<u>Review of:</u> **Status Review of Hawaiian Insular False Killer Whales (*Pseudorca crassidens*) under the Endangered Species Act**

Introduction:

In 2009 a petition was made by the Natural Resources Defense Council to have the 'insular' Hawaiian false killer whale population designated as a DPS and protected under the Endangered Species Act (ESA). In response the NMFS established a BRT to evaluate the merits of this proposal and provide a recommendation, and here I provide peer review of the resulting report. The information provided by the petitioner is summarized in section 1.4 of the BRT report. Here the insular (near shore) population of false killer whales (FKW) is proposed to be unique, and it is noted that a similar distribution off the coast of Costa Rica was described by Acevedo-Gutierrez et al. (1997) as 'pelagic animals that are only foraging near the island'. However, it is not made clear how this distinction was made, or why this could not be the case for the Hawaiian insular population. The petitioners don't cite a more recent study which also reported both pelagic and coastal sightings for this species in Costa Rica (May-Collado et al. 2005).

With respect to the distinction between the Hawaiian insular and pelagic populations, the petitioners note that only mitochondrial DNA data are available, and state that these are sufficient to demonstrate genetic distinction and demographic independence. This is justified by the petitioners in part by speculation about the mode of transmission of behaviors such as alloparenting and foraging, but without any support from published data. No mention is made of whether or not the Hawaiian pelagic and insular populations have been directly compared using genetic data. The ranges of pelagic and insular populations are said to overlap, but it isn't made clear how these sightings are classified as being of one type or the other. Classification by location alone would be circular. If it is by exclusion from the photo-ID catalogue, then it may be a distinction between identified and unknown whales – possibly consistent to some extent with regular compared to infrequent use of the surveyed geographic range, depending on the completeness of the photo-ID catalogue. There is no discussion of the 'catchability' of this species using the photo-ID technique. However, this information is provided in Baird et al. (2005), where 75% of encountered individuals were classified as reliably identifiable using photo-ID.

On the question of abundance and population trends, the petitioners propose 123 as a best estimate census number, but don't address the possibility of mixing with the pelagic animals further, or provide an estimate for the size of the pelagic population. Evidence for a population decline is based on comparison of the 2005 estimate (123 whales) with that provided by five aerial surveys between 1993 and 2003 (from Mobley, reported in Baird 2009), an estimate derived from Mobley et al. (2000) – 121 whales, and survey estimates from 1989 (Reeves et al. 2009). The data provided in Table 3 of Baird 2009 show that there were 18 FKW sightings during 361 hours of survey time over the five year period, and while most of these (17) were during the 1993 and 1995 surveys, it isn't clear to what extent this could be explained as a result of stochastic sampling. At the same time, various other species were seen more consistently, which would support the interpretation of a drop in sightings for FKW between 1995 and 1998. The 1989 sightings were inconsistent with all other sightings purported to be from the insular population with respect to the size of the groups seen – mean group size of 195 during the 1989 surveys and 15 during the subsequent boat surveys (Baird 2009), and 5.1 during the aerial surveys conducted by Mobley et al. (2000). No information is provided on the mean group size observed for surveys of the pelagic population, but the mean size reported in Barlow & Rankin (2007) is 12. Therefore the group sizes reported from the 1989 surveys are anomalous compared to all other published reports.

With respect to risk the petitioners cite mortality and serious injury via fishing gear, overfishing and prey reductions, potential for increased levels of toxic chemicals, ocean acidification, the potential for acoustic impacts as specific habitat-related threats, inadequacies in State and Federal law, and "other factors including risks inherent to small populations and synergistic and cumulative effects of other threats as specific threats to the population's survival." The details provided are primarily indirect (e.g. 4% rate of fin disfigurements proposed to be from long-line fisheries interactions) or anecdotal (e.g. interactions with short-line fisheries). Information on impact from the long-line fishery is unclear in this summary, but clearer in publications by Forney and co-authors (see BRT citation list). Evidence with respect to contaminant levels is stronger, but similar to that observed in other cetacean species (see citations provided). Some aspects of the concerns about Federal law were unclear. For example, why is a take-reduction team from NMFS not being proposed now, if that would have been important in 2008?

In general, there are aspects of petitioner's report summary that suggest the need for stronger support and therefore require careful attention from the BRT team. These include the following main questions:

1) Is a coastal distribution unique or unusual for this species?

2) Is the putative 'insular' Hawaiian stock reproductively isolated from a putative Hawaiian 'pelagic' stock, especially with respect to possible male mediated gene flow? The importance of the latter should be made clear by the BRT team (e.g. few genetic migrants are required to negate the effect of local loss of diversity through drift, and consequently the potential impact of inbreeding depression – male mediated gene flow will not be revealed by investigation of mtDNA). In this context a population size estimate for the pelagically distributed animals is important.

3) How are insular and pelagic animals distinguished during surveys, and may errors associated with this significantly impact on existing abundance and population trend estimates?

4) Of the various risks proposed, which can be quantitatively and reliably assessed?

Review of BRT assessment:

The first part of the BRT report provides a review of the biology and habitat of FKWs. This is in general a thorough and useful review, but there are some points that would benefit from further details. On the topic of abundance and distribution, Barlow & Rankin (2007) have provided a preliminary estimate of the abundance of the Hawaiian pelagic stock (484; CV = 0.93), and given the possibility of connectivity between the insular and pelagic populations, this number is important to further considerations about potential impact, and so should be cited and discussed here. With respect to the uniqueness of the Hawaiian pattern of distribution, the BRT authors should cite and consider further data about the putative coastal population off Costa Rica (e.g. May-Collado et al. 2005), and furthermore consider other potential locations where surveys have not yet been conducted. Little is know about this species beyond a few specific locations, and in this context the lack of knowledge is quite important to the assessment of significance. I also have some specific questions and comments, as follows:

1) page 14: The authors say that the mating system ‘may be polygynous’, but what is that based on? It matters because the effective population size is reduced by reproductive skew. On page 17 the authors note that little is known about breeding behavior in this species.

2) page 16: The authors suggest that the ‘large, dispersed groups in which false killer whales typically occur suggest that this species forages cooperatively’. It isn’t made clear what this inference is based on, and furthermore contradicts the earlier text stating that group size is typically small for this species (e.g. based on quite extensive sightings described in Baird et al. 2005). This matters because cooperative foraging is later implicated in the case for DPS designation.

3) page 16: Data on prey choice are few and mostly indirect, and some of the inference base on estimated dive depth seems poorly supported. The BRT authors should be clearer about these limitations.

4) page 17: The authors note that ‘the only markers examined for FKWs are neutral genetic markers’. However, we don’t really know that mtDNA markers are effectively neutral, even though that’s a frequent assumption. This is a closed circular marker without recombination (or too little to be significant), and so all genes in the mitochondrial genome are closely linked. Selection on any of the 13 functional genes could affect the interpretation of control region comparisons, and in fact some authors suggest that there is extensive evidence for selective sweeps in mtDNA data (e.g. Bazin et al. 2006), though this is not without controversy.

5) page 17-18: The analytical methods used to compare FKW populations (from Chivers et al. 2007; 2010) are based on equilibrium assumptions, and so cannot provide inference about migration rates at different time scales – migration rates are assumed to be at equilibrium and therefore constant.

6) page 18: The authors discuss signals from mtDNA data indicating ‘separation on long time-scales’. Potentially relevant for the case of the insular population is the fact that a local founder event can sample lineages stochastically. If the source population haplotypes sampled during a founder event are rare, then they may appear common and unique in the new founder population, and this can happen very quickly.

The next section reviews a number of case studies representing other cetacean species. Some of these data are quite useful, but I think two opportunities were missed. First, as indicated above, the assessment of nuclear markers (such as microsatellite DNA) to provide some assessment of male-mediated gene flow is essential to the understanding of connectivity among putative populations. Few effective migrants are required to negate the impact of drift and inbreeding, and it doesn't matter if the migrants are male or female. However, all of the examples focus on mtDNA, even for those studies (e.g. Andrews et al. 2010) where the microsatellite DNA data provided some important insight, not indicated from the mtDNA data alone. In the case of the killer whale, this is a good opportunity to illustrate the risks associated with considering mtDNA on its own. There have been a number of killer whale studies using microsatellite DNA, and they include some recent analyses that strongly indicate ongoing gene flow between populations separated by the deepest root in the mtDNA genome tree (e.g. see Pilot et al. 2010). There is sometimes clear disagreement between nuclear and mtDNA derived phylogeography for this (e.g. see figures 1&2 in Pilot et al. 2010) and other species, especially when there are matrifocal social affiliations. Illustration of this in the review portion of the report would support the author's returning to this important point later in their assessment. The second missed opportunity was any reference to the various studies indicating differentiation between coastal and pelagic populations of dolphins, though it's not clear if the same pattern would necessarily be seen for a given species in association with oceanic islands (see Querouil et al. 2007 and citations therein).

The next part of the report provides a very extensive review of the marine habitat. While some of this is useful, there is certainly more information provided than is used in support of arguments later in the text. Some inference relevant to FKWs is provided throughout this section, but the relative strengths and limitations of this inference should be clearer throughout. For example, there are observational records of prey taken, and a table of prey species that may be important is produced (Table 2-3), however it needs to be made clearer how these records were acquired, what the biases may be, and what should be done in future to improve understanding about this (e.g. stable isotope studies).

Section 2.3 reviews questions associated with abundance, population trends and genetic structure. This more detailed review makes clear that the distribution of the putative insular stock extends significantly offshore (up to 112km), and again begs the question of how the insular and pelagic populations can be distinguished in census studies. This is never clearly addressed. A single satellite tracked pelagic individual indicated fairly extensive overlap, but broader inference isn't possible from just one track. With respect to abundance, an historical estimate is provided in section 2.3.2.1, but the authors should be clearer about the limitations of this estimate (due especially to problems associated with the necessary assumptions). Although, as indicated above, the estimated size of the pelagic stock is important to a full assessment, section 2.3.2.2 only mentions that an estimate exists, not what it is. The number (484) should be provided together with some discussion of the relative strength of the data supporting that estimate. In general, section 2.3.2 could be tightened up – statements such as 'insular false killer whales were recognized as separate' and various other unconditional statements without appropriate caveats make the report seem less rigorous than it does in some other sections (e.g. the following section, 2.3.3).

The interpretation of data from the 1989 surveys needs more careful consideration. Later in the report the authors state that errors with respect to species identification seem unlikely due to the experience of the observers. I very much agree with that, however the possibility that these sighting were instead of the pelagic stock in a temporary distribution seems quite likely. For one thing the size of these groups is substantially different that that reported from either the boat or later aerial surveys (and the latter started just a few years later). Preliminary estimates of group size for the pelagic population are also small (mean = 12; Barlow & Rankin 2007), but the sightings could represent atypical temporary aggregations. For another, this was an unusual year in terms of the oceanography. The surveys followed an extended period of cold surface waters following the warm surface waters that prevailed during the 1987/1988 El Nino event. Among all of the other survey periods, only 2000 is somewhat similar (see http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml). This in particular may or may not be relevant, but in general I think a broader assessment and more critical interpretation would be appropriate.

The next section is on the genetic assessment. This is, for the most part, quite clear and carefully written. This is the first section where it is reported that data were analyzed for microsatellite DNA loci and a small sample of putative Hawaiian pelagic animals (N=9). The review is based mostly on data from Chivers et al. (2010), but omits some potentially useful data from that study, such as the data from the program Bottleneck, suggesting a recent population contraction (possibly associated with a founder event), though not supported by neutrality tests (suggesting that it may have been a very recent event). The main aspect of this section that requires more careful consideration (though already mentioned briefly) is about the problems associated with comparisons using Fst or related measures when sample sizes are small. This metric requires accurate estimates of gene frequencies, which is not possible when sample sizes are small. Instead the BRT authors should suggest (or undertake?) the comparison of these putative populations using coalescent methods, which are less dependent on sample size (because a few extant genotypes can still identify a diversity of historical coalescent points). A good choice may be the program IMa, which can also provide estimates of splitting time and directional gene flow (see Hey & Nielsen 2004). With respect to the data presented, some of the further details from Chivers et al (2010) are important and should be included. For example, the p-value for insular vs pelagic is 0.034 without correction for type 1 error (and so only marginally significant), and the magnitude of this Fst is similar to the values comparing the Hawaiian pelagic and Mexican samples (0.017 vs 0.019). Further, the difference between samples from the islands Hawaii and Oahu is again about the same (Fst = 0.027). Much of this could be noise due to small sample sizes, but in general shows much lower differentiation than suggested by the mtDNA data, and could easily be consistent with ongoing male-mediated gene flow (estimates provided in IMa would help assess this). There is also one locus out of Hardy Weinberg equilibrium, and it would be useful to test the magnitude of the Fst values when this locus is omitted. There are some specific comments that could be clearer, for example:

1) If I understand the point, the fact that microsatellite-based Fst values don't easily scale to 1 is what is being referred to in bullet point 2 on page 46, but adjusting Fst values to account for this doesn't really change the inference, just the number.

2) In the next bullet point on the same page it isn't clear why selection is relevant.

3) For the estimate of effective population size (next paragraph) it is quite important to mention the method and discuss the implications of the necessary assumptions. This was done using a program called LDNe which assumes neutrality, a closed population (unlikely), and discreet generations (cetaceans have overlapping generations). The implications of violating these assumptions have not been fully worked out for this test, but some relevant discussion is provided in Waples (2006) and could be cited. Note that the definition of effective population size given in the executive summary could be misleading, as it leaves out the important role of historical demographics. I imagine it was defined as it was to be accessible, but probably better to give the full definition (e.g. as in Chivers et al. 2010).

The next section I want to comment on is about the potential impact of the long-line fishery (section 2.4.1.1). This is critical because it seems to be the fisheries interaction with the clearest, published, quantifiable data on impact. Data in Forney (2009) shows that mortalities associated with this fishery are now relatively rare for this species (2 events in 5 years) within the Hawaiian EEZ, consistent with the institution of the exclusion zone in 1991. An estimate of 194 takes on page 52 must refer mostly to the pelagic population, but a reference and further detail on how this was derived would be useful. The 42 takes referred to in the caption to Figure 2-14 is somewhat confusing – seems to contradict numbers provided in the text. In the summary (section 2.4.1.6) a 'high level of risk' is proposed, but isn't very clear how this is justified now for the insular population in particular, given the exclusion zone for the long-line fishery, and for interactions with other established fisheries, given the lack of quantitative evidence. There are few data on diet, and biases in those datasets are not well known, but the summary assertion (in section 2.4.2.8) of a medium risk seems OK in that context.

Section 2.4.5 considers the possible risks associated with small population size. At the start there needs to be a caveat about the very real possibility of ongoing male-mediated gene flow between the insular and relatively abundant pelagic population. If that is the case, then some of the proposed risks in this section could not be supported. At the same time, the text should say more about the risks associated with population dynamics, especially demographic stochasticity and the associated extinction risks. Here and extensively in the executive summary the authors discuss the Allee effect, but without clearly defining it. The Allee effect is the observation that small populations sometimes show low reproductive and survival rates. There have been a variety of proposed mechanisms, most commonly associated with difficulties in finding a mate in sparse populations (clearly not a problem in this case), but also in relation to other factors, including genetic factors such as inbreeding depression. However, the reference to the 'Allee effect of inbreeding depression' as given in this section, could be misleading or confusing without the initial definition (perhaps best provided in the executive summary). A calculation is provided where the estimate of 123 whales (ignoring the further individuals identified near Kuai) is reduced to 43% based on data from killer whales (not established why this should be particularly relevant here), and the consequent number of animals interpreted in the context of inbreeding depression. There are too many weakly supported steps here, and again, it would only be relevant in the case of complete isolation, which seems unlikely.

Review of BRT findings and conclusions:

The BRT team have provided extensive analyses and go to some lengths to document the discussions involved in reaching their decisions (summarized in the appendices), and should be commended on both points. They used the method of generating 'likelihood points' to consider relative support for various proposals. I'll confess I'm a bit uneasy with this method, as the title implies some association with the mathematical method of the same name, while in fact it's more of a collective, informed, best guess, potentially subject to personal experience and biases. However, the discussion provided does make this fairly clear. At the same time, it is never fully established that the population under investigation is unique, or fully isolated, in decline or under severe risk. Opinions differ, but it remains the case that more data are needed to reach strong conclusions on these points, and it's critical that the text reflects the remaining uncertainties throughout. I am in general agreement on most points where the balance of evidence is discussed, and a conditional conclusion is reached. However, the tone of text should consistently reflect the fact that significant uncertainties remain. In terms of the assignment of threat levels (Table 4-1), I think most are reasonably well justified with the exception of a level of 'severe' applied to interactions with troll, handline, shortline and kakaline fisheries. I don't think the authors have made this case based on the data reviewed. For overall threat with respect to small population size, I don't think that 'high' is realistic given the high (but so far untested) likelihood of continuing gene flow between the insular and pelagic populations. With respect to the PVA, I think the use of the 1989 data probably isn't justified (giving a minimum prior of 470, and therefore assuring a negative population trend for most simulations), as there are too many doubts about the source population for the animals counted. If it is used, a parallel assessment should be done without it. This could be based on data from Mobley et al. (2000) and Baird (2009), and would likely lead to quite different conclusions about extinction probabilities. As also seen in the executive summary, the final sections become less cautious about statements of fact and general conclusions, probably in the interest of brevity and accessibility. However, it would be appropriate to edit these to be more consistent with the tone expressed in the more detailed sections where the relative support for interpretations are more carefully considered.

General Conclusions:

I could support designation of this population as a DPS, based primarily on the telemetry and photo-ID data suggesting some level of dependence on coastal habitat in a region of high human activity, on indications that there is significant genetic differentiation (though probably not isolation), on the possibility that this is relatively unusual behavior for this species (though this is perhaps the weakest point – due to the paucity of relevant information), and on consistent census numbers that suggest relatively few animals in this stock. At the same time, more data are needed to be confident about all of these interpretations. As a precautionary measure it seems appropriate to consider the combined strength of inference, which would put me on about a 6 I think, using the authors scale, in support of the DPS designation. However, I'm not yet convinced that a future risk level of 'high' can be justified by the data presented.

References:

Bazin E, Glemin S, & Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. Science 312:570-572.

Hey J. & Nielsen R. 2004 Multilocus methods for estimating population sizes, migration rates and divergence time, with application to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167, 747-760.

May-Collado L, Gerrodette T, Calambokidis J, Rasmussen K & Sereg I. (2005) Patterns of cetacean sighting distribution in the Pacific Exclusive Economic Zone of Costa Rica based on data collected from 1979-2001. Rev. Biol. Trop. 53: 249-263

Pilot M, Dahlheim ME & Hoelzel AR (2010) Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). J. Evol. Biol. 23:20-31

Querouil S, Silva MA, Freitas L, et al. (2007) High gene flow in oceanic bottlenose dolphins (Tursiops truncatus) of the North Atlantic. Cons. Genet. 8:1405-1419.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Cons. Genet. 7:167-184