Responses to CIE reviewer comments on IR-08-010, Assessment of the population-level impacts of potential increases in marine turtle interactions from a Hawaii Longline Association proposal to expand the Hawaii-based shallow-set fishery

Melissa L. Snover

27 February 2009

The Pacific Islands Fisheries Science Center (PIFSC) submitted the internal report, IR-08-010, “Assessment of the population-level impacts of potential increases in marine turtle interactions from a Hawaii Longline Association proposal to expand the Hawaii-based shallow-set fishery” to the Center of Independent Experts (CIE) for a peer review of this assessment as is common practice amongst NMFS Science Centers. This assessment was conducted to assist the Pacific Islands Regional Office (PIRO) with the re-initiated section 7 consultation on the Hawaii-based shallow-set fishery and the subsequent ESA determinations required for sea turtle interactions anticipated with the expansion of the fishery.

One recommendation of all three reviewers was to include parameter uncertainty in the model, including sex ratio, reproductive values, proportions of populations represented, etc. While this is certainly a valid point and would make the approach more robust and sophisticated, it would not change the conservative end of the spectrum of the results. From Table 5 of the report, at the conservative end of 1% of (1-SQE), the results indicated that less than one adult female should be taken from each of the three populations considered. While these values could certainly go lower in terms of decimal point value, we are still talking about less than one animal. So, inclusion of uncertainty in terms of placing ranges around the model parameters would increase the upper end of what we consider ‘reasonable’ interaction levels, it can’t go lower than less than one adult female on the lower end. It would, however, expand the ranges given in brackets in Table 6 of the report.

I am leaving my position as the Leader of the Marine Turtle Assessment Program at PIFSC for a position with the National Park Service in Alaska, effective 27 February 2009 and will not be able to modify the assessment based on the reviewers recommendations. A working group has been initiated composed of Tomo Eguchi (SWFSC and lead on the group), Selina Heppell (Oregon State University), Jeffery Moore (Duke University), Paul Richards (SEFSC), Heather Haas (NEFSC), and myself (I will continue to participate on my own time). The goal of the group is to develop a method to provide advice to our Regional Offices on appropriate interaction levels of sea turtles in US Fisheries. To date the different regions and science centers have been taking widely different approaches to these questions. The group met 5-7 February 2009 in La Jolla, CA for the first time. The method presented here is one that’s being considered and if the group decides to move forward with it, modifications will be made as per these reviews.

While all three reviewers had a number of critical comments, two reviewers were relatively positive regarding the approach to the assessment. The 3rd review was negative of the approach of Snover and Heppell (in press) to assessing risk in population based on an index, the susceptibility to quasi-extinction (SQE). While we appreciate the reviewers’ concerns, we found this review to be somewhat unhelpful and not consistent with the independent reviews conducted by the journal for the original paper. Here I address the major concerns and recommendations of each of the reviews we received.

W. Don Bowen

Below are Dr. Bowen's primary concerns and recommendations for improvement in bold text, with my comments and responses beneath each.

- **This approach only gives near-term assessments of the status of populations as it is not predictive, therefore the assessment should be updates as new data are received for nesting beach trends and interactions with the fishery**

  I agree with this recommendation and as Dr. Bowen mentions, this is recommended in my assessment

- **Uncertainty in some of the parameters, Dr. Bowen specifically mentions population identity and distribution, and sex ratio, was not included in the assessment and might these uncertainties change the overall conclusions in terms of risk.**

  I agree that there were numerous layers of uncertainty in this assessment and many decisions had to be made regarding the values to be used. For many of the values, there are simply no good data to achieve a reasonable confidence interval and often discussions were held with PIRO staff and decisions made as to what value to use, generally leaning toward being conservative on behalf of turtles, but not unrealistically so. For sex ratios, for example, while many nesting beach studies find sex ratios of 90% female or greater, these studies are usually based on unvalidated nest chamber temperature data. Studies of larger juveniles (i.e through laparospic examination of gonads or from necropsies) generally suggest a female bias, but usually not greater than 60-65%. Hence, if we had gone with the extreme of 90% female or greater, the results may have been different as a much greater proportion of the bycatch would have been considered female, however, this value is likely very unrealistic, at least for North Pacific loggerheads and leatherbacks.

  I do believe, though, that this assessment would benefit from a sensitivity analysis to determine which parameters the results are most sensitive to and to better account for the uncertainty in those values. If this approach is repeated I would strongly recommend that analysis and this is one of the recommendations of Dr. Bowen.

- **Dr. Bowen recommends that the basis for predicted turtle encounter rates in the fishery be included in this assessment.**

  I agree that from an outside perspective, the anticipated take levels used in this report seemed arbitrary. This part of the analysis was not something I was asked to do, but rather consider the issue from the standpoint of 'how many turtles would be too many' which is not an easy question to answer. A more thorough treatment of the anticipated take levels for the fishery can be found in the EIS for the expansion prepared by the Western Pacific Regional Fisheries Management Council. The data for the bycatch in the shallow-set fishery are excellent as there is currently 100% observer coverage and they, in conjunction with other data have been used to develop a tool, Turtle Watch, for the fishery to use to limit their interactions.

- **Did the method of effort-correcting the loggerhead nesting beach time series introduce a bias in the trend?**

  This is a valid concern although it is difficult to address with the data we have available. One suggestion from Dr. Bowen was to only use data from the 33 nesting beaches for which there were data over the entire time series. This isn't possible because we only have total counts for all of Japan from 2000 – 2007, hence the earlier data needed to be adjusted up, to estimate total nests, using the only 2 year of overlapping data we had for the two datasets, 1998-1999. Getting a breakdown of nests/beach from Japanese researchers from 2000 – current would be ideal and we'd be better able to test if the effort-correction method used biased the trend.

- **Include a table of hooking types, or injury categories for turtles interacting with the fishery.**

  An earlier draft of this assessment had those tables included for both loggerheads and leatherbacks, I do not recall my reasoning for excluding them in the final report, though possibly it was because they appeared in a memo as referenced in the report.

Graham Pilling

Dr. Pilling had very useful comments in terms of ways to improve and test the approach of the assessment. There is currently a working group composed of agency and academic scientists developing methods to address the questions of reasonable takes of turtles in ESA Section 7 consultations that can be used across the different regions. The approach taken here is one that is being considered and the suggestions of all the reviewers, but especially Dr. Pilling's will be considered in potentially moving forward with this methodology.

Below are Dr. Pilling's recommendations for improvement copied directly from his report in bold text, with my comments and responses beneath each. Each of the recommendations is expanded upon in Dr. Pilling's report, clearly indicated with the recommendation number in bolded text.

Recommendation 1: Examine the impact of the data time series available or estimated for each population on results, by running retrospective analyses.

In conjunction with the sensitivity analysis suggested above this would be a good approach to assessing the stability of the model outputs

Recommendation 2: Given the uncertainty in parameter values within the model, a Bayesian approach, which can encapsulate uncertainty within model inputs and hence outputs, should strongly be considered.

Yes, a Bayesian approach would be a good way to deal with parameter uncertainty, especially given our lack of data on parameters ranges and the need to use 'judgment calls'. Placing the assessment into a Bayesian framework was discussed at the working group meeting mentioned in the introduction.

Recommendation 3: If a frequentist approach is maintained, sensitivity analyses should be performed to assess the impacts of assumptions made during the parameterisation of the model on future population status, and hence management advice.

Agreed, this was discussed above.

Recommendation 4: Show all consistent historical data available from the populations, even if they are not used within the model runs.

Good point. These will be provided as an appendix.

Recommendation 5: Given uncertainty in the historical time series for loggerheads nesting in Japan, and the effect on overall population growth rate estimates, sensitivity analyses for the data assumed in the historical period should be performed. See also recommendation 1. In addition, consider approaches to weight time series periods relative to the reliability of data.

Dr. Bowen also voiced concerns over the need to effort-correct the Japanese nesting data and the retrospective analysis Dr. Pilling recommends is one way to deal with this.

Recommendation 6: Given the results presented in Snover and Heppell (in press), sensitivity analyses should be performed based on alternative running sum lengths to examine the impact of the settings used on model outputs.

I'm not sure how much of an impact the choice of running sum would have. Two year sum would be unreasonable as it is not biologically realistic and statistically increased the variance in the time series, so it would be 3 or 4 year running sums which may not show a large difference.

Recommendation 7: Re-run the analyses using the peer-reviewed approach of Snover and Heppell (in press), with results based upon 90% of bootstrap distributions.

The analysis was originally run with the 90% of bootstrap distributions as in Snover and Heppell (in press). However, I ran into great difficulty trying to convey this concept to managers and I found if I used the mean, they could call it the mean risk of quasi-extinction and be accurate. Using the mean necessitated revising the critical value for SQE based on the population simulations of Snover and Heppell (in press) and so it was similarly validated and highlighted that the specific definition of SQE could be flexible, so long as it is accompanied with appropriate validation (a criticism we received on the manuscript). The final results of the model did not differ substantially regardless of the definition of SQE.

Recommendation 8: There is a need to identify and agree the parameter values or parameter ranges, and associated uncertainty, in the light of agreed levels of precaution. This will require a meeting of scientists, managers and stakeholders to resolve.

Yes, the values used were primarily selected by PIRO given their desired level of precaution and information that could be found in both peer review and gray literature. Given the time constraints involved in preparing the assessment this was the best approach possible but certainly gaining insights from a broader range of stakeholders would be valuable.

Recommendation 9: Examine the sensitivity of model outputs, in particular with respect to interactions within the fishery, to uncertainty in the sex ratio value for loggerheads and leatherbacks used within the model.

Dr. Bowen had a similar concern, which I addressed above. In general, sea turtle scientists do not believe that sex ratios are as heavily skewed towards females as nesting beach temperature data suggest. However, given the possibility of increasing temperatures from climate change, it likely would be valid to test the model under more skewed sex ratios than the 65% used here as a precaution.

Recommendation 10: Given the uncertainty over the accuracy of values of N_0 for Jamursba-Medi and Japan populations presented, and the estimated value of for Jamursba-Medi, these values should be checked and confirmed as a matter of urgency.

Dr. Pilling presents a table on page 12 of his review indicating N_0 values that he estimated from Figures 1 and 2 of the report. While the values for Costa Rica are similar to what I reported in Table 4 of the report, the numbers for Jamursba-Medi and Japan are very different. This is because I presented numbers of nests per year in Fig. 1 and 2 for Jamursba-Medi and Japan, as this is how the data were reported to me. These were converted to estimates of nesting adult females per year as explained in the report. While this would impact the estimate of N_0, it would not have an impact on $\hat{\mu}$.

Dr. Pilling also reports that he estimated a substantially different value for $\hat{\mu}$ for Jamursba-Medi that what I reported, -0.023 compared to the -0.037 that I reported. In Table 1 of Hitipeuw et al. (2007), two columns of nest numbers are reported, one for actual number counted, and one that is adjusted to estimate total nesting for the season. I mistakenly presented the first column of numbers, actual nest counts, in Fig. 1 of my report, and in fact, these data result in $\hat{\mu}$ = -0.023, precisely what Dr. Pilling reports. In my assessment however, I used the second column of values, estimated nesting for the entire season, which are the appropriate values to use, and this trend results in $\hat{\mu}$ = -0.037. Below I present a corrected Fig. 1.

Recommendation 11: Examine the impact of changes in the proportions of interactions of Jamursba-Medi turtles (rather than taking just the mid-point of the μˆ wide range identified) on the population specific takes for this population and that of Costa Rica.

I agree that at least the high end of the range, that 100% of leatherbacks interacting with the Hawaii-based fisheries may all come from Jamursba-Medi, should be considered to be conservative.

Recommendation 12: Given the level of uncertainty around estimates of SQE, the usefulness of 1-SQE for management should be simulation tested.

Recommendation 13: Identify levels of interactions that lead to significant population pressures, *if* the approach is found to be robust (see recommendation 12).

I combine my response to recommendations 12 and 13. I agree and this is something I felt chould be added to the assessment,. Given the variance, the levels of change in 1-SQE are not statistically significant, that is one of the points of this approach, what are levels of removals from the population that are conservative and shouldn't have a detectable impact on the recovery of the population. However I never took it to the point of considering at what removal level significant impacts would be observed, which is a good suggestion.

E. Rexstad and S.T. Buckland

In this review, the paper in press with Ecological Applications, Snover and Heppell (in press) was considered critically and these reviewers seemed to find it unsatisfactory. Three points were highlighted in their Executive Summary:

> **1) That it is inconsistent to give equal weighting to Type I and Type II errors under the Endangered Species Act.**
>
> This is absolutely true and in rereading it, my wording in the report was not explicit. Snover and Heppell (in press) clearly highlight the usefulness of the SQE metric in term of being able to trade off lowering the risk of making a Type I error (considering a population to not be a risk when it is)with increasing the risk of making a Type II error (considering a population to be at risk when it is not). They specifically recommend lowering the critical threshold value for SQE to be conservative and make this tradeoff.
>
> In the report being reviewed, I had to re-establish a critical value as I redefined SQE from the Snover and Heppell (in press) paper. I found 0.75 to be the point that minimized the two types of error, and so set a range between 0.65 and 0.75 and questionable and did not consider populations to not be as risk unless the critical value fell below 0.65.
>
> Additionally, these starting values for SQE and whether the population was or was not at risk was not considered part of this assessment. All three populations considered had current SQE values well within the at-risk range.

2) They question whether or not the population simulation approach of Snover and Heppell (in press) was appropriate given the dichotomous results for actual quasi-extinction risk.

Without going into excessive detail as the Snover and Heppell (in press) paper was not the subject of this review, I believe that the population simulation approach is appropriate. The inherent assumption of the model is that current conditions remain the same over the next 3 generations or 100 yrs. Given that long of a time frame, a population on a current trajectory is either going to fall below a threshold or it is not. The question is whether 15-30 years of corrupt nest census data can accurately determine whether or not the population is or is not at risk. Several steps were taken to corrupt the simulated nest census data and even so, the diffusion approximation method had a high degree of accuracy. Obviously the assumption that current conditions will remain the same over such a long time period is a faulty one, and this is readily addressed by Snover and Heppell (in press) and is why they emphasize that this approach is NOT predictive of what the population will do, but simply an index of risk based on current trends.

3) Inadequate consideration of uncertainty and a motivation to deliver precise conclusions for use by managers.

The other reviewers had similar concerns over parameter uncertainty and I've addressed this elsewhere. My goal with this approach was to establish a method that incorporates two pieces of information that we generally have for sea turtle populations (and often not much else); population trends in units of adult nesting females and population size. From diffusion approximation we get the value of $\hat{\mu}$ to estimate trend and ultimately SQE as presented in Snover and Heppell (in press). However, this index in insensitive to population size, however the RATE of increase in SQE with increasing adult mortalities is dependent on population size. When taken together we can determine what should be an insignificant removal level for a population of a given trend and size. I delivered a range of what I considered to be appropriately low interaction levels along with other factors to consider to managers but left the interpretation of this range and determination of jeopardy up to them. Table 5 indicates ranges of adult female takes that should be acceptable depending on the level of precaution desired. The actual level of precaution was up to managers. Table 6 simply puts the interaction levels being requested by the fishey into context with the assessment and, again, it was up to managers to decide whether those levels indicated jeopardy or not.

Figure 1. Nest or nester abundance trends for Jamursba-Medi, Papua, Indonesia (Hitipeuw et al. 2007) and for Parque Nacional Las Baulas, Playa Grande, Costa Rica (Tomillo et al. 2007). Corrected figure to Fig. 1 presented in the report IR-08-010 where the data presented for Jamursba-Medi were incorrect. Correct values were used in the assessment.