Review of Long-term Research Plan for the

Eastern Tropical Pacific

Michael C.S. Kingsley

July 2006

Nuuk, Greenland

Report prepared for the University of Miami System for Independent Peer Review.

Contents

Executive Summary

The request is to review a 212-page Proposal for Long-Term Research in the Eastern Tropical Pacific, which is supported by a 100-page report on the preceding phase of research[1], as well as its review documents. The proposal is motivated by a history that started, in the 1960s, with large fatal by-catches of offshore dolphins in the yellowfin tuna fishery in the ETP and continued with the development by the fishery of fishing techniques that much reduced—to the point of eliminating—observable directly fatal interactions with dolphin schools. Research and monitoring of these stocks under U.S. programmes has been somewhat sporadic, but sufficiently maintained to give a picture of the trend in numbers; which has been closer to stationary than to the expected 4%/yr increase. The most recent research programme, run from 1997 to 2002, produced some tens of scientific papers. It confirmed that the two dolphin stocks most reduced by the fishery are still at low levels, but did not explain why, although possible causes were highlighted, among them consequences of chase and capture by the fishery that while not immediately, or observably, lethal might be highly prejudicial to the dynamics of the stock, an ecosystem shift of some kind that had deprived dolphins of some needed resource, or a supposedly inevitable, but not otherwise explained, lag before recovery would begin.

The present proposal is comprehensive, including sections on: Assessment Modelling; Monitoring of Numbers; Stock Structure; Investigations of Coastal Dolphins; Ecosystem studies; Fishery Effect, Reduction of Bycatch (of other species); and Data Management. A question that is not directly enunciated, but which is hinted at in the section on ecosystem studies, is that if dolphins associate with tuna for mutual benefit, repeatedly having the tuna fished out from under them may deprive them of the benefits of the association whatever those might be, without necessarily entailing any other form of interaction with the fishing vessel or a major, or measurable, shift in the surrounding ecosystem. If a school of dolphins has its tuna fished out from under it, how long does it take to acquire a replacement, and how does it fare in the meantime?

The proposal could benefit from some re-working with attention paid to some of the following points:

Focus clearly on work directed to answering the priority questions, and relegate other studies (such as collecting photographs of large whales) to an Appendix;

[1] 'Report of the Scientific Research Program under the International Dolphin Conservation Program Act' NOAA-TM-NMFS-SWFSC-372.

Organise sections so that previous work and relevant results are got out of the way before proceeding to describing the proposed new work and its expected new results[2]. There is a general tendency in all sections to be continually skipping back and forth between past studies, principles underlying proposed new work, and the proposed work itself.

Reduce the amount of duplication between—and within—the different sections.

Prioritize—and delete the unimportant. It is not likely that a relationship between sperm competition and spermatozoon morphology really holds the key to the present stagnation of ETP dolphin stocks.

In general, and where possible, provide more detail on the measurements that will be made, on the sample sizes expected, and on the analysis of the data. A failing of the proposal is that there is, generally, a wealth of information on background and principles, less information on specific topics of investigation, and even less on the measurements that will be made and the analyses that will be carried out on the resulting data[3].

The proposal is in general wordy but non-specific about both methods and expected results; it is puzzling, somehow, that a document that is so long should manage to contain so little of real substance. It is difficult to imagine that if submitted as a proposal to a normal scientific funding agency it would receive a very high rating.

[2] As one example, we find at the start of the section 4. in Ecosystem studies (stable isotope analysis) the words 'Stable isotope analysis has become a common tool. . .' Some 4 pages later, below a heading 'Expected Results', we read that 'The use of stable isotopes. . .has been going on for decades'. This isn't a result, it's not even an expected result, its background information, and information that has already been presented.

[3] In the cases of genetic studies of coastal stock structure, and of stable isotope analysis, the *laboratory methods* are described in some detail, curious exceptions to the general lack of specifics.

Summary of Statement of Work:

For sections within the reviewer's area of expertise, the review should determine whether the scope of the proposal is 'adequate and appropriate'; any areas are addressed in less, or more, depth than needed; whether the approaches proposed are unbiased, whether the proposal represents best available science (and if not, corrective suggestions), and overall comments on strengths and weaknesses.

And, overall, whether the different sections are well integrated.

This review is expected, according to the SOW, to concentrate on the reviewer's area of expertise, and I have considered that, in respect of this topic, to be quantitative methods for monitoring and estimating numbers and for modelling population dynamics.

Sections considered:

- **Assessments of Status and Trends**
- **Abundance and Monitoring of Pelagic Dolphin Populations**
- **Abundance, Structure and Montitoring of Coastal Dolphin Stocks**

Supplementary sections considered

- **Fishery Effects**

Overall assessment:

The overall quality of the proposal is not high. The Introduction is reasonably clear, succinct, and logical. But its quality is not backed up by the quality of the individual sections, and the result is a document that is wordy, repetitious and disorganised, and as a result far longer than necessary. It gives little evidence that any one mind has said 'This is my product and it's going to be good and convincing' and has consequently taken hold of the proposal as a single entity and hammered it into shape. The proposal reads more like a shopping list than a prioritized research proposal. As a shopping list it is extremely comprehensive, perhaps too much so: similar activities recur in different sections of the proposal. The individual proposals bear little evidence of quantitative thinking, and the numbers that one might expect to be of interest, that should inform sampling design and logistics in this proposal, and that one would expect to have surfaced in the course of the IDC research programme do not occur. Examples are the numbers of dolphins per school, the density of schools, the size of the 'calf deficit', the ages that can be

separated by photogrammetry, the e.s.w. achieved by in current line-transect survey methods, &c. There is a section on surveys, for example, that one would expect to start with the basic information that present standard methods steam over *xx* miles of trackline with an e.s.w. of *yy*, tend to make on average *zz* sightings comprising *dd* animals, and produce e.c.v.s of *pp* %, of which *ee* is due to encounter rate and *ll* to uncertainty in the sighting model—but this kind of quantitative background is lacking, here and elsewhere.

While the scope of the proposal is adequate, its integration of the proposal into a coherent whole is not one of its strengths. Some areas are perhaps addressed in more depth than needed. In particular, it is difficult to imagine that the apparent failure of dolphin stocks to recover since the number killed in the tuna fishery was greatly reduced by the development of the 'backing down' technique is due to a progressive worsening of survey methods, and the multiplied efforts to refine detection, and the analysis of detections, would appear to be superfluous. Overall, the individual sections are not well integrated, and give instead the impression of being a set of independent proposals written by scientific teams with different interests, instead of connected contributions to the resolution of a central problem. Integration of the different sections will not happen just by wishing for it, but will instead require specific actions directed to achieving it. I would suggest that one way to do this might be first to write into each section a sub-section describing how it would interact with other sections, and review this set of sub-sections—for example to check that they are consistent with one another[4]. With the basic information on integration ready to hand in that form, it would then be possible both to ensure that it was made appropriately evident, and also to rework the individual sections in such a way that the integration would be strengthened.

Questions of whether the approaches proposed are unbiased are difficult to answer. Bias in estimation is principally a question of estimation methods, and the proposal is to such a degree lacking in detail that it is quite difficult to determine whether bias will necessarily be incurred, or creep in, in the execution of the work that is so generally described here. Whether the proposal is biased in the sense of predisposing the enquiry to ascribe the stagnation of stocks to one cause rather than another is a question more important, more subtle, and more interesting. The section in this proposal concerned with possible fishery effects is the longest, the most replete with new ideas, and the best developed, and gives the impression that there is a prior inclination towards the hypothesis that the repeated chase and encircling of dolphin schools is the most likely cause of their stagnation (with which this reviewer is inclined to agree).

This reviewer finds a certain lack of focus; in a document that purports to be concerned with a single primary topic, namely the non-recovery of eastern spinner and north-eastern offshore spotted dolphins, it is disconcerting to find oneself suddenly immersed in a description of the collection of photographs of large cetaceans, for example. I suggest that collecting all the material that is not directed towards achieving a stated primary objective and putting it in an Appendix labelled 'Research activities not directed to solving the dolphin problem, but which could be prosecuted from the same platforms' would reassure the reader that an overall objective for the programme has directed its content and structure.

One measure that I could not find directly addressed in any of the proposals was that of a condition index for animals in these stocks. In connection with hypotheses related to a change in carrying capacity one would expect to encounter a direct effect related to a later onset of female reproductive activity mediated through a lowered condition index; in connection with a hypothesis of a high rate of abortion because pregnant females were so exhausted by being repeatedly chased that they could not complete their pregnancies, again a lowered condition index should be evident. Yet even in the proposal for handling dolphins in inserting PIT tags there is no mention of measuring girth:length ratios, weight:length ratios, blubber thickness, or other gross measure of condition.

The question of whether the proposal's component sections represent 'Best Available Science' is not obviously answerable. In the first place, the quality of the science may depend in large degree on the quality of the execution, and much of the work in this proposal is so generally described that this particular point is difficult to evaluate. In the second, projects may be very good science, but not relevant to the principal topic of the proposal. However, the proposal as a whole is so comprehensive, and sufficiently general in its wording, that good science could certainly be carried out within its framework.

The question is the proposal as a whole 'Appropriate and Adequate' prompts the reply 'to what?' – presumably to solving the problem of the stagnation of the ETP dolphin stocks. Some sections hold more promise than others of making additional progress, and may be considered more 'appropriate': notably, the furthering of studies on the continuing effects of fishing on dolphins, and also, with less promise, the continuation of ecosystem studies. As to whether the proposal is 'adequate', in the sense that the investigations described are likely, within the (mostly quite short) time frames described, to generate (a) definite answer(s) to the questions; that, it is difficult to be optimistic about. Not because the proposal is

[4] For example, the section 'Towards an ecosystem approach to management' seems to imply that dolphin monitoring cruises routinely collect data for ecosystem studies, but this is difficult to determine from the section on abundance

markedly deficient, but because it is mostly treading paths that have already been explored. One field of investigation that seems not to have been explicitly mentioned might be bracketed under 'Other Human Influences' to include e.g. other fisheries—are there any of significance?—shipping, contaminants, etc. It would be helpful at least to get these out of the way in some orderly fashion.

Assessments of Status and Trends

This is a somewhat wordy and not overly well organized section, and was not easy to review. It was remarkable that after three sentences related to hypotheses and management actions, and before having considered either models or data, a declaration was made that 'we will use Bayesian methods to achieve these objectives'—a statement that at such a point seems not only premature, but misaligned: Bayesian methods fit model parameters to data, or describe the distribution of parameter values that can be considered as fitting a model to data. But they don't, on their own, achieve objectives like those stated—that process draws heavily on the skill of the modeller, the availability of suitable data suitably collected, and the available understanding of the processes driving the system being modelled.

The section is continually skipping between data, models, future work, and past experience; a more orderly organization might have been to give a complete description of previous work and history before proceeding to consider the extensions to the data that the future research will incorporate into models and the extensions of the model structure and function.

The document provides no definition of 'assessment', but does add 'of status and trends'. The present-day standard meaning of 'stock assessment' is constructing a more or less complex mathematical model of the dynamics of a stock–-in other words how its size and structure change over time, possibly in response to external influences; to adjust the parameters of the model so that its behaviour fits some set (series) of data from the past on its trajectory and on the external influences; and to use the model with the adjusted parameter values to predict future trajectories in response to assumed or selected sets of values for external variables. The process of fitting the model to past data may also show whether the model fails to fit the past data, perhaps revealing a fundamental problem with the model (or with the data). However, 'assessment of status and trends' might bear a different interpretation.

The modelling process is an interplay between the amount and quality of data available, the level of understanding of the influence external factors have on the dynamics, the level of understanding of the

and monitoring.

internal dynamic processes, and the computational and model-fitting facilities available. It is pointless to build complex models that include many external variables if their influence on the stock dynamics is poorly understood, or if the observations of them are lacking; or if the number of observations of stock size structure is inadequate to allow many parameters to be fitted; or to build complex models of stock structure if stock structure has not been recorded.

The document proposes 3 sets of questions: The first related to whether presently available 'knowledge and observation' support possible explanations of apparent lack of recovery; the second, very diffuse, related to a) whether the expectations are reasonable, b) whether observations (scil. on stock sizes) are accurate enough to do the job and c) whether the assessment models in use are adequate; and the third, related to predictions of future trajectories and effects on them of management actions

The following is a summary of what the proposal includes; I found it necessary to write it in order to be clear about what the proposal proposes.

Data.

The data proposed for use is (with **new data in bold**):

Estimates of current and historical (fishery) kills

including previously unused results on **calf kills in 1973–90 and 1996–2000**.

Stock-specific abundance estimates

including IATTC data of sightings made by TVOs, as an abundance index

Numbers of dolphin sets—hence possible unobserved fishery effects

a possible index of frequency of encirclement per dolphin;

Purse-seine fishing effort and intensity

including IATTC log-book data

Average school sizes (from what; are school sizes set on the same as in surveys, or do log-books record school sizes?)

Age- and stage-specific mortality estimates

including **photogrammetric estimates of proportions of young animals**

Life history data

including **fœtal mortality**

including new information on **e.g. length of lactation** (to index birth-rates)

'Ecosystem indices' (not necessarily available annually)

including yellowfin tuna unfished biomass

including seabird abundance

including 'data from net tows'

Modelling:

The models proposed will be restricted to a 'reasonable set of candidates' that will include:

Density-independent population-dynamics models without external influences, based only on life-history variables, updating previous models of this kind and the resulting estimates of maximum population growth rates.

Process-based models in which population dynamics will depend on vital rates, and vital rates will be affected by: exposure to sets; fishing mortality of tuna (as a fishery exposure index); yellowfin tuna unfished biomass (ecosystem index); and 'estimated biomasses of other components in the ETP' (this may refer to the tuna index, though).

Extensions of these including observations on a larger set of ecosystem variables, some not available every year.

Extensions of the structure of models of dolphin population dynamics that include the possibility of so-called 'depensatory' effects; i.e. decreased growth rates at low densities, in this case due to lowered birth-rates.

Joint modelling of different (offshore) stocks, allowing both the power of using common values for basic population-dynamics parameters and the flexibility of between-species differences such as different exposures, or susceptibilities, to the tuna fishery, different catch histories, or different ecosystem responses mediated through different diets. Stock pairs mentioned are eastern spinner and northeastern offshore spotted; and eastern spinner and whitebelly spinner.

The research plan also proposes simulation modelling of management processes through 'operational models' that embrace all steps up to and including management action, i.e. models that start with the true dynamics of a stock, proceed through imperfect observations of its size and structure, the imprecise estimation of stock and dynamic parameters through the application of an assessment model, and the

imperfect implementation of management decisions made, according to modelled rules, on the basis of those estimates[5].

Specific actions listed include, in the short term:

- setting up facilities to provide assessment advice from single-stock models of dolphins at short notice;

Then using single-stock models built on these platforms as a springboard for modelling:

- unobserved mortality (in fishery interactions);
- ecosystem effects, considering external 'ecosystem' influences on population dynamics parameters;
- simultaneously, more than one stock or species.

Short-term actions will also include a workshop on TV observer data.

In the medium term

- from joint models of several stocks or species, evaluate alternative hypotheses for delay in the recovery of the stocks; (three classes of hypothesis mentioned relate to persisting fishery effects, ecosystem changes, and 'complex' population dynamics);
- be able to provide scientifically based management advice;
- embed assessment models in operating models, that will then be used to evaluate management decision rules and possible population trajectories;

The longer term activities mentioned include:

- developing assessment models for coastal spotted dolphins
- use operating models to provide advice on how best to acquire data for future management of dolphin stocks.

It is expected that work on developing models will continue until the rate of improvement in the advice developed decreases significantly. There is no statement on how the rate of improvement would be measured.

[5] The research plan draws a parallel with the RMP of the IWC, but a better parallel might be the implementation trials that are used to test the application of the RMP.

Review Comments

The material presented is general enough, while not being very specific, to allow little scope for pointed review. There being little in the way of description of specific analyses makes it difficult to understand, and comment on, how the proposers expect, by analysing *model* outputs, to ascribe particular aspects of observed past or present population trajectories to one hypothetical cause or another.

The classes of hypothesis that this section proposes to examine are basically those that the entire proposal is directed towards, and this section does little—or nothing—to refine or shape them.

The proposers are insistent on using Bayesian methods for fitting models to data sets, and provided due caution is observed this is unexceptionable, but it would not necessarily be wise completely to lose sight of other fitting methods.

The programme of research is ambitious, but not overly so, and appears to proceed logically from step to step. The scope does appear adequate and appropriate, but the level of integration with the programme of ecosystem studies, with which one would expect this programme to be closely coordinated, seems lacking. In terms of bias, it appears that this programme has an inclination to build population-dynamics models of dolphin stocks, and then surround them with ecosystem influences; while the 'Ecosystem Studies' section appears to have a bias to working the other way round, building ecosystem models and working inwards to dolphin dynamics. However, the approach proposed here is likely to keep its feet on the ground and to produce, and maintain, workable models.

Whether this represents Best Available Science is not easy to answer. It seems likely to represent the best use of the available data in the current context. The ecosystem components seem somewhat hit-and-miss. There is little evidence that the assessment modelling has directed data gathering in the ecosystem programme. It would seem appropriate for this modelling programme to be more closely coordinated with the Ecosystem modelling described in the ecosystem methods section.

Abundance and Monitoring of Pelagic Dolphin Populations:

This section is organized under a series of topics, some of which have to do with the monitoring of numbers. However, the section might better be entitled 'At-sea studies of cetaceans and survey methods'.

I summarise the individual topics below, first in term of the past work described, and then in terms of the extensions now proposed.

Past work

1. Survey design.

This section gives a short overview of existing design methods, including a description of how surveys are now stratified on stock boundaries.

2. Aerial photogrammetry.

Existing methods described include: using aerial counts of groups to get correction factors for observers' estimates of group size; using photogrammetric methods both for stock structure, from size of adults, and age structure of stocks, from sizes of all individuals, with corresponding inferences on the population dynamics of stocks.

4. Line transect analyses

Developments to date have included the development of sighting-curve models that include the effect of various factors, presumably improving the description/definition of sighting curves.

5. Closing mode bias.

A bias induced by the use of closing mode is suggested, supported by heuristic reasoning and some (conflicting) results from Antarctic surveys.

6. Echosounder effects

Some preliminary results indicate that 'dolphin school sizes, sighting rates, and distances may be affected by the echosounder'.

7. Unidentified sightings

'Unidentified sightings represent abundance that has to be accounted for'. Not necessarily; line-transect analysis ought to be able to treat unidentified sightings as sightings that were not made, and estimated numbers should not be affected, unless there is a failure to identify schools that are close to the line. The method described as being currently in use, assigning unidentified sightings in proportion to identified ones, is worse than useless, in fact very bad indeed.

8. Passive acoustics

It appears that no work has yet been done on whether passive acoustics could improve or extend surveys. 'The identification of sounds to species, and how production of sounds is related to group size, time of day, activity, &c, is an area of current research'.

9. Additional estimates.

'Data in earlier years were collected somewhat differently.' (It is not quite clear whether this means in the earlier years of the series since 1979, or simply 'before 1979').

10. Using past data to improve current estimates.

Current analyses use information collected in each year to estimate abundance separately for that year. This approach is not best calculated to measure changes in numbers from year to year.

Extensions and Future Research Possibilities

1. Survey design.

It was difficult to discern that this subsection included any specific suggestions for research or development, although one would not expect that survey design has in fact reached such a point of perfection that no development was possible.

2. Aerial photogrammetry.

'Extension' of aerial photogrammetry for both training, and evaluating, observers in the laboratory, including comparing laboratory training situations with field conditions, is proposed.

3. Use of shipboard photographs to differentiate stocks.

Present methods would be continued; but it is not clear what effect this has on the effectiveness of programmes for monitoring numbers.

4. Line transect analyses.

Developments proposed are further elaboration of sighting models. It is suggested that this will improve estimation of dolphin abundance, although present deficiencies are not specified.

5. Closing-mode effects.

Research to evaluate closing-mode bias is proposed, in particular a dedicated cruise and replicated transects. Replication of transects is, however, not very precise and much data would be needed to generate results that would be in any way definitive.

The importance of this topic will depend on the priority given in the monitoring programme to trend vs. absolute numbers.

6. Echosounder effects

'Further investigation is needed to determine whether use of echosounders affects line-transect data, and if so, how much.' How these further investigations might be carried out is not specified; nor, in fact, is the mechanism by which the effect might occur: whether by the dolphins' avoiding the vessel—in which case the effect should be subsumed in other studies of dolphin reaction to survey vessels—or by their spending more, or less, time submerged, or changing porpoising behaviour in a way that affects the estimation of school size.

7. There are suggestions to improve the assignment of unidentified sightings by using acoustic, behavioural or oceanographic observations.

8. Passive acoustics:

There is a suggestion to investigate whether collecting acoustic observations with a towed hydrophone array would improve or extend the visual line-transect data.

9. There is a suggestion that with 'additional work' data from 'earlier years' could produce useful estimates. What sort of additional work is necessary, is not specified at all.

10. Using all data has the potential to improve estimates.

Review comments:

It seems quite unlikely that the apparent lack of recovery of the offshore dolphin stocks is due to a progressive worsening of survey competence at a rate of 4% per year over the last 20–30 years, so to some extent it appears that the elaboration of enquiries into the performance of shipboard surveys exceeds the appropriate.

In spite of this, the proposers have been quite comprehensive in suggesting topics for investigation, but not very specific in suggesting how the investigations would be carried out. In the view of this reviewer, two main points are missing.

The first is that surveys appear to be planned and executed one survey at a time, not as survey series. If the objective is to monitor trends over time, there are good reasons to plan the series as a series. If distribution tends to be stable (so that densities on given transect segments are positively correlated from year to year), repeating survey designs is likely to result in survey errors that are positively correlated between surveys, and this has the potential to improve the detection of trends. The single-survey approach, in which each survey is planned, designed and analysed as though it were a single shot, tends to produce estimates that are uncorrelated between surveys (i.e. surveys have independent errors) and, over time, unbiased estimates of distribution and numbers. However, it is not clear that the unbiased long-term mean should be the primary objective in this survey programme. If the objective is to detect trends, then it may be legitimate to seek positive between-survey error correlations, even though such positively correlated errors do not average out over the series.

Furthermore, in the use of line-transect analysis, a survey programme directed toward comparisons should place high value on standardised methods, using common sighting curves for all surveys (including common corrections for other factors than distance), and collecting distant sightings. Such a programme might also benefit from standardising survey tracks, stratification, and effort allocation. A programme interested less in trend estimation than in long-term data on numbers and distribution will have less interest in distant sightings, will not necessarily wish to standardise sighting curves over survey series, will have no interest in standardising survey tracks, &c. Instead it would prefer negatively correlated selection of survey tracks, and will wish to have maximum accuracy in sightings and group-size estimates in a limited strip near the trackline.

A clearer view of the objective of the monitoring programme would aid in priorising some possible investigations. It would, clearly, much affect the weight to be placed on the simultaneous analysis of multi-year survey data sets, and the way in which such analyses might be structured.

A second point that is not clearly considered is the question of the spatial allocation of survey effort, its effect on the accuracy of surveys, the use of systematic or semi-systematic sampling for improving precision, and the use of appropriate analyses for taking advantage of systematic sampling in estimating precision. I suppose that the standard package DISTANCE is probably used for analyses, and it is my

understanding that DISTANCE is not highly sophisticated in its consideration of the spatial properties of survey design, being largely preoccupied with the determination of sighting curves. The spatial design of surveys is moderately important in getting good results, but the entire proposal is completely mute on this topic[6].

The trend in numbers for these stocks is fitted to the survey results by linear regression, and it appears that in fitting the regression lines the individual surveys are weighted by their realised standard errors. This is probably fallacious, and a more appropriate weighting would probably be by effort, either total trackline or total effective strip area.

The proposal is probably not biased in the sense that the investigations proposed will not necessarily generate, of themselves, increases or decreases in survey estimates. It does not represent Best Available Science, in my view being over-preoccupied with the measurement aspects of surveys and giving inadequate attention to some of the more basic questions related to the objectives of the survey series, the consequences of those objectives for design, and the selection of suitable analysis methods. The allocation of identities to unidentified sightings is likely to over-estimate the precision of estimates and the effects of doing so should at least be investigated, although I would suggest abandoning this practice.

Abundance, Structure and Monitoring of Coastal Dolphin Stocks.

This section of the proposal is short (15 pp.) and quite unspecific about methods and about experimental designs. The topic that receives the most attention and a detailed description is the *laboratory methods* for genetic analysis which, in this reviewer's opinion, is misplaced: this is not the stuff of research proposals.

Aerial survey is not mentioned as a possibility for surveys, although most of the parameters of the problem indicate that parallel, systematic, northeast–southwest-going transect would be an appropriate design, and it is difficult to envisage such a design being efficiently executed by ship. The design for collection of samples for genetic analysis is not discussed; if there is no prior information on stock boundaries, systematic sampling might be appropriate, but if there is some information already available on the likely position of stock boundaries, it would make sense to concentrate sampling near the supposed boundaries, with lower intensity nearer to the centre of the distribution of each supposed stock.

[6] I did eventually find some information on this topic, but in the section on ecosystem research.

The scope of the proposal is not adequate, in that important topics are not discussed and important components of the design of an investigation are not mentioned, and cannot be considered Best Available Science, for much the same reasons. This proposal needs to give some more background on the present state of knowledge and put more reasoning behind its design and methods, and produce some more definite statements of what will be done, where, and why.

Fishery Effects

This lengthy section is set up with 5 main sections:

Fishery Exposure, with sub-sections Exposure Index, Chase and Set Frequency, PIT Tagging

Life History, with sub-sections Updating Estimates of L.H. parameters, Estimating Vital Rates from Aerial Photogrammetry, Estimating V.R. from Blubber Steroids, and Estimating Relative Stress from Blubber Steroids

Physiology, with sub-sections on Drafting Hydrodynamics, Analysis of data from ETP dolphins, each with a complex suite of sub-sub-sections and further sub-divisions

Behaviour, with sub-sections on Responses to Chase, Social disruption, Mating behaviour,

Class-5 fishing.

However, it can be seen as having two components: one, (the first and fifth sub-sections) based on Class-6 observer data, on observer data from Class-5 vessels, on catch sampling from Class-4 and -5 vessels, and on PIT tagging, documents and analyses dolphins' exposure to the fishery, including its spatial distribution and individuals' recurrent exposure, and seeks correlations between values of various life-history variables and fishery exposure. The other comprises a comprehensive review of life history parameters, behaviour, including social behaviour, and physiology, including hydrodynamics of swimming.

The first of these components, the documentation and analysis of fishery exposure, is necessary basic work; that much of the data is over 30 years old is a concern. The investigations into chase and set frequency and the exposure index seem to be to some extent overlapping: the last two paragraphs of the 'Chase and Set Frequency' material duplicate quite closely the 'Exposure Index' material. If the sets going into the 'Exposure Index' are weighted by the size of the school that was set on, it would appear to be appropriate to introduce a corresponding down-weighting according to the local density of dolphins. A dolphin that had ten sets within a short distance in the past week would have a lower individual chance of being set on if there were many other dolphins in the area to share the load. If this index is not corrected for local densities, it is difficult to see that the index will in fact reflect individual dolphin experience. The analysis of correlations between spatially structured indices that are likely themselves to be spatially autocorrelated is not simple, but the problems involved are not mentioned.

It is still not known how severe are the stress effects of chase or capture, yet this proposal does not propose to continue work on this problem. In view of the indications to date that these effects occur, and the apparent inability to be sure how serious they are, this seems a significant omission.

The tagging proposal is interesting, but I would have expected to see at least some back-of-the-envelope calculations on the effort required to place enough tags to get useful data from stocks which although depleted still number in the hundreds of thousands. Mark-recapture estimations work best on small stocks where a reasonable recapture rate can be expected. Careful technical evaluation will be necessary—placing a large number of tags that were found not to be reliably detected by the installed systems would be unfortunate, and it can probably be expected that the performance of the equipment, and the care taken in using it, will alike tend to decline with time.

The second component describes very wide-ranging enquiries into dolphin life history and population-dynamics variables, and more besides. The section on life history parameters is not well written. The first section on updating estimates ends with a 'Timeline' section that says '. . .data are available to proceed with the proposed analyses.' But I couldn't find a coherent description of a suite of proposed analyses in the foregoing material, which is a mixture of what has been done in the past, inferences from previous analyses, proposals for re-starting sampling, and so on. The estimation of vital rates from aerial photogrammetry is misnamed—the proposal is concerned with estimating length structure from aerial photogrammetry. Making inferences from that to vital rates is another stage.

There may be some potential for misinterpretation in the section on Fœtal Mortality. The proposers seem to be suggesting that existing estimates are for 70% fœtal mortality, or one conception in 3 1/3 carried to term, which if true makes it rather surprising that the populations even survive, let alone fail to increase. However, that apart, it is not clear from the proposal what kind of data is available or how it will be analysed to determine the rate of abortion. The proposers mention shortage of food resources in connection with the possibility that pregnant females are expending so much energy in fleeing from purse-seiners that they cannot complete their pregnancies, but there is no suggestion to measure body condition by any method, even for dolphins killed in sets. A standard analysis of growth-rates would also be a useful basic indicator of possible effect of stress due to the fishery.

The sections on estimating population structure through blubber biopsy are straightforward enough, although as above, what is being estimated is population structure, not vital rates directly. The proposers do not discuss the possibility of segregation of the stocks according to sex or pregnancy status, which could induce bias in the estimation of population structure from the projectile sampling.

The work on physiology, mostly of swimming, is interesting, but is unlikely in the near future to contribute much to answering the present questions. Questions of ontogeny would appear to be largely confounded with immediate direct effects of fishery exposure. Although investigations of swimming physiology and energetics are planned, there is no mention of any studies in flume tanks with measurement of gas exchange, heart rate or similar physiological variables. Even small amounts of such data could be invaluable in reinforcing and validating the proposed studies, which will take an approach to the measurement of the physiology of swimming, including such complex subjects as the energetics saving to calves from rafting on the mother, that seem to be largely based on theoretical physical principles. And if it is suggested that the ‘calf deficit’ may be due to separation of calves from mothers in the course of the set, is there no possibility for enhanced observer effort to see whether this is in fact occurring?

The work on mating systems, sperm competition, and its relationship to the construction of spermatozoa, would not appear to be directly relevant to the problem of the stagnation of the offshore dolphin stocks, and might therefore be considered for inclusion in an Appendix of less-relevant investigations.

The studies of fishing practices by vessels under Class 6 intend to find out whether these smaller vessels, by setting on dolphins, are killing some dolphins that are not included in mortality limit quotas. Two lines of investigation are proposed: the first comprises analysis of existing (past) data on observed sets by

Class 5 vessels, and extending this data set by putting observers on Class 5 vessels; the second consists of dockside sampling of catches from Class 4 and 5 vessels with the expectation of being able to infer the frequency of dolphin sets from the composition of the catch.

The studies in this section are very diverse, and it is not possible to give a single answer to such questions as whether they are all adequate and appropriate, for example. The documentation and charting of dolphins' exposure to the fishery and the frequency with which they may be set on is basic work that is worth doing; I suspect a possibility of bias in the intention to weight sets by school size without correcting for density. The investigations of life history variables and vital rates are necessary and appropriate, but might not be adequate, in that the measurement of additional physiological parameters might be needed to complete the information. Sampling bias in the by-catch is described here, but there is also a possibility of bias with respect to projectile sampling, as some classes of animal are often easier to approach than others, and I don't know to what extent this question has been answered for these dolphins.

Appendix I: Bibliography of Material Provided

Anonymous. 2006. Long term research in the eastern tropical Pacific: a proposal from the Southwest Fisheries Science Center, NOAA Fisheries Service. Unpublished. 212 pp.

Dower, J.F. n.d. A review of ecosystem research in the IDCPA Science Report.. Prep. For the UM Independent System fo Peer Reviews. Unpublished. Not paginated.

Drinkwater, K. n.d. Final review of the report of the scientific research program under the International Dolphin Conservation Program Act. Unpublished. 30 pp.

Martineau, D. 2002. Review. 2001 chase encirclement stress studies on dolphins targeted in eastern tropical PacificOcean purse seine operations. Final Report. Unpublished. 31 pp.

McAllister, M. 2002. Review of the International Dolphin Conservation Program Act scientific DRAFT report. Unpublished. 43 pp.

Ortiz, R.M. n.d. Final report on the IDCPA. Prep. for the U. o fMiami Independet System for Peer Review. Unpublished. Not paginated.

Reilly, S.B. and 11 others. 2005. Report of the scientific research program under the Internatoinal Dolphin Conservation Program Act. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-372. 100 pp.

Appendix II:

Consulting agreement between the University of Miami and Michael Kingsley
STATEMENT OF WORK
Eastern Tropical Pacific Dolphin Research Plan

Background

The topic of the review is the evaluation of a long-term research plan to monitor the abundance and environment of several species of tropical pelagic dolphins that are killed in the purse seine tuna fishery of the eastern tropical Pacific (ETP), and the evaluation of reasons for the apparent lack of recovery of depleted stocks. The Southwest Fisheries Science Center (SWC) has been conducting research in the ETP since the 1960's. Research topics through the 1980's ranged from assessing direct dolphin mortality in the fishery to an examination of fundamental aspects of biology and life history, monitoring the numbers and types of dolphins being taken, conducting sighting surveys of dolphin abundance from ships to estimate abundance and trends over time, and collecting data and samples on a broad range of attributes of the physical and biological environment.

In a 1997 amendment to the Marine Mammal Protection Act, Congress directed the National Marine Fisheries Service to undertake a research program to determine, by the end of 2002, whether the fishery was having a "significant adverse impact" on depleted dolphin stocks in the ETP. The research program that the SWC designed included four components: abundance estimation, ecosystem studies, stress and other fishery effect studies, and stock assessment. This research culminated in a Final Science Report (FSR) in 2002 and thirty-four separate science papers to provide information for answering the question posed by Congress. The FSR contained the following primary conclusions: (1) northeastern offshore spotted dolphins were at 20% and eastern spinner dolphins at 35% of their pre-fishery levels of the late 1950's, levels largely unchanged since the 1970s; and (2) neither population is recovering at a rate consistent with these levels of depletion and very substantial reductions in reported kills. Data on the possible causes for the continued depletions were too sparse to be conclusive on possible ecosystem effects, but existing information did not support the occurrence of the 70% reduction in effective carrying capacity that would be required to cause the dolphin stocks to remain stable at such low levels. Data and results on possible indirect fishery effects also were inconclusive, but did disclose a common pattern of separation of cows and nursing calves. More data and studies are needed to bring closure to questions surrounding the lack of substantial progress toward recovery by these severely depleted dolphin stocks. The long-term ETP research proposal describes a program of action directed at this closure.

Reviewer Responsibilities

The Center of Independent Experts (CIE) shall provide four expert reviewers. Each reviewer's duties shall require a maximum of seven days of effort, including time to read relevant documents and to produce an individual written report consisting of their comments and recommendations. No travel is required, so each reviewer shall work from their home location.

Each reviewer's report shall reflect his/her area(s) of expertise, and no consensus opinion (or report) will be required. Further, each reviewer shall only comment on sections within his/her area of expertise.

Expertise needed to review the proposed long-term research plan, including its methods, scope and priorities, includes the following: (1) cetacean biology, (2) line transect-based abundance estimation and stock assessment modeling, (3) biological oceanography and pelagic marine ecology, and (4) population identity – stock structure.

Documents supplied to the reviewers shall consist of the (1) Long-Term Research Proposal in the ETP, (2) 2002 Final Science Report, and (3) CIE reviews of the Final Science Report. The reviewers shall become familiar with the research plan and the background documents.

Specific Reviewer Tasks and Schedule

1. Read and consider the 2002 Final Science Report and CIE reviews of the Final Science Report that provide context and background on research in the eastern tropical Pacific Ocean.

2. Read and analyze the Long-Term Research Proposal for the ETP that describes the SWC's approach to resolve the cause(s) of the apparent lack of recovery by depleted dolphin stocks in the ETP.

3. Specific points to be addressed (at minimum) for sections within each reviewer's area of expertise:
   (a) Is the scope of the proposal adequate and appropriate?
   (b) Are any areas addressed in less, or more, depth than needed?
   (c) Are the approaches proposed unbiased?
   (d) Does the proposal represent Best Available Science? If not, what specifically would be required to meet that designation, in your opinion?
   (e) Comment on the proposal's strengths and weaknesses, and suggest any additional lines of research that appear promising.

4. Specific points to be addressed (at minimum) for all sections:
   (a) Overall, are the individual sections well integrated into the proposal as a whole? If not, what could be done to improve integration?

5. No later than August 1, 2006, submit a written report[7] to the CIE that addresses the points in items 3 and 4 above. See Annex I for additional details on the report outline. Each report shall be sent to Dr. David Die, via email at ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani, via email at mshivlani@rsmas.miami.edu.

Submission and Acceptance of CIE Reports

[7] Each written report will undergo an internal CIE review before it is considered final.

The CIE shall provide the final individual reviewer reports in pdf format for review for compliance with this Statement of Work and approval by NOAA Fisheries to the COTR, Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov), no later than August 15, 2006. The COTR shall notify the CIE via e-mail regarding acceptance of the reviewers' reports. Following the COTR's approval, the CIE shall provide pdf format copies of the reviewers' reports to the COTR.

ANNEX I: REPORT GENERATION AND PROCEDURAL ITEMS

1. The report should be prefaced with an executive summary of comments and/or recommendations.

2. The main body of the report should consist of a background, description of review activities, summary of comments, and conclusions/recommendations.

3. The report should also include as separate appendices the bibliography of materials provided by the Center for Independent Experts and a copy of the statement of work.

Please refer to the following website for additional information on report generation: http://www.rsmas.miami.edu/groups/cimas/Report_Standard_Format.html