opinion;
    - do not prioritise fisheries that use strong effective mitigation;
    - demonstrate that scoring and weightings are logically consistent and defensible;
3. **note** that the results from the framework will be grouped into broad priority categories but an approach to do so is yet to be agreed;
4. **agree** that the SBWG and its members will contribute to intersessional work, both to complete the framework and to provide advice to the AC on its adoption and appropriate use prior to MoP4;
5. **note** that the secondary objectives of the prioritisation framework (including identifying research and capacity building priorities) can be addressed following finalisation of the framework and are therefore not urgent at this time; and
6. **agree** to support the intersessional work of the PWG which will, before MoP4:
    - identify a suitable scoring and weighting regime for the at-sea framework;
    - agree upon a scheme to present the results of the prioritisation process using a simple categorical system; and
    - provide recommendations to the AC for the use and maintenance of the prioritisation framework.

Figures

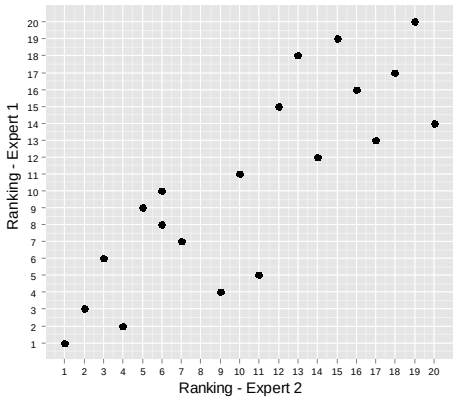

Figure 1. Correlation of records ranked manually by recognised experts in seabird bycatch mitigation.

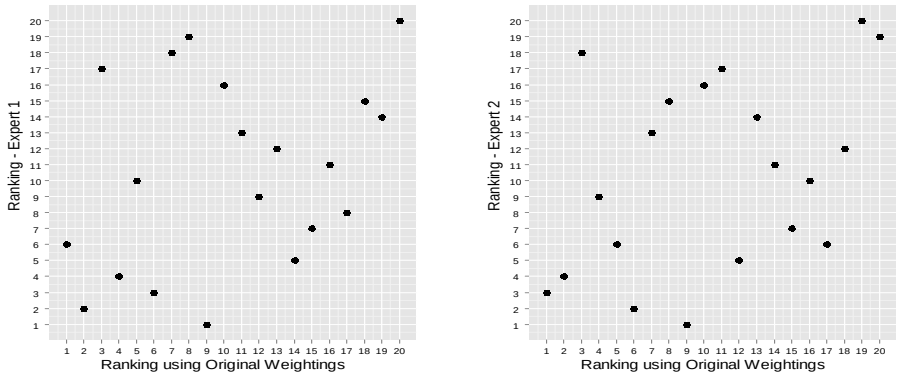

Figure 2. Correlation of records ranked manually by recognised experts in seabird bycatch mitigation with those ranked by the 'original' scoring and weighting regime.

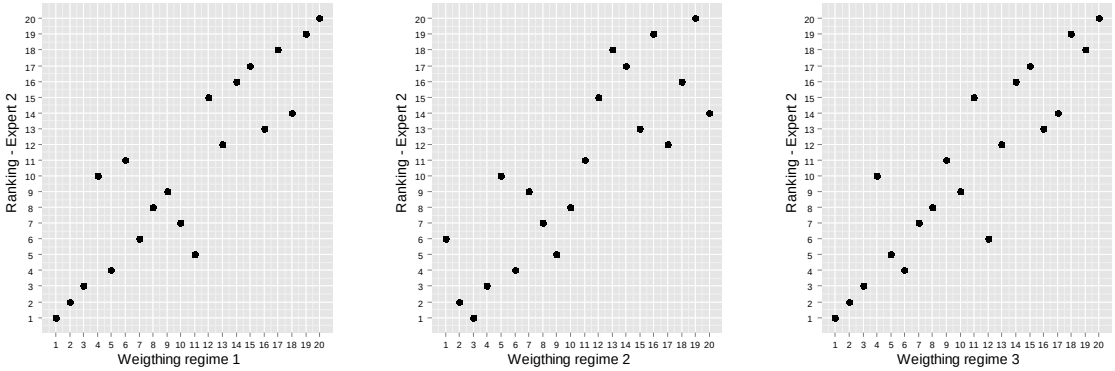

Figure 3. Correlation of records ranked manually by a recognised expert in seabird bycatch mitigation with those ranked by three simulated scoring and weighting regimes (see Table 1).

Table 1. The 'original' scoring and weighting regime plus three simulated scoring weighting regimes that generated rankings that most closely correlated with those of the two experts in seabird bycatch mitigation.

Factors	Levels	Original	Regime 1	Regime 2	Regime 3

Popln size	0-99	5	3	5	4
	100-999	4	1	3	2
	1,000-9,999	3	1	1	1
	10,000-99,999	2	1	1	1
	100,000+	1	1	1	1
Percentage of total global population	0-10%	2	2	5	3
	11-50%	3	3	5	4
	51-100%	4	3	5	4
Popln trend	Steep decline	5	5	5	5
	Decline	4	3	5	1
	Stable	2	1	2	1
	Increase	1	1	1	1
	Steep increase	1	1	1	1
	Unknown	3	4	4	4
Fishery overlap	High	5	3	3	3
	Medium	3	3	1	1
	Low	1	1	1	1
	Unknown	3	3	3	1
Fishery effort	High	5	4	5	3
	Medium	3	1	5	3
	Low	1	1	5	1
	Unknown	3	3	4	2
Risk of fishing method	High	5	4	3	3
	Medium	4	4	2	3
	Low	2	3	2	3
	No	1	3	2	2
	Unknown	3	3	3	4
Use of effective mitigation	High	1	3	1	2
	Medium	3	4	1	3
	Low	5	4	3	3
Likelihood of success if using mitigation	High	5	5	5	5
	Medium	3	5	5	5
	Low or unknown	1	4	3	3
Weightings	**Vulnerability**	**4**	**5**	**5**	**4**
	Threat	**4**	**1**	**2**	**1**
	Likelihood of success	**1**	**5**	**5**	**5**
	Effective mitigation	**5**	**1**	**1**	**1**