

Steller Sea Lion Recovery Plan

Review (July 2007 version)

I.L. Boyd

Sea Mammal Research Unit, University of St Andrews, Scotland

Contents

Basis of the review

Summary and recommendations

Overview

1. *In gross terms, what is happening to the SSL population?*
2. *Is the SSL population still declining?*
3. *Is the SSL population endangered?*
4. *What, in practice, can be done in addition to the conservation measures already in place to promote positive change in the SSL population?*
5. *Are the recovery criteria reasonable and achievable?*
6. *Is the current WDPS population of 45,000 a relatively large number of animals that should allow the RP to be somewhat less conservative?*
7. *Is the range of the SSL contracting from west to east and is this an important problem?*
8. *Should Killer Whale predation be listed as a High threat based on the current evidence, regardless of the inability of the agency to control that threat?*
9. *Will the current management measures ever produce unequivocal results as a test of the competition hypothesis? If so, what sort of results should we look for and how many years must the current management measures remain in place to get those results, one way or the other?*
10. *Has NMFS jumped to conclusions in its determination that nutritional stress is the cause for reduced reproduction and natality and in its assessment of the killer whale threat?*
11. *Has NMFS applied an appropriate weight of evidence approach to the assessment of nutritional stress?*
12. *Are there objective standards in the use of weight-of-evidence evaluations?*
13. *Is the uncertainty over population segmentation such that down-listing is possible sooner rather than later?*

Assessment of revisions made to the Recovery Plan following public consultation in 2007

General issues concerning the approach

Specific issues in the RRP that require comments

Nutritional stress

Killer whales – contrasting treatment to nutritional stress

Adaptive management program to evaluate fishery conservation measures – confused objectives

Population definitions

Declining birth rates
Population viability analysis
Examples of critical questions

Basis of this review

This review was commissioned by the Marine Conservation Alliance. It has been conducted as a peer-review assessment and it is, therefore, a view developed by the author as a practicing scientist in this field. The review should be considered together with the commentary written by the author, again for the Marine Conservation Alliance, about the previous draft version of the Recovery Plan (published in May 2006). Many of the general comments made there, most especially the general issues described in the appendices – (1) A Risk Assessment Framework for the SSL Recovery Plan; (2) Assessment of the effectiveness of past management actions; (3) The “is it food” hypothesis – still stand.

Summary and recommendation

The Revised Recovery Plan (RRP) has presented some important modifications. In some cases these are an improvement (e.g. additional text explaining the construction of the PVA), but some are also important because they could undermine the position of NMFS as a competent body to construct and lead the development of this plan. Issues concerning inconsistency of the analysis of the effects of nutritional stress compared with killer whale predation illustrate a general and worrying lack of objectivity in the underlying approach. In addition, the structure of the RRP does not suggest that NMFS has set clear and transparent policy objectives. If, as I suspect, the objective is to place further restrictions on the fishery then this should be clearly stated and justified against the evidence. In its current form, the plan leaves the strong impression that NMFS is pursuing this policy as an unstated objective, possibly because it does not have evidence properly to back this up but is operating in the “belief” that this is necessary.

Overall, the RRP claims to take a “balance of evidence” approach but in so doing it opens the possibility that NMFS is weighting the assessment to support preconceived notions of underlying mechanisms (especially one that involves the fishery), and this problem has simply been exacerbated by the modifications made to the current version of the RP. Although it may be difficult to accept that over a decade of intensive research has not tended to provide clear evidence supporting particular hypotheses, it would be much more satisfactory for NMFS to admit that the current level of uncertainty is so large that it cannot develop a rationale for distinguishing between several leading hypotheses. In these circumstances, a different approach is required to the management of the SSL. This should be a risk-based approach built around the risk framework developed by the Environmental Protection Agency.

The RRP needs to distinguish policy from the assessment of threats and resulting actions. I suggest that the policy – essentially the objectives for recovery and re-listing - is now articulated reasonably well and with a rationale that can generally be supported by legislation and high-level scientific knowledge (e.g. less than a 1% chance of extinction in 100 years), but this is lost within a mountain of other

irrelevant detail. However, the plan still does a very poor job of justifying the choice of recovery criteria. The RRP also fails to grasp and articulate, in an easily digestible form, the complexity of the knowledge base and to communicate this in a manner that is useful for policy implementation. In fact, it appears that this complexity is being used, on occasions, to build quite unjustified scenarios based upon little evidence (e.g. in the case of nutritional stress). Unfortunately, this problem has been exacerbated by the responses to public comment that have patched up the problems rather than provided a root and branch review. More than even in the previous draft, it constructs some elaborate (perhaps fanciful) arguments around very little data that are associated with a high level of uncertainty. The process by which knowledge is mapped into the policy through the assessment of threats is particularly weak and, as a result, the whole structure of the document needs to be re-worked.

The Recovery Plan illustrates two substantial failures in a best practice approach to developing science-led policy. The first is that NMFS, the policy arm, should not also be making the scientific judgements. Somehow the higher level management within NMFS needs to grapple with this problem because it is creating a fundamental conflict of interest at the staffing level that is responsible for drafting the Recovery Plan. Staff that are making scientific judgements cannot also be making policy! The temptation to manipulate the scientific evidence to support a particular policy view is just too great and I fear that the Recovery Plan is seriously infected with this type of issue. Even if it is not infected in this way, it is now seen to be so, and that is just as bad.

The second failure is that the Recovery Plan has been written by stakeholders all of whom have attempted to manipulate the process in one way or another. This results in the worst of all worlds – a plan that has no purpose other than to address the stand-offs between the different stakeholder groups. Worst of all, NMFS scientists are now seen as one of those stakeholder groups because, whether true or not, they are not seen to be assessing evidence appropriately and they are seen to be manipulating evidence to support a particular policy. This represents a major systemic failure in NMFS and it places NMFS staff in an invidious position.

My recommendation is that NMFS should adopt the Environmental Protection Agency's risk assessment/mitigation approach to managing the SSL. This can be supported by, first, providing a clear statement of its policy with respect to recovery of the SSL which should be independent of the Recovery Plan and should not be produced by the NMFS scientists involved. This should be based upon the legal obligations upon NMFS, but it could also review those obligations in the light of new evidence, especially with respect to the justification for establishing distinct population segments. This policy is the reason for the Recovery Plan in the first place. Second, NMFS should establish an independent panel (mainly of scientists many of whom should not be involved in SSL research) to map the current science into the policy and to produce the plan to achieve the objectives.

Overview

I have responded to the RRP in the context of a set of questions generated by me as a way of attempting to articulate the nature of the current problems with the RRP. There

are a number of important questions concerning the SSL population that the RRP either does not address or, if it does address them, they are lost within the complex text. There are also questions that relate to specific management measures and options, or the way in which evidence has been used to construct the Recovery Plan.

1. In gross terms, what is happening to the SSL population?

The count data for SSL are probably the most reliable indicator of population status. It is unusual for all the SSL population to be plotted together by combining across all the DPSs but, for reasons articulated below, this is worth doing because I suggest there is little justification for considering that the different population segments are distinct. The overall pattern of SSL abundance suggests rapid decline through the 1980s, stability through the 1990s and a slow but consistent increase through since 2000.

Without getting into the detail of the meta-population structure, and assuming no human impacts have occurred, this suggests a density-dependent pattern of change most likely in response to a change in carrying capacity during the 1980s. Some of the condition indicators from the population support this view.

If one then adds in the knowledge that declines in the 1980s were exacerbated by shooting, this probably explains the apparent rapidity of the decline through that period and the likelihood that the decline may have undershot carrying capacity, at least in some regions. Current increasing trends may represent an increase towards a reduced carrying capacity.

Although this overall view is speculative, the lack of evidence of any significant human factors driving the population dynamics in recent years means that we need to place more weight on the potential that changes have been caused by natural processes.

2. Is the SSL population still declining?

The evidence suggests that the population as a whole is not declining. Some local populations may be in decline but this is to be expected in a population like the SSL that is distributed over a very wide range and where conditions will differ geographically.

But, the observation that birth rate may be low is of concern. It is not absolutely certain that birth rate is a problem but, in the absence of other indicators, the focus of concern may be upon the birth and weaning process either because of teratogenic factors like disease or pollution, or because of a possible tendency for mothers to extend the duration of lactation and trade off new births for investment in current offspring. The latter is certainly a possibility and one can find parallels in other mammals – mainly marsupials - that have to cope with uncertainty in environment productivity. If this was the case then it suggests that mothers are experiencing some level of food shortage through lactation.

The suggestion that lactating mothers could be a focus of management action has to be tempered by two factors;

- (1) Our certainty about this being the cause of current population dynamics is fairly low, even if it is the best current model; and
- (2) Assuming there is no pathology involved, the responses of mothers is likely to be adaptive, i.e. mothers are trading off new births for investment in current offspring within the norms of the life-history of the species. This means that it is likely to fall within the boundaries of an “evolutionary stable strategy”¹ and is adaptive in the sense that it helps to maintain the population in the long-term. Consequently, it may not be a feature of concern when placed alongside relatively high juvenile and adult survival rates.

A problem faced by managers is that the raw population data placed within models suggests that the current balance of birth and death rates makes the population unsustainable. But the gross evidence from the population trajectory does not support this view. Either the raw input data are wrong, biased or do not represent the population as a whole – which is possible since they apply mainly to the Central Gulf of Alaska – or the data showing the population trajectory are inaccurate. At present, it is probably not possible to distinguish between these alternatives but the most secure data sets for SSL are likely to be the population counts. Even if these are biased, they are likely to be internally consistent unless biases have changed through time. This means that data about population trajectories are likely to be the most robust, and therefore, believable part of the SSL story.

3. Is the SSL population endangered?

The RRP makes the case that, under the conditions of the ESA, the WDPS of the SSL is endangered. However, the RRP does not make a strong case for the separation of the WDPS from the EDPS – in fact it makes no case for this historical separation. My view is that the weak evidence for this separation undermines the fundamental basis of the ESA classification and the classification should be revisited.

Overall, there are about 110,000 SSL. For at least the past 15 years there has been little evidence that the problems that existed in the 1980s are still present and some recent indicators suggest modest rates of increase. The population of 110,000 animals is similar to that of the Canadian and European populations of Atlantic grey seals (*Halichoerus grypus*), it exceeds the population size of the European harbour seal (*Phoca vitulina*), and is of a similar order of magnitude as distinct coastal harbour seal populations in the northern hemisphere. Although smaller than the population of California sea lions, the current SSL population size is of a similar order of magnitude to that of the California sea lion which is not considered to be endangered.

My view is that it is difficult to justify listing the WDPS as an “endangered” population, at least in the same sense as the Hawaiian monk seal, for example. There is strong evidence that several pinniped populations have recovered from population levels that are two orders of magnitude lower than the SSL population.

4. What, in practice, can be done in addition to the conservation measures already in place to promote positive change in the SSL population?

¹ John Maynard Smith and George R. Price (1973), The logic of animal conflict. *Nature* **246**: 15-18

In its assessment of threats the RRP provides an assessment of the feasibility of mitigation (Table IV-1, p120). Those threats classified as “high” in terms of feasibility for mitigation have received most attention in the past with the introduction of measures to control disturbance, shooting and to reduce the potential impact of fisheries around rookeries.

The “elephant in the room” concerning the RRP, and that is evident from the way in which the evidence has been presented, is that NMFS would like to introduce further controls on the fishery. I would not argue with this as a precautionary measure if it appeared to be proportionate and justified by the evidence. However, as I shall argue in more detail below, it may be neither justified nor proportionate. In summary:

- (1) NMFS has already introduced conservation measures that affect the fishery. Insufficient time has been allowed for these to bed in and it remains possible that some of the positive increase we are seeing at present could be the result of some of these conservation measures;
- (2) Seen in the context of an appropriate risk assessment/mitigation process it would be normal for the effects of current mitigation to be assessed before moving forward to introduce more mitigation – adding one set of measures on top of another would create a mess.
- (3) Indirect effects concerning interactions between SSL and fisheries are likely to be very complex. There is unlikely to be a simplistic connection between the fishery and sea lions (e.g. fish taken by fishermen = fish lost to sea lions) and there is probably just as much chance that well-intentioned manipulation of the fishery will have a negative effect on sea lions.

In summary, since we appear to be witnessing a recovery in the population, I suggest that the most constructive management measure would be to maintain current measures but to review these after each range-wide population survey or if other strong evidence emerges to suggest that they should be revised.

5. Are the recovery criteria reasonable and achievable?

The policy description of “1% probability of extinction in 100 years” is a reasonable threshold to apply for ESA listing. However, there is uncertainty about the meaning of “extinction”. This arises in two areas:

- (i) In current legal terms, since the ESA recognises two distinct population segments, extinction is applied exclusively to each segment. However, I have argued later and in my previous comments that the distinction between population segments is certainly questionable based upon current evidence. If one retains the current segmentation of the population then the PVAs done to date certainly suggest that the WDPS is within the ballpark of the ESA listing criterion. If one considers the population as a single entity (i.e. a single metapopulation) it is highly unlikely that the same ESA listing would apply.
- (ii) The PVA in the RRP assumes extinction to have occurred at a population size of 4,743 sea lions. While one would never suggest that a population of this size is in any way desirable or healthy, there is plenty of evidence showing that several populations of pinnipeds that are now healthy were once reduced to much lower levels than this. Conversely, relatively few have gone extinct.

An important change made to the recovery criteria from the first draft of the RP is that the section requiring “population ecology and vital rates in the U.S. region are consistent with trend...” has been removed. This was a poorly defined recovery criterion for the WDPS which, in essence, left the judgement about whether to down-list in the hands of the biologists when their data are never likely to provide enough certainty for them to be able to make this judgement.

However, the criteria for down-list of the WDPS still seem overly precautionary. For example, if the population remained stable at current numbers for the next 15 years, the PVAs as applied in the RRP would almost certainly show an extremely low probability of extinction and would take the population well above the ESA criteria. However, because the criteria for down-listing require that the population should increase significantly over a 15 year period then the WDPS would not be down-listed in these circumstances. This shows that the recovery criteria being applied are inconsistent with the ESA listing criteria and with the policy apparently being developed by NMFS, such as it is. In my view, the probability of down-listing the WDPS using the criteria provided in the RRP is very low.

6. Is the current WDPS population of 45,000 a relatively large number of animals that should allow the RP to be somewhat less conservative?

Very roughly, if we consider the population sizes of pinnipeds world-wide in 5 broad categories on a logarithmic scale (1 = 100-1,000; 2 = 1000-10,000; 3 = 10,000-100,000; 4 = 100,000-1,000,000; 5 = > 1,000,000) there are probably <10 populations in the truly endangered category 1; there may be <50 in category 2 which could be considered as probably threatened; there are probably <20 in category 3 and <10 in each of categories 4 and 5. Although this assessment is built upon my own experience rather than an objective analysis of the data, I don't think a detailed analysis is likely to come to a very different conclusion. What this suggests is that the WDPS lies well within the normal range of population sizes for pinnipeds on a global scale. Of course some of these may be depleted but, overall, there is not an impression that pinniped populations are depleted on a global scale. Given this, and assurance that the rapid declines that occurred in the 1980s and 1990s have ended, there would be a case for saying that the WDPS is not an endangered species.

7. Is the range of the SSL contracting from west to east and is this an important problem?

Ideally, we would wish to see the geographic extent of the SSL range maintained in the long-term. However, changes in the balance of numbers in different regions within a range over time periods of decades to centuries should be considered to be normal in these types of populations.

The RRP probably places too much emphasis upon the avoidance of local extinctions. While it is reasonable to show some level of precaution towards the SSL as a whole, this does not need to extend to the SSL over its whole range. We do not know why there is an apparent contraction of range from west to east but there will always be an argument between those who “believe” that this is caused by human factors and those who do not choose to believe this. However, this is not an argument for science or

scientists whose only role should be to advise those involved in the debate about the relative strength of evidence to support particular arguments (note, it is this objectivity that has been lost by NMFS scientists).

8. *Should Killer Whale predation be listed as a High threat based on the current evidence, regardless of the inability of the agency to control that threat?*

I cannot understand the logic applied in the RRP for down-grading the potential influence of killer whales. Is this just an inconvenient potential truth for NMFS? NMFS has systematically dismantled the killer whale hypothesis for reasons that are not completely clear and with arguments that verge on advocacy rather than objective assessments of evidence. I can understand this to some extent because those proposing the killer whale hypothesis tend to do the same. But it is NOT the job of NMFS to be the opposition in this case because, as we see in the RRP, it compromises their ability to arbitrate on the issue. As I also suggested in the summary to this document it also compromises NMFS competence to produce this Recovery Plan.

I certainly think that the threat from killer whales has to be considered to be at least as high as those relating to nutritional stress, based on current evidence. There is confusion between the concept of killer whales potentially having a significant impact and the causes of this effect, i.e. megafaunal collapse. This unfortunate confusion between the megafaunal collapse and a potential predator pit for the SSL appears to have led to an extreme polarisation of views. I think there is no doubt that killer whales could have caused the decline, in theory, but there is insufficient evidence either way to decide the relative strength of this hypothesis.

9. *Will the current management measures ever produce unequivocal results as a test of the competition hypothesis? If so, what sort of results should we look for and how many years must the current management measures remain in place to get those results, one way or the other?*

It is likely that current management measures, used in the context of a long-term experiment to manipulate the system to the advantage of the SSL, will never produce unequivocal results but they may eventually indicate a smoking gun. One of the problems with the approach is that it is always possible for those who “believe” there is an effect to argue that the experiments are inadequate if they show no effects. However, this is not a problem that is isolated to SSL-fisheries interactions. It is one that is used continuously by NGOs the world over.

As I have indicated in the past, extending the types of studies begun by Wolf and Mangel would start to define the presence/absence of signals of SSL-fisheries data. In this case, we would test formally for evidence of all the hypotheses simultaneously in all the data we have. This will provide a value for the relative support for each hypothesis from the data. It may, or may not, show a stronger effect of fisheries than other factors. In terms of time scale, I suspect we already have the data to do these analyses but we should be constructing them now and re-running them regularly to examine the emerging pattern. It may then be possible to modify management progressively based upon the emerging results.

We need a mechanism to undertake this work. I suggest it is necessary to engage 3-6 independent scientists/groups. Each needs to be provided with a brief and the same data and each needs to produce their own model structure independently. It is important to do this in order to eliminate model uncertainty through model averaging.

10. Has NMFS jumped to conclusions in its determination that nutritional stress is the cause for reduced reproduction and natality and in its assessment of the killer whale threat?

This is partly dealt with in my response to Q8 above. I suggest that NMFS has not taken a balanced view. NMFS contains some of the main advocates of the “it is not killer whales” hypothesis, if it can be called that. Although there is strength to their arguments, I feel that there has been an unfortunately confusion between the hypothesis of sequential megafaunal collapse and the hypothesis that killer whales could have been implicated as a cause in the decline of SSL. The megafaunal collapse hypothesis invades the territory of the wider ecological impacts of historical whaling, which is an extremely controversial subject. A fairly convoluted sequence of logic is required for this to be upheld and NMFS is probably correct to illustrate the weaknesses associated with using this as the causal mechanism for SSL declines.

However, in my view, NMFS has been guilty of selectively quoting from the literature to support the non-killer whale view. The same is true for those that proposed the killer whale hypothesis and, when read in isolation, both views seem plausible to an outsider. But we know that both can't be correct and, in reality, the truth is probably somewhere between what is being proposed on both sides of the killer whale argument.

NMFS staff should not be allowed, or allow themselves, to get involved in these types of arguments. In the end, NMFS has to sit in judgement on biological arguments and make management decisions. In this case, they are acting as both the judge and the advocate which is morally and procedurally wrong.

11. Has NMFS applied an appropriate weight of evidence approach to the assessment of nutritional stress?

Evidence of NMFS not applying an appropriate weight of evidence approach comes from the fact that much of the research conducted during the past 5-10 years has been specifically aimed at testing the hypothesis that there is evidence of nutritional stress. I think it can be said that, without exception, no study has found support for this hypothesis. There have been those that have found some evidence that the decline has been related to changes in the prey available to SSLs but this is weak and is in anycase one step removed from nutritional stress. Just because no evidence has been found does not mean that nutritional stress is not a potential cause but, relative to other hypotheses, the evidence places less weight upon it than in the past.

NMFS has used three main lines of evidence to support their arguments about nutritional stress:

(i) Changes in the condition of animals between the 1970s and 1980s. Elsewhere in this commentary I say why I do not think it is safe to place much weight upon this as evidence.

(ii) The results of fitting a demographic model to current population structure and the recent population trajectory. I also comment on this Holmes et al. model elsewhere in the commentary but it is worth noting that Holmes et al. themselves only place some form of nutritional stress as one of several possible, and at present equally likely, explanations for this observation. Recent evidence that there may be previously unobserved mortality of pups on rookeries is important in this context.

(iii) The relationship between the fishery activity and the local population trajectory of the SSL as described by Hennen. However, while these results are undoubtedly interesting, this study is flawed in the sense that it was an exploration of data to examine the possibility of a correlation between fishery activity and SSL population dynamics. In other words, it was not a fair test because it neither explored mechanisms for the interaction nor did it test the relative strength of alternative hypotheses (see Wolf and Mangel) or factors that could co-vary with the fishery. Although the Hennen study cannot be dismissed, it must carry relatively little weight in the assessment of evidence.

12. Are there objective standards in the use of weight-of-evidence?

I know of no methods other than to place the evidence in a statistical framework like that used by Wolf and Mangel. This has problems in that the statistical framework needs to be designed carefully so as not to introduce bias and also to capture the subtleties and the complexities of the data. This is why I have suggested that a fair test would involve several groups working independently to produce statistical frameworks with the same information. The funds currently allocated to research would be much better spent doing this than setting new hares running or continuing current relatively unproductive approaches. A further advantage of this approach is that it would help to identify those areas of research that need most attention because one could turn them into a sensitivity analysis highlighting which data are likely to yield the greatest return in terms of improved understanding.

Another type of test of the data, in terms of developing a weight of evidence approach, is roughly along the lines of that used by the Recovery Team, except that in this case the process was compromised for two reasons:

- (i) The members of the recovery team were not independent, i.e. they were advocates for particular views, so the whole process of assessing weight of evidence was flawed (unless an independent judge or jury was also to preside and make the decisions);
- (ii) The analysis produced was subsequently edited by NMFS which, as I have already indicated, does not take an independent view in this case.

13. Is the uncertainty over population segmentation such that down-listing is possible sooner rather than later?

My view is that the criteria probably no longer exist for classifying the WDPS as endangered. So long as the high rates of decline seen through the 1980s and early 1990s are no longer present, the conditions for classifying the population as endangered no longer exist. Nevertheless, it is important to maintain the measures introduced in the early 1990s to protect the SSL so it would be unfortunate if down-listing led to the return of those factors, especially indiscriminate shooting and a

culture of extermination in some quarters. It is essential that we should never allow a return to those days.

Assessment of revisions made to the Recovery Plan following public consultation in 2007

Some important changes have been made to the text and structure of the Revised Recovery Plan (RRP). This has clarified some issues and made the RRC generally more readable and it has improved communication of some important ideas. It has also raised some important problems, especially concerning the inconsistent use of evidence to support the recovery policy developed within the plan.

Some of the following comments are highly critical. I emphasise that this criticism is aimed at the process that allows a plan of the type being proposed to be produced, not the individuals involved.

On a more mundane note, the system used to number paragraphs is irritatingly complex and makes navigation through the document, and also appreciation of one's location in the document, very difficult indeed. This is really simple to fix and would help greatly with the case that is being made by not making reading and digesting the document more complex than it needs to be. This is made more difficult by numbering errors, especially in Section V.

Additional sections to the report in the form of the RRP that are especially important include:

- (i) Nutritional stress, section I.H.5, and Table I-15
- (ii) Introductory explanation to Section III, "Factors potentially influencing the western population".
- (iii) Reassessment of the importance of killer whale predation in Section III.B.1
- (iv) Additional text assessing the possible effects of toxic substances (Section III.B.9) and nutritional stress (Section III.B.11) but also removal of some speculative assessments of nutritional stress.
- (v) Some additional introductory text to the Threats Assessment (Section IV)
- (vi) In Section V.A there is additional explanation of the definition of recovery.
- (vii) There is extensive additional explanatory material concerning the Population Viability Analysis (PVA) in Section V.C.1-3.

Many of these changes strengthen of the document and I am grateful to NMFS for giving attention to my original comments.

In my original commentary, I proposed that the approach taken in the production of the SSL RP is inappropriate and, because of this, the RP contains important logical inconsistencies and encourages polarised views of appropriate actions. The current approach assumes high levels of certainty in the knowledge-base which simply does not exist.

It is entirely possible to deal with high levels of uncertainty using risk-based approaches to management. This type of approach investigates solutions that do not

preclude later action. This has the added advantage that funding is not wasted in a shot-gun approach to solving the problem.

General issues concerning the approach

In my previous review I suggested that “the RP should use a risk assessment framework ... involving regular review and re-assessment of management actions in the light of assessments of their effectiveness... Risk assessment involves a substantially greater level of managerial effort throughout implementation but it has the large advantage that it does not commit the Agency to a single course of action at the outset and it is based upon a *balance of evidence*. Perhaps the most frustrating aspect of the current RP is that the Agency is already using a *de facto* risk assessment approach but the RP fails to recognise this or to develop it into a formal mechanism. Much unnecessary tension could be avoided by doing so.”

“...Implicit within the RP is the idea that a rational basis for management requires detailed knowledge and this can be achieved through high levels of investment, especially in research. My thesis here is that some of the vital information required by the RP may never be forthcoming – some of the scientific problems simply cannot be solved with our current technology and capability and they are very unlikely to be solved in the near future and certainly not within the lifetime of the current PR [sic]. The RP is knowledge-hungry to an extent that makes it unsustainable. Although admission of failure to solve some of the most pressing scientific problems may be a profoundly depressing scenario (especially for those toiling on SSL research), there are pragmatic ways of dealing with this issue.”

Fundamentally, these comments still stand. While I appreciate that changing the structure of the plan, and the implementation process, may not be possible at this stage, I was hoping that NMFS would acknowledge the strengths of the Environmental Protection Agency’s risk assessment/mitigation approach. With rather few changes, this could be implemented in this case and it would guide and focus the research effort being used to support the RRP outcomes. However, this would require a root and branch change in philosophy.

The risk assessment/mitigation approach would allow a variable level of management/mitigation depending upon feedback from indices of progress towards objectives. The research would then be focussed upon developing and maintaining appropriate indices.

Specific issues in the RRP that require comments

1. Nutritional stress

Considerable additional effort has been made to rationalise the concept of nutritional stress. This has been subdivided into acute and chronic forms of nutritional stress. My assessment is that this extended discussion simply deepens the doubts that exist about the nutritional stress hypothesis – even though its intention seems to be the opposite. The evidence leans heavily upon two shot samples of SSL from the 1970’s and the 1980. As I emphasised as strongly as I felt I could in my previous commentary, these samples have to be viewed with a great deal of caution. Anybody who has sampled

pinnipeds in this way understands how easy it is to build in biases, based upon a wide range of factors that lead to inconsistent and sometimes undetectable forms of bias. For example, I understand that the body measurements of SSL were not taken in identical ways between the two periods. On balance, these two samples, and the supposed differences between them, provide little evidence for nutritional stress affecting Steller sea lions and have to be weighted accordingly.

The new section of the RRP on nutritional stress spins a complex story around nutritional stress involving backdated growth through the lifetime of these animals to critical periods in life-histories of these animals. I simply cannot accept that this is justified. We have no life-history data for these individuals and we have no data about the levels of food supply through these periods. An example of the poor logic that has been spun is present in the following quotation from the concluding parts of the arguments put forward for nutritional stress (p42): “Females during the summer breeding season (on rookeries) appear to be able to attain [presumably ‘obtain’ is meant here] adequate energy to nurse their pups. However, pregnant females with and without pups **may** [my emphasis] be experiencing chronic nutritional stress after leaving the rookery as evidenced by decreased pregnancy rates of lactating females (Pitcher et al. 1998), and decreased natality rates overall (Holmes and York 2003, Fay 2004, Holmes et al. in review).”

An analysis of this statement shows how hollow it is. The information from Pitcher et al. (1998) was based upon the 1970s and 1980s samples discussed above and has to be given relatively little weight. The other papers quoted are examples of modelling exercises that fit different scenarios to data and that tend to show that part of the reason for recent population declines has been a low birth rate (See Appendix 1 for comments on the Holmes et al. paper that is in review). At best, they can be viewed as fairly circumstantial evidence supporting low birth rates but they say absolutely nothing about the causes of the low birth rates. They provide no evidence for a nutritional cause of low birth rates.

It would be possible to extract many statements of this type that have been built upon very little evidence. Indeed, the RRP itself is inconsistent in the weight it places on evidence. For example, on p29 there is the following statement “The studies [referring to most of those used above] attempting to estimate past demographic rates were motivated in part by the hope that these could shed light on the various possible causes for the changes in vital rates responsible for the population decline. In this, the retrospective studies have been largely inconclusive.” This means that, having admitted that most retrospective analyses have been of little help, many of these analysis are then used in later parts of the RRP to justify a particular position especially about the effects of nutritional stress and also when assessing the levels of threat. In the same place, the RRP then goes on to say, “One exception is the study of Hennen (2006) which found an association between rate of by-rookery decline and the fishing activity around the respective rookeries...” However, there is no in-depth analysis of the inherent weaknesses in the Hennen study which appeared to have been designed from the start to investigate relations with fisheries, meaning that it began with a biased view and was probably constructed (most likely inadvertently) to show a positive result. Again, the conclusion from Hennen (2006) is greatly weakened by an uncritical and selective type of quotation of evidence. For example, it is not clear why the study referred to by Wolf on p28 is not also an exception.

An enormous amount of time, effort and resource has been expended on the nutritional stress hypothesis. While there are a broad range of additional scenarios that could involve nutritional stress, it may be impractical to investigate these in a way that will allow us to draw anything other than flimsy conclusions about their role in the population dynamics of the SSL.

My suggestion is that the section of the RRP describing nutritional stress says more about current internal agendas in NMFS than about what we actually know about the influence of nutritional stress on Steller sea lions.

2. Killer whales – contrasting treatment to nutritional stress

The RRP has also added a new analysis of the killer whale issues. However, the way in which killer whales have been dealt with is in stark contrast to the analysis of nutritional stress. Again, this inconsistency of approach probably reflects internal tensions and agendas in NMFS rather than genuine differences in the level of certainty that can be applied to the problem of nutritional stress as opposed to killer whales.

I am not very knowledgeable about the details of the killer whale data but, for example, reference on p85 to Maniscalco et al (in press) – which is not referred to in the list of references - states that predation events were lower than expected by Williams et al. (2004). Rather than then concluding that the observations of Maniscalco et al (in press) might be under-estimates, the RRP concludes that Williams et al. (2004) was wrong. I cannot understand why William's estimates have been dismissed so easily and much greater weight given to the data from Maniscalco et al. (in press). Although I have not had the pleasure of reading Maniscalco et al (in press), the estimates of killer whale predation events are very likely to be much more uncertain than the energy consumption estimates provided by Williams, which are based upon an extensive data set from marine mammals and a strong underpinning theoretical understanding which is, to a great extent, backed up by empirical observation. Through the use of fairly robust scaling relationships, the energy consumption estimates are easily scaled to killer whales from other marine mammal species.

On p88, the suggestion from Williams et al (2004) that a population of 170 mammal-eating killer whales could have caused the decline in SSL abundance is incorrect. Williams et al. suggested that fewer than 27 mammal-eating killer whales could have caused the decline.

Most of the discussion of killer whales shows that there are very high levels of uncertainty, that there is a strong case to suggest that killer whales could be responsible for declining Steller sea lion populations and that a precautionary approach would be to retain this as a potentially important factor regulating Steller sea lions. Unfortunately the predation hypothesis has become mixed up with the causes (sequential mega-faunal collapse) which is unhelpful, but it does not reduce the likelihood of the hypothesis.

Probably the worst aspect of this analysis is that NMFS appears to have gone out of its way to counter the killer whale argument put up by Williams et al (2004). Instead of conducting a balanced analysis of the pros and cons and assessing the overall level of uncertainty across all arguments, it has seen its role to put up counter-arguments. Given the current levels of uncertainty it seems perverse to come to any strong conclusions either way in a document such as the RRP.

Overall, the balance of evidence put forward as supporting nutritional stress is probably weaker than the balance of evidence supporting killer whale predation effects and yet the RRP comes to quite different conclusions about them as threats. The RRP chooses to place Environmental Variability and Competition with Fisheries and “potentially high” threats because of the effects they may have upon the nutrition of SSL. In contrast, the RRP has reduced the threat from killer whales to “Medium”. I can see no rationale in the RRP for viewing all of these threats as being anything other than highly uncertain.

3. Adaptive management program to evaluate fishery conservation measures – confused objectives

NMFS appears to be confusing two issues in its description of “adaptive management” (p5 & p160). The first is the need to be as responsive as possible to circumstances and to introduce mitigation of potential impacts on SSL as soon as they have been identified. The second is the need to develop an improved knowledge of how SSL respond to different potential impacts through experimental manipulations. Unfortunately, for NMFS, these are likely to be quite different activities. One is a best-practice, and often precautionary, response to a set of circumstances while the other is a designed approach to obtaining new knowledge.

I suggest that NMFS began adaptive management in the early 1990s and has continued that since with additional restrictions being placed upon the fishery. It is possible that we may only now, some 15 years on, be seeing some of the rewards from the measures introduced in the early 1990s mainly to eliminated depredation of the SSL. The long time lags there are between introducing management measures and measurable effects are a feature that cannot be overcome. They derive from a combination of the relatively slow dynamics of large mammal populations and the inherent imprecision of the measurements of population change. Although many of the management measures introduced to date seem sensible, some such as prevention of shooting are obvious whereas others such as the reduction of fishing activity in the vicinity of rookeries is precautionary because the evidence to support them is comparatively poor. Nevertheless, they are proportionate even though it may never be possible to determine if they are being effective.

The suggestions raised by Bowen et al. (2001) and the NRC (2003), also subsequently by Wolf et al. 2006, were of an entirely different nature. Their point was that if one wanted to clearly distinguish between different hypotheses it may be necessary to conduct an experiment. Notwithstanding the fact that the types of models proposed by Wolf et al. (2006) use natural variability to move us along this track, there would unquestionably be advantage in this approach. However, the practicalities of implementation make these types of experiments almost impossible to conduct. The text on p160 says “Given signs of recovery in the western DPS, it is important to take

this opportunity to implement an adaptive management program to test the underlying hypotheses of the conservation measures”. THIS WOULD BE THE MOST FOOLISH RESPONSE TO A POSITIVE RESULT IN ANY EXPERIMENT. NMFS needs to understand very clearly that it is already in the middle of an experiment of sorts and to change the variables in mid stream would be a folly of gigantic proportions. It would simply muddy the waters for years to come.

My suggestion is that NMFS needs to review the current management measures for their likely relevance to SSL conservation but that, in general, they should be retained and reviewed after each range-wide survey. We know there are long time-lags in the system and fiddling with management on short time-scales will not help to clarify issues. NMFS needs to understand also (not just state this in the RRP) that the cause of the current population dynamics are almost certainly multi-factorial and experiments aimed at dealing with single factors, e.g. fisheries, will likely turn out to be inconclusive.

In the meantime, much more could be done by NMFS to undertake state-space, multi-factor modelling, like that begun by Wolf et al. (2006) to partition the variance in responses between different factors.

4. Population definitions

Additional explanations have been provided to underpin the issues about population structure on page 11. At present NMFS defines two “Distinct population segments”, one in the west and the other in the east. For the current RRP to be useful the validity of the definition of a Western DPS needs to be justified because if one considers the SSL population overall (western DPS plus eastern DPS) a very different demographic and ecological picture emerges.

What the additional text on p11 reveals is a greater level of complexity than previously thought, perhaps not surprisingly. The text reveals that there is movement of breeding females from the western DPS to the eastern DPS. The narrative goes on to suggest that “This has potential long term implications to the viability of these populations and their management”, but does not go on to elaborate what these implications are (I suspect this is because they are too difficult for NMFS to contemplate because it cuts across the current ESA classification). The narrative continues, “It is possible that we are witnessing in real-time a very infrequent event in which female sea lions from one population cross over and breed in another”. The important question raised here is how plausible is it that this is a rare event and how plausible is it that this event does not also occur in the opposite direction?

Of course, it is possible that this event is a one-off occurrence that may never be repeated but the assumption really has to be that, given the low level of previous knowledge, these results are indicative of the normal rate of population introgression. Applying the same logic to some results in a manner that would be less convenient for NMFS would lead one to the conclusion, for example, that the difference in the apparent body condition and pregnancy rate of SSL in the shot samples from the 1970s and 1980s was abnormal and should be discounted in a similar way. However, considerable weight is placed by NMFS on this result when arguments are put

forward about the effects of nutritional stress (see above). There is, therefore, an inconsistency in how evidence is being used to support a particular point of view.

As I proposed in my commentary on the previous draft of the Recovery Plan, the most parsimonious explanation for the population structure of the SSL is that the population should be viewed as occupying a linear habitat (the coastline), that inter-rookery distance is the most important factor regulating introgression and that there is probably comparatively little evidence that supports the clear division between the western DPS and the eastern DPS. In fact, the current description of the genetics and observed movement provided in the RRP does not support the current population definitions and makes little overt attempt to defend them. Based on the current evidence it seems much more likely that the SSL population is contiguous but with clines in genetic diversity through the long coastal range.

The implications for management are that it would be much more sensible to consider the western and eastern DPSs as a single management unit with meta-population structure divided by rookery, or some logical grouping of rookeries that show congruent dynamics. I suspect this would lead to a very different management scenario.

5. Declining birth rates

The RRP has been modified to include the results of a new study by Holmes et al. (in press). A detailed critique of this paper is provided in Appendix 1. Overall, the paper suggests there is a case for saying that the declines in the SSL population in the Central Gulf of Alaska have been driven by low birth rates. This is an interesting result, although it builds on previous population analyses from this group so it is not a complete surprise.

The Holmes et al. (in press) analysis employs a formal “balance of evidence” approach. Accepting that any decline of a population has to be caused by reduced birth rates or increased death rates or some combination of both, the approach examines to what extent there is evidence for any of these from the data for SSL in the Central Gulf of Alaska. On balance, the results push us in the direction of looking towards lower birth rates. However, this still does not preclude other possible explanations. It simply says that the balance of probability suggests that lower birth rates could have been a major driver. Note also that it says nothing about current or future causes; the analysis is necessarily retrospective.

I would suggest that this is one of the stronger lines of evidence provided within the RRP, mainly because it relies on data from direct counting of sea lions. However, its implications are much less clear although it makes a case for a shift in the focus of investigation of causation. Teratogenic toxicity, either from some form of infection (*brucella*, for example) or from chemical toxicity may need additional focussed study. The lactation physiology of females SSL may also require some research because we need to know the extent to which females may sacrifice a reproductive attempt to sustain suckling offspring. There is also a strong possibility that low birth rate is being confused with low perinatal/pup survival rate, especially for rookeries which are swept by waves on occasions. For example, it might be interesting to examine the effects of storms as a co-variate in the Holmes et al model.

One final issue that I suggest needs some further thought with respect to the types of models fitted by Holmes et al. is the extent to which the model structure may itself introduce bias. The probability surfaces explored to derive the kind of results coming from the Holmes et al. study can be complex and it is often not very clear the extent to which quite different positions on the probability surface, or assumptions attached to the calculation of joint probabilities, can affect the results. In addition, Holmes et al use maximum likelihood to explore this surface which is not a very sophisticated method and involves underlying distributional assumptions. This type of model uncertainty can be dealt with up to a point by sensitivity analyses but, ultimately, we probably need to see several independent groups of scientists tackle the same question using different approaches before we can build confidence in the output.

6. Population viability analysis

A substantial section of new text has been added to explain the PVA more clearly (p130-131). The limitations of the PVA are also listed on p132-133.

The additional explanation is helpful and, in addition, a study by Loughlin² and a report from the Marine Mammal Commission³ has suggested that PVAs are now a fairly standard part of the development of the recovery plans for endangered species. The RRP suggests that the PVA has not been used in the development of the plan, although evidence remains that it has had a strong influence.

My view is that the general utility of PVA is now largely undermined. It is necessarily an unsophisticated projection of the past into the future. Although it may give comfort to some to rationalise management based upon a PVA, it is not a replacement for risk-based management using adaptive methods – which is, in effect, what NMFS has been practicing for many years.

Annex 1

Age-structured modeling provides evidence for a 28-year decline in the birth rate of western Steller sea lions

E. E. Holmes, L. W. Fritz, A. E. York and K. Sweeney

The paper explores the underlying drivers of the population dynamics of Steller sea lions in the Central Gulf of Alaska. It tests a range of different scenarios against the available data both through the formal fitting of models and through informal comparison between model results and ad-hoc data sets. Overall, the paper concludes that there has been a progressive decline in the birth rate within the Steller sea lion population.

The analysis conducted in this paper is sophisticated and probably robust. The authors have included sensitivity analyses and a range of checks to increase their confidence in the conclusions. The approach taken was to construct a broad range of different

² Loughlin, T.R. (2007) Review and comparison of recovery criteria in the 2006 draft revised Steller sea lion recovery plan. Report to the North Pacific Fisheries Management Council.

³ Report of the Workshop on Assessing the Population Viability of Endangered Marine Mammals in U.S. Waters, 13-15 September 2005, Savannah, Georgia. Marine Mammal Commission.

population models that included many possible combinations of vital rates (birth and death rates in relation to age) and to examine which of the combinations gave the most convincing fit to the data. Because model complexity often leads to better fits, the authors used standard methods to penalise model complexity to help define the model that gave the best fit to the data.

All population models end up being a balancing of birth and death rates but changes in either can lead to increasing or declining population trajectories. However, it is sometimes difficult to distinguish between the relative effects of birth and death rates and this can only normally be resolved by examining the relative number of individuals in different age classes. For example, if a population has high birth rates and low survival rates then it will have a relatively large number of young animals. Fitting different models of the balance between birth and death rates to actual data is one approach to suggesting where the causes of changes in the trajectory of the population as a whole. In this case, the fitting process has generally pointed towards birth rates as the most likely source of the recent Steller sea lion population dynamics, at least in the Central Gulf of Alaska.

It should be appreciated that the conclusion of this study suggests that the weight of current evidence is that the birth rates are constraining Steller sea lion population growth in the CGA. This differs from proof of such an effect but it helps to focus attention on specific causes and perhaps suggests research that could be carried out and some potential management actions. It also does not distinguish between different causes of declining birth rate.

Furthermore, it is also important to appreciate that these models contain assumptions. They are imperfect views of the world but they are probably better at assimilating and synthesising across a complex set of data than less formal, ad-hoc approaches to data interpretation. Nevertheless, their complexity is such that their underlying characteristics are often difficult to understand and sometimes it is possible to inadvertently introduce bias into model structure that can lead it towards one conclusion or another.

For example, one important assumption of this approach is that neither sex ratio nor the proportion of non-pups hauling out has changed systematically across the period represented by this study. This assumption is recognised in the text of the methods. However, the text also makes the slightly odd statement that if this assumption is violated then the apparent stabilisation in the population since 2000 is “illusory”. This statement is correct only under some circumstances.

Holmes and York also justify these assumptions (p15) but this justification does not account for systematic age-specific changes in behaviour, so the validity of these assumptions still need to be considered when interpreting the model outputs. It is difficult without running the whole model to fully appreciate the sensitivity there is to these assumptions but it does seem that relatively small changes in haulout behaviour between different age classes could lead to the observed effects. Although, on balance, such effects are not very likely, that they remain as a possibility is important.

These models are also often very sensitive to the starting conditions for population projection and, while the approach adapted here did make every effort to deal with this, ultimately it is limited by the data that are available.

Even taking into account all these issues, it seems to me that this paper builds a strong case for a declining birth rate in the Central Gulf of Alaska. Perhaps of more interest is the discussion of the potential causes which covers a wide range of different effects and, unlike the RRP, does not focus upon nutritional stress.