SUMMARY REPORT:

REVIEW OF

ACOUSTIC INTEGRATION MODEL (AIM)

25-27 SEPTEMBER, 2006

WASHINGTON, D.C.

Prepared by

Patrick Cordue
Fisheries Consultant
New Zealand

for

University of Miami
Independent System for Peer Review

11 December 2006

EXECUTIVE SUMMARY

A three person CIE Review Panel meet from September 25-27, 2006 at Marine Acoustic Inc.'s offices in Arlington VA. The purpose of the meeting was to review Marine Acoustic Inc.'s Acoustic Integration Model (AIM). The review was initiated by NMFS who required an independent peer review of AIM.

AIM is a software package. Its primary use has been to develop specific application models which were used to predict the average number of marine mammals which would be exposed to sound levels above a given threshold. Such estimates are currently needed in permit applications for any activity which could adversely impact the sound environment of marine mammals.

The three terms of reference required that the Panel evaluate whether AIM correctly implements the models and data upon which it is based; whether animal movements are adequately simulated; and whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development and evaluation.

The Panel agreed that AIM appears to be correctly implemented. However, all panelists had recommendations for further testing to be undertaken. They also agreed that animal movement appears to be appropriately modelled within AIM given the inadequacies of the available data.

With regard to whether AIM satisfies the CREM guidelines there was some diversity of opinion. This is understandable given that the CREM guidelines are not *directly* applicable to AIM since it is not an application model (but a tool for developing such models).

One of the requirements of the CREM guidelines is for the "model" to have undergone "adequate" peer review. The panelists were split on this question. NMFS clearly thought that an independent peer review was required and hence they initiated this review. The Panel have now reviewed AIM (in what appears to be the first independent peer review), but it is not for them to judge whether their review was an "adequate peer review".

The Panel did agree that the principles of credible science had been addressed during the development of AIM. They agreed that AIM is a useful and credible tool for developing application models. The need for expertise in the use of AIM was noted (e.g., in the choice of transmission loss model); as was the absence of appropriate uncertainty and sensitivity tests in the current applications of AIM. It follows, that the Panel agree that the use of AIM can lead to models which will meet the CREM guidelines. However, such models, at this stage, would need to be evaluated on a case-by-case basis (i.e., merely using AIM is not sufficient; it must be used appropriately for the specific application).

BACKGROUND

A three person CIE Review Panel meet from September 25-27, 2006 at Marine Acoustic Inc.'s offices in Arlington VA. The purpose of the meeting was to review Marine Acoustic Inc.'s Acoustic Integration Model (AIM). The review was initiated by NMFS who required an independent peer review of AIM.

AIM is a software package. Its primary use has been to develop specific application models which were used to predict the average number of marine mammals which would be exposed to sound levels above a given threshold (see Anon. 2006 b & c). Such estimates are currently needed in permit applications for any activity which could adversely impact the sound environment of marine mammals (under the Marine Mammal Protection Act). AIM-models are mechanistic in nature. A virtual environment is created within which sound sources and animals are placed. The specific circumstances are modelled: sound source properties and movements; animal dive and swim patterns; and environmental conditions affecting transmission loss. A complete sound exposure history is constructed for each animal placed within the virtual environment (Frankel et al. 2002).

The three terms of reference required that the Panel evaluate whether AIM correctly implements the models and data upon which it was based; whether animal movements are adequately simulated; and whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development and evaluation.

The meeting was chaired by a CIE appointed moderator who was also responsible for the production of this summary report. The report is the moderator's attempt to summarize the independent views of the panelists. Points of agreement and disagreement are noted. This report is designed to be self-complete, but does not cover all of the material in each panelists report. Readers who require more detail are advised to consult the individual reports. The three panelists are specialists in their own fields and their reports reflect this: Dr Wayne Getz (modeling), Dr Michael Porter (acoustics), and Dr Jeanette Thomas (marine mammals).

REVIEW ACTIVITIES

Meeting Preparation

The main background documents and presentations were distributed to the Panel and moderator in advance of the meeting (Appendix 1). A conference call was held on Friday September 22, 2006 for the parties to discuss the supplied material and meeting logistics. During the call some aspects of AIM were clarified (e.g., its modular nature) and some preliminary model runs were discussed in general terms. The role of the moderator was also clarified.

Meeting Attendance

A brief narrative of the meeting is given below.

25 September

The meeting was convened at 9.00 am and began with a round of introductions. Clayton Spikes, the Chief Operating Officer of Marine Acoustics Inc. (MAI), gave a brief introduction to the review and welcomed participants. Ken Hollingshead, from NMFS, covered, in more detail, the reasons for the review and discussed the terms of reference.

The formal presentations then began and they followed the agenda with the exception that the presentation on example AIM scenarios was not given (see Annex IV in Appendix 2). The panelists considered the scenarios to be too complex to be of use in determining whether AIM was performing properly.

The presentations gave a good over-view of AIM, its history, and the testing that it had been subjected to. The presentation on the CREM guidelines served to introduce a discussion of the CREM guidelines document and the applicability of the guidelines to AIM. The moderator requested that participants, and panelists in particular, consider the nature of AIM with regard to CREM terminology. Was AIM a "model" or a "model framework"? The text under TOR 3 (see Appendix 2) required the reviewers to address various points which were applicable to models, but not to model frameworks.

The moderator also drew the meetings attention to another problem with TOR 3. A requirement of the CREM guidelines was that the "model" had received adequate peer review. There appeared to be some circularity in the request for a peer review, by NMFS, which included a requirement to evaluate whether AIM had had adequate peer review. It was certainly not for the panelists to comment on whether their peer review of AIM was adequate.

The meeting closed for the day at approximately 5 pm.

26 September

The meeting resumed at 9.00 am. The Panel and moderator had discussed the TOR on Monday evening and had agreed to leave them as they were. There was also a brief pre-meeting discussion on the TOR between the moderator, a panelist, and Ken Hollingshead, and the same conclusion was reached. With regard to the nature of AIM, as a "model" or a "model framework", it was agreed to leave this interpretation to each panelist individually.

The bulk of the day was devoted to AIM model runs (with only Adam Frankel, Jack Buchanan, the moderator and the Panel present). The first runs examined animal movement within the model. The objective was to determine how quickly the spatial-distribution of animals exhibited the expected properties of a stationary distribution. Histograms of animal position in latitude, longitude, and depth were visually compared over time. In the first run, done in a deep ocean environment, the spatial distribution quickly appeared to become stationary. In the second run, shoreline was included and two different types of animals were incorporated – inshore and offshore. A problem was seen with the depth distribution of the offshore animals in that they were diving to depths beyond their allowed maximum. This problem-run was set aside while runs exploring the transmission loss models were pursued.

Dr Porter supplied the suggestions for these runs. They were aimed at checking that the Bellhop and PE models gave the expected results under a variety of circumstances – where he knew what to expect. Some of the results were not quite as he expected. He asked for the model output and undertook to investigate the results further that evening. Also, Jack Buchanan undertook to investigate the problem with animals exceeding their maximum dive limits.

The meeting concluded for the day at 6.45 pm.

That evening, the moderator supplied the panelists with suggested headings and sub-headings for their reports. The objective was to encourage some uniformity of structure to enable the three reports to be more easily summarized. The moderator emphasized that the headings were only suggestions and that the panelists could use or modify them as they wished. As agreed earlier, the decision on the nature of AIM, as a model framework, or a model, was left to individual panelists to determine.

27 September

The meeting resumed at 9.00 am with an update on the previous days model runs and the subsequent investigations. The bug which had allowed the animals to exceed their maximum dive depths had been found and corrected by Jack Buchanan. Dr Porter had satisfied himself that the transmission loss models were giving the expected results except for one case, where he had determined that the Bellhop output file was being mis-interpreted. The magnitude of the transmission loss was being determined from only the real part of the complex representation (i.e., only the first number was being used – whereas, both numbers were required to compute the magnitude). When this error was corrected the results were satisfactory.

Further model runs were conducted by the sub-group. The animal only runs were checked with the corrected code. The depth distributions were as expected. In the second run the spatial distribution took some time to become stationary. This lead to the conclusion that a “burn-in” period was required for AIM model runs so that the results did not depend on the initial (arbitrary) distribution of the animals. A problem was also detected with the individual movements of some inshore animals. One or two of them were seen to be “trapped” near their initial location. This problem was fixed when the prescribed reflection behaviour was modified. It was suggested that some standard diagnostics be routinely produced and displayed to alert users to such problems.

Near the end of the meeting there was discussion on the absence of variability in the sound velocity profiles. AIM uses a database that provides 30-year monthly averages. In reality, the sound velocity profiles are constantly changing. It was suggested that ducts could periodically occur in some environments, but these may be completely absent from the averaged environment. This was not to say that ducts did not occur in some averaged environments – indeed the discussion arose because of an example AIM run where a surface duct was present.

The meeting closed at 2.45 pm.

Post Meeting Activities

Before the meeting had closed there had been three post-meeting activities agreed to. MAI would supply a "timeline" for the development of AIM, including presentations given and papers written; MAI would supply documentation describing the quality control and project management procedures that had been in place when the code for AIM was written; and Dr Porter would supply specifications for more AIM runs to check on the implementation of the transmission loss models (and MAI would perform and supply the results of such runs).

We also agreed that there would be email exchanges between the panelists and MAI if the panelists needed clarification on any points which arose while they were writing their reports.

There were various email exchanges, which were copied to all panelists, the moderator, and CIE. One further run was performed by MAI, at Dr Porter's request, to check whether the 20logR (BAM), PE, and Bellhop models gave the same results when perfectly absorbing bottom and surface losses were specified. The "timeline" document was supplied (Anon. 2006b). The quality assurance documentation (Anon. 2006d & 2006e) was not received until after the deadline for panelists to submit their reports, but it just confirmed what panelists had been advised of during the meeting.

Before production of the summary report, the moderator requested clarification from the reviewers on some of their conclusions with regard to the terms of reference. The additional information was to ensure that the summary report accurately reflected the panelists opinions.

SUMMARY OF FINDINGS

This section is organized according to the TOR provided in the SOW (Appendix 2).

a. Assess whether the AIM implementation is correct.

The Panel were able to address this question from the presentation of previous test results (Frankel 2006, Shores 2006), the conducting of test runs during the review meeting, and examination of results from a further run after the review meeting.

The existing test runs are described in detail in Frankel (2006), Shores (2006) and Thomas (2006). The results of the acoustic runs conducted during the review showed only minor problems with the implementation of the acoustic models (Porter 2006). It was generally agreed by the Panel that AIM appears to be correctly implemented. However, it was noted, that in the time available for the review, it was not possible to conduct sufficient tests to be absolutely sure that the implementation was correct.

All panelists had recommendations for future testing.

Relevant extracts:

- Each of the components of AIM need to be tested or "exercised" by running a set of simulations to insure that AIM is performing as expected (Thomas 2006).
- Several issues arose in these brief tests that raise questions about whether more careful testing should be done and carefully documented on the use of these models (Porter 2006, re the acoustic models).
- The best way to verify the implementation of AIM is to compare its performance with a software platform such as ESME over a set of exercises that covers the full range for which AIM is designed to provide assessments (Getz 2006).

b. Assess the animal movement simulation within AIM.

The animal movement capabilities of AIM were evaluated through a presentation of existing test runs (Frankel 2006) and by conducting test runs during the review meeting (Thomas 2006).

Within AIM, dive and movement characteristics are species specific. Various parameters can be set to cover a wide variety of dive patterns (allowing species-specific depth distributions). When data are available for a particular species, parameters can be chosen to tune the simulated behaviour to the observations (Frankel 2006). AIM also allows "aversions" to be programmed (e.g., animals may react to a sound-source, or move away from water that is too shallow). "Attractions" – being negative-aversions are also possible. AIM does not allow interaction between animals – so aggregation behavior cannot be modeled explicitly (but it would be possible to assign "pod" sizes to individual animals and post-process AIM simulation results on that basis).

The existing test runs are described in detail in Thomas (2006) and Frankel (2006). The results of the runs conducted during the review showed only minor problems (Thomas 2006). It was generally agreed by the Panel that the animal movement methods used in AIM were appropriate given the level of available data. The qualifier is important here. The Panel did not perceive a problem with AIM's animal movement methods. They do acknowledge a problem with the absence of the type of data needed to realistically simulate animal movement within AIM.

Relevant extracts:

- At this point in time, I believe the reliability of AIM to assess the exposure hazard of marine mammals to anthropogenic sound is more limited by the realism of the animate engine module of AIM than the sound propagation modules … animal behavior is far more complicated than behavior of physical systems (Getz 2006).
- … requires that aggregative social, feeding, or predator avoidance behavior of individuals be taken into account. In the absence of data that allows aversion parameters to be set that would simulate such behavior, plausible scenarios need to be investigated under "what if …?" scenarios that assumed that individuals aggregate for various reasons (Getz 2006).
- … the approach used in AIM based on simple statistical characterizations is a very reasonable approach given the amount of data currently available (Porter 2006).
- … MAI provided a document by Frankel and Vigness-Raposa (2006) that tabulates the best available data on marine mammal behavior that is the basis for user input to AIM …this is an accurate summary of currently available data (Thomas 2006).

c. Assess whether AIM meets the CREM guidelines for model development and evaluation.

It was noted by the Panel that the CREM guidelines were supplied in a November 2003 draft document, and yet "are accorded the status of providing a standard that should be meet" (Getz 2006). This caused some disquiet amongst the Panel but it was acknowledged that the "spirit" of the guidelines represent some "important principles on minimal standards" and raise some "excellent questions" which can be used to discuss the strengths and weaknesses of AIM.

This term of reference presupposes that AIM is a *model* in the CREM terminology. However, in the CREM guidelines it is clear that a *model* is an application model which is derived from a *modeling framework*, when specific parameter values are supplied (Pascual et al. 2003, page 9). Only one reviewer explicitly addressed this issue, but the assumption that AIM is more than an application model is implicit in each panelists report. Indeed, it is argued that AIM is even something more than a modeling framework: "… it is flexible enough to embrace several different sound propagation frameworks, with the best implementation depending on bathymetric specifications. AIM implemented for a specific set of conditions yields a model that can be evaluated using the CREM guidelines…" (Getz 2006).

The five bullet points of this TOR (see Appendix 2) were used to develop appropriate questions/headings given that AIM is not an application model, but a tool for developing application models.

Have the principles of credible science been addressed during the development of AIM?

The Panel agreed that this was generally true. Relevant extracts:

- Generally an excellent job has been done in applying the best science in the development of AIM (Porter 2006).
- The animat, or simulated marine mammals, used by AIM are based on the best available data on dive profiles of free-ranging marine mammals (Thomas 2006).
- The modules used by AIM to calculate the spatiotemporal dynamics of sound intensities and frequencies have a long history of development, evaluation, and testing in the marine acoustics literature and appear to be the best available to date (Getz 2006).

However it was noted that the systems that AIM models are applied to are very complex, relative to the current "state of knowledge". From Porter (2006), "… it is important to realize that the best science is woefully inadequate here. It is widely recognized that the core weakness in assessing impacts of sound is the lack of knowledge of marine mammal populations. This will improve in the coming decades but policy makers must be aware of the great uncertainty in this area."

The questions/headings (below) address individual aspects of "credible science".

Has there been adequate peer review of AIM?

To a certain extent this question is answered by the existence of the CIE review since NMFS required an "independent peer review of AIM" (see SOW, Appendix 2). However, the Panel showed a divergence of opinion on this question. One panelist was satisfied that the existing published papers, conference presentations, and contributions to environmental impact statements constituted adequate peer review (Thomas 2006). It was noted that the "adequacy" of the review depended upon the "importance of the resulting policy decisions" (Porter 2006). Also, it was suggested (with good reason) that this CIE review "appears to be the first independent peer-review of AIM" (Getz 2006).

Thomas (2006) argued that the cumulative activities surrounding AIM since its inception, constituted adequate peer review. Conference presentations were made: "Presentations of components of the AIM model were given at well respected scientific meetings, such as the Acoustical Society of America and the International Whaling Commission." Papers were published: "Aspects and background information used in the development of AIM were published in respected peer-reviewed journals, such as *Animal Conservation, Reports of the International Whaling Commission, Journal of the Acoustical Society of*

America, and *IEEE Oceans*". AIM was also reviewed in applications: "AIM was used to model received levels for marine mammal exposures to noise in the Environmental Impact Statement for the Navy's SURTASS-Low Frequency Sonar program. EIS are highly reviewed by both specialists employed by the applicant and by the Office of Protected Resources, Permit Office, of the National Marine Fisheries Service. EIS procedures include publication in the Federal Register and a period for public comment. The EIS review had no negative comments that would indicate problems with the application of AIM to accurately determine received levels by marine mammals."

Getz (2006) commented on the current review, noting that "this review is much more thorough than any review that publications arising out of AIM would get if they were submitted to the peer-reviewed literature". However, he also noted that "the best way to review AIM is to compare its performance against EMSE Workbench: when output differs by more than a percent or two then both platforms need to be checked for likely sources of error…".

Does the conceptual model have an adequate theoretical basis?

There was general agreement that the best available data and models have been incorporated into AIM. However, it was again noted, that the "best available" is not very good in a number of areas.

Relevant extracts:

- AIM is generally using the best available knowledge; however, the limitations of that knowledge are widely recognized (Porter 2006)
- Environmental databases, especially for the geoacoustic properties of the bottom are recognized as weak links in the modeling process (Porter 2006)
- The behavior of real animals … are, for the most part, not well enough empirically established to have a theoretical basis for implementation (Getz 2006)
- … by compartmentalizing the task into the sound source, propagation path, and receiver characteristics, AIM developers have been able to employ the best theories, available data, best models, … (Thomas 2006)

Has the conceptual model been adequately implemented?

The Panel agreed that this was generally true. Limitations in implementing the conceptual model were largely data driven.

Relevant extracts:

- … discussed many ways that AIM could be adapted … to make the animat behavior more real world … These adaptations are largely limited by the lack of detailed behavioral data on free-ranging marine mammals (Thomas 2006).

- As with all implementations, some problems often arise in setting up parameters, defining aversions correctly, and output should never be accepted without verification that it conforms to expectations (Getz 2006).
- The space is defined in terms of site specific bathymetry data with temperature and salinity profile specified by 30-year monthly averages … Conceptually, this does not capture the full range of monthly variation … in averaging, the user misses the full range of variation, which would certainly affect computations designed to calculate events occurring in tails of distributions (Getz 2006).

Were adequate techniques and procedures used for code verification?

The Panel gave a qualified “yes” in response to this question. The Panel requested a written summary of quality assurance and project management procedures which were in place during the software development. This was not received until after the panelists reports had been written. However, the moderator confirms that the documentation was in agreement with a verbal description which was given at the review meeting.

As one panelist aptly described, “AIM … has followed a typical development of scientific software in which features are added as needed and with restricted funding” (Porter 2006). The question of how much verification is necessary was raised: “Obviously, the code would not meet NASA standards for rocket or module control or nuclear plant standards for process control; but the cost of checking all code at the highest level is not justified in this case” (Getz 2006). As noted, under TOR 1, the panelists each had recommendations for further testing.

Relevant extracts:

- There was limited information to answer this question. The Review Panel could not and did not verify programming code. At the Review Meeting, Jack Bucanan of MAI provided an oral report of the chronology of the steps taken to verify code during AIM development. This presentation indicated a logical method of code development with in-house checks (Thomas 2006).
- It is not surprising that interactive demos of a complex piece of scientific software such as AIM revealed flaws. However, the particular problems that showed up in our testing suggest to me that greater effort should have been spent on code verification. The author does recognize that limits in this area are probably largely attributable to limits in the funding (Porter 2006).
- For its purpose, code verification appears to be followed at the level expected. Fortunately, however, a similar, independently developed software platform exists–the ESME workbench mentioned above. These two platforms should be tested against each other. This is the cheapest, quickest, and, in fact, best way to verify at this point whether or not the procedures and code in AIM are performing as expected (Getz 2006).

AIM model applications

The Panel had little opportunity to evaluate AIM's performance in particular studies. Much of the CREM guidelines, and the questions posed under TOR 3 in the SOW (Appendix 2), can only be answered for specific application models. However, much of what had been done in specific applications was discussed in general terms during the meeting. Also, much of what was done can be deduced from a knowledge of AIM, given that it was used to derive the application models. The following comments therefore broadly apply to existing AIM applications while not applying to any specific application.

Was the choice of model appropriate given the quantity and quality of available data?

The collective comments of the Panel indicate that an AIM application model is an appropriate choice for predicting the acoustic exposure of marine mammals. However, it is up to the user of AIM to make appropriate choices for a given application. AIM models can be relatively simple (e.g., a single generic species, with a "20logR" sound source) or very complex (multiple species, multiple sound sources using PE or Bellhop transmission loss models, with aversions and attractions specified). The level of complexity needs to be tailored to the available data.

How closely did the model simulate the system of interest?

As one panelist noted, "this is hard to say" (Getz 2006). Another agreed, "validating the end-to-end process is extremely difficult" (Porter 2006). The knowledge of marine mammals was identified as the weakest component: "… the animal animats are unlikely to behave anything like the real systems because ecological and sociological components of the behavior of individuals are ignored (Getz 2006). However, it was acknowledged that within AIM, "intelligent decisions have been made throughout and generally represent the current best knowledge" (Porter 2006).

Were adequate sensitivity and uncertainty analyses performed?

Two panelists answered this question with regard to AIM-model applications. From the modeling specialist, familiar with statistical modeling in broad biological applications came a succinct answer: "No" (Getz 2006). The acoustic expert noted, with regard to providing error bounds on estimates, "that the acoustic modeling community has historically not graduated to this stage either, partly because of the computational cost of doing such analyses. In fact, this motivated a recent ONR effort on 'Capturing uncertainty'" (Porter 2006).

The CREM guidelines go into quite some detail on the types of uncertainty and sensitivity analyses which are required for an application model to meet the guidelines. There are few restrictions within AIM which would prevent suitable analyses being performed. However, it is abundantly clear that these analyses have not been performed in applications to date.

Relevant extracts:

- AIM is generally using the best standard practices. However, for important policy decisions, AIM should be run with several best/worse case scenarios to provide a rough sense of the uncertainty in the results (Porter 2006).
- The implementations were appropriate, but do not appear to have gone far enough to evaluate the sensitivity of results to parameter uncertainties, were incomplete with regard to finding ranges that bound the results, and did not attempt to find out more about the tails of distributions and the probability of rare events (Getz 2006).
- … this is not necessarily inappropriate to generate a baseline analysis, providing this analysis is augmented by scenarios where assumptions regarding behaviour that lead to highly aggregated population distributions are also investigated in comparative "what if …?" types of analyses (Getz 2006).
- The application of AIM to particular problems, and its use in the EIS process has not been particularly innovative and has not exploited the potential utility of the AIM software platform (Getz 2006).

RECOMMENDATIONS

The Panel's main recommendations are given below.

a. Further testing and validation

- The best way to verify the implementation of AIM is to compare its performance with a software platform such as ESME over a set of exercises that covers the full range for which AIM is designed to provide assessments (Getz 2006).
- If the model is to be used for particularly important policy decisions, an investment is needed to benchmark the code and provide transparent documentation that demonstrates better benchmarking has been done (Porter 2006).
- Each of the components of AIM need to be tested or "exercised" by running a set of simulations to insure that AIM is performing as expected (Thomas 2006).
- The animat movement simulations of AIM should be thoroughly tested to insure that each of the input variables related to animat movement are biologically reasonable (Thomas 2006).

b. Suggested enhancements

- Getz (2006): the utility of AIM should be increased in terms of facilitating,

    - analyses based on the construction of full distributions of possible events and the extraction of probabilities associated with the tails of these distributions rather than the current focus on averages;
    - sensitivity and uncertainty analyses;
    - comparative analyses and addressing conditional (what if?) questions.
- Porter (2006):
    - Modify BAM to include surface and bottom bounce paths in the near field;
    - Use the RAM option in NPSE to provide wide-angle capability
    - Use the coherent option as the default in BELLHOP.
- Thomas (2006):
    - Could AIM incorporate a user setting of existing ambient noise levels over a certain frequency range that would make the prediction of received levels at a given range more accurate?
    - Could AIM develop a way that the user could input species-specific underwater vocalization rates, frequencies, and amplitudes typical of the time and season of users operations?
    - Could the surface time be used to imitate haul out periods in pinnipeds?
    - Could AIM incorporate a "generic deep scatter layer" into the sound propagation models that would allow the user to input the size, density and depth of the DSL, as well as the source level of the DSL?
    - AIM could incorporate "daytime animats" and "night time animats".

c. Application of AIM

- Use a burn-in period to stabilize animat behaviour (Porter 2006).
- Porter (2006): in cases where the highest quality predictions are important,
    - do calculations at multiple frequencies;
    - do sensitivity studies to provide error bars on model predictions.
- Getz (2006):
    - an output file of critical assumptions should be generated with each simulation;
    - critical output should be reported minimally in terms of both averages and confidence intervals;
    - the sensitivity of key measure, such as Level A and B takes, to the most uncertain or speculative parameters in the model should be reported;
    - results from simulations that take the full spectrum of uncertainty into account (e.g., moving beyond analyses based on 30-year monthly averages to include inter and intra-day variation, allowing for the aggregation of pods of individuals) should be implemented whenever feasible to help assess the probabilities of rare events that become increasingly likely over large spatial and temporal scales of operation.

REFERENCES

(see Appendix 1 for further references)

Getz, W. 2006: Review of Marine Acoustic Inc.'s (MIA) Acoustic Integration Model (AIM). *For* University of Miami Independent System for Peer Review. xx p.

Porter, M. 2006: Review of Acoustic Integration Model (AIM) 25-27 September, 2006 Washington, D.C. *For* University of Miami Independent System for Peer Review. xx p.

Thomas, J.A. 2006: Review of the Acoustic Integration Model (AIM) for determining noise exposure levels for marine mammals. *For* University of Miami Independent System for Peer Review. xx p.

APPENDIX 1: MATERIAL PROVIDED

In addition to the documents, extracts, and PowerPoint presentations which were distributed in electronic and/or hardcopy, the results of some AIM model runs were also distributed in the form of "screen grabs" (captured in a Word document) and .png files.

Documents

Anon. 2006 a: LFA Observation techniques notes. Marine Acoustics, Inc. 1 p.

Anon. 2006 b: AIM timeline. Marine Acoustics, Inc. 4 p.

Anon. 2006 c: Annual Report No. 4: Operation of the Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar onboard the R/V *Cory Chouest* and USNS IMPECCABLE (T-AGOS 23) under the National Marine Fisheries Service Letters of Authorization of 12 August 2005. Maritime Surveillance Systems. 67 p.

Anon. 2006 d: AIM bug fixes. Marine Acoustics, Inc. 14 p.

Anon. 2006 e: Software development process. Marine Acoustics, Inc. 2 p.

Pascual, P. et al. 2003: Draft guidance on the development, evaluation, and application of regulatory environmental models. The Council for Regulatory Environmental Modeling. Draft, November 2003. 60 p.

Frankel, A.S. 2006: Acoustic Integration Model (AIM) Internal Review Document. Marine Acoustics Inc., proprietary document (draft, 14 September 2006). 50 p.

Frankel, A.S. & Buchanan, J.M. 2006: Acoustic Integration Model© (AIM) Users Manual. Marine Acoustics, Inc. (draft, April 2006). 58 p.

Frankel, A.S., Ellison, W.T., and Buchanan, J. 2002. Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts. *Oceans 2002 MTS/IEEE 3*: 1438-1443.

Frankel, A.S & Vigness-Raposa, K. 2006: Marine Animal Behavioral Analysis. Marine Acoustics Inc., Tech. Memo. 63 p.

Marine Acoustics, Inc., Patent Application, 29 Oct 04. Method for modeling the effect of a stimulus on an environment. Marine Acoustics, Inc. proprietary document. 42 p.

Document extracts

Anon. 2006 f: Summary of SURTASS LFA sonar operations for 16 August 2002 to 15 August 2006. 2 p.

Lecky, J.H. 2006: Taking and importing marine mammals; taking marine mammals incidental to the U.S. Navy operations of Surveillance Towed Array Sensor System Low Frequency Active Sonar. Extract from Federal Register. 4 p.

Various authors and dates: Excerpts related to AIM from responses to comments in major regulatory publications. 16 p.

PowerPoint presentations

Frankel, A.S., & Ellison, W.T. 2004: AIM modeling of oil industry activities to derive marine mammal take estimates. IMEMS 2004 conference.

Frankel, A.S., & Ellison, W.T. 2006: The Acoustic integration Model (AIM): application issues for risk assessment. Navy environmental planning and natural resources workshop; risk assessment technologies, Virginia Beach, VA, May 2006.

Messegee, J. 2006: History of marine mammal exposure. AIM CIE review 25-27 September 2006.

Shores, L. 2006: Examination of AIM-generated PE files. AIM CIE review 25-27 September 2006.

APPENDIX 2: STATEMENT OF WORK

Consulting Agreement between the University of Miami and Patrick Cordue

Review of Acoustic Integration Model © (AIM)

The National Marine Fisheries Service (NMFS) requires an independent peer review of the Acoustic Integration Model© (AIM), which shall assess whether AIM correctly implements the models and data upon which it is based, whether animal movements are adequately simulated within AIM, and whether AIM meets the Council for Regulatory Monitoring guidelines for models, which primarily involve scientific credibility.

Background

Minimizing and mitigating the potential effect of sound upon the environment is an increasing concern for many activities. Naval operations, seismic exploration, vessel and aircraft operations, certain construction activities, and scientific investigations now need to consider the potential effects underwater acoustic sources have on marine life. Marine mammals are usually the primary concern, due to their widespread distribution and excellent hearing ability, although impacts on fish are increasingly being considered as well. Predicting the exposure of marine mammals is complicated by their diving behavior and, in some cases, long-range migrations, which causes them to “sample” many depth strata within the water column.

Acoustic propagation and sound received levels are a function of water depth, range from the source, and a host of sound source and environmental variables. This, combined with the variable diving behavior of different species, makes for a very complex problem. The Acoustic Integration Model© (AIM) addresses these specific complications. A principal component of AIM is a movement simulation engine. Both sound sources and animals, collectively addressed as “animats,” are programmed to move in location and depth over time in a realistic fashion. Animal movement is based on documented regional and seasonal behavioral data for each species evaluated. Acoustic sources and receivers are programmed to move through a virtual acoustic environment, based on external environmental databases and radiated sound fields created from a choice of several propagation models (*e.g.,* Parabolic Equation [PE], Bellhop, etc.). The integration component of the AIM engine then predicts the exposure level of each simulated animal at successive operator-selected time steps. Furthermore, each animal can evaluate its environment at each time step, and can be programmed to alter direction or diving behavior in response to variables, such as sound level or sea depth. AIM allows the user to predict the effects of different operational scenarios and animal responses, thereby allowing the selection of an alternative that produces the least impact and still meets operational requirements.

AIM is a proprietary model owned by Marine Acoustics, Inc (MAI). Its value in predicting the acoustic exposure of animals has been demonstrated in earlier documents. However, the continued use of the model to provide acoustic exposure and impact predictions for regulatory assessment purposes requires that the model be reviewed independently, so that NOAA and other federal agencies can comply with the Data Quality Act.

Reviewer Requirements

The Center for Independent Experts (CIE) shall provide three panelists and a moderator for the review of AIM. Expertise in underwater acoustics, modeling, and marine mammalogy is required. The underwater acoustician should be familiar with propagation-loss models. Ideally, the acoustician will have experience or knowledge of the Bellhop and Navy Standard Parabolic Equation (PE) models, as these are the two main propagation models incorporated into AIM. The modeler should be familiar with individual-based models, preferably those dealing with animal behavior, and the integration of multiple data streams (*e.g.*, multiple databases). The modeler should be able to understand the dynamic interactions of databases. The marine mammalogist should have experience in marine mammal behavior, including diving behavior of more than one species. The moderator should have a reasonable level of scientific and technical understanding, with a reasonable degree of knowledge and experience in at least two of the three scientific categories (underwater acoustics, modeling, marine mammalogy).

The review will be organized around a three-day meeting at MAI during the week of 25 September 2006, to be scheduled based on the availability of all pertinent personnel. The moderator shall be required for a maximum of 12 days for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, and completing the summary report. The three panelists shall be required for a maximum of 12 days each for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, completing their individual reports, and reviewing the draft summary report. Thus, a total of 48 reviewer days is required.

Review Process

The panelists and moderator shall review the documents listed in Annex I prior to attending the three-day meeting at MAI, 4100 N Fairfax Drive, Suite 730, Arlington, VA 22203. After the MAI meeting, the three panelists shall write individual reports addressing all the Terms of Reference given below. The moderator shall consolidate the individual reports into a summary report. Details of the panelists' and moderator's reports are given in Annex II and III, respectively.

<u>CIE – MAI Conference Call – September 21, 2006:</u> The CIE panel and MAI team will discuss via conference call the details of the upcoming meeting, and the CIE panel will raise questions concerning background documents, specifications for trial runs, and other review-related material, including logistics. The call is tentatively set for 5 pm EDT on Thursday, September 21, 2006, but participants and timing may be changed due to individuals' availability.

MAI Meeting - Day 1: Begin with an introduction and a "charge" to the panel, which includes laying out the roles of the participants and the terms of reference (see below), and a time table. AIM presentations will be made by scientists from MAI, with question and answer sessions as needed. Presentations shall include:

1) Introduction to the AIM approach and the software, including data input requirements;

2) Review of results of internal testing of the software;

3) Overview of the process used to derive animal behavior parameters from data and the scientific literature.

Following the conference call and these presentations, the moderator and panelists shall meet to develop a set of test runs for days 2 and 3. The purpose of this evaluation is to see how the AIM responds to a set of inputs that are designed to test the model. It must also be noted that in assessing the functioning of the model, it is important to acknowledge the differences and the roles of the external components (*e.g.*, animal input parameter values, propagation models) and the internal components of the model (*e.g.,* Animat Builder, Movement Simulator, Integration Engine). The distinction is drawn here to emphasize that the values of behavioral parameters can and should change when new data are available, and that AIM can utilize that new data. Valid input data *(e.g.*, animal behavior parameters) are critical for valid predictions from a model run for a specific scenario. However, a valid model can be provided with invalid animal behavior inputs, and still produce accurate outputs. The purpose here is to test the internal components of the model, including how AIM handles the input data. Example scenarios for devising these runs are provided in Annex IV.

MAI Meeting - Days 2-3: Dedicated to the CIE panel working with MAI scientists to perform AIM model runs, so the panel will have sufficient information on the input data, execution parameters, and model outputs for writing their respective review reports. The CIE panel shall, with the assistance of MAI scientists as required, design simulations and request that the MAI scientists create input files to represent these simulations during the course of the review. Projects can be created in a few minutes. Because AIM is a working model (not yet streamlined and simplified for public use), requiring expertise and familiarity with data input procedures and model execution techniques, MAI scientists will perform the model runs under the oversight of the CIE panel. The number and complexity of simulations to be run during the evaluation period will have been discussed in the conference call and finalized on Day 1. To run the models, MAI scientists will require sufficient time to research the values of the basic parameters (i.e., beam pattern information for source, or behavioral parameters for different species). The input files will then be run, and the inputs and outputs will be provided to the CIE panel for their analyses and evaluations.

Terms of Reference

The CIE panel shall complete the following tasks, and document their results in the individual panelist and summary reports.

1. Assess whether the AIM implementation is correct.
2. Assess the animal movement simulation within AIM.
3. Assess whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development.

<u>1. AIM Implementation</u>

Details relevant to the topics described below are given in the Robustness Review Document, which addresses operation of the AIM model.

- Does AIM accurately and efficiently implement the propagation models? Identify any errors in the implementation. The propagation models implemented in AIM include Bellhop (Porter, 1992) and Navy PE (Zingarelli, 1999). These models were created by other individuals and organizations. *The propagation models themselves are not the subject of the CIE review, but rather the implementation of these models in AIM.*

- Does AIM correctly handle the input values to the models? If not, identify any errors. For example, are acoustic source level and frequency values properly transferred through the model components?

- Does AIM correctly and efficiently extract data from databases? If not, identify any errors. AIM uses the GDEM-V (v 2.6) database for sound velocity profiles and the NOAA ETOPO2 database for bathymetry.

<u>2. Simulation of Animal Movement</u>

- Does the ANIMAT movement model in AIM adequately simulate animal behavior? Comment on the strengths and weaknesses of the modeling approach, and suggest possible improvements.

  The review panel shall devise one or more approaches for addressing this issue. One approach that shall be considered is to evaluate, given appropriate input values, how closely the modeled animal movements mimic the known responses of free-ranging animals. *The species-specific values used in the models are not the focus of the review, but rather the ability of the ANIMAT model to simulate movement.*

3. CREM Guidelines

The panel shall assess whether AIM meets the Council for Regulatory Environmental Monitoring (CREM) guidelines for model evaluation, which are summarized below. Some of the points listed below will have been addressed by the reviewers as part of their comments on Terms of Reference 1 and 2 above. Each reviewer shall ensure that clear answers are provided for the CREM guidelines, though extensive repetition of technical comments is not required.

- Have the principles of credible science been addressed during model development?
- Is the choice of model supported given the quantity and quality of available data?
- How closely does the model simulate the system of interest?
- How well does the model perform?
- Is the model capable of being updated with new data as it becomes available?

Schedule of Activities

The schedule of activities, including timelines (all in 2006) and identification of responsible parties, is provided in the following table.

Activity and Responsible Party	**Date**
NMFS provides background documents (Annex I) to moderator, panelists, and CIE	September 11
Moderator and panelists participate in a conference call with MAI to discuss technical and logistical details (depending on availability of participants). This call shall be arranged by the CIE.	September 21
Moderator and panelists read background documents	September 24
Moderator and panelists meet at MAI to test AIM model	3 days during week of September 25
Panelists write draft individual reports (Annex II); moderator begins summary report (Annex III)	October 2-13
Panelists provide draft individual reports to moderator for summarization and to CIE for review	COB October 13
Moderator provides draft summary report to CIE	October 27
CIE approves final individual reports, submits them to the moderator and the NMFS COTR	October 27
NMFS COTR approves individual reports; CIE provides final pdf versions to NMFS COTR and to moderator	November 1
Moderator provides draft summary report to CIE	November 8

CIE approves final summary report, submits it to NMFS COTR	November 22
NMFS COTR approves summary report, CIE provides pdf version to NMFS COTR	November 30

Submission and Acceptance of CIE Reports

The CIE shall provide the final individual and summary reports for review and approval to the NMFS COTR, Dr. Stephen K. Brown via e-mail (Stephen.K.Brown@noaa.gov), according to the schedule above. Approval by the COTR shall be based on compliance with this Statement of Work. The COTR shall notify the CIE via e-mail regarding acceptance of the reports. Following the COTR's approval, the CIE shall provide pdf-formatted copies of the reports to the COTR via e-mail.

ANNEX I: Documents to be reviewed in preparation for the AIM review.

Document Titles
1. IEEE article *Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts*, Adam S. Frankel, William T. Ellison, and Jacquin Buchanan.
2. Presentation given by Dr. Adam S. Frankel at the International Marine Environmental Modeling Seminar (IMEMS) 2004 Conference: *AIM modeling of oil industry activities to derive marine mammal take estimates*.
3. Acoustic Integration Model (AIM) Robustness Review Document
4. Acoustic Integration Model (AIM) Users Manual.
5. The Acoustic Integration Model (AIM): *Applications to predicting and reducing acoustic exposures of marine mammals*, Adam S. Frankel, William T. Ellison. Abstract and presentation for the Navy Environmental Planning and Natural Resources Symposium, May, 2006.
6. Draft Guidance on the Development, Evaluation and Application of Regulatory Environmental Models, The Council for Regulatory Environmental Modeling (CREM)

Document	Document Type	Number of Pages	Degree of difficulty
1. IEEE Article	PDF	6	Low-medium
2. IMEMS Presentation	Powerpoint	28 slides	Low
3. AIM Robustness Review Document	Word/PDF	51	Medium-high
4. AIM User's Manual	Word/PDF	56	Medium
5. EPNR presentation	Powerpoint	33 slides	Medium
6. CREM guidelines	PDF	60	High

ANNEX II: Panelist Report Generation and Procedural Items

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

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

3. The report shall also include as separate appendices the bibliography of materials provided by NMFS for the review and all additional references cited, 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

ANNEX III: Moderator's Summary Report Generation and Process

1. The summary report shall include an overview of the review process.

2. The summary report shall provide a synopsis of the three panelist reports.

3. Points of agreement and disagreement among the panelists shall be documented.

4. The summary report shall also include as separate appendices copies of each of the panelists' reports.

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

ANNEX IV: Example AIM Scenarios. These examples illustrate the types of scenarios that AIM can address. They cover three main sources of anthropogenic noise. The panel shall create their own scenarios to be sure that their questions are addressed and answered. Development of scenarios will be addressed during the conference call and the first day of the meeting.

1. A vessel equipped with a 400 Hz sonar source is operating off the North Carolina coast in May. The sonar signals are broadcast once a minute with a source level of 210 dB re 1 µPa at 1 meter. The ship is to move in a 40 km square sawtooth pattern, with individual north-south legs spaced 5 km apart. The question to be addressed is what is the acoustic exposure of the offshore bottlenose dolphins if the operation goes forward? Alternatively, how would the acoustic exposure be altered if the exercise where conducted in January?
2. A supertanker with a propulsion system that produces a constant 205 dB re 1 µPa at 1 meter with a peak frequency of 10 Hz is transiting off Monterey Bay in June. The ship is paralleling the shoreline at a distance of three miles offshore. Blue whales are feeding within and outside the bay. What is the sound exposure of these whales to the sound of the tanker?
3. A seismic exploration vessel towing an industry standard airgun array is performing a high-resolution survey along the continental shelf south of the mouth of the Mississippi River. Each of the survey lines are 50 km long and spaced 1 km apart. The total survey area is 20 x 50 km. The airgun array is fired every 10 seconds. What is the exposure of sperm whales during the survey?

APPENDIX 3: INDIVIDUAL REVIEW REPORTS

Review of Marine Acoustics Inc.'s (MAI) Acoustic Integration Model (AIM)

Wayne Getz
Dept. Environmental Science, Policy and Management
University of California at Berkeley
CA 94720-3114
getz@nature.berkeley.edu

Executive Summary

1. AIM version 2.0 is a simulation platform developed by Marine Acoustics Inc. (MAI) primarily to evaluate the impact of anthropogenic acoustic signals on marine mammals.
2. AIM includes a sound propagation module that has the flexibility to use various well-established sound-propagation/transmission-loss models (e.g. NSPE v5.0, Bellhop) in conjunction with methodologies to account for the affects of the ocean surface and floor on transmission loss and duct propagation, or can rely entirely on user provided sound-field input.
3. The heart of AIM is an animat engine that moves sound-source animats and individual animal animats around in a virtual space.
4. The virtual space is either a user defined input space or constructed using data from the US Navy's Digital Bathymetric Data Base with Variable resolution (DBDB-V) and their Generalized Digital Data Model with Variable Resolution (GDEM-V) to input temperature and salinity profiles that conform to 30-year monthly averages for the region of interest. It also has the capability of using other data bases such as the National Oceanic and Atmospheric Administration's (NOAA) Earth Topography (ETOPO5) data base that specifies both land and ocean elevations to a resolution of 5 minutes of arc.
5. The sound-source animats are simply a user-specified stationary or moving sound source with particular sound generating profiles (dB levels, frequencies, source durations, and duty cycles).
6. The animal animats are stochastically generated movement trajectories using either user-specified movement parameters (in terms of minimum and maximum times, speeds, and headings for diving and other movement related behaviors) or parameters obtained from a "Marine Mammal Database" (MDB) that is being continuously enlarged by MAI scientists as new information becomes available in the marine mammal scientific literature.
7. AIM, together with at least one other comparable platform known as ESME[1], provides state-of-the-art software tools for calculating, among other things, the

[1] Effects of Sound on the Marine Environment: a software workbench developed by the Naval Research Laboratory.

distribution of the cumulative exposure of individuals belonging to a suite of user specified species of marine mammals, where these individual animals move through the virtual acoustic-space generated by the sound-source animats and modified by the marine bathymetric configuration.

8. AIM has a relatively open software architecture that allows it to be continuously improved. AIM currently provides an effective tool to "convolve" distributions of moving animals within a static or time-varying sound field in a bathymetrically realistic environment.
9. The complexity of gluing together the different component parts of AIM, together with the proprietary nature of the software, imply that the only way to evaluate the correctness of the AIM implementation is through evaluations of the output it produces to a suite of canonical test runs. Beyond this, the existence of the ESME Workbench provides an opportunity for both platforms to be tested against one for accuracy of implementation.
10. As with almost all software platforms and models, in specific applications they are easily misused when unwarranted confidence is placed on the output they generate. Thus it is desirable for software platforms such as AIM to develop components and present output in ways that mitigate misuse and improve their utility. In particular, I recommend that:
    a. an output file of critical assumptions should be generated with each simulation specifying some of the key assumptions specific to that simulation and, as appropriate, these assumptions should be prominently featured when reporting results in EIS documents;
    b. critical output should be reported minimally in terms of both averages and confidence intervals rather than just averages;
    c. the sensitivity of key measures to the most uncertain or speculative parameters in the model, such as Level A and B takes, should be reported;
    d. results from simulations that take the full spectrum of uncertainty into account (e.g. moving beyond analyses based on 30-year monthly averages to include inter and intra-day variation, allowing for the aggregation of pods of individuals) should be implemented whenever feasible to help assess the probabilities of rare events that become increasingly likely over large spatial and temporal scales of operation.
11. In conclusion:
    a. The best way to verify the implementation of AIM is to compare its performance with a software platform such as ESME over a set of exercises that covers the full range for which AIM is designed to provide assessments.
    b. The utility of AIM should be increased in terms of facilitating analyses based on the construction of full distributions of possible events and the extraction of probabilities associated with the tails of these distributions rather then the current focus on averages.
    c. The utility of AIM should be increased in terms of facilitating sensitivity and uncertainty analyses.
    d. The utility of AIM should be increased in terms of facilitating comparative analyses and addressing conditional (what if?) questions.

Structure of this Report

This report is split into two parts. Part I provides the context for addressing the National Marine Fisheries (NMFS) terms of reference (TOR) for the review panel conducting this review. The TOR themselves are addressed in Part II. The reason for this separation is that in my view the TOR cannot effectively be addressed without providing a proper context. In the context of the Council for Regulatory Environmental Modeling (CREM) document referred to in the TOR, the software product being reviewed is both a modeling framework when in software platform guise and a model when in an application guise. Further, its validity as a scientific tool lies not only in the correctness of its software implementation (whether in the component modules, the code gluing these modules together, or the data used to set parameter values in general and in specific applications), but in the way analyses obtained from the model are used to fashion scientific reports. I will attempt to clarify this in the "Platform and Implementation" section of Part I, which will then provide a proper context for my response to the specific TOR headings in Part II.

PART I: Setting the Context

Background

The Center for Independent Experts (CIE), administered by the University of Miami, was charged by the National Marine Fisheries Service to convene a panel of three experts and a moderator to review the current implementation of AIM (Acoustic Integration Model), as developed by Marine Acoustics, Inc. (MAI). The three experts were chosen for their particular expertise in acoustics, modeling, and marine mammalogy respectively. This particular review focuses on modeling issues, covering areas implied by the NMFS specification that the "modeler should be familiar with individual-based models, preferably those dealing with animal behavior, and the integration of multiple data streams (e.g., multiple databases)." See Appendix 2 for more details. Beyond this aspect, other aspects of the charge to the panel are commented on in more depth in the other two expert reports, independently written by the acoustician and marine mammalogist.

AIM v2.*date* is part of a concerted development since 2001 by MAI of AIM version 1.0, which itself grew from 1997 and 2001 as a platform for integrating various component programs used to handle marine environmental and bathymetric data, simulate sound propagation in marine environments and move animal "animats" around in this environment to calculate their exposure to sound. See AIM Timeline (Anon 2006b) for more details. The current version of AIM is regarded by MAI as an appropriate tool for both evaluating the impact of noise on marine mammals and, more generally, for modeling the effect of a stimulus on the environment, as outlined in the MAI "AIM Patent Application" of 29 October, 2004.

AIM is currently used by the US Navy and other organizations contracting MAI to carry out simulation studies required for the development of Environmental Impact Statements (EISs) that need to be submitted to the National Marine Fisheries Service (NMFS). If approved, NMFS then issues Letters of Authorization (LOA) to carry out sound generation activities in marine environments. These activities are required by law to conform to the provisions of Congressional acts such as the National Environmental Policy Act (NEPA), the Marine Mammal Protection Act (MMPA), and the Endangered Species Act (ESA). In particular, NMFS has a US Government mandate to ensure that the thresholds associated with "Level A Takes" (injury and death) and "Level B Takes" (harassment) are not likely to exceed whatever the current standards may be.

NMFS also monitors the activities of the Agencies in question, so if it appears that their activities are exceeding Level A and B Take thresholds, then the activities of the sound generating parties may be shut down for a period of time. Thus both NMFS and other agencies or companies seeking permits rely on AIM or similar tools, such as the Naval Research Labs (NRL) "Effect of sound on the marine environment" (ESME) workbench (Shyu and Hilson, 2006), to provide reliable assessments of the effects of acoustic activity on marine mammals, so that sound generating activities can be safely (for the marine mammals) and optimally/cost-effectively (for the agencies and companies) implemented.

The Specific Charges to the Panel are to:

1. Assess whether the AIM implementation is correct.
2. Assess the animal movement simulation within AIM.
3. Assess whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development.

The CREM guidelines are laid out in a Draft Document, dated November 2003, entitled: Draft Guidance on the Development, Evaluation, and Application of Regulatory Environmental Modeling. For brevity, this document will be referred to herein as the "CREM Guidelines." CREM is an arm of the Environmental Protection Agency (EPA), and the principal authors of the CREM Guidelines are three individuals out of the Office of Scientific Policy, Office of Research and Development of the EPA. Thus the task appears to be well defined, as bounded by the CREM Guidelines and the questions articulated under Specific Charges 1 and 2. The task, however, is complicated by the fact that AIM is neither a modeling framework nor a model, as defined in the CREM Guidelines, but plays both roles. AIM is an application software platform with plug-in components and an associated marine mammal database (MMDB).

The physical process plug-in components are used to generate time-varying fields that provide information on the intensity and frequency of sound waves propagating through the marine environment at node points on a grid that spans the oceanographic environment under consideration. The primary sound-field computation component is a propagation model with auxiliary procedures for determining sound-velocity profiles, taking care of bottom and surface loss reflectance processes and handling bathymetric data. Possible implementations of the propagation model include PE and Bellhop or combinations of these models together with simpler procedures for calculating the sound energy field. AIM is sufficiently flexible to accommodate other plug-in modules for determining this sound field. In this sense, these plug-in modules are external to the central AIM module that is the focus of this "modeling expert" review.

The central AIM component is an animal and sound-source animat engine: software generated animats (elements, objects) that are simulated to move around in computer-generated virtual space and either behave as noise point sources or exhibit the movement (swimming, diving, hauling onto land) behavior of various marine mammals that are being modeled. Each animat has associated with it during the simulation a time series record of exposures to noise at each time step of the AIM simulation.

The appropriateness of the different sound-field-generating plug-in modules is dealt with in depth by the acoustic specialist member of the review panel and will not be evaluated in depth here. AIM also allows for inputs that set the scene with regard to seasonal factors that influence ocean conditions and behavior of marine mammals. Currently, the oceanographic information that feeds into the models is monthly averages for the last 30 years.

In the sense that the AIM software platform has a versatile animal and sound-source animat engine that handles input from user-selected plug-in physical process modules, and that AIM can be applied to any segment of ocean (excluding at this time areas under ice because of the lack of an ice absorption/reflectance module) AIM is a modeling framework, as defined in the CREM Guidelines. However, as soon as it is used as a tool and applied to a particular problem with area and season specific data, and a particular selection and distribution of marine mammals, AIM is a model.

The marine data base has been compiled since 2000/2001 as a text document in a format that matches the input required to define the movement behavior of the animat representing individuals belonging to the species in question, extensively referenced to the current literature used to obtain the parameter values entered into the data base. Incorporation of data into the model, however, has not been automated which may be an advantage in that the user is required to look at the data and, where necessary, change parameter values to reflect the latest data or special constraints.

Platform and Implementation

A scientifically relevant and valid analysis of a complex environmental problem, such as the impact of anthropogenic sounds on marine mammals in natural environments, has three hierarchical components that all need to conform to certain "best practice" standards before the analysis can be regarded as appropriate for making decisions relating to the protection of human and animal health, safeguarding the natural environment, or using natural resources in a sustainable (or at least, a wise) manner. The three hierarchical levels that need to be assessed are:

1. The veracity and suitability of the software platform used to carry out the analysis.
2. The correctness and appropriateness of the implementation of the software

analysis for the system at hand.

3. The pertinence and comprehensiveness of the analysis, particularly as it pertains to addressing the management questions of concern.

In going up this chain from 3 to 1, each succeeding level depends critically on the level that precedes it. However, in going down the chain from 1 to 3, each level stands on its own in terms of its qualities—i.e. we have necessity but not sufficiency in a high quality of performance as we progress down the chain. Also, an additional link in the quality of the analyses undertaken at the third level back to the first level does exist in terms of the software platform being set up to facilitate good practice at the third level, such as running scenario, sensitivity, and uncertainty analyses, the latter two of which are discussed in quite some detail in the CREM Guidelines.

At the software platform level, one needs to be concerned about the veracity of the computer code and suitability of the model framework. In terms of code veracity one needs to answer the question: Does the code actually do what it was designed to do? This is very difficult to ascertain if the code is not checked directly or, when checked, if the code is extensive. Confidence can only be built up by comparing output produced by the code over a range of test cases that collectively exercise all parts of the code and for which the correct outcome is already known. The best-case situation for verification and validation of code is to have other platforms that can be used to compare output in test cases. The suitability issue pertains to the concordance of the theoretical underpinnings of the software design and equations and rules from which the code is generated and the class of questions that can be addressed. Thus, for example, a software platform that has been set up to model the growth of stands of coniferous trees may not be suitable for modeling the growth of stands of deciduous trees. As a link to Part II, it will become apparent that AIM is currently at the leading edge of software tools for estimating the impacts of anthropogenic sound on marine mammals, but greater confidence in its performance will come by comparing the output it generates to that generated by the ESME Workbench.

At the model implementation level, the first issue is whether specific modules, routines, utilities that are part of the software platform have been correctly selected or called, as well as if they have been strung together correctly so that the software platform carries out the intended computations and simulations. The appropriateness question pertains to the concordance of theoretical underpinnings of the software design and code and the questions being addressed in the specific application. For example, if it is important to know what the impacts of sound might be on a species that has a strong fission-fusion social structure (e.g. animals aggregate in large groups during particular periods of time and break up into smaller groups at other periods of time) and the platform only allows for animals (or groups of animals) to be represented by individual, independently moving animats, then the platform is limited in its ability to assess the impacts of sound on this species.

Finally, even if the software package is powerful and suitable, and the implementation is appropriate and correct, the simulations that are undertaken may not be very informative for the question at hand. To what degree should the software platform be designed to both facilitate and encourage selection of simulation runs that promote the utility of the output in its application to addressing ranges of questions? Thus one may want to carry out sensitivity, uncertainty, and scenario analyses. One may want to ask questions that relate not only to central tendencies of distributions produced in simulations, but also something about the tails of these distributions. These aspects of software evaluation lie beyond the issues of veracity of the code and framework or correctness of an implementation of the platform with regard to modeling specific systems. They relate to the utility of the code in facilitating certain types of studies. In many cases, as with AIM, software platforms themselves arise out of a need to efficiently glue together physical and biological modules in a way that facilitates running many different case studies accurately and repeatedly. The platform should also provide utilities to store and link inputs and outputs for further analyses to be carried out at a future date. Thus, extending the platform to facilitate more comprehensive types of analysis at the implementation level is a step that now should to be taken to improve the utility of AIM as a self-contained platform.

Central Tendencies and Tails of Distributions

An important dichotomy in the way one endeavors to understand a process is in terms of central tendencies versus extreme events. Both provide important, but rather different types of information. Thus they are not mutually exclusive but rather complementary ways of viewing the process of interest. In the context of harvesting healthy, robust populations, for example, a deterministic modeling framework may be regarded as adequate for addressing the question of what are the take levels that maximize yield. These are the kinds of models that were used in the 1970's to find maximum sustainable harvesting levels in large fisheries. Deterministic models, however, are not suitable for answering questions relating to the extinction of threatened species. Further, even stochastic models that focus on averages rather than rare events are not going to effectively estimate relevant quantities such as the period of time for which one can be at least 95% confident that a particular population is not going to go extinct. One may also want to know how the length of this period is influenced by certain mitigating interventions that are being planned. The latter requires being able to calculate the probabilities of events occurring in tails of distributions.

The need to efficiently and correctly calculate information that relates to events in tails of distributions requires setting up an appropriate model framework from the start and developing a software platform that facilitates this type of analysis. As will become clear in the next section, AIM is computationally oriented to calculating averages rather than events in the tails of distributions. Consequently, under circumstances where AIM correctly predicts the average rates of Level A and B takes, it will overestimate the number of times these takes involve a single animal and concomitantly underestimate the number of times these takes will involve several animals. Marine mammal A level takes involving several have been reported to occur on a number of occasions (Balcomb, K. C. III and D. E. Claridge. 2001), as well as B level takes involving more than 100 individuals (Southhall et al. 2006). In terms of the press and the media, rare events involving many animals are more likely to grab headlines and public interest than if the same events occurred more frequently but only involving single or a couple of animals. Hence the imperative is to accurately characterize the distributions of events associated with Level A and B takes themselves, most easily in the form of histograms, rather than just the central tendencies and average rates of these distributions.

PART II: Addressing the Terms of Reference

Review Activities

Documents were provided to reviewers several days prior to the review panel meeting at MAI on September 25-27, 2007. Besides Panel Members and MAI staff, Ken Hollingshead of NMFS (Silver Spring, MD) also attended this meeting. After the meeting, the panelists went back to their respective institutes to work on their reports. More information on these activities and measures taken to ensure the independence of the reports of the three Panelists are dealt with in the Moderator's Summary Report.

Summary of Findings

TOR 1: Assess whether the AIM implementation is correct

Implementation of propagation models

Ocean acoustic propagation models have been around for several decades, and current implementations of the Parabolic Equation (e.g. the version maintained by the Naval Research Laboratory) and the Gaussian beam tracing approach (e.g. Bellhop) have been thoroughly worked over and evaluated in terms of bathymetric and other environmental conditions under which they perform well or poorly. Further, surface and ocean floor reflectance properties depend on surface winds, floor sediments and sound frequencies. Given all these complexities, a software platform such as AIM relies on a sophisticated user (someone with a strong background in sound propagation in marine environments) to ensure that AIM is appropriately implemented with regard to ocean acoustics propagation computations. This requirement could be weakened with the addition of an expert system front end that would query the user on the points regarded by experts as relevant to having AIM produce the best possible implementation of its various modules for the problem under consideration. At this time, a suitable Ocean Acoustic Propagation Expert System front end is not available, thereby constraining the use of AIM to individuals with an appropriate level of expertise. Not being an expert myself in this area, I cannot articulate what this level should be or provide insightful comments on how well the AIM implementation uses the different marine acoustic propagation models and with boundary condition (surface and ocean floor) modifications needed to capture reflectance affects. In our examination of AIM-generated acoustic field output and the ensuing discussion, however, I was struck by how easily a novice could make important mistakes if someone with the appropriate expertise does not examine the AIM acoustic field output carefully.

Since our panel included such expertise in the person of Dr. Michael Porter, I will leave it to him to provide a much more insightful review of the AIM implementation of the acoustic propagation component of its calculations.

Implementation of conceptual model

The conceptual model consists of the AIM module that creates a virtual space in which the sound and animal animats are moved according to input specifications for the sound animats and numerically generated rule-based animal animat trajectories (i.e. the trajectories are not generated by solving a set of equations, but by implementing a set of rules). This will be discussed further under TOR 2. The space is defined in terms of site-specific bathymetry data with temperature and salinity profiles specified by 30-year monthly averages for the month in question. Thus, as time ticks by for any run of the model, temperature and salinity conditions will discretely switch from one profile to another as the time clock moves from one month to the next. Since the time clock is generally run in half minute time steps, the salinity and temperature conditions will switch from one set to another after 86,400 time steps (for a 30 day month). Conceptually, this does not capture the full range of monthly variation. Thus the issue arises as to how best to deal with this. If only monthly data are available, then the user is forced either to follow this implementation (i.e. change conditions one a month) or the user can smooth out this change by using the monthly averages as pivots with daily interpolation between pivots. This latter approach is likely to be overkill, but what is troublesome with the monthly average approach is that, in averaging, the user misses the full range of variation, which would certainly affect computations designed to calculate events occurring in tails of distributions. This is such an important point that I have discussed under its own heading in Part I and will discuss more fully under TOR 2.

Although the review panel was presented with comprehensive verbal descriptions of the structure of AIM during our three-day meeting with MAI scientists, and several documents contained further descriptive information on the structure of AIM, it remains essentially a black box (at least to this reviewer). This can only be remedied by being able to look at the code or to subject AIM to a rather comprehensive set of exercises that would take weeks to accomplish rather than days. The best way to evaluate a black box is to have two block boxes independently developed under the same conceptual design and then to see if they both produce the same or very similar output. If they do, then one has evidence that they both are essentially correct in terms of the conceptual implementation, although they could both be seriously deficient in terms of using the same, inappropriate external models, such as the same acoustic propagation module that may be inappropriate for the problem at hand. If they are different, then tests on where they differ are diagnostic of points in the architecture where at least one of the software platforms is malfunctioning. These points can then be checked in both platforms to see which one is faulty.

Fortunately, another software platform exits that is designed to answer the same set of questions motivating the development of AIM. This platform mentioned in Part I of this report and referred to in the literature as the ESME (Effects of Sound on the Marine Environment) Workbench, is being developed by NRL (Naval Research Laboratory) scientists, and is sufficiently advanced to undertake the type of simulations that AIM has been used for in preparation of EISs submitted to NMFS over the past few years (Skyu and Hillson, 2006). Thus, at this point in time, a comprehensive, comparative analysis of output from both platforms over a range of problems that span a variety of situations expected to arise in practice is the swiftest and surest way to uncovering any implementation errors that may exist in either platform. I highly recommend that such comparisons be undertaken as the best way to address the question of whether or not the conceptual component of AIM has been correctly implemented in the software platform.

Data extraction

As with the conceptual implementation, the best way to test whether or not AIM is using input data from the various data sources it relies upon to construct the virtual space as a replica of a particular real location (both with respect to bathymetry and to marine ocean conditions for place and season being considered), run the propagation model, run the sound and animal animats correctly, extracting the correct species data from the marine mammal data base for the latter, and deliver the output correctly, is to compare the output generated by the AIM and ESME platforms.

Without the benefit of this comparison, at a superficial level AIM appears to run correctly, although again an expert is needed to look over the data to see whether it is reasonable. Beyond the issue of whether or not AIM correctly produces the output in terms of the veracity of its implementation based on a rather large number of simplifying assumptions, is the issue of how useful a tool it is to carry out various kinds of analyses (uncertainty, sensitivity, scenario comparisons). As discussed more fully below, I think that AIM can be extended to become a more useful tool in informing managers of the potential impacts of sound on marine mammals in ocean environments.

TOR 2: Assess the animal movement simulation within AIM

At MAI/panel meeting, MAI scientists presented the results of their "internal robustness testing" of the AIM animat engine by providing illustrations of animat trajectories generated under a variety of canonical situations. For example, they checked to ensure that animats behaved as expected when turning and course variation parameters were selected to produce known output results. They also checked to see whether or not animats conformed to appropriate depth distributions and behaved sensibly with respect to reflecting boundaries (e.g. ocean floor, shallow areas, imaginary boundaries in mid-ocean). They tested to see whether aversion components of the animats worked correctly with respect to physical (propensities to move towards or away from sound, shorelines, ocean floors etc.) and biological objects (group or schooling behavior, etc.). They also generated diving patterns to make sure that the trajectories looked similar to empirically measured trajectories used to extract the data (e.g. minimum and maximum lengths of dives, ranges of dive depths, etc.) Visual inspection of output was used to evaluate whether or not animat trajectories conformed to expectations or looked sufficiently similar to trajectories plotted from empirical data.

At this point in time, I believe the reliability of AIM to assess the exposure hazard of marine mammals to anthropogenic sound is more limited by the realism of the animat engine module of AIM than the sound propagation modules. This conclusion is based on the following: 1.) the sound propagation simulators used by the AIM sound propagation module have been extensively used over several decades and the output has been tested by placing microphones at various points in a real environment to obtain empirical data that can be compared to predicted results (e.g. McDonald et al., 1999; Hursky et al., 2004); 2.) animal behavior is far more complicated than behavior of physical systems.

Enlarging on this second point, the animat engine is designed to generate trajectories of biological organisms that cannot be precisely simulated because trajectories reflect individual behavior that exhibits both intra-specific and, even more so, inter-specific variation. Thus all that can be generated is a trajectory that belongs to a class of trajectories associated with individuals in the population under consideration. Each trajectory is a sequence of points providing position in space (three coordinates) at close enough intervals of time so that positions between points can be reconstructed using simple interpolations (lines or splines of desired order). The longer the trajectory, the more points in the time series defining the trajectory and hence the more information needed to characterize the trajectory. Thus the question arises: can a few essential features of the trajectory be extracted in way that allow typical trajectories to be reconstructed from these features, and are these features sufficient with regard to the problem at hand—which in this case is moving animats around a sound field in a realistic way—so that the sound exposure hazard can be realistically estimated?

Among that features that AIM uses to define trajectories are population specific values (different populations of the same species in different regions may have somewhat different parameters) for different classes of "dive type" as well as aversion/attraction behaviors. Each dive type has a probability of occurrence attached to it. For example, individuals in population *X*, may exhibit 3 dive types that, at a specified time of day, are likely to occur with probabilities p_1, p_2 and p_3. Each dive type has its own set of parameters, viz.: Top Depth, Bottom Depth, Least Time (for a dive), Greatest Time (for a dive), Dive Down Angle, Dive Up Angle, Heading Variance (maximum turn rate in degrees per Turn Time), Variance Turn Time (see previous parameter), Bottom Speed (lowest speed of movement), Top Speed (highest speed of movement), Speed-Alpha and Speed-Beta (handles how to distribute actual speed between Top Speed and Bottom Speed), Weighting (how often a particular dive occurs), Initial Heading (initial course, default random). At the end of each dive, individuals return to the surface to breathe before making another dive. Aversions are specified in terms of rules that relate an animates response to one for more of the following stimuli: Land (proximity), Sonar Signal, Sonic Boom, Sound Received Level, Latitude, Longitude, Sea Depth. Parameters used to describe the aversion behavior include: Data Type (see stimulus types listed above), Value (along with a response direction threshold of "greater than" or "less than"), Reaction Angle (change of course measure), Delta Value (behavior switches off when the threshold value has changed in the appropriate direction by this amount), Delta Seconds (length that each behavioral response lasts).

The marine mammal data base contains sets of parameters of the type listed above that have been extracted from the literature for a host of marine mammal species, and currently includes the following categories, some of which cover several species when individual species parameters are not known: Fin Whale, Gray Whale, Sei/Bryde's Whale, Blue Whale, Minke Whale, Humpback Whale (Migrating, Feeding, Winter Grounds: Singer, Calf, Adult), Right Whale, Bowhead Whale, Sperm Whale, Beaked Whales, Dwarf and Pygmy Sperm Whales (Kogia spp.), Blackfish Category (False Killer Whale, Pygmy Killer Whale, Melon-headed Whale, Pilot Whale), Killer Whale, Risso's Dolphin, Bottlenose dolphin, Stenella (spinner, spotted and striped dolphins), Fraser's dolphin, Rough-toothed dolphin, Common dolphin, Tucuxi , Harbor Porpoise, Dall's Porpoise, Lagenorhynchus species, Right Whale Dolphins, Beluga, Hawaiian Monk Seal, California Sea Lion, Steller Sea Lion, Northern Fur Seal, Guadalupe Fur Seal, Bearded Seal, Phagophilic *Phoca* spp. (Ringed, Spotted and Ribbon Seals), Northern Elephant Seal, Harbor Seal, Harp Seal, Hooded Seal, Walrus, and Sea Otters.

At this point in time the "aversions" (negative aversions are attractions) that have been implemented in various tests and analyses relate to bathymetric and physical environment criteria, but not to social or ecological criteria (e.g. presence of other species or of biological resources). Thus the distribution of individuals in the AIM constructed virtual marine environment is stochastically determined by individual movement and marine condition parameters that, from the individual's point of view, ignore the presence of conspecifics and organisms from other species. Implementation of the AIMS animat module from a theoretical point of view, appears to be an ergodic process: the final distribution of individuals in the virtual space is independent of initial conditions and, if you have sufficient animats being simulated, a distribution will be constructed that after an initial burn-in period is stationary over time. We verified this by simulating two different scenarios at our meeting and, in each case, looking at histograms of the distribution of individuals with respect to depth, latitude, and longitude at several different points in time (see figures in Appendix 3). Thus, in the absence of the implementation of "aversions" in response to sound or the presence of other animates, the AIM animat engine can be used to construct the ergodic distribution arising from a purely animat description of the problem. In this case, the computation time required to carry out certain types of sensitivity and scenario analyses can be greatly reduced for when only the convolution of the animat distribution with sound field intensity and frequency distributions are required. This approach, however, will not generate the time histories of the cumulative exposure of individual animats to sound; although it is a quick way to compare relative affects with regard to stationary distributions that ignore ecological (location of resources, predators, or other factors that cause locations of individuals to deviate from purely physically and surface/dive-behavior determined factors modifying an otherwise random distribution of positions) and social factors.

Ultimately, however, it is important to know what the cumulative effects of sound exposure are on individuals, and to do so in the context of evaluating the properties of extreme events, even though rare. This kind of evaluation requires that aggregative social, feeding, or predator avoidance behavior of individuals be taken into account. In the absence of data that allows aversion parameters to be set that would simulate such behavior, plausible scenarios need to be investigated under "what if … ?" scenarios that assume that individuals aggregate for various reasons. Again, what is important here is not actual numbers, but comparative changes in the probability of extreme events compared with the baseline analysis of "no aggregating" behavior.

TOR 3: Assess whether AIM meets the Council for Regulatory Monitoring guidelines for model development

Applicability of CREM Guidelines to AIM

The CREM Guidelines (which disturbingly appear still to be in a draft state—we have the November 2003 version—and yet are accorded the status of a providing a standard that should be met) idealize a step of procedures that the development of environmental modeling software is unlikely to follow to the letter unless the development was designed with the guidelines in mind. However, the guidelines embody some important principles on minimal standards that should be followed in obtaining reliable models, even though a lack of adherence is not indicative of bad software. The CREM Guidelines discuss both modeling frameworks and models, although AIM does not fit comfortably within these definitions. AIM, as a software platform is somewhat more general than a model framework: it is flexible enough to embrace several different sound propagation frames works, with the best implementation depending on bathymetric specifications. AIM implemented for a specific set of conditions yields a model that can be evaluated using the CREM guidelines, as I will endeavor to do here.

Have the principles of credible science been addressed during the development of AIM?

- **Does the conceptual model have an adequate theoretical basis?**

  The modules used by AIM to calculate the spatiotemporal dynamics of sound intensities and frequencies have a long history of development, evaluation, and testing in the marine acoustics literature and appear to be the best available to date.

  The theory behind the animat engine component is that realistic trajectories can be generated from a set of parameters, described elsewhere in this report, that are embedded in a set of rules. This approach does not allow verification of elementary cases as easily as a more mathematical compact Markov transition modeling approach. However, to the extent that phenomenological physical characteristics can be extracted from empirical trajectories independent of all but bathymetric constraints and spatiotemporal variables designating a particular activity as suitable for the time and space concerned, the algorithms driving the animat engine (Monte Carlo methods using uniform distributions except for speed which is assumed to be truncated Gaussian) are adequate for the task at hand. Aversion structures are needed and implemented to ensure physical constraints are not violated. The behavior of real animals of the different species model are, for the most part, not well enough empirically established to have a theoretical basis for implementation.

- **Has the conceptual model been adequately implemented?**

  I will leave the answer to this question to Michael Porter to address for the propagation model component of AIM. For the most part, the conceptual model appears to have been correctly implemented for the Animat Engine. As with all implementations, some problems often arise in setting up parameters, defining aversions correctly, and output should never be accepted without verification that it conforms to expectations.

- **Were adequate techniques and procedures used for code verification?**

  The code has been developed over a number of years during which it has been continuously tested and evaluated. Obviously, the code would not meet NASA standards for rocket or module control or nuclear plant standards for process control; but the cost of checking all code at the highest level is not justified in this case. For its purpose, code verification appears to be followed at the level expected. Fortunately, however, a similar, independently developed software platform exists—the ESME Workbench (Shyu and Hilson, 2006) mentioned above. These two platforms should be tested against one another. This is the cheapest, quickest and, in fact, best way to verify at this point whether or not the procedures and code in AIM are performing as expected. The comparison of performance should be conducted over a set of exercises that covers the full range for which AIM has been designed to carry out analysis and develop impact assessments.

- **Has there been adequate peer review of AIM?**

  This evaluation appears to be the first independent peer-review of AIM. This review is much more thorough than any review that publications arising out of AIM would get if they been submitted to the peer-reviewed literature (e.g. of the type that ESME got when published IEEE J. Ocean Engineering by Shyu and Hilson, 2006). However, I stress again that the best way to review AIM is compare its performance against ESME Workbench: when output differs by more than a percent or two then both platforms need to be checked for likely sources of error which, when fully understood, should lead to one or both being modified to correct any errors.

AIM model applications

We did not get the opportunity to carry out in-depth evaluations of AIM's performance in particular studies, although we did get a written record of comments and responses that were part of the process of evaluation of the Navy's EIS analysis pertaining to the use of its Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar.

- **Was the choice of model appropriate given the quantity and quality of available data?**

  The implementations were appropriate, but do not appear to have gone far enough to evaluate the sensitivity of results to parameter certainties, were incomplete with regard to finding ranges that bound the results, and did not attempt to find out more about the tails of distributions and the probability of rare events as discussed earlier in this report.

- **How closely did the model simulate the system of interest?**

  This is hard to say. With correct implementation, the accuracy of the sound propagation simulations is determined by the accuracies of the well-accepted Bellhop and Parabolic Equation models. The sound animats are easily verified to be correct, but the animal animats are unlikely to behave anything like real systems because ecological and sociological components of the behavior of individuals are ignored. This is not necessarily inappropriate to generate a baseline analysis, provided that this analysis is augmented by scenarios where assumptions regarding behavior that lead to highly aggregated population distributions are also investigated in comparative "what if …?" types of analyses.

- **How well did the model perform?**

  The model appears to perform well if the focus is on calculating long-term averages. The platform is not set up to directly facilitate other types of analyses relating to the calculation of events in the tails of distributions or to carry out comparative studies where the relative effects of different scenarios, some of which pertain directly to management questions, are compared to each other and to the baseline results.

- **Were adequate sensitivity and uncertainty analyses performed?**

  No.

- **Are the models transparent and the results reproducible?**

  The transparency of AIM is limited by the propriety of its software, but the results are easily reproducible when carrying out average calculations.

- **What processes or procedures exist to enable ongoing model evaluation and improvement?**

  The developers of AIM did not articulate any particular processes or procedures that they have in place to evaluate and improve AIM, other than to continue as they have done over the past five years to make incremental improvements as needed. It is not clear, however, whether or not MAI is interested in investing in further development of AIM over the next five years at the same intensity or levels as during the past five years

Additional comments

AIM was initially developed as a platform that would facilitate ease of use and replicability in bringing together different sound propagation and animat movement components to calculate the exposure of marine animals to anthropogenic sounds and, in particular, calculate Take A and B rates during the operation of sound arrays by the Navy or sound blasts by different groups involved in exploration of the oceans and exploitation of marine resources. With this same philosophy, AIM needs to be further developed to facilitate ease of use and replicability in carrying out uncertainty and sensitivity analyses, providing easy comparisons among scenario analysis and with baselines runs.

Conclusions

AIM is at the cutting edge of software platforms for evaluating the impacts of anthropogenic sounds on marine mammals and animals. It has at least one competitor: the Naval Research Labs (NRL) “Effect of sound on the marine environment” (ESME) workbench (Shyu and Hilson, 2006). This is indeed fortunate because our confidence in the performance of either of these two platforms can be greatly enhanced by comparing simulation results from both and investigating sources of error or poor performance when they disagree by a few percent. The application of AIM to particular problems, and its use in the EIS process has not been particularly innovative and has not exploited the potential utility of the AIM software platform. This utility can be greatly enhanced by the addition of procedures and output control in AIM to facilitate uncertainty, sensitivity, and scenario analyses.

Recommendations

1. AIM and ESME should undergo a series of comparative performance analyses that put both platforms through their paces over a range of problems for which both platforms were designed to produce results.
2. AIM should be enhanced to improve its ability to undertake simulations that include the full range of probabilities that are likely to be encountered in everyday situations in the marine environment (namely by including daily variation in the physical environment, as well as a richer suite of animal behaviors in relation to ecological and sociological factors).
3. AIM should be enhanced to calculate probabilities involved with events in the

tails of distributions rather then focus just on average or mean rates.

4. AIM should be enhanced to facilitate uncertainty, sensitivity, and scenario analyses and produce output that allows for quick visual comparisons of different runs.

References

Balcomb, K. C. III and D. E. Claridge. 2001. A mass stranding of cetaceans caused by naval sonar in the Bahamas. Bahamas J. Sci. 2:2-12.

Haw-Jye Shyu and Roger Hilson, 2006. A software workbench for estimating the effects of cumulative sound exposure in marine mammals, IEEE J. Ocean Engineering 31(1):8-21,

Hursky, P., M. B. Porter, and M. Siderius, V. K. McDonald, 2004. High-frequency (8–16 kHz) model-based source localization. Acoustical Society of America 43:3021-3032.

McDonald V. K., J.A. Rice, Michael B. Porter, Paul A. Baxley, 1999. Performance measurements of a diverse collection of undersea, acoustic, communication signals", Proceedings of I.E.E.E. Oceans'99,

Southall B. L., R. Braun, F. M. D. Gulland, A. D. Heard, R. W. Baird, S. M. Wilkin, and T. K. Rowles, 2006. Hawaiian Melon-headed Whale (Peponacephala electra) Mass Stranding Event of July 3-4, 2004, NOAA Technical Memorandum NMFS-OPR-31.

Bibliography of distributed material

In addition to the documents, extracts, and PowerPoint presentations which were distributed in electronic and/or hardcopy, the results of some AIM model runs were also distributed in the form of "screen grabs" (captured in a Word document) and .png files.

Documents

Anon. 2006 a: LFA Observation techniques notes. Marine Acoustics, Inc. 1 p.

Anon. 2006 b: AIM timeline. Marine Acoustics, Inc. 4 p.

Anon. 2006 c: Annual Report No. 4: Operation of the Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar onboard the R/V Cory Chouest and USNS IMPECCABLE (T-AGOS 23) under the National Marine Fisheries Service Letters of Authorization of 12 August 2005. Maritime Surveillance Systems. 67 p.

Pascual, P. et al. 2003: Draft guidance on the development, evaluation, and application of regulatory environmental models. The Council for Regulatory Environmental Modeling. Draft, November 2003. 60 p.

Frankel, A.S. 2006: Acoustic Integration Model (AIM) Internal Review Document. Marine Acoustics Inc., proprietary document (draft, 14 September 2006). 50 p.

Frankel, A.S. & Buchanan, J.M. 2006: Acoustic Integration Model© (AIM) Users Manual. Marine Acoustics, Inc. (draft, April 2006). 58 p.

Frankel, A.S., Ellison, W.T., and Buchanan, J. 2002. Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts. Oceans 2002 MTS/IEEE 3: 1438-1443.

Frankel, A.S & Vigness-Raposa, K. 2006: Marine Animal Behavioral Analysis. Marine Acoustics Inc., Tech. Memo. 63 p.

Marine Acoustics, Inc., Patent Application, 29 Oct 04. Method for modeling the effect of a stimulus on an environment. Marine Acoustics, Inc. proprietary document. 42 p.

Document extracts

Anon. 2006 d: Summary of SURTASS LFA sonar operations for 16 August 2002 to 15 August 2006. 2 p.

Lecky, J.H. 2006: Taking and importing marine mammals; taking marine mammals incidental to the U.S. Navy operations of Surveillance Towed Array Sensor System Low Frequency Active Sonar. Extract from Federal Register. 4 p.

Various authors and dates: Excerpts related to AIM from responses to comments in major regulatory publications. 16 p.

PowerPoint presentations

Frankel, A.S., & Ellison, W.T. 2004: AIM modeling of oil industry activities to derive marine mammal take estimates. IMEMS 2004 conference.

Frankel, A.S., & Ellison, W.T. 2006: The Acoustic integration Model (AIM): application issues for risk assessment. Navy environmental planning and natural resources workshop; risk assessment technologies, Virginia Beach, VA, May 2006.

Messegee, J. 2006: History of marine mammal exposure. AIM CIE review 25-27 Sept. 2006.

Shores, L. 2006: Examination of AIM-generated PE files. AIM CIE review 25-27 Sept. 2006.

Statement of Work

Consulting Agreement between the University of Miami and Wayne Getz

STATEMENT OF WORK
Review of Acoustic Integration Model © (AIM)

The National Marine Fisheries Service (NMFS) requires an independent peer review of the Acoustic Integration Model© (AIM), which shall assess whether AIM correctly implements the models and data upon which it is based, whether animal movements are adequately simulated within AIM, and whether AIM meets the Council for Regulatory Monitoring guidelines for models, which primarily involve scientific credibility.

Background

Minimizing and mitigating the potential effect of sound upon the environment is an increasing concern for many activities. Naval operations, seismic exploration, vessel and aircraft operations, certain construction activities, and scientific investigations now need to consider the potential effects underwater acoustic sources have on marine life. Marine mammals are usually the primary concern, due to their widespread distribution and excellent hearing ability, although impacts on fish are increasingly being considered as well. Predicting the exposure of marine mammals is complicated by their diving behavior and, in some cases, long-range migrations, which causes them to "sample" many depth strata within the water column.

Acoustic propagation and sound received levels are a function of water depth, range from the source, and a host of sound source and environmental variables. This, combined with the variable diving behavior of different species, makes for a very complex problem. The Acoustic Integration Model© (AIM) addresses these specific complications. A principal component of AIM is a movement simulation engine. Both sound sources and animals, collectively addressed as "animats," are programmed to move in location and depth over time in a realistic fashion. Animal movement is based on documented regional and seasonal behavioral data for each species evaluated. Acoustic sources and receivers are programmed to move through a virtual acoustic environment, based on external environmental databases and radiated sound fields created from a choice of several propagation models (*e.g.,* Parabolic Equation [PE], Bellhop, etc.). The integration component of the AIM engine then predicts the exposure level of each simulated animal at successive operator-selected time steps. Furthermore, each animal can evaluate its environment at each time step, and can be programmed to alter direction or diving behavior in response to variables, such as sound level or sea depth. AIM allows the user to predict the effects of different operational scenarios and animal responses, thereby allowing the selection of an alternative that produces the least impact and still meets operational requirements.

AIM is a proprietary model owned by Marine Acoustics, Inc (MAI). Its value in predicting the acoustic exposure of animals has been demonstrated in earlier documents.

However, the continued use of the model to provide acoustic exposure and impact predictions for regulatory assessment purposes requires that the model be reviewed independently, so that NOAA and other federal agencies can comply with the Data Quality Act.

Figures and related material

In Figs 1-3, we compare the histogram of frequencies of individual animats at different depths for the following two runs:

1. 5454 animats running for 500 minutes of simulation time (30 min step time). The population is confined to a deep sea, flat floor, 1-degree latitude by 1-degree longitude box with reflecting boundaries. The simulation rapidly settles into its final stationary distribution, as seen in terms of the average distribution of individuals during the first 104.5 minutes is the same as the second and the 14th block of 104.5 minutes of data.
2. The second run has 375 inshore and 555 offshore animats (kept inshore and offshore through appropriate aversions specifications). The inshore animats, however, have not settled down by 2nd time block (Fig. 2), nor by the 3rd, although the three histograms are generally quite similar. The offshore population, however, settles into its final distribution more rapidly (Fig. 3).

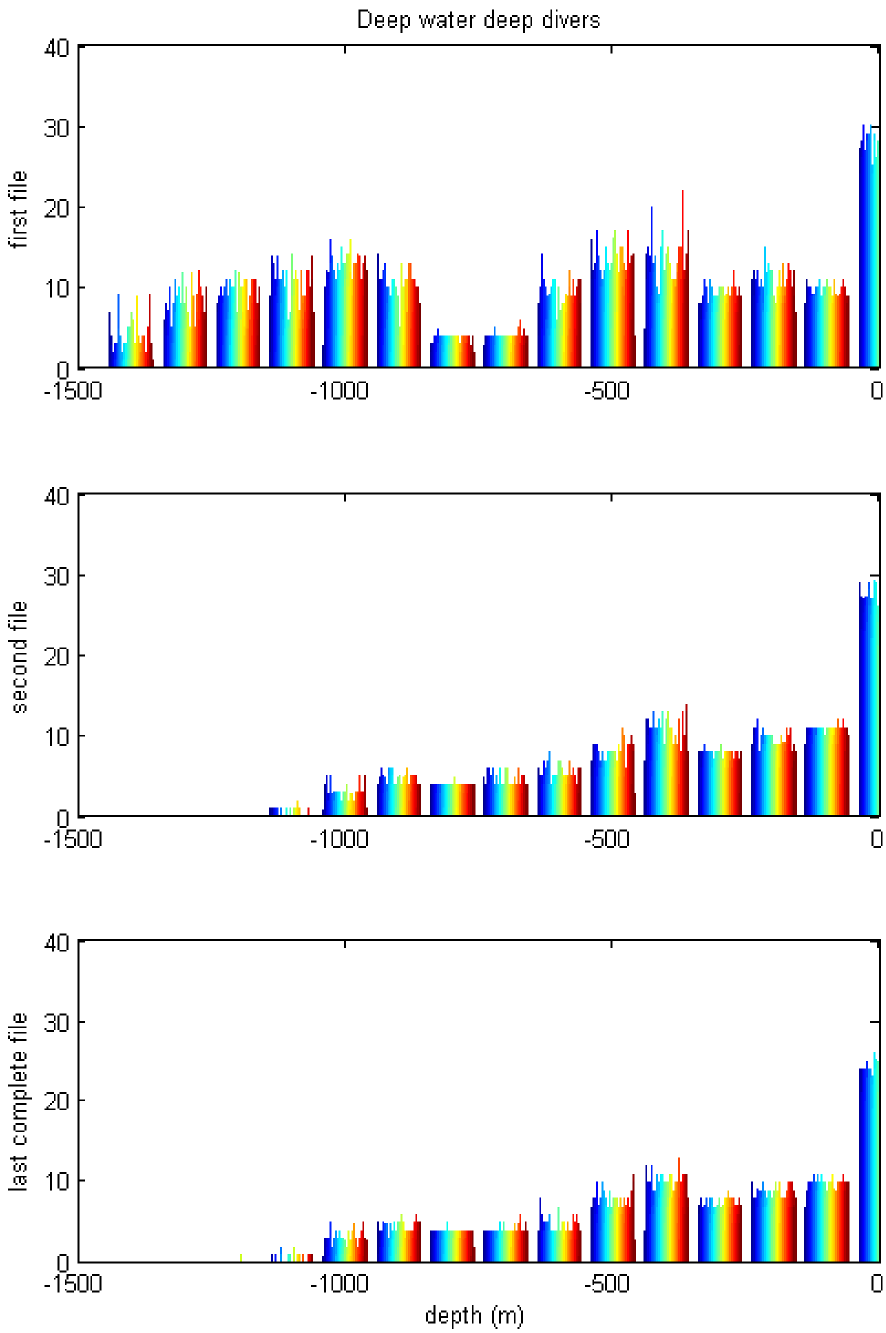


Fig. 1. Results from First Run

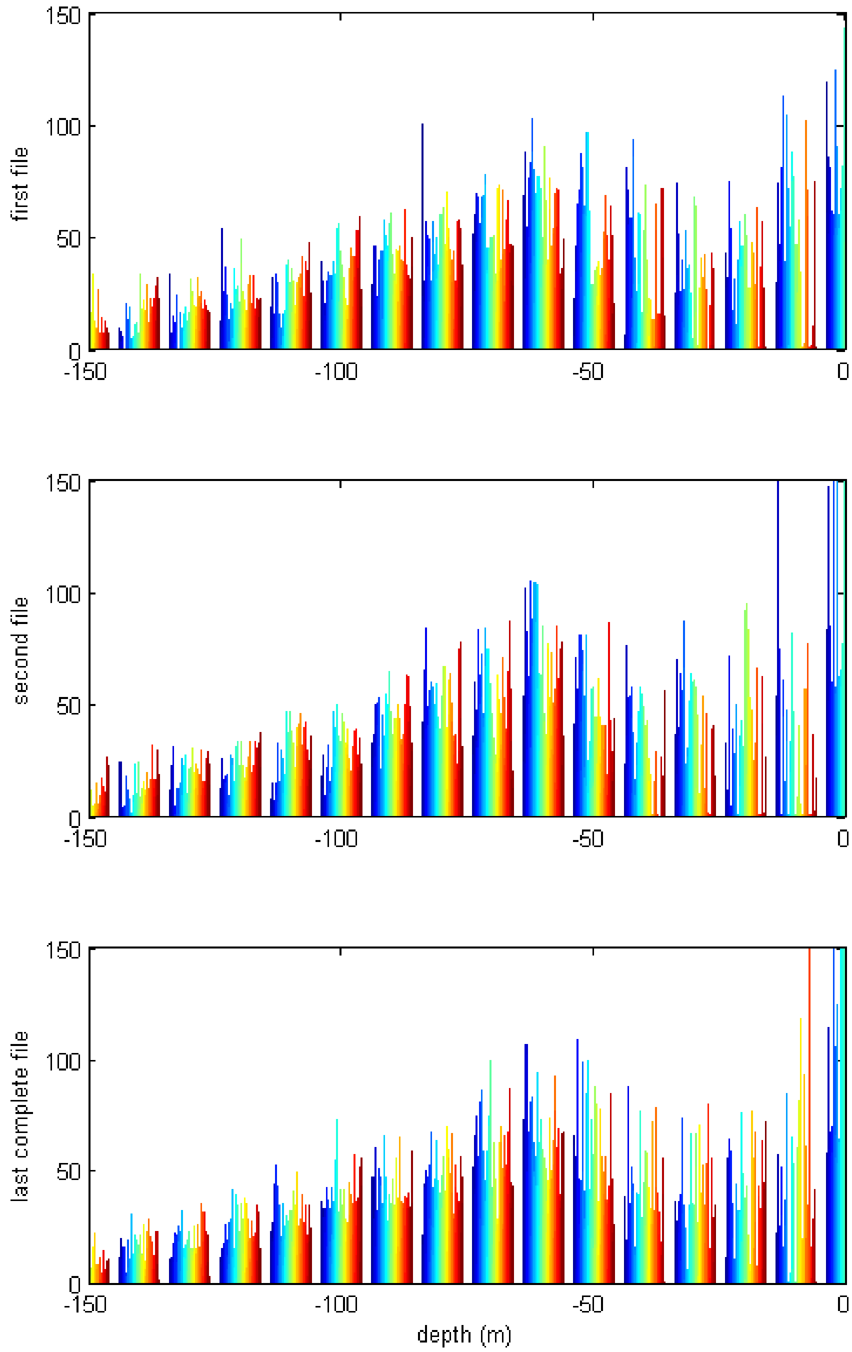


Fig. 2. Results from Second Run: Inshore group.

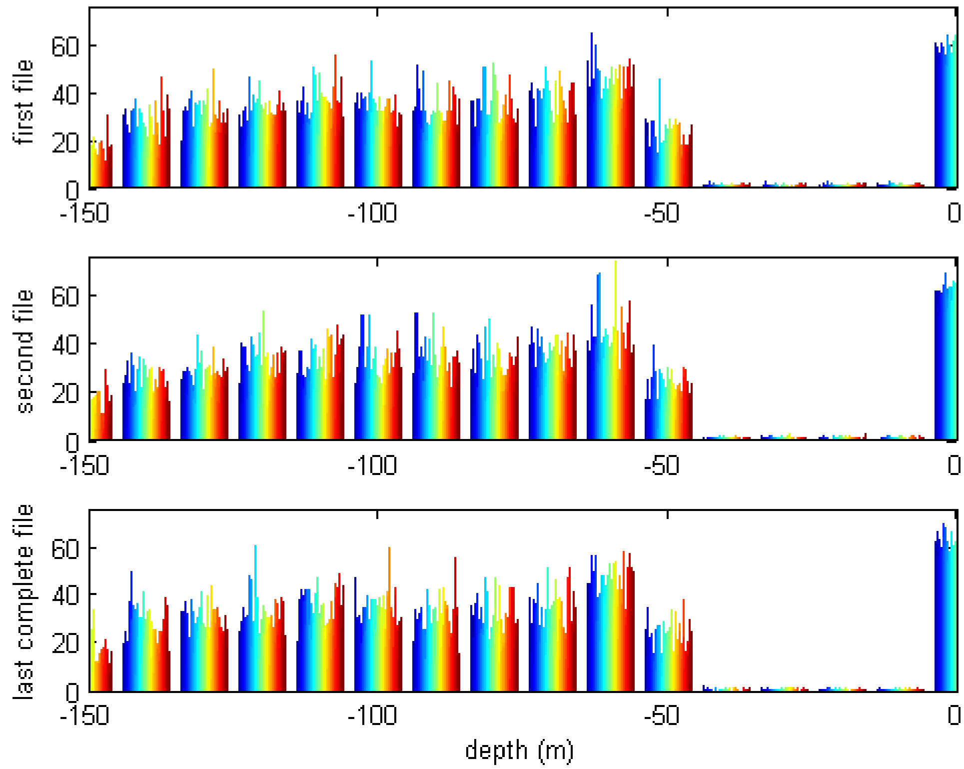


Fig. 3. Results from Second Run: Offshore group.

Acoustic Integration Model (AIM) Review

Dr. Michael Porter

Executive Summary

Marine Acoustics Inc. has developed a model called AIM (Acoustic Integration Model) that is used to simulate exposure levels for marine mammals to acoustic sources of interest. Examples of the latter include SONAR systems and airguns used for seismic exploration. Fundamentally the tasks in this review involve an assessment of the quality of this process. The “Terms of Reference” (TOR) defining the goals of the review address this task both narrowly (have the propagation models been implemented correctly?) and broadly (does the conceptual model have an adequate theoretical basis?). Briefly, I consider that AIM provides a very reasonable scientific approach to the problem it addresses and generally seems to have been correctly implemented.

Considering the Terms of Reference in sequence: TOR1: The acoustic models generally seem to have been incorporated correctly. TOR2: I'm not an expert on mammal movement models; however, I believe the approach used in AIM is a very reasonable start on a difficult problem. TOR3: The AIM software generally meets the standards of the CREM guidelines.

That said, there is clearly room for improvement in all these areas. Although AIM presents a smart, professional interface, it has followed a typical development of scientific software in which features are added as needed and with restricted funding. The CREM guidelines generally envision an idealized process that requires significant funding. Model documentation and benchmarking are typically the first casualties of tight budgets. How much effort should be spent in this area is really more of a programmatic consideration than a science one, as it can go on indefinitely providing continued improvements with continued expenditures. However, considering the public concerns related to policy decisions based on AIM results, I think that this is the area deserving greatest attention. Test cases done during the review revealed a few minor problems in AIM. Those test cases, and more, ideally should be part of a document that convinces an outsider that the software is performing correctly.

Background

One of the acoustic models used in AIM is BELLHOP, which was developed by this author and which is publicly distributed through the Ocean Acoustics Library (OALIB). OALIB is supported by the U.S. Office of Naval Research to provide acoustic modeling software to the research community. In accordance with the terms of reference, this review will generally not address the quality of the acoustic models but rather will focus on their correct use in AIM. The author has provided similar input to the ESME Workbench (Effects of Sound on the Marine Environment), which is a software package developed by ONR that provides similar capabilities to AIM.

It is important to note that AIM is used for a fairly complicated task. The quality of the resulting model predictions is highly dependent on the skill of the user. AIM draws upon a variety of inputs (bathymetry, ocean sound speed profiles, source waveform characteristics, sea state, ocean-bottom properties). It uses sophisticated propagation models with various strengths and weaknesses. It incorporates mammal distribution and movement models. Finally, it calculates an exposure metric.

Considering this chain, we note the following challenges: 1) Environmental information is notoriously difficult to obtain. In particular, geoacoustic models (essentially databases of how reflective the ocean bottom is to sound) are widely recognized as a major limitation in navy applications of acoustic models. 2) Modeling sound propagation in the ocean is fairly challenging as exact solutions to the wave equation that governs sound propagation are generally not practical. Instead a variety of approximate techniques are used. The choice of any given model and the proper use of a model require skill, especially if the model is used for a wide variety of source types. 3) There is very limited information on the distribution of marine mammals. (This is widely considered to be the biggest challenge in predicting exposure levels.) 4) There is a limited understanding of what levels of sound are harmful or disturbing to different species.

In short, the quality of predictions resulting from AIM depends on research in a variety of areas that is ongoing.

Review Activities

A core part of this review consisted of a 3-day meeting at MAI in which MAI personnel demonstrated and explained AIM. In this process they responded to detailed questions from the review panel and applied AIM to test cases defined by the review panel. This report is a summary of those findings.

Summary of Findings

TOR 1: Assess whether the AIM implementation is correct (focusing on acoustic models)

AIM uses three principal acoustic models:

1) The Basic Acoustic Model (BAM),
This model is basically just a simple formula that uses spherical spreading (20 log r) in the near field and then switches smoothly to cylindrical spreading (10 log r) in the far-field.

2) Navy Standard Parabolic Equation (NSPE),
The NSPE is actually a wrapper for two different PE models and allows the user to choose between the old NSPE, based on code developed by Harvey Brock and a more recent code developed by Michael Collins and called RAM. Both of these models have had important modifications for use in the NSPE and therefore differ from other versions in use.

3) BELLHOP.
This is a Gaussian beam based model developed by this author.

A number of tests were done during the review to assess the use of these acoustic models. We summarize the highlights here. Figure 1 shows the transmission loss (TL) for "Test Case 1" as calculated by the NSPE. Note the sudden change in behavior at a range of about 1 km. This results from using BAM in the near field. This in turn was done because of the narrow angle limitations of the NSPE. (The term narrow angle is used because the PE is derived in an approximation that is increasing accurate towards horizontal propagation. The model is generally started with a beam-like source in the near field, which generates energy that is predominantly horizontally propagating.)

This example shows two flaws in the use of the acoustic models. First, BAM ignores the field due to the reflection of the source in the ocean surface. The total field is the combination of the direct sound from the source with its image in the surface. This leads to an interference pattern frequently called the Lloyd mirror pattern, with signal level 6 dB higher than that resulting from just the direct sound.

Secondly, the narrow-angle PE used here is rarely used in the research community as wide-angle PE's are readily available that not only are more accurate but generally faster. This is the RAM option which is also available in the NSPE and generally used by the fleet (personal communication, R. Zingarelli) in place of the narrow angle PE.

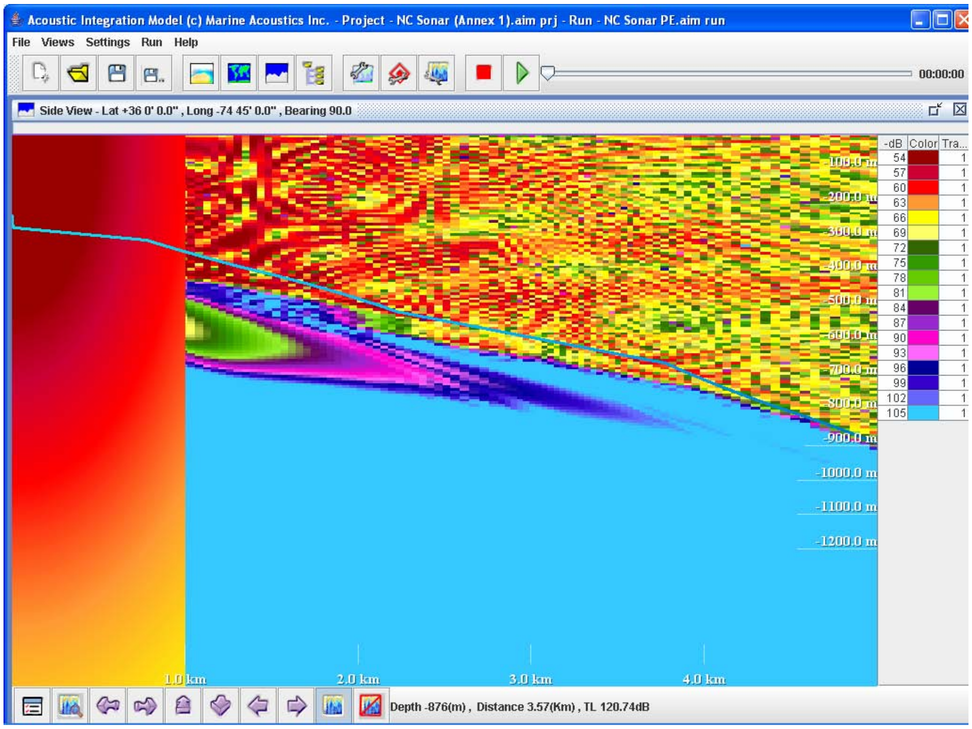


Figure 1: NSPE transmission loss. BAM is used in the near field, producing an abrupt transition.

With some adjustments in the sampling of the field, AIM produced the BELLHOP result shown in Figure 2. Here we can clearly see the Lloyd mirror patter that is missing from BAM. This test case revealed that the AIM model had only been using the real part of the complex pressure field so that AIM was underestimating the sound level. However, BELLHOP was apparently normally being used in a "semi-coherent" option where the imaginary part vanishes. In the course of the review, MAI corrected this error resulting in the plots here.

Figure 3 shows the NSPE result with the near-field BAM option disabled so that we can see more clearly the NSPE field. Here you can see the narrow angle limitation of this particular PE. For the purposes of the AIM application, BELLHOP and the NSPE are producing fields that are qualitatively similar and probably sufficiently accurate. However, we can see discrepancies between the models that can probably be removed with some improvements to the way the models are used. In particular, we see that the PE has a strong beam in the sediment that is refracted back into the ocean some distance down range. BELLHOP could be modified to handle this beam shift; however, the beam shift in the NSPE is probably a result of an artificial sediment layer that may itself not be representative of the real sediment. Figures 4 and 5 simply extend the view of Figures 2 and 3 showing that the models produce reasonable agreement at longer ranges.

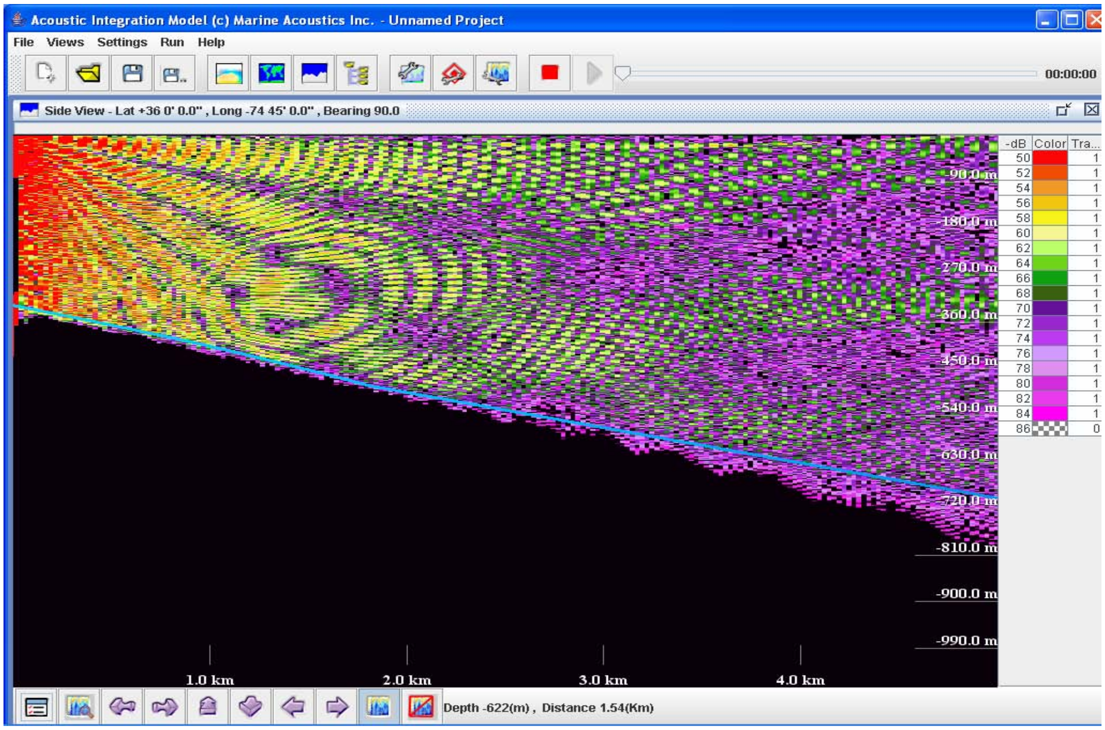


Figure 2: BELLHOP transmission loss showing the Lloyd mirror pattern in the near field.

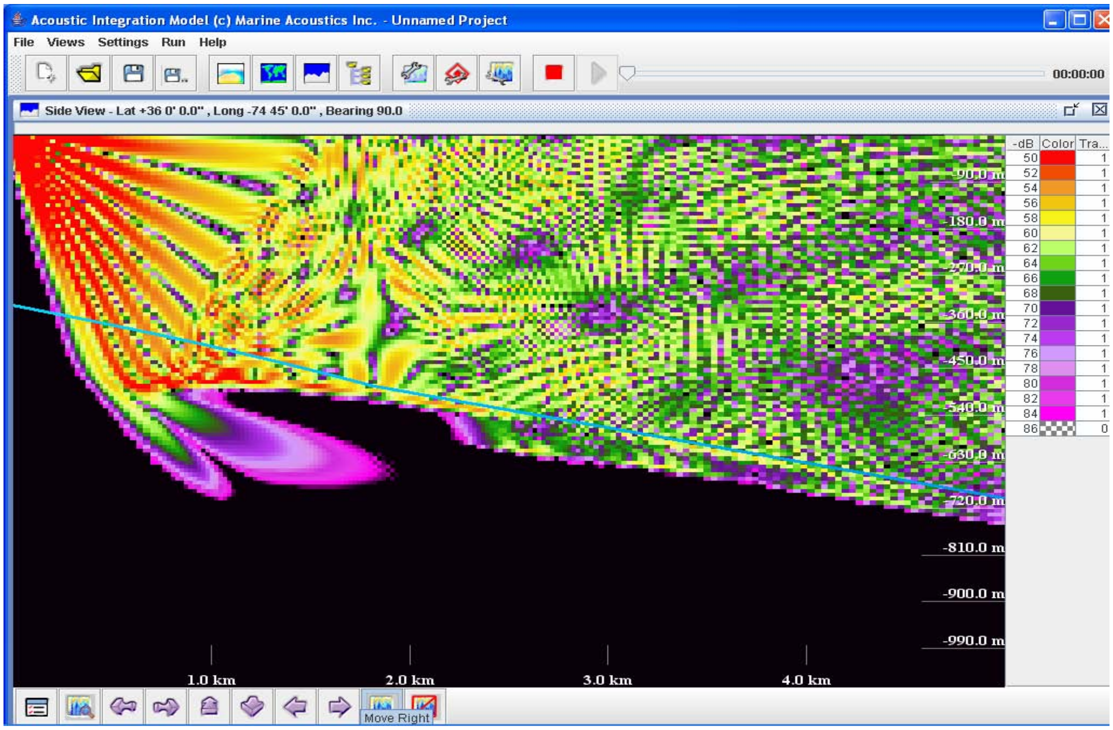


Figure 3: NSPE transmission loss

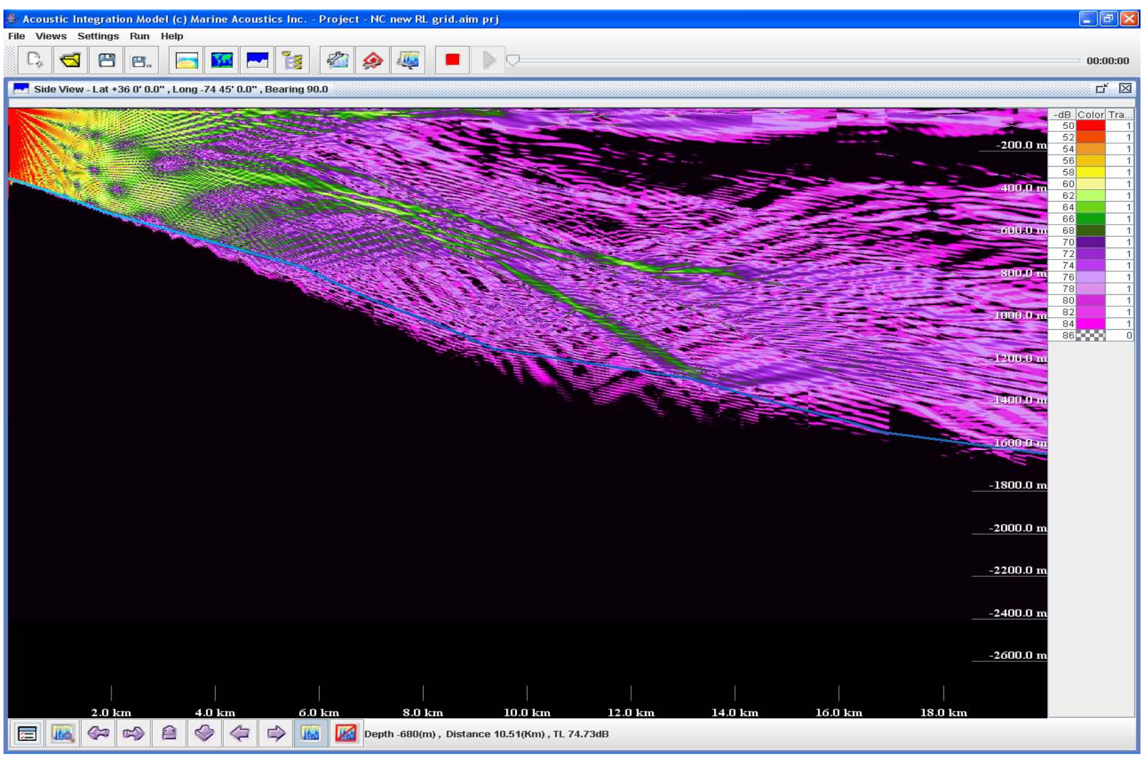


Figure 4: BELLHOP transmission loss.

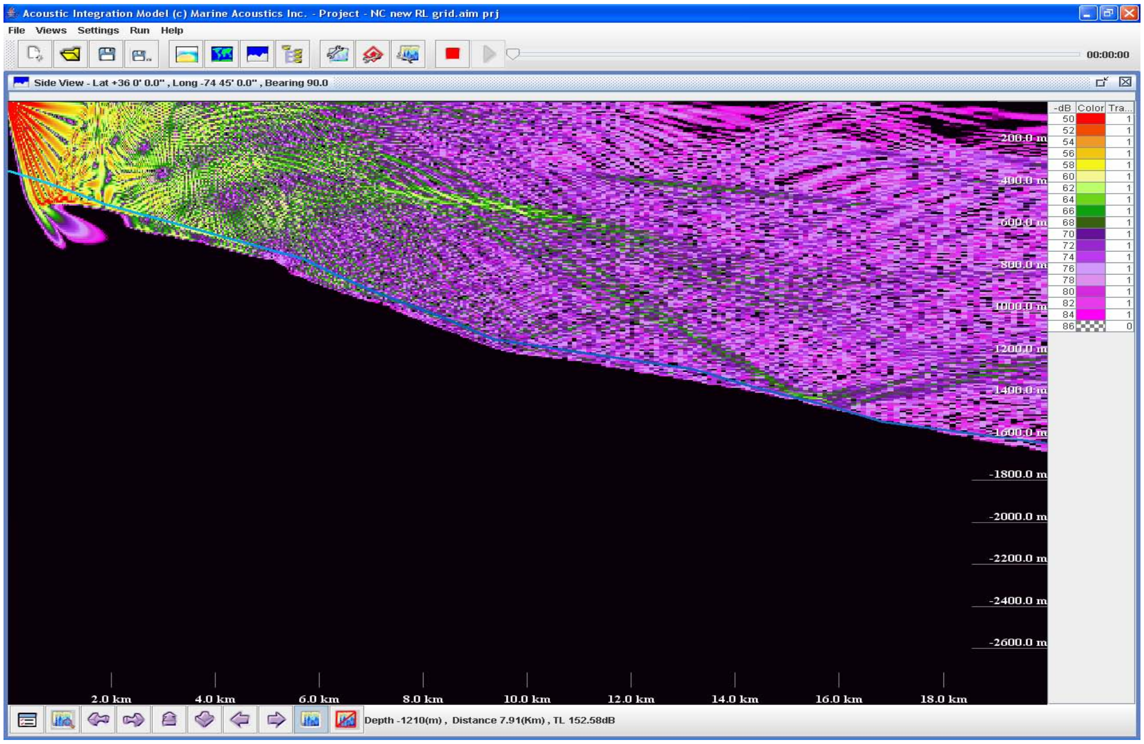


Figure 5: NSPE transmission loss.

A number of other tests were done giving some confidence that the acoustic models were extracting environmental information from the databases in a reasonable way and that the TL output was carried through to the animats consistently.

Summary: Several issues arose in these brief tests that raise questions about whether more careful testing should be done and carefully documented on the use of these models. Several changes were suggested by these tests:

1) Modify BAM to use a Lloyd mirror solution in the near field
2) use the RAM option in NSPE to provide a wide-angle capability
3) Use the coherent option in BELLHOP and correct the code so that it extracts the full complex pressure field, rather than just the real part (done during the review).

The recommendations about modifying BAM to produce the Lloyd mirror solution are based on the observation that the PE is producing a coherent field. There is an inconsistency in a fundamental inconsistency in the process. However, it's not entirely clear if the coherent field provides the correct metric for exposure. This issue deserves careful consideration in AIM.

Finally, we note that the acoustic model calculations are generally done at a single carrier frequency. There are fairly common circumstances (Porter, 1985) in which interference effects can produce large areas that are either in the acoustic shadow or fully "'lit' with great sensitivity to the carrier frequency. Predictions of animal exposure will be sensitive to this effect. The resolution is to run the model over several frequencies.

TOR 2: Assess the animal movement simulation within AIM

There is currently a great deal of research on mammal movement with extensive tagging in which sensors are attached to marine mammals to observe their movements. Much will likely be learned in coming years. However, the approach used in AIM based on simple statistical characterizations is a very reasonable approach given the amount of data currently available.

Testing focused on some simple metrics on histograms of the depth distributions of the animals. An example of one such test is shown in Figure 6. For this simulation, 5454 animats were tracked for 500 minutes of simulation time. Two populations of dolphins (offshore and coastal) were simulated with the populations separated by a depth boundary and bounded on one side by the coast. The panels in Figure 6 from top to bottom show 1) the initial distribution, 2) the distribution during the second time frame, and 3) the final distribution. Of interest is the fact that the distribution is stable over time. This result was obtained after correcting an error in the model, detected during the review. The error resulted in some animals diving too deep and the histograms failing to stabilize.

While there is no indication of any errors in this component of AIM, two recommendations can be made. First, a “burn in” period is suggested to allow the animats to attain stabilized movement patterns. Secondly, a set of test cases should be developed for which analytic solutions to the stable depth distributions are available.

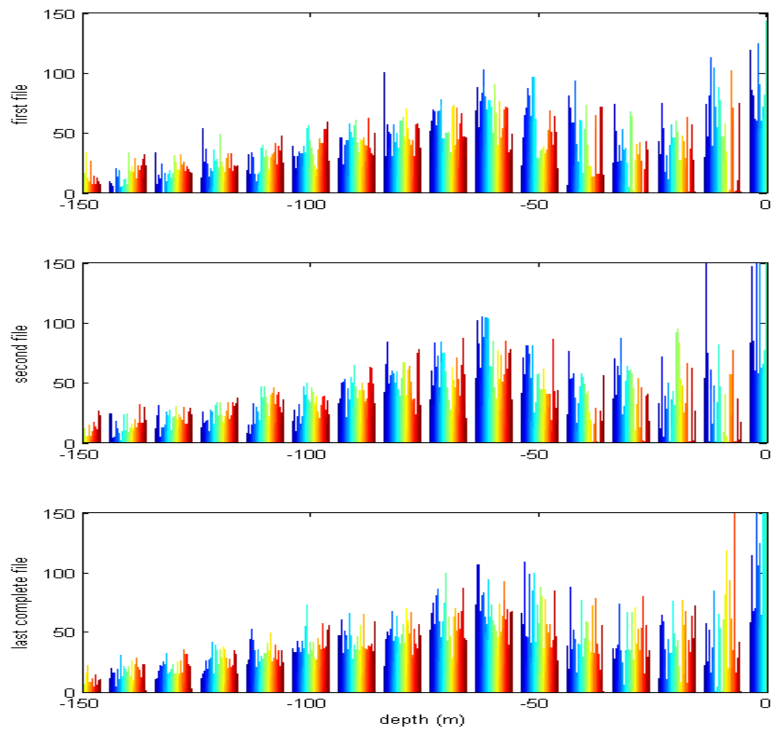


Figure 6: Histograms showing depth distributions of animats at various stages. Note that the distributions rapidly stabilize.

TOR 3: Assess whether AIM meets the Council for Regulatory Environmental Monitoring guidelines for model development

Applicability of CREM guidelines to AIM

This reviewer has struggled with the term 'adequate' that is continually raised in the CREM guidelines. What constitutes adequate testing, documentation, etc.? This question can really only be answered in the context of the importance of the model results in terms of impact on policy decisions. Therefore, the question of 'adequacy' is not just a science issue.

Another issue in considering the CREM guidelines is that the ‘boundaries’ of AIM are somewhat nebulous. AIM relies on a number of databases to characterize the ocean the sediment below it. It relies on operator searches of the literature to estimate marine mammal populations and in many cases to substitute specialized geoacoustic, bathymetric, and oceanographic databases for specific areas. It relies on operator expertise to incorporate the sound source correctly and to interpret the exposure metrics. The quality of the overall process depends in a critical way on the expertise of the users.

As AIM continues to be used and extended, it may well require less expertise. For instance, libraries of common sound sources will be built up and tested so that future modelers do not need to repeat that effort. However, models such as AIM are used also for sound sources used in research. The waveforms in that application are constantly changing to follow the research. Therefore, AIM will never be a fully turnkey product.

In summary, the fundamental question raised by TOR 3 cannot be answered in a simple pass/fail sense. However, we can take advantage of the excellent questions raised by the CREM guidelines to discuss the strengths and weaknesses of AIM with respect to each metric. In the following sub-sections we shall do exactly that, addressing specific questions raised in the CREM guidelines.

Have the principles of credible science been addressed during the development of AIM?

> Generally an excellent job has been done in applying the best science in the development of AIM. However, it is important to realize that the best science is woefully inadequate here. It is widely recognized that the core weakness in assessing impacts of sound is the lack of knowledge of marine mammal populations. This will improve in the coming decades, but policy makers must be aware of the great uncertainty in this area.

Does the conceptual model have an adequate theoretical basis?

> Again the answer is yes and no. AIM depends on metrics on reasonable exposure levels for marine mammals; however, our understanding of what types and levels of sound are disturbing to various species is very limited. Knowledge of marine mammal populations and movement patterns is similarly limited. Environmental databases, especially for the geoacoustic properties of the bottom are recognized as weak links in the modeling process. Research in new techniques to rapidly survey new areas has been ongoing for many years and will undoubtedly continue. Finally, research is ongoing on more accurate ways to model boundary scatter. In short, AIM is generally using the best available knowledge; however, the limitations of that knowledge are widely recognized.

Has the conceptual model been adequately implemented?

> Our tests revealed a number of flaws in AIM, several of which were corrected on the spot. Other issues should be addresses as indicated under the sections addressing TOR 1 and 2.

Were adequate techniques and procedures used for code verification?

It is not surprising that interactive demos of a complex piece of scientific software such as AIM revealed flaws. However, the particular problems that showed up in our testing suggest to me that greater effort should have been spent on code verification. The author does recognize that limits in this area are probably largely attributable to limits in the funding.

Has there been adequate peer review of AIM?

Adam Frankel is well known in the marine mammal community and has presented his work in numerous professional meetings. MAI has also presented their work to ESME. Again, the question of adequacy here depends on the importance of the resulting policy decisions.

Was the choice of model appropriate given the quantity and quality of available data?

AIM has included several acoustic models so that an appropriate acoustic model can be chosen for the given application. The field of underwater acoustic modeling goes back many decades and there has been a sizable investment in developing these models. While every institution tends to have its in-house models that are supported by local experts, the choices used in AIM are widely recognized.

The user community is generally well aware that the environmental data is frequently the limiting factor in acoustic model accuracy. The databases used are therefore very important. While different researchers favor different databases for different parts of the world, the choice in AIM would be considered reasonable.

How closely did the model simulate the system of interest?

AIM attempts to simulate an extremely complicated situation with the quality largely limited by our knowledge of marine mammal population densities. Validating the end-to-end process is extremely difficult; however, one can validate each step in the sequence. Intelligent decisions have been made throughout and generally represent the current best knowledge.

How well did the model perform?

Our tests revealed a number of flaws in AIM that were generally promptly corrected. The user interface seemed to be responsive and the code operated smoothly. However, given the errors that showed up in our testing, further benchmarking is warranted.

Were adequate sensitivity and uncertainty analyses performed?

In modeling very complicated systems, the first goal is to assemble all the components and make the best prediction. Once that goal is achieved, one progresses to try to put error bars on the predictions, typically through a Monte Carlo process. Aim does not do the latter; however, it should be noted that the acoustic modeling community has historically not graduated to this stage either, partly because of the computational cost of doing such analyses. In fact, this motivated a recent ONR effort on “Capturing Uncertainty”. Again, AIM is generally using the best standard practices. However, for important policy decisions, AIM should be run with several best/worst case scenarios to provide a rough sense of the uncertainty in the results.

Are the models transparent and the results reproducible?

AIM has the option of using acoustic models (NSPE) with a restricted distribution. However, there are publicly distributed models that are arguably better. It would be difficult for an outsider to reproduce the precise AIM predictions; however, there is probably enough information to produce an alternative, credible simulation and proceed to dispute the results.

What processes or procedures exist to enable ongoing model evaluation and improvement?

Model evaluation and improvement at MAI is likely driven by future funding. As MAI does studies for varying customers, they will surely continue to enhance AIM, particularly with regard to new acoustic sources. Adam Frankel is active in the marine mammal community and will undoubtedly continue to improve the marine mammal movement models and incorporate the latest knowledge on population densities.

Conclusions

AIM provides an intelligent approach to the problem of predicting the acoustic exposure of marine mammals. Generally reasonable choices have been made in the acoustic models, and databases. However, it is important to realize that this problem is extremely difficult and fundamentally limited by our lack of knowledge about marine mammals.

Recommendations

The sections on acoustic modeling and animal movement make a number of specific recommendations for improvements. The discussions on the CREM guidelines raise further issues. We could suggest an unbounded amount of further work here to improve this process; however, one has to recognize practical funding limits. Generally, a reasonable job has been done on AIM for the way it is understood to being used. However, there is room for improvement and the recommendations are summarized here:

- Modify BAM to include surface and bottom bounce paths in the near field.
- Use the RAM option in NSPE to provide a wide-angle capability
- Use the coherent option in BELLHOP and correct the code so that it extracts the full complex pressure field, rather than just the real part (done during the review).
- Use a burn-in period to stabilize animat behavior.
- In cases where the highest quality predictions are important
    - do calculations at multiple frequencies.
    - do sensitivity studies to provide error bars on model predictions.

If the model is to be used for particularly important policy decisions, an investment is needed to benchmark the code and provide transparent documentation that demonstrates better benchmarking has been done.

References

Additional references are provided in following section.

Bibliography of distributed material

In addition to the documents, extracts, and PowerPoint presentations which were distributed in electronic and/or hardcopy, the results of some AIM model runs were also distributed in the form of "screen grabs" (captured in a Word document) and .png files.

Documents

Anon. 2006 a: LFA Observation techniques notes. Marine Acoustics, Inc. 1 p.

Anon. 2006 b: AIM timeline. Marine Acoustics, Inc. 4 p.

Anon. 2006 c: Annual Report No. 4: Operation of the Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar onboard the R/V *Cory Chouest* and USNS IMPECCABLE (T-AGOS 23) under the National Marine Fisheries Service Letters of Authorization of 12 August 2005. Maritime Surveillance Systems. 67 p.

Pascual, P. et al. 2003: Draft guidance on the development, evaluation, and application of regulatory environmental models. The Council for Regulatory Environmental Modeling. Draft, November 2003. 60 p.

Frankel, A.S. 2006: Acoustic Integration Model (AIM) Internal Review Document. Marine Acoustics Inc., proprietary document (draft, 14 September 2006). 50 p.

Frankel, A.S. & Buchanan, J.M. 2006: Acoustic Integration Model© (AIM) Users Manual. Marine Acoustics, Inc. (draft, April 2006). 58 p.

Frankel, A.S., Ellison, W.T., and Buchanan, J. 2002. Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts. *Oceans 2002 MTS/IEEE 3*: 1438-1443.

Frankel, A.S & Vigness-Raposa, K. 2006: Marine Animal Behavioral Analysis. Marine Acoustics Inc., Tech. Memo. 63 p.

Marine Acoustics, Inc., Patent Application, 29 Oct 04. Method for modeling the effect of a stimulus on an environment. Marine Acoustics, Inc. proprietary document. 42 p.

Document extracts

Anon. 2006 d: Summary of SURTASS LFA sonar operations for 16 August 2002 to 15 August 2006. 2 p.

Lecky, J.H. 2006: Taking and importing marine mammals; taking marine mammals incidental to the U.S. Navy operations of Surveillance Towed Array Sensor System Low Frequency Active Sonar. Extract from Federal Register. 4 p.

Various authors and dates: Excerpts related to AIM from responses to comments in major regulatory publications. 16 p.

PowerPoint presentations

Frankel, A.S., & Ellison, W.T. 2004: AIM modeling of oil industry activities to derive marine mammal take estimates. IMEMS 2004 conference.

Frankel, A.S., & Ellison, W.T. 2006: The Acoustic integration Model (AIM): application

issues for risk assessment. Navy environmental planning and natural resources workshop; risk assessment technologies, Virginia Beach, VA, May 2006.
Messegee, J. 2006: History of marine mammal exposure. AIM CIE review 25-27 September 2006.
Shores, L. 2006: Examination of AIM-generated PE files. AIM CIE review 25-27 September 2006.

Statement of Work

Subcontract between the University of Miami and Heat, Light, and Sound Research, Inc. (Michael Porter) Statement of Work Review of Acoustic Integration Model © (AIM)

The National Marine Fisheries Service (NMFS) requires an independent peer review of the Acoustic Integration Model© (AIM), which shall assess whether AIM correctly implements the models and data upon which it is based, whether animal movements are adequately simulated within AIM, and whether AIM meets the Council for Regulatory Monitoring guidelines for models, which primarily involve scientific credibility.

Background

Minimizing and mitigating the potential effect of sound upon the environment is an increasing concern for many activities. Naval operations, seismic exploration, vessel and aircraft operations, certain construction activities, and scientific investigations now need to consider the potential effects underwater acoustic sources have on marine life. Marine mammals are usually the primary concern, due to their widespread distribution and excellent hearing ability, although impacts on fish are increasingly being considered as well. Predicting the exposure of marine mammals is complicated by their diving behavior and, in some cases, long-range migrations, which causes them to "sample" many depth strata within the water column.

Acoustic propagation and sound received levels are a function of water depth, range from the source, and a host of sound source and environmental variables. This, combined with the variable diving behavior of different species, makes for a very complex problem. The Acoustic Integration Model© (AIM) addresses these specific complications. A principal component of AIM is a movement simulation engine. Both sound sources and animals, collectively addressed as "animats," are programmed to move in location and depth over time in a realistic fashion. Animal movement is based on documented regional and seasonal behavioral data for each species evaluated. Acoustic sources and receivers are programmed to move through a virtual acoustic environment, based on external environmental databases and radiated sound fields created from a choice of several propagation models (*e.g.,* Parabolic Equation [PE], Bellhop, etc.). The integration component of the AIM engine then predicts the exposure level of each simulated animal at successive operator-selected time steps. Furthermore, each animal can evaluate its environment at each time step, and can be programmed to alter direction or diving behavior in response to variables, such as sound level or sea depth. AIM allows the user to predict the effects of different operational scenarios and animal responses, thereby allowing the selection of an alternative that produces the least impact and still meets operational requirements.

AIM is a proprietary model owned by Marine Acoustics, Inc (MAI). Its value in predicting the acoustic exposure of animals has been demonstrated in earlier documents.

However, the continued use of the model to provide acoustic exposure and impact predictions for regulatory assessment purposes requires that the model be reviewed independently, so that NOAA and other federal agencies can comply with the Data Quality Act.

Reviewer Requirements

The Center for Independent Experts (CIE) shall provide three panelists and a moderator for the review of AIM. Expertise in underwater acoustics, modeling, and marine mammalogy is required. The underwater acoustician should be familiar with propagation-loss models. Ideally, the acoustician will have experience or knowledge of the Bellhop and Navy Standard Parabolic Equation (PE) models, as these are the two main propagation models incorporated into AIM. The modeler should be familiar with individual-based models, preferably those dealing with animal behavior, and the integration of multiple data streams (*e.g.*, multiple databases). The modeler should be able to understand the dynamic interactions of databases. The marine mammalogist should have experience in marine mammal behavior, including diving behavior of more than one species. The moderator should have a reasonable level of scientific and technical understanding, with a reasonable degree of knowledge and experience in at least two of the three scientific categories (underwater acoustics, modeling, marine mammalogy).

The review will be organized around a three-day meeting at MAI during the week of 25 September 2006, to be scheduled based on the availability of all pertinent personnel. The moderator shall be required for a maximum of 12 days for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, and completing the summary report. The three panelists shall be required for a maximum of 12 days each for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, completing their individual reports, and reviewing the draft summary report. Thus, a total of 48 reviewer days is required.

Review Process

The panelists and moderator shall review the documents listed in Annex I prior to attending the three-day meeting at MAI, 4100 N Fairfax Drive, Suite 730, Arlington, VA 22203. After the MAI meeting, the three panelists shall write individual reports addressing all the Terms of Reference given below. The moderator shall consolidate the individual reports into a summary report. Details of the panelists' and moderator's reports are given in Annex II and III, respectively.

<u>CIE – MAI Conference Call – September 22, 2006:</u> The CIE panel and MAI team will discuss via conference call the details of the upcoming meeting, and the CIE panel will raise questions concerning background documents, specifications for trial runs, and other review-related material, including logistics. The call is tentatively set for 4 pm EDT on Friday, September 22, 2006, but participants and timing may be changed due to individuals' availability.

MAI Meeting - Day 1: Begin with an introduction and a "charge" to the panel, which includes laying out the roles of the participants and the terms of reference (see below), and a time table. AIM presentations will be made by scientists from MAI, with question and answer sessions as needed. Presentations shall include:

1) Introduction to the AIM approach and the software, including data input requirements;

2) Review of results of internal testing of the software;

3) Overview of the process used to derive animal behavior parameters from data and the scientific literature.

Following the conference call and these presentations, the moderator and panelists shall meet to develop a set of test runs for days 2 and 3. The purpose of this evaluation is to see how the AIM responds to a set of inputs that are designed to test the model. It must also be noted that in assessing the functioning of the model, it is important to acknowledge the differences and the roles of the external components (*e.g.*, animal input parameter values, propagation models) and the internal components of the model (*e.g.,* Animat Builder, Movement Simulator, Integration Engine). The distinction is drawn here to emphasize that the values of behavioral parameters can and should change when new data are available, and that AIM can utilize that new data. Valid input data *(e.g.*, animal behavior parameters) are critical for valid predictions from a model run for a specific scenario. However, a valid model can be provided with invalid animal behavior inputs, and still produce accurate outputs. The purpose here is to test the internal components of the model, including how AIM handles the input data. Example scenarios for devising these runs are provided in Annex IV.

MAI Meeting - Days 2-3: Dedicated to the CIE panel working with MAI scientists to perform AIM model runs, so the panel will have sufficient information on the input data, execution parameters, and model outputs for writing their respective review reports. The CIE panel shall, with the assistance of MAI scientists as required, design simulations and request that the MAI scientists create input files to represent these simulations during the course of the review. Projects can be created in a few minutes. Because AIM is a working model (not yet streamlined and simplified for public use), requiring expertise and familiarity with data input procedures and model execution techniques, MAI scientists will perform the model runs under the oversight of the CIE panel. The number and complexity of simulations to be run during the evaluation period will have been discussed in the conference call and finalized on Day 1. To run the models, MAI scientists will require sufficient time to research the values of the basic parameters (i.e., beam pattern information for source, or behavioral parameters for different species). The input files will then be run, and the inputs and outputs will be provided to the CIE panel for their analyses and evaluations.

Terms of Reference

The CIE panel shall complete the following tasks, and document their results in the individual panelist and summary reports.

1. Assess whether the AIM implementation is correct.
2. Assess the animal movement simulation within AIM.
3. Assess whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development.

1. AIM Implementation

Details relevant to the topics described below are given in the Robustness Review Document, which addresses operation of the AIM model.

- Does AIM accurately and efficiently implement the propagation models? Identify any errors in the implementation. The propagation models implemented in AIM include Bellhop (Porter, 1992) and Navy PE (Zingarelli, 1999). These models were created by other individuals and organizations. *The propagation models themselves are not the subject of the CIE review, but rather the implementation of these models in AIM.*
- Does AIM correctly handle the input values to the models? If not, identify any errors. For example, are acoustic source level and frequency values properly transferred through the model components?
- Does AIM correctly and efficiently extract data from databases? If not, identify any errors. AIM uses the GDEM-V (v 2.6) database for sound velocity profiles and the NOAA ETOPO2 database for bathymetry.

2. Simulation of Animal Movement

- Does the ANIMAT movement model in AIM adequately simulate animal behavior? Comment on the strengths and weaknesses of the modeling approach, and suggest possible improvements.

> The review panel shall devise one or more approaches for addressing this issue. One approach that shall be considered is to evaluate, given appropriate input values, how closely the modeled animal movements mimic the known responses of free-ranging animals. *The species-specific values used in the models are not the focus of the review, but rather the ability of the ANIMAT model to simulate movement.*

3. CREM Guidelines

The panel shall assess whether AIM meets the Council for Regulatory Environmental Monitoring (CREM) guidelines for model evaluation, which are summarized below. Some of the points listed below will have been addressed by the reviewers as part of their comments on Terms of Reference 1 and 2 above. Each reviewer shall ensure that clear answers are provided for the CREM guidelines, though extensive repetition of technical comments is not required.

- Have the principles of credible science been addressed during model development?
- Is the choice of model supported given the quantity and quality of available data?
- How closely does the model simulate the system of interest?
- How well does the model perform?
- Is the model capable of being updated with new data as it becomes available?

Schedule of Activities

The schedule of activities, including timelines (all in 2006) and identification of responsible parties, is provided in the following table.

Activity and Responsible Party	Date
NMFS provides background documents (Annex I) to moderator, panelists, and CIE	September 11
Moderator and panelists participate in a conference call with MAI to discuss technical and logistical details (depending on availability of participants). This call shall be arranged by the CIE.	September 21
Moderator and panelists read background documents	September 24
Moderator and panelists meet at MAI to test AIM model	3 days during week of September 25
Panelists write draft individual reports (Annex II); moderator begins summary report (Annex III)	October 2-13
Panelists provide draft individual reports to moderator for summarization and to CIE for review	COB October 13
Moderator provides draft summary report to CIE	October 27
CIE approves final individual reports, submits them to the moderator and the NMFS COTR	October 27
NMFS COTR approves individual reports; CIE provides final pdf versions to NMFS COTR and to moderator	November 1
Moderator provides draft summary report to CIE	November 8

CIE approves final summary report, submits it to NMFS COTR	November 22
NMFS COTR approves summary report, CIE provides pdf version to NMFS COTR	November 30

Submission and Acceptance of CIE Reports

The CIE shall provide the final individual and summary reports for review and approval to the NMFS COTR, Dr. Stephen K. Brown via e-mail (Stephen.K.Brown@noaa.gov), according to the schedule above. Approval by the COTR shall be based on compliance with this Statement of Work. The COTR shall notify the CIE via e-mail regarding acceptance of the reports. Following the COTR's approval, the CIE shall provide pdf-formatted copies of the reports to the COTR via e-mail.

ANNEX I: Documents to be reviewed in preparation for the AIM review.

Document Titles
1. IEEE article *Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts*, Adam S. Frankel, William T. Ellison, and Jacquin Buchanan.
2. Presentation given by Dr. Adam S. Frankel at the International Marine Environmental Modeling Seminar (IMEMS) 2004 Conference: *AIM modeling of oil industry activities to derive marine mammal take estimates*.
3. Acoustic Integration Model (AIM) Robustness Review Document
4. Acoustic Integration Model (AIM) Users Manual.
5. The Acoustic Integration Model (AIM): *Applications to predicting and reducing acoustic exposures of marine mammals*, Adam S. Frankel, William T. Ellison. Abstract and presentation for the Navy Environmental Planning and Natural Resources Symposium, May, 2006.
6. Draft Guidance on the Development, Evaluation and Application of Regulatory Environmental Models, The Council for Regulatory Environmental Modeling (CREM)

Document	Document Type	Number of Pages	Degree of difficulty
1. IEEE Article	PDF	6	Low-medium
2. IMEMS Presentation	Powerpoint	28 slides	Low
3. AIM Robustness Review Document	Word/PDF	51	Medium-high
4. AIM User's Manual	Word/PDF	56	Medium
5. EPNR presentation	Powerpoint	33 slides	Medium
6. CREM guidelines	PDF	60	High

ANNEX II: Panelist Report Generation and Procedural Items

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

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

3. The report shall also include as separate appendices the bibliography of materials provided by NMFS for the review and all additional references cited, 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

ANNEX III: Moderator's Summary Report Generation and Process

1. The summary report shall include an overview of the review process.

2. The summary report shall provide a synopsis of the three panelist reports.

3. Points of agreement and disagreement among the panelists shall be documented.

4. The summary report shall also include as separate appendices copies of each of the panelists' reports.

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

REVIEW OF THE ACOUSTIC INTEGRATION MODEL (AIM) FOR DETERMINING NOISE EXPOSURE LEVELS FOR MARINE MAMMALS

Jeanette A. Thomas, Ph.D.
Department of Biological Sciences
Western Illinois University-Quad Cities
3561 60th St.
Moline, IL 61265

13 October 2006

Submitted to:
The Center for Independent Experts
4600 Rickenbacker Causeway
Miami, Fl 33140

Executive Summary

Worldwide, there are growing concerns about the possible detrimental effects of anthropogenic noise on marine mammals. There are a variety of human-generated noise sources in the ocean, such as sounds from military activities, shipping, fishing, recreational vehicles, oceanographic studies, marine construction, and noise associated with exploration, production and transport of oil & gas. In 1997, because of the need to predict underwater noise levels received by marine mammals exposed to anthropogenic noise, Marine Acoustics Inc. (MAI) started developing the Acoustic Integration Model (AIM).

AIM is an application tool to help "users" predict the zone of influence of their sound source on the species of marine mammals typically present in a particular area and time of year. The components of AIM basically provide inputs at the Source-Path-Receiver levels. AIM creates a virtual environment through which animats mimic the dive behavior of specific marine mammal species and can be used to estimate the received levels at any point in time or at any location between the sound source and the animat. Input data for AIM can be pulled from long-term databases, the peer-reviewed literature, or input by the user for very specific cases. AIM draws on two areas that have been studied on a long-term basis: a) change in oceanographic features such as water temperature and salinity by depth over time and bottom contour characteristics, and b) models of sound propagation as a function of these oceanographic features. AIM inputs data from two readily available worldwide, oceanographic databases: GDEM provides water temperature, salinity data by depth and ETOPO2 provides bottom topography. Having access to these data, the user can select one of three sound propagation models employed by AIM; the "Cookie Cutter" transmission loss equation, the Parabolic Equation (PE) and BELLHOP. Glue code was written to interface the access to the GDEM data base and the ETOPO2 bathymetry data with the selected sound propagation model. The novel application of AIM is the incorporation of animats that are programmed to mimic the dive behavior of marine mammals based on actual data collected on a given species.

Currently, AIM has been used to estimate the noise fields of a sound source in three Environmental Impact Statements (EIS), the US Navy LFA (SURTASS Low Frequency Array) sonar tests, and by Minerals Management Service to predict the noise exposure of marine mammals to explosive removal of offshore structures and air gun array operations. In addition, MAI plans to use AIM to assess shipping traffic noise in Stellwagen Bank National Marine Sanctuary in cooperation with National Marine Fisheries Service (NMFS) and to address the cumulative effects of multiple seismic operations in the Gulf of Mexico.

NMFS requires an independent peer-review of AIM, which will assess whether AIM correctly implements the models and data upon which it is based, whether animal movements are adequately simulated within AIM's animat program, and

whether AIM meets the Council for Regulatory Monitoring guidelines for models, which primarily involve scientific credibility.

The Center for Independent Experts (CIE) was contracted by NMFS to find reviewers and coordinate a review process for AIM. Dr. Patrick Cordue, of Innovative Solutions, Ltd., was the Moderator of the review panel. Dr. Wayne Getz, Department of Environmental Science, Policy, & Management at the University of California, Berkeley, Dr. Michael Porter, Heat, Light, and Sound Research, Inc., and Dr. Jeanette Thomas, Department of Biological Sciences, Western Illinois University-Quad Cities, were the members of the Review Panel. The panelists were selected based on their respective expertise in underwater acoustics, modeling, and marine mammalogy.

A three-day meeting was held at MAI in Arlington, VA from 25 to 27 September 2006. Representatives from MAI gave presentations and handouts that described: 1) the history of marine mammal exposures estimation methods for underwater sounds, 2) an overview of AIM operation and features, 3) tests conducted by MAI to test the robustness of AIM, and 4) an evaluation of AIM relative to the Draft Guidance on the Development, Evaluation, and Application of Regulatory Environmental Models (CREM).

AIM has drawn on years of published studies on noise in the ocean to input the parameters that best describe sound's sources; yet, AIM also allows the user to input their own, perhaps proprietary, signal characteristics. In addition to the option of using the basic, generalized "Cookie Cutter Model" of sound propagation (20 log r), which assumes spherical spreading out to 10 km then cylindrical spreading thereafter, AIM allows the user to select two other sound propagation models. PE is best for low frequency signals and BELLHOP is more appropriate for high frequency sounds. Both PE and BELLHOP have has been used extensively and tested by the US Navy. The animat, or simulated marine mammals, used by AIM are based on the best available data on dive profiles of free-ranging marine mammals. AIM has the ability to update code for any of these components or allow users to input their specific data at any of these components.

AIM was developed through a scientifically credible manner. Personnel at Marine Acoustics Inc. have reviewed the structure and performance of AIM throughout its development since 1997. They have adapted AIM to be specific for different user needs, as projects were provided to them. Developers of AIM have made presentations at 2 scientific meetings, published 7 white-paper reports, published 6 peer-reviewed reports, and have a pending patent for AIM.

AIM has received sufficient peer review in a variety of formats; presentations of the program at the Acoustical Society of America and the International Whaling Commission; publications in respected peer-reviewed journals, such as *Animal Conservation, Reports of the International Whaling Commission, Journal of the*

Acoustical Society of America, and *IEEE Oceans*, through three Environmental Impact Statements, which are highly reviewed by both specialists, by the Office of Protected Resources, Permit Office, of the National Marine Fisheries Service, and published in the Federal Register and a period for public comment . The EIS review produced no negative comments that would indicate problems with the application of AIM to accurately determine received levels by marine mammals.

At the Review Meeting, MAI gave a presentation that indicated a logical method of code development with in-house checks. Although the Review Panel requested a written summary of the procedures used by AIM developers to verify code, none was provided by the time of this report.

The Review Panel examined the application of AIM in four ways: 1) by runs conducted before the Review Meeting and discussion by MAI of these results using PowerPoint presentations, 2) by comparisons of AIM predicted received levels with empirical data collected from the sound source, 3) by on- the- spot runs requested to demonstrate AIM performance at the Review Meeting, and 4) by a post-meeting request to run AIM using a simple test case model to compare the output of all three propagation models used by AIM (Cookie Cutter, PE, and BELLHOP).

In any model, the output of a model is only as good as its inputs and data processing by each component. Each of AIM's components is based on the best available data; however, data inputs to AIM are not very site/time specific. For example, although the oceanographic data input to AIM is the best available data, it is based on a monthly average, which does not provide a very sensitive measure. Data input for animats is based on only a very few studies of a few species and the examples tested in this review were only on entirely marine cetaceans, no tests on amphibious pinnipeds were conducted. As data collection abilities improve, so will AIM's accuracy.

The most supportive runs of AIM were:
1) Five Tests Cases of the AIM/PE propagation models simulating variations in sound source frequency, sound source depth, bottom contour, and latitude/longitude. The same 5 Test Cases were re-run, using AIM/BELLHOP propagation model. MAI's presentation of these test cases demonstrated AIM's ability to accurately use the PE or BELLHOP models to calculate TL, AIM's ability to accurately input SVP data from GDEM and input bathymetry data from ETOPO2. AIM accurately detected and adapted the TL when an up slope or down slope bottom was incorporated. AIM accurately notified the user when the signal frequency was inappropriate for the propagation model specified by the user. AIM accurately placed the sound source at the user specified depth, latitude and longitude.

2) Test runs on an artificial situation with a totally absorbing bottom were requested by Dr. Michael Porter after the meeting. AIM predicted that the

number of Level A takes (i.e. exposures to the user specified 180 dB level) was 2 for each of the three propagation models. The number of Level B takes (160 dB) was 35, 39, and 36 for the Basic 20 log r, BELLHOP, and PE propagation models, respectively. Basic 20 log r produced a lower number number of takes than the other two models. MAI attributed the differences in Level B takes due to the fact that AIM could not be programmed for this artificial situation of a completely absorbing surface. MAI noted that when the bootstrap-t resampling parameter was applied the variance between the numbers of takes was reduced to less than one animat among models.

At the meeting, the on-the-spot test cases worked well, with a few exceptions: 1) Two groups of animats were programmed to dive at two different depth ranges, but the outputs from AIM showed some diving at all depths. This problem was corrected by a code change at the meeting and a re-run showed animats diving at two distinct depths. 2) In a test run of the BELLHOP model it was discovered that AIM was reading only the real part of the BELLHOP pressure field. Later, AIM code was corrected and subsequent runs showed AIM's propagation models using PE and using BELLHOP were in good agreement. 3) In one run to examine animat behavior, a close inspection of individual animats showed that occasionally an animat could become "stuck" in a location, such as a bay. This was explained by the programmed aversion angle of the animat to be 180 degree reversal in course, changing the course to 135 degrees eliminated the problem.

In this report, I suggest an additional 14 ways that AIM could be refined or applied to examine the behavior of marine mammals around anthropogenic noise. These additions relate to incorporating an ambient noise or animal-produced background factor that would reflect possible masking. Also, incorporating "generic ducts" or "generic deep scattering layers" into the model would help simulate the marine mammal's real world encounters. Other additions relate to incorporating the time-variable behavior of marine mammals, such as diel haul out patterns, day/night differences in behavior, and seasonal differences in behavior by some species.

I also recommend some additional test cases that would examine the behavior of animats to ensure they do not become "cornered" in the user-specified grid, to understand how animal swim speed affects received levels, and to ensure that animats initial placement on the grid is random, regardless of the grid size and shape.

I recommend that AIM should be thoroughly tested to ensure that each of the input variables related to animat movement is biologically reasonable. Some sample tests could include: 1) examine what happens if the user specifies a dive time that is shorter than the time it takes an animat to move at a specified speed to and from a selected depth, 2) given the dive pattern programmed for a particular species, examine whether the animat performs with surface respirations and resting times typical of that species, 3) ensure that AIM alerts the user that depth selections are not proper for a species-specific dive profile, 4) examine how the probability of exposure changes if the surface time is varied, 5) document how the probability of exposure changes if the number of dives per time is varied, 6) explore how the probability of exposure changes if the animat remained near the surface in a duct; especially if the sound source was in deep water, 7) review how the probability of exposure changes if AIM was run with two species separately versus with two species simultaneously exposed to the same sound, and 8) give some thought as to whether the input variables for individual animats are appropriate for an animat that represented a pod or subgroup.

I also emphasize that to verify AIM estimated received levels it is important to obtain more empirical data from on-site receivers and compare the AIM estimated received levels with the actual received levels at the site.

Some adaptations of AIM are limited by the lack of detailed behavioral data on free-ranging marine mammals, especially at empirically measured received levels. AIM will continue to be perfected as it is used in applications with different sound types, in different marine environments, and with different species of marine mammals.

I appreciate the opportunity to review AIM. It is a greatly needed tool that has many applications in the field of anthropogenic noise on marine life. I applaud MAI for taking the initiative to develop AIM and wish them continued success in upgrading it as more data becomes available.

Background

Worldwide, there are growing concerns about the possible detrimental effects of noise on marine mammals (Richardson et al., 1995; National Research Council 1994, 2000, & 2003; Cox et al., 2006). There are a variety of human-generated noise sources in the ocean, such as sounds from military activities, shipping, fishing, recreational vehicles, oceanographic studies, marine construction, and noise associated with exploration, production and transport of oil & gas (Greene and Richardson 1988; Richardson et al. 1990a, b; HESS 1999; Harris et al. 2001; Blackwell et al., 2004a). Sound in water has a much greater transmission distance than in air, so the effects of noise need to be examined at long ranges from the source. Anthropogenic noise varies in signal frequency characteristics, source level, and whether it is transient or continuous (Harris 1998). For example, seismic exploration includes repeated pulses that can ensonify an area for hours or days. Noise from shipping is almost continuous is some areas of the ocean (Mansfield 1983). Fishing contributes sounds from speed boats, winches, generators, and engines to the ambient environment. Some gill-net fishers use acoustic pingers to deter marine mammals from stealing fish from their nets (Kastelein et al. 2000, 2001). Some aquaculture operations broadcast sound to keep pinnipeds away. The military uses short, low frequency tonal sonar signals for testing and training (Frantzis 1998; US Department of Navy 2001). Marine construction produces transient pile-driving sounds (Blackwell et al. 2004b). Recreational vehicles, such as powerboats and jet skis, produce noise that is especially concentrated near the water surface. Oceanographers produce sonar signals to investigate sound propagation characteristics in the water column (Bowles et al. 1994). Marine mammal scientists use playbacks of animal sounds to understand the behavior of animals in the wild (Frankel and Clark 1998; Croll et al. 2001; Lessage et al. 1999; Tyack 1998).

Marine mammals include a variety of taxa: cetaceans (porpoise, dolphins, toothed and baleen whales), sirenians (manatees and dugongs), pinnipeds (sea lions, seals, fur seals, and walrus), and two carnivores (sea otters, family Mustelidae, and polar bears, family Urisdae). Cetaceans are entirely marine (except for a few river dolphin species); sirenians inhabit shallow waters in both marine and freshwater environments. Pinnipeds are amphibious, spending much of their life hauled out of water to rest, rear pups, and breed. Sea otters inhabit intertidal areas and rarely haul out on land. Polar bears are predominantly terrestrial, but feed primarily on pinnipeds and will enter the water for short periods of time.

In 1972, the US Congress passed the Marine Mammal Protection Act (MMPA) that states it is illegal to “take” a marine mammal. Two levels of “take” were defined; level A, which results in death or injury, and level B, which is harassment or the potential to harass. As a result, any activity that could result in a Level A or Level B take of marine mammals needs to apply for a permit or authorization to conduct their activities. The MMPA is periodically reviewed and special conditions for Native American exceptions, public display, fishing, and military activities have been incorporated since the Act was passed.

The US government divided the oversight of the various marine mammal species between two organizations: the Department of Commerce, specifically National Marine Fisheries Service (NMFS) and the Department of Interior, or the US Fish & Wildlife Service (USFWS). The Department of Interior is responsible for polar bears, walrus, sea otters, and manatees. The remaining species are monitored by NMFS. Because NMFS has jurisdiction over the majority of marine mammal species and many anthropogenic activities are related to commerce, most of the policy and research related to the effects of noise on marine mammals has been conducted by NMFS (National Marine Fisheries Service 1995). As a result, there is little information related to the effects of noise on sea otters, walrus, or polar bears. Some data exist on the hearing and effects of noise on manatees.

To examine the possible effects of anthropogenic noise on marine mammals, a sound must be examined at the Source-Path-Receiver levels.

Source-Details about the frequency, amplitude, and time features of a sound must be well-known and available to evaluate possible detrimental effects from anthropogenic noise on marine mammals. Hearing sensitivity in marine mammals is a U-shaped curve, with maximum sensitivity in the middle of the curve and sensitivity declining below and above the frequency range of best hearing. Specifically, data on whether the frequency of the sound is within a species' hearing range and how many dB above threshold is needed to evaluate whether there is a potential for harm (National Research Council 2000, 2003). Even though some anthropogenic noise is high in amplitude it may be sufficiently low in frequency to not be heard well by some species of marine mammals. If the signal is within the hearing range of a species the number of dB above threshold is important to know. For example, humans (Mills et al 1979; Nielsen et al, 1986; Kryter 1994) exposed to a signal at 100 dB re 1 ɥPa above threshold would be at risk for temporary threshold shift (TTS) and at levels of 120 dB re 1 ɥPa could be at risk for permanent threshold shift (PTS).

The time domain of a signal is important because the mammalian ear is adapted to respond to sound with a certain rise time. If the rise-time of the signal is too rapid, hearing damage may occur. However, there is variability in the ability to withstand rise-times in marine mammals, particularly when comparing entirely marine species versus amphibious species that need to listen in both air and water. In the echolocating odontocetes, their ears are highly adapted to receiving high amplitude, short duration pulses (Au 1993). In humans, TTS can be recoverable if there is adequate time between noise exposures, so the duty cycle of the sound is important. Over the lifetime of a human, the cumulative effects of noise exposure can eventually lead to permanent hearing problems (Kryter, 1970; 1994). There is concern that marine mammals may have similar problems.

Path- Marine mammals are adapted to feed and reproduce in a variety of marine niches; so, some species live in shallow water/coastal areas, others in mid-water continental shelf areas, and others in deep water. Marine mammals occupy habitats in polar continents, temperature waters, equatorial waters, and even some inland seas, rivers, and lakes. As a result, the potential for noise exposure needs to be understood at a particular location.

In addition, some marine mammals exhibit temporal variation in their distribution. For example, gray whales (*Eschritius robustus*) and humpback whales (*Megaptera novaeangliae*) migrate between warm tropical waters to breed and cool polar waters to feed. Some species exhibit diurnal variation in their locations. For example, spinner dolphins (*Stenella longirostris*) spend daytime in sheltered bays around the island of Hawaii and move off shore in the evening to forage (probably reflecting the diurnal movements of their prey). Pinnipeds exhibit seasonal, daily, and hourly variations in their distribution. For example, an elephant seal (*Mirounga angustirostris*) will return to rookeries each spring to give birth and breed and exhibit a preferred time of day to haul out on the beach. During the non-breeding season, the animals forage over large areas and their haul out pattern is very different. To examine the effects of potential noise exposure requires a clear understanding of how a particular species' distribution varies by season, day, and hour (Costa et al. 2003).

Receiver-

The effect of noise on a marine mammal depends on its ability to hear a particular frequency/amplitude sound, the animal's behavior at the time of exposure, the duration/amplitude of the exposure, and the animal's previous experience with that noise.

<u>Hearing Abilities--</u>

The underwater hearing abilities have been tested in a few marine mammals. Using the responses of a trained animal, underwater hearing has been tested in these cetaceans: bottlenose dolphins, *Tursiops truncatus (Johnson 1967);* belugas, *Delphinapterus leucas (Awbrey et al., 1988; Johnson et al. 1989; Johnson 1992);* false killer whales, *Pseudorca crassidens (Thomas et al., 1988*); killer whales, *Orcinus orca* (Bain et al., 1993); Risso's dolphin, *Grampus griseus* (Nachtigall et al., 1995); harbor porpoise, *Phocoena phocoena* (Andersen, 1970); Chinese River dolphin, *Lipotes vexillifer* (Wang et al., 1992); Amazonian River dolphin, *Inia geoffrensis* (Jacobs and Hall, 1972); striped dolphin, Stenella coeruleoalba, (Kastelein et al. 2003), and Pacific white-sided dolphin, *Lagenorhynchus obliquidens* (Tremel et al., 1998).

Underwater hearing abilities have been tested in these pinniped species: California sea lions, *Zalophus californianus* (Schusterman 1974; Southall et al. 2000, 2001, 2003, 2004); harbor seals, *Phoca vitulina* (Kastak and Schusterman, 1996); elephant seals, *Mirounga angustirostris* (Kastak and Schusterman 1999);

ringed seals, *Phoca hispida (Terhune and Ronald, 1972)*; harp seals, *Pagophilus groenlandicus* (Terhune and Ronald 1971; 1975); and walrus *,Odobenus rosmarus (Kastelein et al.,2002).* These studies were time-consuming, typically were on a single animal, and responses were influenced by the animal's attention to the hearing task.

More recently, hearing has been tested at the electrophysiological level using auditory brainstem responses (ABR). Electrodes placed on the animal's skull receive hundreds of responses to a test tone at the neurological level, they are averaged, and a hearing curve can be produced in a short a time. Species tested to date include: the spinner dolphin (*Stenella longirostris*); bottlenose dolphin (*Tursiops truncatus*); false killer whale, *Pseudorca* crassidens, (Yuen et al., 2005); beluga whale, *Delphinapterus leucas, (Supin et al., 2001*); common dolphin (*Delphinus delphis*); Amazon River dolphin (*Inia geoffrensis*); killer whale, *Orcinus orca, (Szymanski et al. 1995; 1999)*; Pacific white-sided dolphins, *Lagenorhynchus obliquidens;* Risso's dolphin (Nachtigall et al., 2005); and the harbor porpoise (*Phocoena phocoena*) (see Bullock et al. (1993) and Supin et al., 2001). It should be emphasized that the accuracy of these audiograms varied among studies, ranging from tests of 1-dB to 6-dB steps; most often 3-dB steps have been used as an amplitude difference at which an individual can reliably distinguish hearing versus not hearing a test signal.

Because of their taxonomic diversity, marine mammals exhibit a variety of hearing abilities and behavioral responses to noise (Richardson et al., 1995). In fact, Ketten (1994, 1997) recommended that species be examined relative to "ear type", rather than strict taxonomic groups. Essentially, there are high-frequency toothed whales, mid-frequency toothed whales, low-frequency baleen whales, pinnipeds in water and pinnipeds in air. Pinnipeds typically have more sensitive hearing under water than in air.

Recent studies by Schlundt et al. (2000) and Finneran et al. (2002b) examined the possible amplitude levels at which exposure to noise could result in Temporary Threshold Shift (TTS) in bottlenose dolphins. TTS indicates a reduction in hearing sensitivity that is recovered over time, and the greater the amplitude and duration of exposure, the longer the recovery time. If exposures are repeated or of sufficiently high amplitude, the hearing loss can be a PTS or not recoverable. Because of the MMPA restrictions and concerns by Animal Rights groups it is unlikely that experiments will be intentionally conducted to document the level of noise exposure needed to induce PTS. This makes setting a specific criterion for limiting noise exposure to marine mammals to prevent PTS difficult. Current criteria are based on interpretations of behavioral responses by animals to high amplitude noise or anatomical evidence of injury or death.

The possible responses of marine mammals to explosions or transient impulses requires a better understanding of the marine mammal's ear mechanics when exposed to short duration, rapid rise-time signals. Finneran et al. (2000) examined the effects of impulsive noise on belugas and bottlenose dolphins.

Finneran et al. (2003) examined behavioral responses of California sea lions to single underwater impulses from an arc-gap transducer.   Criteria for a “take” of marine mammals might need to be different for an impulsive noise versus a continuous sound.

<u>Behavior at the time of Exposure--</u>
If exposed to noise, animals have the choice to leave the area, alter their behavior in response to the noise, or ignore the noise. The response depends on the animal’s previous experience with the noise. Richardson et al., (1995) described a variety of responses by marine mammals to anthropogenic noise. Some animals avoid an area or dramatically change their course to avoid anthropogenic noise. Some animals initially respond to a noise, but become habituated or accustomed to the noise with repeated exposures.   In some situations, particularly in important feeding or breeding areas, animals seem to ignore the noise, perhaps because the area is so important to their fitness.

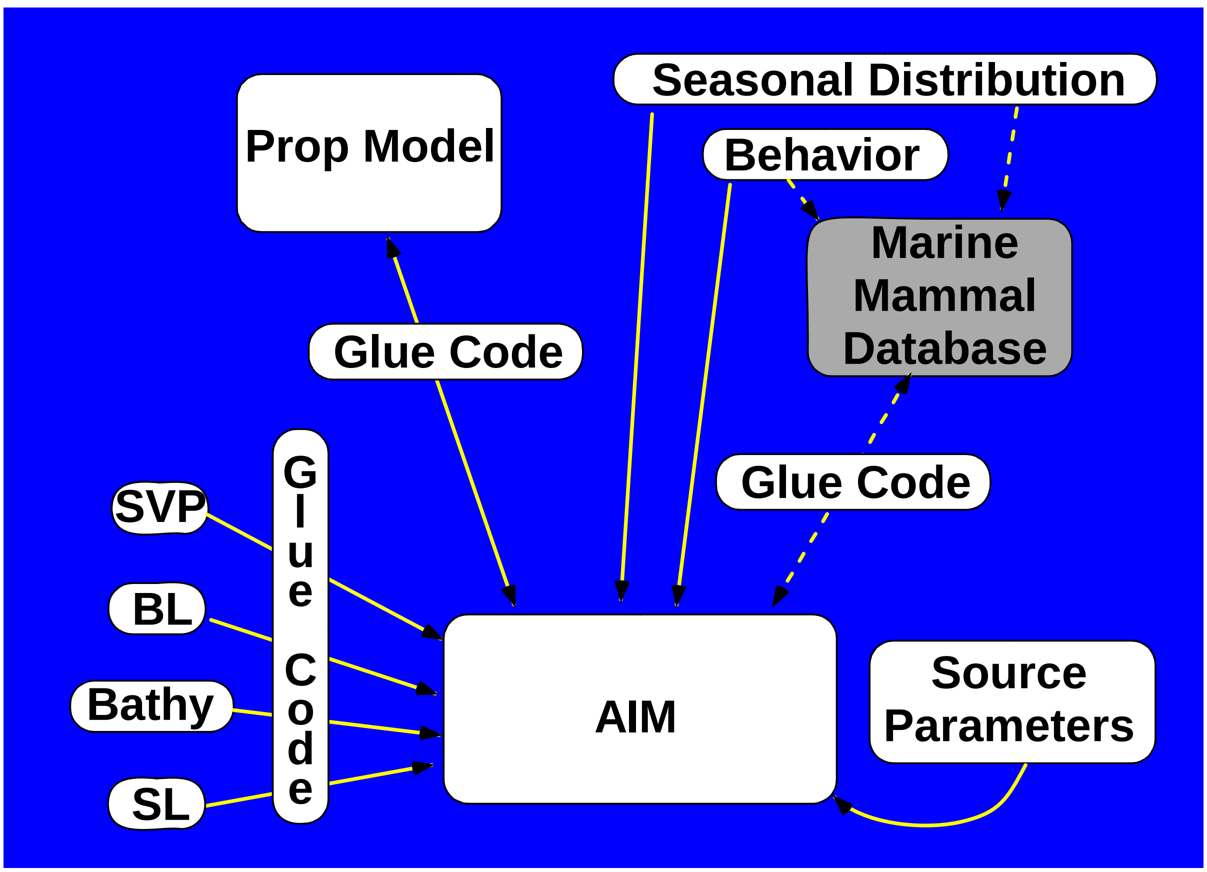
Three situations add to the problem of interpreting whether marine mammals are detrimentally affected by noise: 1) many of the species produce communication and/or echolocation sounds well above 180 dB re 20 ųPa, 2) anthropogenic noise has evolved into a “dinner bell” for some species that learn to approach the sound of a winch hauling in fish, and 3) some species of delphinids voluntarily expose themselves to loud noise when they ride the bow of a vessel for hours to days at a time.

<u>Duration/Amplitude of the Exposure—</u>
Although a transient noise may generate a startle response, change in respiration rate, or deviation in movement in an animal, if only heard sporadically, these short-term responses probably have little potential to affect the animal’s biological fitness. However, if an animal is exposed on a regular basis to anthropogenic noise of sufficient amplitude there are concerns about stress that could affect an animal’s ability to adequately feed or breed. In more extreme situations, there are potentials for TTS or for repeated TTS to develop into PTS. In extreme cases, the beaching and subsequent death of beaked whales was proposed to have been caused by exposed to high amplitude US Navy and NATO sonar signals (Frantzis 1998; NOAA and US Navy 2001).

To prevent needless exposure of marine mammals to potentially harmful anthropogenic noise, a good method or tool for being able to estimate exposure levels by distance from a noise source and the number of individuals of a given species that would be exposed at a certain location, season, day, or hour is needed. Such a tool would be of great value for regulatory agencies to be able to issue operation permits or restrict operations in a consistent manner, for sound users to be able to detect and mitigate noise exposures, and for researchers to understand marine mammal behavior,

Goals of the Acoustic Integration Model (AIM)

Because of the need to predict underwater noise levels for marine mammals exposed to anthropogenic noise, Marine Acoustics Inc. (MAI) started developing the Acoustic Integration Model (AIM) in 1997. AIM was developed as an application tool to help "users" predict the zone of influence of their sound source in a particular area and on the species of marine mammals typically present in a particular area and time of year. Figure 1 shows the components of AIM, which basically provide inputs at the Source-Path-Receiver levels. Currently, AIM has been used to estimate the noise fields of a sound source in three Environmental Impact Statements (EIS), the US Navy LFA (SURTASS Low Frequency Array) sonar tests [3], [13], [14], and was used by Minerals Management Service to predict the noise exposure of marine mammals to explosive removal of offshore structures and air gun array operations [2]. In addition, MAI plans to use AIM to assess shipping traffic noise in Stellwagen Bank National Marine Sanctuary in cooperation with NMFS and to address the cumulative effects of multiple seismic operations in the Gulf of Mexico [2].



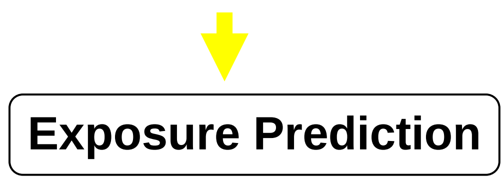


Figure 1. Framework of the AIM model.

Review Activities

Moderator and Review Panel--

Given this AIM model has the potential to become the "industry standard" for permit applications submitted to NMFS, the Center for Independent Experts (CIE) was contracted by NMFS to find reviewers and coordinate a review process for AIM. Dr. Patrick Cordue, of Innovative Solutions, Ltd., was the Moderator of the review panel. Dr. Wayne Getz, Department of Environmental Science, Policy, & Management at the University of California, Berkeley, Dr. Michael Porter, Heat, Light, and Sound Research, Inc., and Dr. Jeanette Thomas, Department of Biological Sciences, Western Illinois University-Quad Cities, were the members of the Review Panel. The panelists were selected based on their respective expertise in underwater acoustics, modeling, and marine mammalogy.

Ken Hollingshead was the representative for National Marine Fisheries Service, and Stephen Brown was the NMFS Contract Technical Representative. Presentations and background documents on AIM were provided by the following MAI representatives: Clay Spikes, Adam Frankel, Jim Messegee, Lee Shores, Steve Labak, and Jack Buchanan.

Review Process—

On 22 September 2006, a conference call was held with the Moderator, Review Panel, Ken Hollingshead, Clay Spikes and Adam Frankel to inform the Panel of available background documents for the meeting, discuss logistics of the meeting, review the meeting agenda, and provide an opportunity for the Review Panel to ask questions about the review process.

<u>Review Day 1 (25 September 2006)—</u>
Representatives from MAI gave presentations and handouts that described [16-19]: 1) the history of marine mammal exposures estimation methods for underwater sounds, 2) an overview of AIM operation and features, 3) tests conducted by MAI to test the robustness of AIM, and 4) an evaluation of AIM relative to the Draft Guidance on the Development, Evaluation, and Application of Regulatory Environmental Models (CREM).

<u>Review Day 2 (26 September 2006)—</u>
Jack Buchanan of MAI gave a presentation on the history of developing AIM. To evaluate whether AIM meets the CREM guidelines, the Review Panel requested that MAI provide: 1) a time-line of AIM development that showed how components were incorporated into AIM and 2) documentation of ways that AIM's programming code was verified and how AIM was tested for robustness and sensitivity throughout its development. The time-line was emailed to the Moderator and Review Panel after the meeting [2]. At the time of this report, materials for item 2) have not been received.

The Review Panel asked for an on-the-spot demonstration of AIM operation using two test cases:

Run 1) No sound source, only a test of 10,000 animats selected to imitate the dive pattern of sperm whales (which have 3 different dive types), run the simulation for 500 times, with an artificial bottom contour (a hypothetical 5 by 5 degree grid (or 300 by 300 mile) box named "Atlantis") in an open ocean environment. Eighteen runs were examined. Run time was 5 minutes 48 seconds. Histograms of the number of animats by depth and by latitude and by longitude were reviewed for initial runs and the last and next to last run. Adam Frankel pointed out that the last run had a different number of data files than the previous runs, so it was not a good comparison [10].

Run 2) Same as Run 1, except that it incorporated animats in both offshore and in a complex near shore environment. This run had 1028 animats inshore that dove up to 50 m depths for (75 seconds per dive) and 375 animats that were offshore and dove from 50 to 150 m depths (for 75 seconds per dive). The run was simulated 500 times. Run time was 48 seconds. Eight data files were generated and histograms of the number of animats by depth and by latitude and by longitude were reviewed for initial runs and the last and next to last run. A problem was detected with a few animats diving deeper than specified, in both near shore and offshore environments [10].

Run 3) Test case generated by MAI (see Statement of Work [Appendix II] and [10]). A 400-Hz sonar projected in North Carolina, an area with abundant bottlenose dolphins. The source level was 210 dB re 20 ųPa, and the ship was moving in a 40-km saw-tooth pattern and conducting survey tracks at 5 km apart. This run was to examine the traditional "Cookie Cutter" sound propagation model of a sound less than 10 km from the source as 20 log range (i.e., assumed spherical spreading) and at ranges greater than10 km estimate propagation as 10 log range or sometimes estimate as 15 log range (i.e., assumed cylindrical spreading). This run used 2,000 animats in a 200 nautical mile box and resulted in 8 output data files. This run produced 720 estimates of a received level of 1,972 animats up to 10 km from the source (i. e., used 20 log range) and at ranges greater than 10 km provided a histogram of the number of animals exposed by 5-dB bins, which could be expanded between 150- and 165 dB. The run estimated that if 160 dB was an exposure criterion, two animats would be exposed. AIM provided results in a top view and a side view [10, A, B] and a zoom in view [10C].

Run 4) Same as Run 3, except used PE propagation model with an image out to 5 km [10D & H], and BELLHOP with an image out to 5 km with semi-coherent mode [10 G & I] and BELLHOP with an image output to 5 km with the coherent mode [10 J]. The image [10 K & M] updated the color schemes for PE and BELLHOP runs to be comparable. Graph [10 L] provided the best comparison of the bottom loss calculations by AIM for PE versus BELLHOP, which were nearly identical. BELLHOP run out to longer distance of 20 km showed the appearance of a duct [10 N & O]. A similar duct was detected using the PE propagation model out to 20 km from the source [10 P].

Review Day 3 (27 September 2006)—

Rerun of Run2) Based on overnight fixes to AIM code, Run 2 from yesterday was demonstrated for some animats in the 0-50 m depth and some in the 50-150 m depth. This run took 4.5 minutes and the problem with a few animats diving deeper than specified was fixed. Run showed 375 animats near shore and 555 off shore, and none were below 150 m. Fifteen data files were generated and histograms of the number of animats by depth and by latitude and by longitude were reviewed for initial runs and the last and next to last run. Data for inshore and offshore animats were examined separately. Looking at near shore data only, the distribution of animats by received level looked very similar between the 2^{nd}, 3^{rd}, 13^{th} and 14^{th} files.

For this run, the Review Panel asked to view the behavior of individual animats. From this examination, at least one animat was "stuck" inside a bay along the coastline. This is believed to have resulted from the boundary behavior specified in the animat's behavior of 180 degree movement change when it encounters a boundary, like the shoreline. By changing the aversion angle to 135 degrees the single animat became "unstuck".

The Review Panel requested that Run 2 of BELLHOP be conducted again, with the coherent versus semi-coherent settings of AIM. This run was conducted and emailed to the Review Panel after the meeting. This rerun indicated that the existing AIM software as of 26 September 2006 was reading only the real part of the BELLHOP pressure field. After AIM code was corrected, the AIM's propagation models using PE and using BELLHOP were in good agreement. Unfortunately, the code corrections were not a version of AIM at the meeting site that would allow further test runs, except by email.

The Review Panel requested that a dataset be selected and AIM run using the three different propagation models selectable in AIM: 20 log range, PE, and BELLHOP to test for comparable results. These runs were emailed to the Review Panel after the meeting.

Run 5) Rerun of Run 2 to compare performance of BELLHOP versus Parabolic Equation (PE) propagation models. BELLHOP seemed to produce an accurate propagation path; however, at less than 1 km from the source the PE model produced an unexpected propagation pattern that transmitted into the bottom, according to the display.

Run 6) Rerun of Run 2, but incorporated a surface duct and expanded the view to detect surface ducting fairly far from the source.

Review Activities After the Meeting—

On 9 October, Dr. Michael Porter requested a 7-step test case with the goals to: 1) verify that all 3 acoustic propagation models produced identical results in a simple case, 2) verify that the near-field solution used with the PE model smoothly blends into the PE results, and 3) verify that the acoustic field produced by all 3 models is processed in an identical way by AIM (i.e. that animats receive the same exposure level with the 3 methods). The simple case was to: 1) set up a fictitious Level A criterion, 2) disable the boundary reflections in PE and BELLHOP by putting in a lot of reflection loss, 3) have an ideal bathymetry, 4) make the sound velocity profile iso-velocity, 5) set the parameters so that 20 log r is used to infinity, 6) produce Transmission Loss by range and depth with all 3 models on the same slice, and 7) produce exposure metric on all three models.

Dr. Porter believed that steps 2-4 should make all 3 models produce identical propagation as a 20 log r field.

Comparison of the output for all three models relative to the selected Level A criteria should produce identical results.

Adam Frankel conducted the 9 October 2006 requested test run and reports his findings [11]. The run was based on a sonar simulation outlined in Example 1 of Annex IV of the CIE Statement of Work [Appendix II]:

A stationary 400-Hz sonar source operating off the North Carolina coast in May. Broadcasts were produced once per minute with a source level of 230 dB re 1 ɥPa. An isovelocity sound profile is used. The bottom loss was 40 dB for all angles. The surface wind was 365 kt was used to maximize the surface loss. The animats were bottlenose dolphins, 3910 animats, step time was 30 sec, and the model ran for 12 hours in each of the 3 models. However, AIM can not specify a surface loss greater than 11 dB using the PE model. BELLHOP's surface reflection values were high, starting at 0.87 and decreasing. In summary, AIM could not specify a perfectly absorbing surface; this should increase propagation of sound. So, it is not possible to directly compare the two models (PE or BELLHOP) with 20 log r.

Frankel [11] reported the transition from 20 log r to PE is not smooth. AIM cannot be used to force PE to calculate transmission loss using a total absorbing bottom or surface without creating a new version of software. PE models showed some energy reflecting off the bottom and surface that is not accounted for in the 20 log r model –compare [11.1] and [11.3].

Interestingly, in the table of [11], the number of Level A takes (180 dB) was 2 as predicted in each of the three propagation models. The number of Level B takes (160 dB) was 35, 39, and 36 for the Basic 20 log r, BELLHOP, and PE models, respectively. Basic 20 log r produced slightly less number of takes than the other two models. Frankel felt that the differences in Level B takes were due to the fact that AIM could not be programmed for this artificial situation of a completely absorbing surface.

Frankel [11] noted that when the bootstrap-t resampling parameter was applied, the variance between the numbers of takes was reduced to less than one animat among models.

Summary of Findings

TOR 1: Assess whether the AIM implementation is correct

AIM is an application tool that inputs data about the sound source, propagation path, and receiver that allows the user to look forward in time to predict the received levels of a particular sound source by distance away from that source. AIM creates a virtual environment through which animats mimic the dive behavior of specific marine mammal species and can be used to estimate the received levels at any point in time or at any location between the sound source and the animat. Input data for AIM can be pulled from long-term databases, the peer-reviewed literature, or input by the user for very specific cases. AIM draws on two areas that have been studied on a long term basis: a) change in oceanographic features such as water temperature and salinity by depth and bottom contour characteristics over time, and b) modeling sound propagation patterns as a function of these oceanographic features. The novel application of AIM is to incorporate the behavior of marine mammals into the model.

Implementation of propagation models--

One of the major contributions of AIM is the ability to use site-specific Sound Velocity Profile (SVP) data and a user-selected propagation models to predict received levels at various distances from a source. In past EIS and research studies, scientists have simply used the “Cookie Cutter Model” of assuming underwater sound transmission was best predicted by using 20 log Range at distances less than 10 km from the source and 10 log range at distances further than 10 km from the source. However, this approach has received criticism for being too general and not being able to detect unpredictable sound levels that can occur when the SVP are included in the model. Depending on the SVP, sounds can propagate further than expected, as in a surface duct, and can create acoustic shadow zones where they should be heard but are not, or create a convergent zone where sound paths intersect to create unexpected areas of high noise.

SVP depends on the salinity, temperature and depth at a give location. Conductivity, temperature, depth (CTD) profiles can be collected with expendable units or equipment mounted on permanent buoys. There are general trends about SVP for tropical, temperature, and polar regions, but the SVP can vary by time of day, season, or in el Niño versus la Niña years. So, the best input of for AIM is SVP data that are collected as close to the time and place of the sound projection. Collection of these data at sea is costly and therefore not necessarily available to all users. Some users may have proprietary oceanographic data that they alone would use as an input to AIM.

AIM uses the Generalized Digital Environmental Model (GDEM), 30-year old, data base of depth, water temperature, salinity, and the resulting sound velocity profile that is easily accessible and covers most of the world. GDEM provides average water temperature and salinity data by depth for a given month of a given year. Lee Shores reported that GDEM is considered the best available sound speed database and is well documented and updated by the US Navy [19].

Bottom contour affects sound propagation, so AIM uses the ETPOP2 bathymetry database, a readily available NOAA data source for bottom contour characteristics. AIM retrieves the closest point on the ETOPO2 database and generates the range versus depth dataset. Lee Shores stated it was the best publicly available source for bathymetric data.

AIM produces graphical 3-dimensional results and tabular values by depth, latitude, and longitude. Inputs of source depth, location and bathymetry can not be entered into the same grid, so AIM will shift these values to the nearest available grid point, thus creates an offset characteristic of AIM. Lee Shores said this should not be a problem if the user is aware of the offset, knows the SVP and bathymetric features, and knows enough about basic acoustic propagation to identify if the offset is causing an unusual propagation pattern.

AIM allows the user to select among three methods to predict the received level by distance from the source: 1) the Cookie Cutter Model, 2) the Parabolic Equation (PE) model long used and perfected by the US Navy, and 3) the BELLHOP model (Porter, 1987). The PE model works best at low frequencies

The way that sound propagates depends on its duration, amplitude, frequency, and duty cycle. AIM allows the user to input these specific features of the sound source: source level, beam pattern, frequency, duration, and movement parameters.

Currently, NMFS evaluates whether a “take” occurs based on a single exposure to a sound. AIM allows the ability to input multiple exposures.

Implementation of conceptual model--

Currently, AIM is best run by personnel at MAI who are very familiar with the programming and user-specified settings of AIM.

To create AIM, Marine Acoustic Inc. (MAI) assembled a highly qualified group of computer programmers, acoustic propagation experts, modelers, and marine mammalogists. Much of the acoustic propagation model and the oceanographic database are “off the shelf”, currently available for use in a variety of applications other than AIM. It is the integration of data on marine mammal dive behavior and the “glue code” that links the components together that makes AIM a new application tool.

Data extraction--

The Review Panel examined the Data extraction abilities of AIM in four ways: 5 test cases of the PE propagation model presented at the Review meeting, 5 test cases of the BELLHOP propagation model presented at the Review meeting, by requesting special on the spot runs of AIM to be observed at the Review Meeting, and runs of AIM requested after the Review Meeting.

1) Five Test Cases of the PE Propagation Model--

At the review (reference presentation), Jim Messegge provided a summary of the history and limitations of the PE propagation model [18]. PE calculates Transmission Loss (TL) of a sound source for use by AIM using these input variables: sound speed vs. depth profile at a given range along a track, bottom depth by range along a track, sediment thickness and bottom loss along a track, surface loss versus range along a track, and source frequency, depth and vertical beam pattern. He pointed out the limitations of PE as being most applicable for low frequency signal sources, unexpected results may occur if rapid changes in environmental variables are input, it should be used in water depths of greater than 2 wavelengths of the sound source, and uses a very basic treatment of bottom sediment thickness. PE is a state-of-the-art TL model that has been well tested, documented, and updated by the US Navy.

So, the reviewers need to examine whether AIM faithfully replicates results of TL as used by PE in other applications. At the review, Lee Shores provided evidence from 5 test cases that compared PE model output run on the OAML system versus with the PE model output generated by AIM [19].

Test case 1-- intentionally ran a high-frequency sound source of 900 Hz at 60 feet deep, with 6 elements spaced at 0.8 m apart, with SVP data from April and the source and receiver at 00 21.208 S and 000 12.106 W. PE does not operate properly at high frequencies. AIM correctly notified the user that the source frequency was too high for PE to reliably be used. Once the user lowered the test frequency to 500 Hz there was excellent agreement between PE run on OAML versus run on AIM, regardless of the bearing from the sound source. This test case verified that AIM correctly calculated SVP from GDEM, the beam pattern generated by the 6-element array, and the ability of a user to input a bottom loss curve to the model.   Only minor differences appeared at the 270 degree bearing from the sound source and at TL levels less than 80 dB.

Test case 2—ran a very low frequency source of 50 Hz at a depth of 60 feet, with 10 elements space at 20 m apart using SVP data from December at 00 22.200 N/179 47.100 W. This tested the ability of AIM to extract data from tracks crossing 0 degrees N and 180 degrees W, tested whether the beam pattern generated from the 10 element spacing was greater than half wavelength, and whether the user could input a bottom loss curve with predictable results. The tests showed AIM made minor adjustments to latitude/longitude of sound source because of the AIM offset characteristic (AIM output was 00 20.0N/179 45.0W). AIM made minor adjustment to source depth (AIM offset the source to 68.9 ft). When bottom loss was inaccurately input at greater than 1 dB/degree, AIM identified the error and requested the user to correct the bottom loss value. The beam patterns of individual elements were reproduced accurately. This test case also supported the ability of the user to input wind speed and AIM to correctly calculate TL.

Test case 3—ran a 1000 Hz signal from a 60 ft deep, omni directional source, using SVP data from October at 25 00.4N/122 20.3E. The bathymetry for this run was an up slope. The test showed that latitude/longitude offset of the source was accurate (AIM location at 25 0.0N/122 20.0E). AIM made a minor adjustment to the source depth (AIM offset the source to 72 ft). AIM correct extracted SVP data for the correct location and season. AIM bathymetry data matched that from ETOPO2. This test case did show that AIM underestimated TL by a minor amount above 80 dB.

Test case 4—ran a 3000 Hz signal at a depth of 20 feet from an omni directional source using July SVP data at 33 07.2N/078 10.6W. The bathymetry for this run was a down slope. The test showed that latitude/longitude offset of the source was accurate (AIM location at 33 05.0N/078 10.0 W). AIM made a minor adjustment to the source depth (AIM offset the source to 23 ft). AIM correct extracted SVP data for the correct location and season. AIM bathymetry data matched that from ETOPO2. This test case also verified that AIM correct set the “Horran” value in this shallow water situation.

Test case 5—ran a 800 Hz omni directional signal at a depth of 20 ft. at 21 43.1N/158.47.6W and this sound was projected near an island. Rather than using GDEM database for SVP calculations the user input SVP data, wind speed, and bottom low. Bathymetry was taken from ETOPO2. This test case showed that latitude/longitude offset of the source was accurate (AIM location at 21 45.0 N/158 50.0 W). AIM made a minor adjustment to the source depth (AIM offset the source to 23 ft). The user input of SVP, wind speed and bottom loss were all translated by AIM correctly.

In summary, Lee Shore's presentation of these 5 test cases demonstrated AIM's ability to accurately use the PE model to calculate TL, AIM's ability to accurately input SVP data from GDEM and input bathymetry data from ETOPO2.

In a second presentation by Lee Shores at the Review Meeting, he compared the ability of AIM to accurately apply the BELLHOP propagation model.

Five test cases of the BELLHOP Propagation Model—

At the review [19], Lee Shores provided a summary of the history and limitations of the BELLHOP propagation model, created by M. Porter (1987) to provide fast ray tracings of TL of high frequency sound sources. BELLHOP calculates Transmission Loss (TL) of a sound source for use by AIM using these input variables: sound speed vs. depth profile at a given range along a track, bottom depth by range along a track, and bottom loss along a track, surface loss versus range along a track, and source frequency, depth and vertical beam pattern. He pointed out the limitations of BELLHOP as being range independent, in that results are valid for constant environments or short ranges. The user input step size must be small to prevent irregular ray coverage and BELLHOP should be used in water depths greater than 10 wavelengths of the sound source frequency. BELLHOP is a state-of-the-art TL model that has been well tested, documented and updated by the US Navy.

So, the reviewers needed to examine whether AIM faithfully replicates results of TL as used by BELLHOP in other applications. At the review, Lee Shores provided evidence from 5 test cases that compared BELLHOP model output run on the OAML system versus with the BELLHOP model output generated by AIM. In his presentation, Lee Shores listed 15 constants that AIM sets in running BELLHOP (see presentation use this page as a figure). As with the PE propagation model, AIM does not allow the source depth, location and bathymetry data to be in the same grid points, so AIM creates an offset that must be recognized and understood by the user. Results from the AIM run of BELLHOP are computed in three dimensions: depth, latitude, longitude.

Test Case 1-- Same scenario as Test Case 1 for AIM/PE model evaluation, except ran at 500 Hz sound source at 60 feet deep, with 6 elements spaced at 0.8 m apart, with SVP data from April and the source and receiver at 00 21.208 S and 000 12.106 W. User input bottom loss and default wind speed was used. The test showed that latitude/longitude offset of the source was accurate (AIM location offset to 00 20.0S/000 10.0W). AIM made a minor adjustment to the source depth AIM offset the source to 68.8 ft). AIM correctly input user specified bottom loss. Examination of the beam pattern of two of the six sources using GBEAM adapted by AIM versus beam patterns predicted by use of other software (CASS/GRAB) indicated negligible differences.   Surface loss due to wind speed was examined by the AIM calculation using GBEAM software and by using the Bechmann-Spezzichino surface loss model of CASS/GRAB and there was little difference between the two outputs. The AIM out of SVP from BELLHOP was essentially the same as the GDEM generated SVP profiles for the same data. There were a few small differences in bathymetry output from AIM/BELLHOP compared to the ETOPO2, but are believed to be caused by difference in data input methods (closest point vs 4-point average).

Test Case 2-- similar run to Test Case 2 for AIM/ PE model evaluation-- ran a very low frequency source of 50 Hz at a depth of 60 feet, with 10 elements space at 20 m apart using SVP data from December at 00 22.200 N/179 47.100 W. User defined the bottom loss and a default wind speed was used. The test showed that latitude/longitude offset of the source was accurate (AIM location offset to 00 20.0 S/000 10.0 W). AIM made a minor adjustment to the source depth (AIM offset the source to 68.8 ft). AIM correctly input user specified bottom loss. Examination of the beam pattern of two of the six sources using GBEAM adapted by AIM versus beam patterns predicted by use of other software (CASS/GRAB) indicated negligible differences. AIM correctly input user specified bottom loss. The AIM out of SVP from BELLHOP was essentially the same as the GDEM generated SVP profiles for the same data. There were a few small differences in bathymetry output from AIM/BELLHOP compared to the ETOPO2, but these are believed to be caused by difference in data input methods (closest point vs 4-point average). This test case confirmed the ability of AIM/BELLHOP to accurately adjust for user input wind speed.

Test Case 3--similar to run of Test Case 3 for AIM/PE model evaluation-- ran a 1500 Hz signal from a 60 ft deep, omni directional source, using SVP data from October at 25 00.4 N/122 20.3 E. The bathymetry for this run was an up slope. The test showed that latitude/longitude offset of the source was accurate (AIM location at 25 0.0 N/122 20.0 E). AIM made a minor adjustment to the source depth (AIM offset the source to 68.8 ft). AIM correctly input approximated the user specified bottom loss, but resulted in a step-like curve below -10 dB. The AIM out of SVP from BELLHOP was essentially the same as the GDEM generated SVP profiles for the same data, except it extended the SVP beyond the bottom depth. There was a negligible difference in bathymetry output from AIM/BELLHOP compared to the ETOPO2 and AIM correctly described an upslope environment. Some difference were observed between AIM's bottom loss curve versus BELLHOP's curve using "standalone" inputs; however this was a very shallow water environment and results of any model are often variable. Standalone refers to a run of a model, like PE or BELLHOP outside of AIM with the goal to test the propagation model performance outside AIM to the output generated by AIM using the same data

Test Case 4-- similar run to Test Case 4 of AIM/PE model evaluation-- ran a 3000 Hz signal at a depth of 20 feet from an omni directional source using July SVP data at 33 07.2 N/078 10.6 W. The bathymetry for this run was a down slope. The test showed that latitude/longitude offset of the source was accurate (AIM location at 33 05.0 N/078 10.0 W). AIM made a minor adjustment to the source depth (AIM offset the source to 22.9 ft). AIM correctly input approximated the user specified bottom loss, but resulted in a step-like curve below -10 dB. The AIM out of SVP from BELLHOP was essentially the same as the GDEM generated SVP profiles for the same data, except the SVP extended beyond the bottom depth. There was a negligible difference in bathymetry output from AIM/BELLHOP compared to the ETOPO2 and AIM correctly described a down slope environment. Some large differences were observed between AIM's bottom loss curve versus BELLHOP's curve using "standalone" inputs; however, this was a very shallow water environment and results of any model are often variable. This was not an acceptable match. BELLHOP was sensitive to AIM induced source depth offset and/or bottom loss precision in this very shallow water environment.

Test Case 5-- similar to run of Test Case 5 of AIM/PE model evaluation-- ran a 800 Hz omni directional signal at a depth of 20 ft. at 21 43.1N/158.47.6W and this sound was projected near an island. Rather than using GDEM database for SVP calculations the user input SVP data, wind speed, and bottom low. This test case showed that latitude/longitude offset of the source was accurate (AIM location at 21 45.0 N/158 50.0 W). AIM made a minor adjustment to the source depth (AIM offset the source to 23 ft). The user input of SVP, wind speed and bottom loss were all translated by AIM correctly. There was a very small difference in precision of AIM/BELLHOP to calculate bottom loss. Bathymetry was taken from ETOPO2; three of four radials from the sound source were correctly predicted by AIM/BELLHOP; however, there was a small difference at the 000 degree radial from the source (perhaps due to the differences of input as the closest point (by AIM) versus the 4-point average). There was good agreement between AIM's bottom loss curve versus BELLHOP's curve using "standalone" inputs.

Summary of Data Extraction--

Exercising the PE model run by AIM and the BELLHOP model run by AIM in a variety of locations, bottom contours, source frequencies, and water depths tested AIM's ability to adapt to changing environments. In all cases AIM was able to extract sufficiently accurate SVP data, provide an acceptable source depth, provide an acceptable source latitude and longitude, re-create the beam pattern of individual elements in the sound source. Also, user input for SVP and wind speed were acceptably translated by AIM. While bottom loss was less accurate in shallow water environments, this is a problem that many propagation models encounter.

TOR 2: Assess the animal movement simulation within AIM

Features of the AIM Animat Simulation of Marine Mammal Behavior

Distribution of marine mammals—

AIM allows the input of the specific species of marine mammals that might be present in an area of noise exposure. Also, the expected density of a particular species in the area of interest and the time of year are input to AIM. These data are gathered from a long-term database from NMFS, from continued reading of the peer-reviewed literature by MAI personnel, or can be input by the user.

One advantage of AIM is that it allows the user to interpret the noise exposure from a single sound source on more than one species. Sometimes the species may be different enough in their densities, dive behavior, or hearing abilities that one species may be consider at risk of a "take" whereas other species in the same area are not.

Behavior of marine mammals--

Data on the behavior of free-ranging marine mammals in response to noise is very limited. Ship time is expensive, encounters with some species are totally opportunistic, many species are difficult to capture, and present methods of attaching data collection devices are not very successful. The best data on behavior of free-swimming marine mammals are collected in one of three ways:

1) Shore-based tracking of marine mammal movements using theodolites in conjunction with underwater acoustic monitoring with hydrophones allows the ability to synchronize the movement path and dive behavior of an animal at the time of a noise exposure.
2) Time-depth recorders (TDR) attached to an animal collect a dive history or a record of swimming at various depths and locations. The unit can be retrieved from the animal and data downloaded or can be linked by a satellite tag and remotely collected when the animal surfaces. Recently, some "acoustic, TDR" tags incorporated an audio-sensor to record the sounds to which animals are exposed and these data are synchronized with their dive profile.
3) A critter cam is a video camera with audio recording abilities that is housed in a waterproof case that can be worn or carried by a trained marine mammal.

Because tag attachment is more secure and long-lasting in pinnipeds, most reliable TDR data are from pinnipeds. Whales, like the gray, bowhead, and humpback whales that migrate in predictable locations each year, have been successfully studied using theodolite tracking.

AIM uses the best currently available data on marine mammal dive behavior and responses to noise. MAI continues to monitor the peer-reviewed literature for updated studies and incorporates new data into the AIM model as applicable.

Fairly detailed dive profile data are available on elephant seals (Costa et al. 2003), sperm whales, humpback whales, gray whales, fin whales, bowhead whales, right whales, and bottlenose dolphins. AIM imitates movement of marine mammals in four dimensions: three-dimensional space and time by use of “Animats” that are programmed to mimic the dive profiles of a given species of marine mammals [5,6,7]. The user can specify whether the dive profile is divided into phases (like shallow versus deep dive), the relative use of time in these phases, the maximum/minimum time at the surface, and the time between surfacings to breathe. The angle that the marine mammal dives after a surfacing can be selected. The animals’ maximum/minimum swim speed can be selected. The user can specify boundaries of the animats, like water surface, land, bottom, or depth. The user can specify the “heading variance” or the change in course that the animat should make when it reaches a boundary. In the case of migratory animals, like gray whales, humpbacks, or bowheads, where their movement path is designed to cover a great distance, the user can specify a slight heading variance, such as 10 degrees. In contrast to a foraging animal that would stay in the same general area could be assigned a heading variance of 45 to 60 degrees to simulate this scenario. The user can specify aversions or restrictions to the animat movement pattern that could be used to restrict its movement, ex. to avoid a drill platform, a sea mound in a particular area, or an island [16,17].

At the Review [8], MAI provided a document by Frankel and Vigness-Raposa (2006) that tabulates the best available data on marine mammal behavior that is the basis for user input to AIM. The list is limited to species in US waters. This is an accurate summary of currently available data. In some cases more detail is available, such as being able to describe humpback whale behavior during migration, feeding, by singing whales in wintering areas, by a calf, and by an adult. This is an accurate catalog of available data on marine mammals relative to anthropogenic noise.

I suggest removing the name “blackfish” on page 23. This is not a taxonomic rank, includes species from four different genera, and rather is a common name used by fishers to describe mid-sized black whales. The Marine Animal Behavioral Analysis [8] includes the sea otter and walrus, even though NMFS does not have jurisdiction over these species. If AIM wants to include marine mammal species that are not monitored by NMFS, they should consider adding the growing body of literature on the behavior and movements of the Florida manatee. Unlike cetacean dorsal fins, the manatee’s spoon-shaped fluke provides an ideal attachment site for a satellite-liked TDR. Long term movements of manatees along the Atlantic and Gulf of Mexico coasts of Florida would provide a unique opportunity the test AIM in a shallow water and continental shelf environment. In addition, there is great concern about the effects of anthropogenic noise and injury by boat propellers on this endangered species.

Evaluating the Robustness of the Animat Performance of AIM

The robustness of the animat perform of AIM was evaluated in three ways: 1) presentation at the Review that demonstrated the behavior of individual animats, 2) presentation at the Review on a set of hypothetical animat exposures based on situations encountered by specific marine mammals and 3) by simulations requested to be run on the spot at the Review.

<u>Behavior of Individual animats—</u>

At the Review meeting [16,17], Adam Frankel demonstrated the performance of individual animats located near Hawaii's, with a bottom depth of 19,000 ft, in March, with SVP input from GDEM and bathymetry input from ETOPO2. The AIM extraction of SVP and bathymetry was in good agreement. Animats were examined by four selectable features:

Dive Depth and Dive Angle--Three animats were set to dive at different depths and with different dive angles. Each animat maintained its own depth and the tracks of movement reflected the angles of dive specified for a given animat.

Linearity of Travel—Four animats were set to move to travel at a different initial course and then turn at a specified time. The resulting linearity index for each animat was calculated. The animat that set out at 0 degrees and at every 300 seconds was not to change course produced a straight line track and a linearity index of 1.0. The animat that started at 90 degrees and changed course every 90 seconds produced a linearity index of 0.39 and exhibited the most irregular movement path. Another animat started an initial course at 30 degrees and changed path every 300 second, produced a linearity index of 0.94. Another animat started the course to 60 degrees and changed path every 150 seconds, produced a linearity index of 0.64. The tracks of these last two animats were as expected; both headed in the approximate same direction, with the latter animat having a more jagged path because of its changing course more often.

Aversion Angle—This selectable feature in AIM was tested by programming animats with different aversions (180, 135, 90, 45, and 0 degrees) to the 20 0.0 N latitude line. Tracks of individual animats corrected displayed their programmed aversion angles. Another example allowed animats to have multiple aversions; in this case some animats were to stay in shallow water and some to stay in deeper water. The output from this simulation showed a clear distinction between inshore and offshore animat locations.

Density over Time—AIM sets animats to occupy a certain grid and does not allow animats to leave the grid, rather when encountering a boundary, like land or a latitude or longitude the animat reverses direction. Adam Frankel showed a plot of the animate density over time in 6 different test areas. The normalized density of animats was centered on one over a 24-hour test period. The normalized density of animats varied from 0.92 to 1.1 among areas, but over time the normalized density converged around 1.0 for each area. So, it is easy to conclude that the number of animats in an area is remaining stable over time.

Species-specific dive behavior—AIM allows the input of dive profile (time versus depth history) for a particular species. There are sufficient data for a few species of marine mammals to allow AIM to generate a “typical” dive profile for a species. At the Review, Adam Frankel showed three examples of species-specific dive profiles based on original data versus the AIM animat simulation of this pattern: beaked whales (*Ziphius*), foraging right whales (*Eubalaena*), and spotted dolphins (*Stenella*). AIM animats adequately simulated the details or patterns of the original dive profiles for each species. In addition, AIM animats produced a very similar histogram of the number of dives by depth. However, it would have been useful to statistically compare the distributions. The dive profile for the foraging right whales was very interesting because it showed consecutive dives with time spent at a depth about 120 m, even though the bottom was deeper. This could indicate the depth of the copepods layer on which the right whales were feeding.

Number of Animats Needed—In AIM the user must specify the number of animats in the area. This can be based on “real world” densities of a species in the area of interest at the time of year desired to operate a sound source. At the review, Adam Frankel produced a simulation of the number of exposures by received Levels for 3 different densities of animats (0.01, 0.10 and 1.0). At the lowest density, no animats were exposed above195 dB. At the highest density a few animats were exposed at levels up to 216 dB and animats at the intermediate density were exposed up to 213 dB. So, the density of animats does affect the predicted number of exposures, and a very low density may imply no animats exposed.

A boot-strap-t procedure was describe by where the user could predict the number of animats to run to ensure enough samples were taken and provide a distribution of number of exposures and be able to calculate a 95% confidence interval around the distribution. This seems like a good procedure to estimate the number of animats needed for a run, but it was unclear to me whether this was a separate program or an optional calculation made within AIM.

<u>Sensitivity Study of AIM animat model-</u>-

In the Review [16,17] Adam Frankel described two sensitivity studies conducted with AIM animats - two runs in two different environments (a downward refracting and a ducted area). Fast, shallow-swimming bottlenose dolphin animats and slow, deep-diving sperm whale animats were run with densities of 0.1 animats/km^2. Course deviation was 30 degrees over 5 minutes. The sound source was stationary and acoustic exposures were compared using Kolmogorov-Smirnov tests. Three variables were examined for the two animat species in the two environments: course change, speed, and depth. Statistically significant differences in the distribution of both species in both environments were found by speed and depth. However, for course change only the sperm whale in the ducted environment behaved significantly different.

<u>Hypothetical Animat Exposures—</u>

At the Review [16,17], Adam Frankel provided a "real world" test case of the animat exposure to an array of seismic air guns, which operates with repeated short, impulsive, broadband signal and produces a three-dimensional beam pattern. He noted that AIM operates in only a two-dimensional beam pattern; however, this is believed to be sufficient for predicting RL by range from the source. The output of AIM was compared to three existing empirical data bases: EARS buoy, d-tag, and streamer data. The EARS buoy collected data of 5 seismic survey tracks run between 90.4 and 90.3 degrees longitude and 27.61 and 27.68 latitude. The track distances from EARS buoy were 0, 500, 2000, and 5,000 m. The RL by range in meters from the source were calculated by AIM/PE, AIM/BELLHOP and compared to the received levels by the EARS buoy for each track. For each track, the expected decline in RL by range was predicted by AIM/PE and AIM/BELLHOP.

Empirical Data from EARS buoy
The EARS buoy was an automated bottom-mounted recorder that can be retrieved for data collection.

Line 0--The best match of AIM received levels compared to empirical data was near the source (Line 0); however, at distances of less than 500 m from the source, AIM/PE and AIM/BELLHOP predicted slightly lower received levels than EARS buoy measured. For line 0, at 2,000 m from the source the AIM/BELLHOP generated received levels that were higher than empirical data reported. At 2,500 m from the source AIM/PE was a better predictor of received level and tracked the empirical data well.

Line 500 m—At distances of < 1,000 m from the source, the AIM/BELLHOP generated received levels were much lower than those measured by the EARS buoy. The match improved at 1,500 to 2,500 m from the source but at greater distances continued to underestimate received levels by as much as 15 dB at 4500 m from the source. For this line, PE was a much better predictor of the received level and tracked changes in the empirical data fairly well.

Line 1,000 m—Both AIM/BELLHOP and AIM/PE calculated received levels agreed with empirical received levels up to 3,500 m from the source. At greater distances the AIM/BELLHOP model consistently underestimated the received levels, by as much as 12 dB at 4,500 m from the source. In contrast, the AIM/PE generated values at all distances from the source.

Line 2,000 m—At distances < 3,000 m from the source AIM/PE underestimated the received level. At distances < 3,800 m AIM/BELLHOP estimated the actual received levels fairly well. At distances > 3,500 m from the source AIM/PE provided fairly good agreement with empirical measurements of received level by the EARS buoy.

Line 5,000 m –At this range, models were not run at closer than 5,000 m from the source. The AIM/BELLHOP model consistently under estimated the received level by as much as 10 dB at 5,000 meters from the source. The AIM/PE model under estimated received levels at distances of 5,000 to 5,200 meters from the source. At greater distances the AIM/PE model was quite variable, but at least occasionally calculated received levels in line with the empirical data.

Empirical Data from Kondor and D-tag

The *M/V Kondor* is a seismic industry vessel that towed an industry standard air gun array. The D-tag, equipped to sample audio frequencies at 96 kHz sampling rate, was attached to a surface buoy and was deployed at a depth of 150 m to record sounds from the *M/V Kondor* as it passed the buoy.

At distances of less than 1,500 m from the source, both AIM/BELLHOP and AIM/PE predicted received levels in line with those measured empirically. At greater distances AIM/PE was not a good predictor; predicting too low received levels between ,1500 and 3,500 m and predicting too high levels at distances greater than 4,000 m from the source. AIM/BELLHOP produced a more gradual received level curve that predicted the received level fairly well out to 4,000 m, but thereafter predictions were lower than the observed values from *Kondor* and D-tag.

Empirical Data from Veritas Vantage Streamer

The *M/V Veritas Vantage* was a seismic industry vessel and monitored signals levels were taken from hydrophones within the array or streamer.

In this test, it is clear that the AIM/BELLHOP calculations adequately predict the received levels measured by the streamer. AIM/PE under estimates received levels at ranges < 4,000 m from the source, and it over estimates received levels at ranges greater than 4,000 m.

It would have been good to use some statistical analysis to compare the AIM generated received levels versus the empirically obtained received levels. The reviewers only could visually observe comparisons and try to make conclusions. Based on empirical data, there was no clear trend about when one model would be preferred over another.

In summary, the tests of AIM received levels by range compared to empirically collected data varied by range from the source. Most often PE was a better predictor of received levels, except in the case of data from the *Veritas Vantage* streamer, where BELLHOP seemed to provide a very good fit. Neither BELLHOP nor PE worked well in shallow water environments. As Dr. Frankel indicated, MAI should explore the use of other propagation models that may be more suited for the air gun arrays.

TOR 3: Assess whether AIM meets the Council for Regulatory monitoring guidelines for model development

Applicability of CREM guidelines to AIM

The Acoustic Integration Model (AIM) is an application tool that inputs data about an underwater sound source, the propagation path, and receiver that allows the user to look forward in time to predict the received levels by distance away from that source. AIM creates a virtual environment through which animats mimic the dive behavior of specific marine mammal species and can be used to estimate the received levels at any point in time or at any location between the sound source and the animat. Input data for AIM can be pulled from long-term databases, the peer-reviewed literature, or input by the user for very specific cases. AIM draws on two areas that have been studied on a long-term basis: a) change in oceanographic features such as water temperature and salinity by depth and bottom contour characteristics over time and b) modeling sound propagation patterns as a function of these oceanographic features. AIM inputs data from two readily available worldwide, oceanographic databases (Figure 1): GDEM provides water temperature, salinity data by depth and ETOPO2 provides bottom topography. Having access to these data, the user can select one of three sound propagation models employed by AIM; the “Cookie Cutter” transmission loss equation, the Parabolic Equation (PE) and BELLHOP. Glue code was written to interface the access to the GDEM data base and the ETOPO2 bathymetry data with the selected sound propagation model. The novel application of AIM is the incorporation of animats that are programmed to mimic the dive behavior of marine mammals.

AIM is undergoing the CREM review to ensure that there are no features that would make it consistently liberal or conservative in estimating received levels from a sound source.

Have the principles of credible science been addressed during the development of AIM?

Yes, AIM broke down the task of predicting received levels by marine mammals near a sound source by examining the characteristics of three components: the sound source, propagation path, and receiver. AIM has drawn on years of published studies on noise in the ocean to input the parameters that best describe sounds sources; yet, AIM also allows the user to input their own, perhaps proprietary, signal characteristics. In addition to the option of using the basic, generalized “Cookie Cutter Model” of sound propagation (20 log r), which assumes spherical spreading out to 10 km then cylindrical spreading thereafter, AIM allows the user to select two other sound propagation models; PE which is best for low frequency signals and BELLHOP which is more appropriate for high frequency sounds. Both PE and BELLHOP have has been used extensively and tested by the US Navy. The animat, or simulated marine mammals, used by AIM are based on the best available data on dive profiles of free-ranging marine mammals. AIM has the ability to update code for any of these components or allow users to input their specific data at any of these components.

Personnel at Marine Acoustics Inc. have reviewed the structure and performance of AIM throughout its development since 1997 [2]. They have adapted AIM to be specific for different user needs, as projects were provided to them. Developers of AIM have made presentations at 2 scientific meetings, published 7 white-paper reports, published 6 peer-reviewed reports, and have a pending patent for AIM [2].

Does the conceptual model have an adequate theoretical basis?

Yes, by compartmentalizing the task into the sound source, propagation path, and receiver characteristics, AIM developers have been able to employ the best theories, available data, best models, and tests to evaluating each independently. Then, the performance of the AIM system was tested as an integrated whole.

Has the conceptual model been adequately implemented?

Yes, given that AIM is based on the best available data, this conceptual model has been adequately implemented. MAI and the review panel discussed many ways that AIM could be adapted for a variety of assessment tasks and to make the animat behavior more real world (also see my Additional Comments below). These adaptations are largely limited by the lack of detailed behavioral data on free-ranging marine mammals, especially at empirically measured received levels. AIM will continue to be perfected as it is used in applications with different sound types, in different marine environments, and with different species of marine mammals.

Were adequate techniques and procedures used for code verification?

There was limited information to answer this question. The Review Panel could not and did not verify programming code. At the Review Meeting, Jack Bucanan of MAI provided an oral report of the chronology of the steps taken to verify code during AIM development. This presentation indicated a logical method of code development with in-house checks. Unfortunately, although the Review Panel requested a written summary of the procedures used by AIM developers to verify code, none was provided by the time of this report

Has there been adequate peer review of AIM?

Yes, the theories and conceptual models supporting the development of AIM have been reviewed since its inception by personnel at Marine Acoustics Inc. National Marine Fisheries Service (NMFS) has played an important role in consultation about AIM development through discussions with MAI about the need for such a model, the types of underwater sounds sources that marine mammals encounter, situations that seem to have indicated a detrimental effect by anthropogenic noise on marine mammals, and by providing updates to the known distributions of marine mammals in specific areas.

Presentations of components of the AIM model were given at well respected scientific meetings, such as the Acoustical Society of America and the International Whaling Commission. This provided attendees at the conferences to the features and goals of AIM and the opportunity to gain feedback from the audience about the model.

Aspects and background information used in the development of AIM were published in respected peer-reviewed journals, such as *Animal Conservation, Reports of the International Whaling Commission, Journal of the Acoustical Society of America,* and *IEEE Oceans* [2].

AIM was used to model received levels for marine mammal exposures to noise in the Environmental Impact Statement for the Navy's SURTASS-Low Frequency Sonar program [3, 13, 14]. EIS are reviewed by both specialists employed by the applicant and by the Office of Protected Resources, Permit Office, of the National Marine Fisheries Service (NMFS). EIS procedures include publication in the Federal Register and a period for public comment [15]. The EIS review had no negative comments that would indicate problems with the application of AIM to accurately determine received levels by marine mammals.

In addition from 2001 to 2003, developers of AIM received funds from the US Air Force Small Business Innovative Research (SBIR) to develop an AIM-based Sonic Boom application for predicting effects of supersonic aircraft operations on marine mammals. They received an SBIR grant from NAVAIR and a STTR from the Office of Naval Research. Development of AIM components were in collaboration with the Cornell University Bioacoustics Laboratory and the North Pacific Acoustic Laboratory, both highly respected facilities specializing in marine mammals and underwater acoustics [2].

From 2003 to 2004, AIM was used for two Minerals Management Service (MMS) projects to predict exposures of marine mammals to 1) explosive removal of offshore structures and 2) to air gun array operations [2].

In 2005-2006, AIM modeled fish school movements with respect to the development of a biologically based sonar system.

In 2006, AIM is working on modeling ship traffic noise in the Stellwagen Bank National Marine Sanctuary. A proposal was just accepted to employ AIM to address cumulative impacts on marine mammals of multiple seismic operations in the Gulf of Mexico [2].

Lastly, AIM is under review by the Center for Independent Experts [this report].

All the above activities required reviews and consultations with a variety of marine mammalogists, modelers, and specialists in underwater acoustics.

AIM model applications

The application of AIM was tested in four ways: 1) by runs conducted before the Review Meeting and presentations of these results made using PowerPoint presentations, 2) by comparisons of AIM predicted received levels with empirical data collected from the sound source, 3) by on- the- spot runs requested to demonstrate AIM performance at the Review Meeting, and 4) by a post-meeting request to run AIM using a simple test case model to compare the output of all three propagation models used by AIM (Cookie Cutter, PE, and BELLHOP).

Previous text in this report described the presentations given at the review that described 5 test cases of the AIM/PE model, 5 test cases of the AIM/BELLHOP model {pp. 12-17}, a series of runs that tested the robustness and sensitivity of animat behavior {pp.20-22}, and 3 "real world" scenarios that used the entire AIM model the signal source, propagation path, and animat behavior to produce the probability of noise exposure by range from the source.

At the meeting, the Review Panel also requested two runs of animat behavior without a sound source and a set of runs to compare the output of sound propagation using BELLHOP versus PE on the same data set.

Previous text in this report {22-23} described the presentations given at the review AIM predictions that compared to empirical data collected in three different situations: from EARS, a stationary buoy recording as a ship towing seismic air guns passed by at different distances, from the *Kondor* and D-tag, and from a ship, the *Veritas Vantage* towing an acoustic streamer

Was the choice of model appropriate given the quantity and quality of available data?

Yes, the 5 Tests Cases of the AIM/PE propagation model simulated a variation in sound source frequency, sound source depth, bottom contour, and latitude/longitude. The same 5 Test Cases were run, but using AIM/BELLHOP propagation model. This allowed the reviewers to easily compare the results from different sound types, in varied environmental conditions, in different locations and between the two propagation models. These runs allowed the reviewers to examine the performance of AIM without the animat component. The propagation models were chosen partly because of their accessibility and because of their long term use by the Navy and oceanographers.

Presentations given at the meeting provided simple runs of the animat component of AIM to examine how the individual behavior of animats changed with depth, course change angle, speed, and turning time. It also demonstrated the behavior of animats near boundaries.

How closely did the model simulate the system of interest?

Both the AIM/PE and AIM/BELLHOP models of propagation performed adequately. As expected AIM/PE operated best at low frequencies.

The data from empirical measures of the sound source by on-site monitoring systems versus the predicted received level estimated by AIM were quite variable. At some distances from the sound source received levels were best estimated by the Cookie Cutter model, at other distances by the PE model, and at other distances by the BELLHOP model. Most often PE was a better predictor of received levels, except in the case of data from the *Veritas Vantage* streamer, where BELLHOP seemed to provide a very good fit. Neither BELLHOP nor PE worked well in shallow water environments. This emphasizes the importance of having a knowledgable user to interpret AIM output. It would have been good to use some statistical analysis to compare the AIM generated received levels versus the empirically obtained received levels. The reviewers only could visually observe and make comparisons and try to reach conclusions. There was no clear trend about when one model would be preferred over another.

How well did the model perform?

In any model, the output of a model is only as good as its inputs and data processing by each component. Each of AIM's components is based on the best available data; however, data inputs to AIM are not very site/time specific. For example, although the oceanographic data input to AIM is comprised of the best available data, the data are based on a monthly average, which does not provide a very sensitive measure. Data input for animats is based on only a very few studies of a few species and the examples tested in this review were only on entirely marine cetaceans, no tests on amphibious pinnipeds were conducted. As data collection abilities improve, so will AIM's accuracy.

To me, the most "telling", supportive test runs were:

1) Lee Shore's presentation of these 5 test cases [19] demonstrated AIM's ability to accurately use the PE model to calculate TL, AIM's ability to accurately input SVP data from GDEM and input bathymetry data from ETOPO2. See specifics of performance in pp. 12-14 of this report.

2) Lee Shore's presentation of these 5 test cases [19] demonstrated AIM's ability

to accurately use the BELLHOP model to calculate TL, AIM's ability to accurately input SVP data from GDEM and input bathymetry data from ETOP. See specifics of performance in pp. 15-17 of this report.

3) a run on an artificial situation with a totally absorbing bottom [11]. The table of the number of Level A takes (i.e. exposures to the user set 180 dB level) that AIM predicted was 2 as estimated by each of the three propagation models. The number of Level B takes (160 dB) was 35, 39, and 36 for the Basic 20 log r, BELLHOP, and PE models, respectively. Basic 20 log r produced slightly less number of takes than the other two models. Frankel attributed the differences in Level B takes was because AIM could not be programmed for this artificial situation of a completely absorbing surface. Frankel [11] noted that when the bootstrap-t resampling parameter was applied the variance between the numbers of takes was reduced to less than one animat among models. This demonstrates good agreement between AIM received level estimates using the three different propagation models.

The inputs at various levels of the model can be quite variable over time and by location. A better question might be: what types of received level estimates should AIM be programmed to produce?, i.e., a single number, a single number $\pm$ X dB, a mean with confidence interval, or a maximum/minimum range? Hearing in marine mammals has been tested in 3-dB steps because it is a level that marine mammals reliably report that they "hear" a difference. So, it may be that at minimum the received levels produced by AIM could be bounded by $\pm$ 3 dB.

Were adequate sensitivity and uncertainty analyses performed?

Yes, some tests of AIM's sensitivity were presented to the Review Panel; however, see Recommendations below for additional runs and checks that should be conducted.

Perhaps the largest concern with using output from AIM is that it provides a single received level at a distance from the source and a single number of animats that would be exposed at a specified "take" level. Given the variability of input data that exists along all three components of the model (source, path, receiver), it would be better to give a range of received levels at a certain distance from the source and a maximum & minimum number of animats that would meet the user-selected "take" level.

Are the models transparent and the results reproducible?

Yes, each part of AIM (animat, PE, BELLHOP, Cookie Cutter, GDEM, ETOPO2) seemed to function reliably and consistently between runs. User inputs are clearly defined. The Review Panel was not provided a way to verify code, nor would I, as a Review Panel member, be qualified to review the code. In two sample runs at the Review Meeting, questionable output was found. However, both situations were quickly fixed by code modification during the meeting.

What processes or procedures exist to enable ongoing model evaluation and improvement?

AIM is being adapted and refined as it is applied to variety of sound sources (continuous vs transient, low vs high frequency), in different marine environments (deep vs shallow water, temperate, tropical, vs polar), and with new species of marine mammals. The strong feature about AIM is that its component design allows it to be adapted to each new situation. In recent work with the US Air Force, AIM has even been used with airborne sound sources.

Additional comments

The patent description for AIM reports several possible ways that AIM could be used to model more than received level by range. I suggest some additional or refinements of AIM for applications to studying marine mammals:

1) AIM does not incorporate ambient noise levels into the model. Could AIM incorporate a user setting of existing ambient noise levels over a certain frequency range that would make the prediction of received levels at a given range more accurate?

2) Similarly, marine mammals are vociferous and certain species at certain times of year contribute significantly to the ambient noise. For example, humpback whales during the singing/breeding season produce high amplitude sounds within the frequency ranges of anthropogenic noise sources. Could AIM develop a way that the user could input species-specific underwater vocalization rates, frequencies, and amplitudes typical of the time and season of the users operations? This could be added on as a possible ambient noise masker.

3) Could the surface time be used to imitate haul out periods in pinnipeds? For example, over a 24-hour period some pinnipeds have a predictable preferred haul out time. Could AIM program blocks of time at certain times of the day and model the number of exposures over a 24-hour day that would reflect that pinnipeds are more likely to be in the water at certain hours?

4) Can AIM model swim speed versus received level to estimate how quickly an animat could leave a received level to move to a tolerable, lower received level?

5) Marine mammals often swim in upper ducts because they either feed on items in the mixed layer or their prey feed on items near the surface. Is there a way to insert a “generic duct” into the sound propagation field produced by AIM to model how this would affect received levels by range? Also, could the user specify the location of the sound source as being within the duct vs. below the

duct, and specify the animats as being within the duct or below the duct?

6) Movements of marine mammals in the water column are often affected by organisms of the Deep Scatter Layer (DSL) (see figure of right whale foraging, [16]). The DSL migrates up and down in the water column on a diurnal basis. The DSL produces noise (such as from snapping shrimp) that could mask a marine's mammal ability to hear an anthropogenic noise source. In addition, the DSL can be an acoustically reflective surface in the path of an anthropogenic noise source. Could AIM incorporate a "generic deep scatter layer" into the sound propagation models that would allow the user to input the size, density and depth of the DSL, as well as the source level of the DSL?

7) In addition to being able to model the received level by range for certain sounds, AIM should be able to predict the zone of audibility for different types of sounds.

8) It would be useful for AIM to incorporate an m-weighting hearing function for the 5 ear types described by Ketten (1994; 1997). With this ability, AIM could predict the received level at an animat, reference a particular species' hearing curve, and calculate for a given frequency the number of dB the source exceeds the hearing threshold. The user could specify a criterion of X dB above threshold that was not to be exceeded, and then calculate the number of animats that exceed that criterion, or the number subjected to TTS or PTS.

9) The critical ratio and critical bandwidth are known for a few species of toothed whales and pinnipeds. If AIM could incorporate the critical ratio and critical bandwidth characteristics for a given species into the model, it should be able to predict the distances at which marine mammal communication or echolocation signals could be masked by certain noise types. This might be specifically applied to biological noise sources, such as snapping shrimp.

10) AIM could be used to calculate the long-term cumulative exposures of noise on animats or the "dose" an animal receives during its typical day or during a whole breeding season. For example, NIOSHA (National Institute for Occupational Safety and Health, 1998) provides guidelines for human exposure to noise in the workplace by calculating an Leq (noise exposure equivalent) statistic. With the Leq, it is possible to determine an average noise exposure over the workday, even though the worker was exposed to high noise levels for 1 hour and quiet office conditions for 7 hours. NIOSHA then sets criteria for the daily Leq value and advises employees to adjust work schedules to stay under that Leq. Perhaps an Leq function for different marine mammal ear types could be used by AIM.

11) AIM could incorporate "daytime animats" and "night time animats". For some species of marine mammals, like spinner dolphins that spend the daytime resting in bays, then move away from the Hawaiian Islands to forage at night, the

behavior of the animals should be described differently. This also could be important to understand the received levels of animats over a 24-hour period, when the sound source is shut down at night.

12) Currently AIM inputs monthly SVP data for a specific location from a long term averaged database; however, it might be useful to adapt AIM to identify El Niño and La Niña years and run comparisons of sound propagation at a location during these two extreme conditions. This would provide the user with two estimates of received levels by range from the source.

13) Since AIM is a virtual tour of sound propagation, it can provide a record of animat behavior during pre-exposure, baseline ambient sounds, record of animat behavior during noise exposures, and a record of animat behavior post-exposure. Thus, AIM could not only predict the number of exposures to a certain received level, but also plot the return to baseline conditions after a noise exposure.

14) Since AIM calculates a linearity index and allows the selection of swim speed, it also could calculate the linear distance moved. These parameters might be used to model the energy expenditure by a particular species of animat. This might help evaluate extra energy expenditures an animal incurs when exposed to noise, i.e., relate to fitness.

Recommendations

Each of the components of AIM need to be tested or "exercised" by running a set of simulations to ensure that AIM is performing as expected. Specifically, AIM should be "exercised" to test animat behavior without a sound source, especially near boundaries. AIM programs an animat to turn about at 180 degrees when it hits a specified boundary, like a latitude, or longitude. Some sample tests of animat behavior without a sound source could include:

1) Flat bottom, open ocean condition, latitude and longitude boundaries, varied by water depth. This would examine whether there might be "corner effects" where because of the right angle boundaries and a particular course change angle an animat might get stuck in a corner (i. e., continue to bounce off the bottom and a latitude/longitude boundaries). If this occurs, could AIM alert the user that an animat is likely to get "cornered" or, alter code to prevent the cornering of an animat?

2) Examine the "stickiness" problem that was seen during the Review meeting, where an animat gets stuck in a repeated bouncing pattern inside a bay. A set of simulations should be run that change the dimensions of bays along a coastline and change the animat's course angles to determine if there are certain scenarios that an animate is "stuck" in a bay. If so, can AIM alert the user that an

animat is likely to get stuck or change the code to prevent getting stuck?

3) Once the speed for an animat is set it can not be changed within a run. AIM should be exercised to understand how an animal's swim speed affects received levels. If one animat swims twice as fast as another animat, but they both spend the same amount of time underwater, is their probability of noise exposure the same?

4) AIM user's guide states that animats are placed initially at random on the grid. Some simulations should be run using different grid sizes and shapes, and the results statistically test to confirm that this is true.

The animat movement simulations of AIM should be thoroughly tested to ensure that each of the input variables related to animat movement are biologically reasonable. Some input variables may have a large affect on received levels and other input variables may be of minimal, or even insignificant, effects on received levels. Some sample tests could include:

1) What happens if the user specifies a dive time that is shorter than the time it takes an animat to move at a specified speed to and from a selected depth? Does AIM notify the user of an incorrect input error?

2) If an animat is programmed to spend X% of time in deep water and X% of time in shallow water and given the dive pattern programmed for a particular species, does the animat perform with surface respirations and resting times typical of that species?

3) Does AIM have the ability to alert the user that depth selections are not selected properly for the species-specific dive profile selected?

4) How would the probability of exposure at certain received level change if the surface time was varied? Surface time could be viewed as a period of mitigation or relief from the sound. Some dolphin species respond to disturbance by milling at the surface.

5) How would the probability of exposure at certain received level change if the number of dives per time was varied?

6) How would the probability of exposure at certain received level change if the animat remained near the surface in a duct, especially if the sound source was in deep water?

7) How would the probability of exposure at a certain received level change if AIM was run with two species separately versus with two species exposed to the same sound during the same time and in the same place? In other words, are the two species modeled separately with no interaction? Can animats of two

different species occupy the same location?

8) Rather than an animat being an individual animal, it could be a pod or subgroup of animals (like a mother/calf pair or a bachelor herd of sea lions). Some thought should be given to if the input variables for animats as individuals would be appropriate for an animat that represented a pod or subgroup. Since many marine mammals travel or feed in social groups, the likelihood of an individual being exposed may be the same as the likelihood of the pod being exposed, and if one member of the pod is exposed, then all are likely to be exposed.

To verify AIM estimated received levels, it is important to obtain more data from on-site receivers and compare the AIM estimated received levels with the actual received levels at the site.

Acknowledgments

I appreciate the opportunity to review AIM. It is a greatly needed tool that has many applications in the field of anthropogenic noise on marine life. I applaud MAI for taking the initiative to develop AIM and wish them continued success in upgrading AIM as more data becomes available.

References (Additional references are provided in Appendix I)

Anderson, S. (**1970).** Auditory sensitivity of the harbour porpoise, *Phocoena phocoena*.
Invest. Cetacea 2, 255-258.

Au, W. W. L. (**1993**). The Sonar of Dolphins (Springer-Verlag, New York).

Au, W. W. L., P. E. Nachtigall, and J. L. Pawloski. (**1997**). Acoustic effects of the ATOC signal (75 Hz, 195 dB) on dolphins and whales. J. Acoust. Soc. Amer. **101**, 2973-2977.

Awbrey, F. T., J. A. Thomas, and R. A. Kastelein. (**1988**). Low frequency underwater hearing sensitivity in belugas, *Delphinapterus leucas*. J. Acoust. Soc. Amer. **84**, 2273-2275.

Bain, D. E., B. Kriete, and M. E. Dahlheim. (**1993**). Hearing abilities of killer whales (*Orcinus orca*) J. Acoust. Soc. Amer. **94(Pt. 2)**, 1829.

Blackwell, S. B., C. R. Greene Jr., and W. J. Richardson. (**2004a**). Drilling and Operational sounds from an oil production island in the ice-covered Beaufort Sea. J. Acoust. Soc. Amer. **116**, 3199-3211.

Blackwell, S. B., J. W. Lawson, and M. T. Williams. (**2004b**). Tolerance by ringed seals (*Phoca hispida*) to impact pipe-driving and constructions sounds at an oil production island. J. Acoust. Soc. Amer. **115**, 2346-2357.

Bowles, A. E., M. Smultea, B. Würsig, D. P. DeMaster and D. Palka. (**1994**). Relative abundance and behavior of marine mammals exposed to transmissions from the Heard Island Feasibility Test. J. Acoust. Soc. Am. **96**, 2469-2484.

Costa, D. P., D. E. Crocker, J. Gedamke, P. M. Webb, D. S. Houser, S. B Blackwell, D. Waples, S. A. Hayes, and B. J. Le Boeuf. (**2003**). The effect of a low-frequency sound source (acoustic thermometry of the ocean climate) on the diving behavior of juvenile northern elephant seals, *Mirounga angustirostris*. J. Acoust. Soc. Amer. **113**, 1155-1165.

Cox, T. M., T. J. Rage, A. J. Read, E. Vos, R. W. Baird, K. Bal comb, J. Barlow, J. Caldwell, T. Cranford, L. Crum, A. D'Amico, G. D'Spain, A. Fernández, J. Finneran, R. Gentry, W. Gerth, F. Gulland, J. Hildebrand, D. Houser, P. D. Jepson, D. Ketten, C. D. MacLeod, P. Miller, S. Moore, R. D. Mountain, D. Palka, P. Ponganis, S. Rommel, T. Rowles, B. Taylor, P. Tyack, D. Wartzok, R. Gisner, J. Mead, and L. Benner. (**2006**). Understanding the impacts of anthropogenic sound on beaked whales. J. Cet. Res. Manag.

7, 177-187.

Croll, D. A., C. W. Clark, J. Calambokidis, W. T. Ellison, and B. R. Tershy. (**2001**). Effects of anthropogenic low-frequency noise on the foraging ecology of *Balaenoptera* whales. Anim. Conserv. **4**, 13-27.

Finneran, J. J., R. Dear, D. A. Carder, and S. H. Ridgway. (**2003b**). Auditory and behavioral responses of California sea lions (*Zalophus californianus*) to single underwater impulses from an arc-gap transducer. J. Acoust. Soc. Amer. **114**, 1667-1677.

Finneran, J. J., D. A. Carder, and S. H. Ridgway. (**2002a**). Low frequency acoustic pressure, velocity and intensity thresholds in a bottlenose dolphin (*Tursiops truncatus*) and white whale (*Delphinapterus leucas*). J. Acoust. Soc. Am. **111**, 447-456.

Finneran, J. J., C. E. Schlundt, R. Dear, D. A. Carder, and S. H. Ridgway. (**2002b**). Temporary shift in masked hearing thresholds (MTTS) in odontocetes after exposure to single underwater impulses from a seismic water gun. J. Acoust. Soc. Amer. **111**, 2929-2940.

Finneran, J. J., C. E. Schlundt, D. A. Carder, J. A. Clark, J. A. Young, J. B. Gaspin, and S. H. Ridgway. (**2000**). Auditory and behavioral responses of bottlenose dolphins (*Tursiops truncatus*) and white whales (*Delphinapterus leucas*) to impulsive sounds resembling distant signatures of underwater explosions. J. Acoust. Soc. Amer. **108**, 417-431.

Frankel, A. S. and C. W. Clark. (**1998**). Results of low-frequency playback of M-sequence noise to humpback whales, *Megaptera novaeangliae*, in Hawai'i. Can. J. Zool. **76**, 521-535.

Frantzis, A. (**1998**). Does acoustic testing strand whales? Nature **392**, 29.

Greene, C. R., Jr. and W. J. Richardson. (**1988**). Characteristics of marine seismic survey sounds in the Beaufort Sea. J. Acoust. Soc. Am. **83**, 2246-2254.

Harris, C. M. (1998). Handbook of Acoustical Measurements and Noise Control, 3rd Ed. (Acoustical Society of America, Huntington, NY).

Harris, R. E., G. W. Miller, and W. J. Richardson. (**2001**). Seal responses to air gun sounds during summer seismic surveys in the Alaskan Beaufort Sea. Mar. Mamm. Sci. **17**, 795-812.

HESS. (**1999**). High Energy Seismic Survey review process and interim operational guidelines for marine surveys offshore Southern California. Rep. from High Energy Seismic Survey Team for Calif. State Lands Comm. and U.S. Minerals Manage.

Serv. [Camarillo, CA] (39 p. plus appendices). Available at: www.mms.gov/omm/pacific/lease/fullhessrept.pdf.

Johnson, C. S. (**1992**). Detection of tone glides by the beluga. Pp. 241-248 *in*: J. A. Thomas, R. A. Kastelein, & A. Ya Supin. Supin (eds.) Marine Mammal Sensory Systems (Plenum Press, New York).

Johnson, C. S., M. W. McManus, and D. Scar. (**1989**). Masked tonal hearing thresholds in the beluga whale. J. Acoust. Soc. Am. **85**, 2651-2654.

Kastak, D. and R. J. Schusterman. (**1999**). In-air and underwater hearing sensitivity of a northern elephant seal (*Mirounga angustirostris*). Can. J. Zool. **77**, 1751-1758.

Kastak, D. and R. J. Schusterman. (**1998**). Low-frequency amphibious hearing in pinnipeds: Methods, measurements, noise, and ecology. J. Acoust. Soc. Am. **103**, 2216-2228

Kastak, D. and R. J. Schusterman. (**1996**). Temporary threshold shift in a harbor seal (*Phoca vitulina*). J. Acoust. Soc. Am. **100**, 1905-1908.

Kastelein, R. A., M. Hagedoorn, W. W. L. Au, and D. deHaan. (**2003**). Audiogram of a striped dolphin (*Stenella coeruleoalba*). J. Acoust. Soc. Amer. **113**, 1130-1137.

Kastelein, R. A., P. Mosterd, B. van Santen, M. Hagedoorn, and D. deHaan. (**2002**). Underwater audiogram of a Pacific walrus (*Odobenus rosmarus divergens*) measured with narrow-band frequency-modulated signals. J. Acoust. Soc. Amer. **112**, 2173-2182.

Kastelein, R. A., H. T. Rippe, N. Vaughan, C. Staal, and N. M. Schooneman. (**2001**). The influence of three acoustic alarms on the behavior of harbor porpoises (*Phocoena phocoena*) in a floating pen. Mar. Envt. Res. **52**, 351-371.

Kastelein, R. A., H. T. Rippe, N. Vaughan, N. M. Schooneman, W. C. Verboom, and D. de Haan. (**2000**). The effects of acoustic alarms on the behavior of harbor porpoises in a floating pen. Mar. Mamm. Sci. **16**, 46-64.

Ketten, D. R. (**1997**). Structure and function in whale ears. Bioacoustics. Int. J. Anim. Sound Rec. **8**, 103-135.

Ketten, D. R. (**1994**). Functional analyses of whale ears: adaptations for underwater hearing. IEEE Proc. Und. Acoust. **1**, 264-270.

Kryter, K. D. (**1994**). The handbook of hearing and the effects of noise. Academic Press, New York). 673 pp.

Kryter, K. D. (**1970**). Effects of Noise on Man. (Academic Press, New York).

Lessage, V., C. Barrette, M. C. S. Kingsley, and B. Sjare. (**1999**). The effects of vessel noise on the vocal behavior of belugas in the St. Lawrence River Estuary, Canada. Mar. Mamm. Sci. **15**, 65-84.

Male, C. I., P. R. Miles, C. W. Clark, P. Tyack, and J. E. Bird. (**1983**). Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior. BBN Rep. 5366. Rep. from Bolt Beranek and Newman Inc., Cambridge, MA, for U.S. Minerals Manage. Serv., Anchorage, AK. NTIS PB86-174174.

Mansfield, A. W. (**1983**). The effects of vessel traffic in the Arctic on marine mammals and recommendations for future research. Can. Tech. Rep. Fish. Aquat. Sci. N0. 1186.

Mills, J. H., R. M. Gilbert, and W. Y. Adkins. (**1979**). Temporary threshold shifts in humans exposed to octave bands of noise for 16 to 24 hours. J. Acoust. Soc. Amer. **65**,

Nachtigall, P. E., M. Yuen, T. A. Mooney, and K. A. Taylor. (**2005**). Hearing measurements from a stranded infant Risso's dolphin, Grampus griseus. J. Exp. Biol. **108**, 4181-4188.

Nachtigall, P. E., A. Ya. Supin, J. L. Pawloski, and W. W. L. Au. (**2004**). Temporary threshold shifts after noise exposure in the bottlenose dolphin (*Tursiops truncatus*) measured using auditory evoked potentials. Mar. Mamm. Sci. **20**, 673-687.

Nachtigall, P. E., J. L. Pawloski, and W. W. L. Au. (**2003**). Temporary threshold shifts and recovery following noise exposure in the Atlantic bottlenosed dolphin (*Tursiops truncatus*). J. Acoust. Soc. Amer. **113**, 3425-3429.

Nachtigall, P. E., W. W. L. Au, J. L. Pawloski, and P. W. B. Moore. (**1995**). Risso's *dolphin (Grampus griseus*) hearing thresholds in Kaneohe Bay Hawaii. Pp. 49-53 *in*: R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (eds.) Sensory Systems of Aquatic Mammals (De Spil, Woerden, Netherlands).

National Institute for Occupational Safety and Health (NIOSH).(**1998**). Criteria for a recommended standard: occupational noise exposure. U.S. Department of Health and Human Services (NIOSH) pub. # 98-126.

National Marine Fisheries Service (NMFS). (**1995)**. Small takes of marine mammals incidental to specified activities; offshore seismic activities in southern California; Notice of issuance of an incidental harassment authorization. Fed. Regist. **60** (200, 17 Oct.), 53753-53760.

National Research Council (NRC). (**2003**). Ocean noise and marine mammals (The National Academies Press, Washington D.C.).

National Research Council (NRC). (**2000**). Marine mammals and low-frequency sound (The National Academies Press, Washington D.C.).

National Research Council (NRC). (**1994**). Low-frequency sound and marine mammals: Current knowledge and research needs (The National Academies Press, Washington D.C.).

Nielsen, D. W., M. J. Bauman, and D. K. Brandt. (**1986**). Changes in auditory threshold during and after long-duration noise exposure: species differences. Pp. 281-293 *in* (R. J. Salvi, D. Henderson, R. P. Hamernik, and V. Colletti, eds.) Basic and Applied Aspects of Noise-Induced Hearing Loss (Plenum, New York).

NOAA and U.S. Navy. (**2001**). Joint interim report/Bahamas marine mammal stranding event of 14-16 March 2000. National Oceanic and Atmospheric Administration (U. S. Department of Commerce) and U.S. Navy, [Washington, D.C.]. 61 p.

Richardson, W. J., C. R. Greene, C. I. Malme, and D. H. Thomson (eds.). (**1995**). Marine mammals and noise. (Academic Press, New York). 576 pp.

Richardson, W. J., B. Würsig, and C. R. Greene, Jr. (**1990a**). Reactions of bowhead whales, *Balaena mysticetes*, to drilling and dredging noise in the Canadian Beaufort Sea. Mar. Env. Res. **29**, 135-160.

Richardson, W. J., C. R. Greene Jr., W. R. Koski, C. I. Malme, G. W. Miller, M. A. Smultea, and B. Würsig. (**1990b**). Acoustic effects of oil production activities on bowhead and white whales visible during spring migration near Pt. Barrow, Alaska -- 1989 phase. LGL Ltd. report for U. S. Minerals Management Service. 284 p. NTIS Pb91-105486.

Richardson, W. J., B. Würsig, and C. R. Greene, Jr. (**1986**). Reactions of bowhead whales, *Balaena mysticetus*, to seismic exploration in the Canadian Beaufort Sea. J. Acoust. Soc. Amer. **79**, 1117-1128.

Richardson, W. J., M. A. Fraker, B. Würsig, and R. S. Wells. (**1985**). Behaviour of bowhead whales, *Balaena mysticetus* summering in the Beaufort Sea: reactions to industrial activities. Biol. Conserv. **32**, 195-230.

Schlundt, C. E., J. J. Finneran, D. A. Carder, and S. H. Ridgway. (**2000**). Temporary shift in masked hearing thresholds (MTTS) of bottlenose dolphins and white whales after exposure to intense tones. J. Acoust. Soc. Amer. **107**, 3496-3508.

Schusterman, R. J. (**1974**). Auditory sensitivity of a California sea lion to airborne sound. J. Acoust. Soc. Am. **56**, 1248-1251.

Southall, B. L., R. J. Schusterman, D. Kastak, and C. R. Kastak. (**2004**). Underwater hearing thresholds in pinnipeds measured over a 6-year period. J. Acoust. Soc. Amer. **116 (4, pt. 2)**, 2504.

Southall, B. L., R. J. Schusterman, and D. Kastak. (**2003**). Auditory masking in three pinnipeds: aerial critical ratios and direct critical bandwidth measurements. J. Acoust. Soc. Am. **114**, 1660-1666.

Southall, B. L., R. J. Schusterman, D. Kastak, and C. R. Kastak. (**2001**). Pinniped hearing and anthropogenic noise. J. Acoust. Soc. Am. **110 (5, pt.2)**, 2722

Southall, B. L., R. J. Schusterman, and D. Kastak. (**2000**). Masking in three pinnipeds: underwater, low-frequency critical ratios. J. Acoust. Soc. Am. **108**, 1322-1326.

Supin, A. Ya., V. V. Popov, and A. M. Mass. (**2001**). The Sensory Physiology of Aquatic Mammals (Kluwer Academic Publishers, Boston, MA). 332 pp.

Szymanski, M. D., D. E. Bain. K. Kiehl, S. Pennington, S. Wong, and K. R. Henry. (**1999**). Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms. J. Acoust. Soc. Amer. **106**, 1134-1141.

Szymanski, M. D., D. E. Bain and K. R. Henry. (**1995**). Auditory evoked potentials of a killer whale (*Orcinus orca*). Pp. 1-10 *in*: R.A. Kastelein, J.A. Thomas and P.E. Nachtigall (eds.) Sensory systems of aquatic mammals (De Spil Publ., Woerden, Netherlands). 588 p.

Terhune, J. M. and K. Ronald. (**1975**). Underwater hearing sensitivity of two ringed seals. Can. J. Zool. **53**, 227-231.

Terhune, J. M. and K. Ronald. (**1972**). The harp seal (*Pagophilus groenlandicus* (Erxleben, 1777). III. The underwater audiogram. Can. J. Zool. **50**, 465-469.

Terhune, J. M. and K. Ronald. (**1971**). The harp seal (*Pagophilus groenlandicus* (Erxleben, 1777). X. The air audiogram. Can. J. Zool. **49**, 385-390.

Thomas, J. A., N. Chun, W. W. L. Au, and K. Pugh. (**1988**). Underwater audiogram of a false killer whale (*Pseudorca crassidens*). J. Acoust. Soc. Am. **84**, 936-940.

Thomas, J. A., J. L. Pawloski, and W. W. L. Au. (**1990a**). Masked hearing abilities in a false killer whale (*Pseudorca crassidens*). Pp. 395-404 *in*: J. A. Thomas and R. A. Kastelein (eds.) Sensory Abilities of Cetaceans (Plenum Press, New York).

Thomas, J, P. Moore, R. Withrow, and M. Stoermer. (**1990b**). Underwater audiogram of a Hawaiian monk seal (*Monachus schauinslandi*). Low-frequency amphibious hearing in pinnipeds: methods, measurements, noise, and ecology. J. Acoust. Soc. Amer. **87**, 417-420.

Tremel, D. P., J. A. Thomas, K. Ramirez, G. S. Dye, W. A. Bachman, A. N. Orban, and K. K. Grimm. (**1998**). Underwater hearing sensitivity of a Pacific white-sided dolphin, *Lagenorhynchus obliquidens*. Aquat. Mamm. **24**, 63-69.

Tyack, P. L. (**1998**). Acoustic communication under the sea. Pp. 163-220 *in* Animal Acoustic Communication (S. L. Hopp, M. J. Owren, and C. S. Evans, eds.), Springer-Verlag, Berlin.

U. S. Department of Navy. (**2001**). Final overseas environmental impact statement and environmental impact statement for Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar (Vol. 1). Available from U. S. Department of the Navy.

U.S. Marine Mammal Protection Act of **1972** (16 U.S.C., §1361).

Wang, D., K. Wang, Y. Ziao, and G. Sheng. (**1992**). Auditory Sensitivity of a Chinese River Dolphin (*Lipotes vexillifer*). Pp. 213-221 *in*: J. A. Thomas, R. A. Kastelein, and A. Ya. Supin (eds.) Marine Mammal Sensory Systems (Plenum Press, New York).

Yuen, M. E., P. E. Nachtigall, M. Breese, and A. Ya. Supin. (**2005**). A comparison of Behavioral and AEP audiograms of a false killer whale (*Pseudorca crassidens*). J. Acoust. Soc. Amer. **118**, 2688-2695.

BIBLIOGRAPHY OF SUPPLIED MATERIAL

DOCUMENTS

[1] Anon. 2006 a: LFA Observation techniques notes. Marine Acoustics, Inc. 1 p.
[2] Anon. 2006 b: AIM timeline. Marine Acoustics, Inc. 4 p.
[3] Anon. 2006 c: Annual Report No. 4: Operation of the Surveillance Towed Array Sensor System Low Frequency Active (SURTASS LFA) Sonar onboard the R/V *Cory Chouest* and USNS IMPECCABLE (T-AGOS 23) under the National Marine Fisheries Service Letters of Authorization of 12 August 2005. Maritime Surveillance Systems. 67 p.
[4] Pascual, P. et al. 2003: Draft guidance on the development, evaluation, and application of regulatory environmental models. The Council for Regulatory Environmental Modeling. Draft, November 2003. 60 p.
[5] Frankel, A.S. 2006: Acoustic Integration Model (AIM) Internal Review Document. Marine Acoustics Inc., proprietary document (draft, 14 September 2006). 50 p.
[6] Frankel, A.S. & Buchanan, J.M. 2006: Acoustic Integration Model© (AIM) users Manual. Marine Acoustics, Inc. (draft, April 2006). 58 p.
[7] Frankel, A.S., Ellison, W.T., and Buchanan, J. 2002. Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts. *Oceans 2002 MTS/IEEE 3*: 1438-1443.
[8] Frankel, A.S & Vigness-Raposa, K. 2006: Marine Animal Behavioral Analysis. Marine Acoustics Inc., Tech. Memo. 63 p.
[9] Marine Acoustics, Inc., Patent Application, 29 Oct 04. Method for modeling the effect of a stimulus on an environment. Marine Acoustics, Inc. proprietary document. 42 p.
[10] Frankel. A. S. Screen shots associated with Test Runs conducted on 26 September 2006, with annotated features.
[11] Frankel, A. S. Report on the attempted testing proposed by Dr. Michael Porter to compare underwater acoustic propagation models.
[12] Center for Independent Experts Statement of Work, September 2006

DOCUMENT EXTRACTS

[13] Anon. 2006 d: Summary of SURTASS LFA sonar operations for 16 August 2002 to 15 August 2006. 2 p.
[14] Lecky, J.H. 2006: Taking and importing marine mammals; taking marine mammals incidental to the U.S. Navy operations of Surveillance Towed Array Sensor System Low Frequency Active Sonar. Extract from Federal Register. 4 p.

[15] Various authors and dates: Excerpts related to AIM from responses to comments in major regulatory publications. 16 p.

POWERPOINT PRESENTATIONS

[16] Frankel, A.S., & Ellison, W.T. 2004: AIM modeling of oil industry activities to derive marine mammal take estimates. IMEMS 2004 conference.
[17] Frankel, A.S., & Ellison, W.T. 2006: The Acoustic integration Model (AIM): application issues for risk assessment. Navy environmental planning and natural resources workshop; risk assessment technologies, Virginia Beach, VA, May 2006.
[18] Messegee, J. 2006: History of marine mammal exposure. AIM CIE review 25-27 September 2006.
[19] Shores, L. 2006: Examination of AIM-generated PE files. AIM CIE review 25-27 September 2006.

Statement of Work

STATEMENT OF WORK
Review of Acoustic Integration Model © (AIM)

The National Marine Fisheries Service (NMFS) requires an independent peer review of the Acoustic Integration Model© (AIM), which shall assess whether AIM correctly implements the models and data upon which it is based, whether animal movements are adequately simulated within AIM, and whether AIM meets the Council for Regulatory Monitoring guidelines for models, which primarily involve scientific credibility.

Background

Minimizing and mitigating the potential effect of sound upon the environment is an increasing concern for many activities. Naval operations, seismic exploration, vessel and aircraft operations, certain construction activities, and scientific investigations now need to consider the potential effects underwater acoustic sources have on marine life. Marine mammals are usually the primary concern, due to their widespread distribution and excellent hearing ability, although impacts on fish are increasingly being considered as well. Predicting the exposure of marine mammals is complicated by their diving behavior and, in some cases, long-range migrations, which causes them to "sample" many depth strata within the water column.

Acoustic propagation and sound received levels are a function of water depth, range from the source, and a host of sound source and environmental variables. This, combined with the variable diving behavior of different species, makes for a very complex problem. The Acoustic Integration Model© (AIM) addresses these specific complications. A principal component of AIM is a movement simulation engine. Both sound sources and animals, collectively addressed as "animats," are programmed to move in location and depth over time in a realistic fashion. Animal movement is based on documented regional and seasonal behavioral data for each species evaluated. Acoustic sources and receivers are programmed to move through a virtual acoustic environment, based on external environmental databases and radiated sound fields created from a choice of several propagation models (*e.g.,* Parabolic Equation [PE], BELLHOP, etc.). The integration component of the AIM engine then predicts the exposure level of each simulated animal at successive operator-selected time steps. Furthermore, each animal can evaluate its environment at each time step, and can be programmed to alter direction or diving behavior in response to variables, such as sound level or sea depth. AIM allows the user to predict the effects of different operational scenarios and animal responses, thereby allowing the selection of an alternative that produces the least impact and still meets operational requirements.

AIM is a proprietary model owned by Marine Acoustics, Inc (MAI). Its value in predicting the acoustic exposure of animals has been demonstrated in earlier documents. However, the continued use of the model to provide acoustic exposure and impact predictions for regulatory assessment purposes requires that the model be reviewed independently, so that NOAA and other federal agencies can comply with the Data Quality Act.

Reviewer Requirements

The Center for Independent Experts (CIE) shall provide three panelists and a moderator for the review of AIM. Expertise in underwater acoustics, modeling, and marine mammalogy is required. The underwater acoustician should be familiar with propagation-loss models. Ideally, the acoustician will have experience or knowledge of the BELLHOP and Navy Standard Parabolic Equation (PE) models, as these are the two main propagation models incorporated into AIM. The modeler should be familiar with individual-based models, preferably those dealing with animal behavior, and the integration of multiple data streams (*e.g.*, multiple databases). The modeler should be able to understand the dynamic interactions of databases. The marine mammalogist should have experience in marine mammal behavior, including diving behavior of more than one species. The moderator should have a reasonable level of scientific and technical understanding, with a reasonable degree of knowledge and experience in at least two of the three scientific categories (underwater acoustics, modeling, marine mammalogy).

The review will be organized around a three-day meeting at MAI during the week of 25 September 2006, to be scheduled based on the availability of all pertinent personnel. The moderator shall be required for a maximum of 12 days for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, and completing the summary report. The three panelists shall be required for a maximum of 12 days each for reviewing documents prior to the MAI meeting, traveling, attending the meeting at MAI, completing their individual reports, and reviewing the draft summary report. Thus, a total of 48 reviewer days is required.

Review Process

The panelists and moderator shall review the documents listed in Annex I prior to attending the three-day meeting at MAI, 4100 N Fairfax Drive, Suite 730, Arlington, VA 22203. After the MAI meeting, the three panelists shall write individual reports addressing all the Terms of Reference given below. The moderator shall consolidate the individual reports into a summary report. Details of the panelists' and moderator's reports are given in Annex II and III, respectively.

<u>CIE – MAI Conference Call – September 21, 2006:</u> The CIE panel and MAI team will discuss via conference call the details of the upcoming meeting, and the CIE panel will raise questions concerning background documents, specifications for trial runs, and other review-related material, including logistics. The call is tentatively set for 5 pm EDT on Thursday, September 21, 2006, but participants and timing may be changed due to individuals' availability.

<u>MAI Meeting - Day 1:</u> Begin with an introduction and a "charge" to the panel, which includes laying out the roles of the participants and the terms of reference (see below), and a time table. AIM presentations will be made by scientists from MAI, with question and answer sessions as needed. Presentations shall include:

1) Introduction to the AIM approach and the software, including data input requirements;

2) Review of results of internal testing of the software;

3) Overview of the process used to derive animal behavior parameters from data and the scientific literature.

Following the conference call and these presentations, the moderator and panelists shall meet to develop a set of test runs for days 2 and 3. The purpose of this evaluation is to see how the AIM responds to a set of inputs that are designed to test the model. It must also be noted that in assessing the functioning of the model, it is important to acknowledge the differences and the roles of the external components (*e.g.*, animal input parameter values, propagation models) and the internal components of the model (*e.g.*, Animat Builder, Movement Simulator, Integration Engine). The distinction is drawn here to emphasize that the values of behavioral parameters can and should change when new data are available, and that AIM can utilize that new data. Valid input data *(e.g.*, animal behavior parameters) are critical for valid predictions from a model run for a specific scenario. However, a valid model can be provided with invalid animal behavior inputs, and still produce accurate outputs. The purpose here is to test the internal components of the model, including how AIM handles the input data. Example scenarios for devising these runs are provided in Annex IV.

MAI Meeting - Days 2-3: Dedicated to the CIE panel working with MAI scientists to perform AIM model runs, so the panel will have sufficient information on the input data, execution parameters, and model outputs for writing their respective review reports. The CIE panel shall, with the assistance of MAI scientists as required, design simulations and request that the MAI scientists create input files to represent these simulations during the course of the review. Projects can be created in a few minutes. Because AIM is a working model (not yet streamlined and simplified for public use), requiring expertise and familiarity with data input procedures and model execution techniques, MAI scientists will perform the model runs under the oversight of the CIE panel. The number and complexity of simulations to be run during the evaluation period will have been discussed in the conference call and finalized on Day 1. To run the models, MAI scientists will require sufficient time to research the values of the basic parameters (i.e., beam pattern information for source, or behavioral parameters for different species). The input files will then be run, and the inputs and outputs will be provided to the CIE panel for their analyses and evaluations.

Terms of Reference

The CIE panel shall complete the following tasks, and document their results in the individual panelist and summary reports.

4. Assess whether the AIM implementation is correct.
5. Assess the animal movement simulation within AIM.
6. Assess whether AIM meets the Council for Regulatory Monitoring (CREM) guidelines for model development.

<u>1. AIM Implementation</u>

Details relevant to the topics described below are given in the Robustness Review Document, which addresses operation of the AIM model.

- Does AIM accurately and efficiently implement the propagation models? Identify any errors in the implementation. The propagation models implemented in AIM include BELLHOP (Porter, 1992) and Navy PE (Zingarelli, 1999). These models were created by other individuals and organizations. *The propagation models themselves are not the subject of the CIE review, but rather the implementation of these models in AIM.*

- Does AIM correctly handle the input values to the models? If not, identify any errors. For example, are acoustic source level and frequency values properly transferred through the model components?

- Does AIM correctly and efficiently extract data from databases? If not, identify any errors. AIM uses the GDEM-V (v 2.6) database for sound velocity profiles and the NOAA ETOPO2 database for bathymetry.

<u>2. Simulation of Animal Movement</u>

- Does the ANIMAT movement model in AIM adequately simulate animal behavior? Comment on the strengths and weaknesses of the modeling approach, and suggest possible improvements.

  The review panel shall devise one or more approaches for addressing this issue. One approach that shall be considered is to evaluate, given appropriate input values, how closely the modeled animal movements mimic the known responses of free-ranging animals. *The species-specific values used in the models are not the focus of the review, but rather the ability of the ANIMAT model to simulate movement.*

3. CREM Guidelines

The panel shall assess whether AIM meets the Council for Regulatory Environmental Monitoring (CREM) guidelines for model evaluation, which are summarized below. Some of the points listed below will have been addressed by the reviewers as part of their comments on Terms of Reference 1 and 2 above. Each reviewer shall ensure that clear answers are provided for the CREM guidelines, though extensive repetition of technical comments is not required.

- Have the principles of credible science been addressed during model development?
- Is the choice of model supported given the quantity and quality of available data?
- How closely does the model simulate the system of interest?
- How well does the model perform?
- Is the model capable of being updated with new data as it becomes available?

Schedule of Activities

The schedule of activities, including timelines (all in 2006) and identification of responsible parties, is provided in the following table.

Activity and Responsible Party	Date
NMFS provides background documents (Annex I) to moderator, panelists, and CIE	September 11
Moderator and panelists participate in a conference call with MAI to discuss technical and logistical details (depending on availability of participants). This call shall be arranged by the CIE.	September 21
Moderator and panelists read background documents	September 24
Moderator and panelists meet at MAI to test AIM model	3 days during week of September 25
Panelists write draft individual reports (Annex II); moderator begins summary report (Annex III)	October 2-13
Panelists provide draft individual reports to moderator for summarization and to CIE for review	COB October 13
Moderator provides draft summary report to CIE	October 27
CIE approves final individual reports, submits them to the moderator and the NMFS COTR	October 27
NMFS COTR approves individual reports; CIE provides final pdf versions to NMFS COTR and to moderator	November 1
Moderator provides draft summary report to CIE	November 8
CIE approves final summary report, submits it to NMFS COTR	November 22

NMFS COTR approves summary report, CIE provides pdf version to NMFS COTR	November 30

Submission and Acceptance of CIE Reports

The CIE shall provide the final individual and summary reports for review and approval to the NMFS COTR, Dr. Stephen K. Brown via e-mail (Stephen.K.Brown@noaa.gov), according to the schedule above. Approval by the COTR shall be based on compliance with this Statement of Work. The COTR shall notify the CIE via e-mail regarding acceptance of the reports. Following the COTR's approval, the CIE shall provide pdf-formatted copies of the reports to the COTR via e-mail.

ANNEX I: Documents to be reviewed in preparation for the AIM review.

Document Titles

1. IEEE article *Application of the Acoustic Integration Model (AIM) to predict and minimize environmental impacts*, Adam S. Frankel, William T. Ellison, and Jacquin Buchanan.
2. Presentation given by Dr. Adam S. Frankel at the International Marine Environmental Modeling Seminar (IMEMS) 2004 Conference: *AIM modeling of oil industry activities to derive marine mammal take estimates*.
3. Acoustic Integration Model (AIM) Robustness Review Document
4. Acoustic Integration Model (AIM) users Manual.
5. The Acoustic Integration Model (AIM): *Applications to predicting and reducing acoustic exposures of marine mammals*, Adam S. Frankel, William T. Ellison. Abstract and presentation for the Navy Environmental Planning and Natural Resources Symposium, May, 2006.
6. Draft Guidance on the Development, Evaluation and Application of Regulatory Environmental Models, The Council for Regulatory Environmental Modeling (CREM)

Document	Document Type	Number of Pages	Degree of difficulty
1. IEEE Article	PDF	6	Low-medium
2. IMEMS Presentation	PowerPoint	28 slides	Low
3. AIM Robustness Review Document	Word/PDF	51	Medium-high
4. AIM user's Manual	Word/PDF	56	Medium
5. EPNR presentation	PowerPoint	33 slides	Medium
6. CREM guidelines	PDF	60	High

ANNEX II: Panelist Report Generation and Procedural Items

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

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

3. The report shall also include as separate appendices the bibliography of materials provided by NMFS for the review and all additional references cited, 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

ANNEX III: Moderator's Summary Report Generation and Process

1. The summary report shall include an overview of the review process.

2. The summary report shall provide a synopsis of the three panelist reports.

3. Points of agreement and disagreement among the panelists shall be documented.

4. The summary report shall also include as separate appendices copies of each of the panelists' reports.

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

ANNEX IV: Example AIM Scenarios. These examples illustrate the types of scenarios that AIM can address. They cover three main sources of anthropogenic noise. The panel shall create their own scenarios to be sure that their questions are addressed and answered. Development of scenarios will be addressed during the conference call and the first day of the meeting.

4. A vessel equipped with a 400 Hz sonar source is operating off the North Carolina coast in May. The sonar signals are broadcast once a minute with a source level of 210 dB re 1 µPa at 1 meter. The ship is to move in a 40 km square saw tooth pattern, with individual north-south legs spaced 5 km apart. The question to be addressed is what is the acoustic exposure of the offshore bottlenose dolphins if the operation goes forward? Alternatively, how would the acoustic exposure be altered if the exercise where conducted in January?
5. A supertanker with a propulsion system that produces a constant 205 dB re 1 µPa at 1 meter with a peak frequency of 10 Hz is transiting off Monterey Bay in June. The ship is paralleling the shoreline at a distance of three miles offshore. Blue whales are feeding within and outside the bay. What is the sound exposure of these whales to the sound of the tanker?
6. A seismic exploration vessel towing an industry standard air gun array is performing a high-resolution survey along the continental shelf south of the mouth of the Mississippi River. Each of the survey lines are 50 km long and spaced 1 km apart. The total survey area is 20 x 50 km. The air gun array is fired every 10 seconds. What is the exposure of sperm whales during the survey?