

*Marine Acoustics, Inc.*  
*Responses to:*

**Summary Report:  
Review of  
Acoustic Integration Model<sup>©</sup> (AIM)  
25-27 September, 2006  
Washington, D.C.  
by  
Center for Independent Experts (CIE)  
Review Panel**

April 2007

Table of Contents

<u>SECTION</u>	<u>PAGE</u>
I. EXECUTIVE SUMMARY.....	3
II. GENERAL DISCUSSION.....	3
III. SUMMARY OF FINDINGS.....	4
a. <u>Assess whether the AIM implementation is correct</u> .....	4
b. <u>Assess the animal movement simulation within AIM</u> .....	5
c. <u>Assess whether AIM meets CREM guidelines for model development and evaluation</u> .....	6
d. <u>AIM model applications</u> .....	10
IV. PANEL RECOMMENDATIONS.....	11
a. <u>Further testing and validation</u> .....	11
b. <u>Suggested enhancements</u> .....	12
c. <u>Application of AIM</u> .....	13
V. NOTES FROM INDIVIDUAL PANELIST’S REPORTS..	14
a. <u>Dr. Thomas</u> .....	14
b. <u>Dr. Porter</u> .....	21
c. <u>Dr. Getz</u> .....	24
VI. REFERENCES.....	27

## **I. EXECUTIVE SUMMARY**

The CIE Review Panel, hereafter referred to as the Panel, was tasked by NOAA's National Marine Fisheries Service (NMFS) to conduct an independent review of MAI's Acoustic Integration Model<sup>®</sup> (AIM). The Panel, consisting of a chairman and three experts in the fields of: 1) mathematics; 2) modeling; 3) acoustics; and 4) marine mammalogy, was convened at the MAI offices in Arlington, Virginia in September 2006. This document provides MAI responses to the CIE Summary Report of December 11, 2006.

The criteria (terms of reference) for the Panel's review was to evaluate whether AIM correctly implements the models and data on which it is based; whether animal movements are correctly implemented; and whether AIM meets the Council for Regulatory Environmental Monitoring (CREM) guidelines.

As stated in the Chairman's Executive Summary, the Panel:

- a. Agreed that AIM appears to be correctly implemented;
- b. Agreed that the animal movement appears to be appropriately modeled; and
- c. Agreed that the principles of credible science had been addressed during the development of AIM and that AIM is a useful and credible tool for developing application models.

In regard to whether AIM satisfies the CREM guidelines, the Panel had some diversity of opinion. They were split on whether AIM has had adequate peer review and they were of the opinion that their review was the initial peer review. MAI concurs with this and notes that this CIE review was convened to specifically address the requirement for peer review of AIM.

## **II. GENERAL DISCUSSION**

MAI appreciates the professionalism and hard work expended by the CIE Review Panel and intends to seriously address all Panel comments. The following comments and recommendations, and responses thereto, are presented to preclude the possibility of a casual reader misinterpreting the Panel's recommendations, or be left with the false impression that analyses performed to date using AIM are somehow invalid or inaccurate.

The primary use of AIM is to produce a statistically valid estimate of the potential to affect marine mammals during the execution of an underwater acoustic operational or research project—that is, the number of potential takes of marine mammals. While the recommendations of the panel are very helpful and can certainly lead to a more robust implementation of AIM, their effect on the actual take quantity estimates, including past analyses as well as those in progress, is at most second order. Moreover, many of the recommendations do not apply to AIM as a model, but rather apply to post-run analysis of the AIM results. This is particularly true of the many recommendations that are in the

categories of error analysis or statistical analysis. Moreover, MAI uses AIM to answer questions from a customer and the customer dictates the level of effort to be expended in exercising AIM to produce those answers. While MAI has done statistical analysis of the input and output of AIM runs, this is not often desired by the customer due to resource constraints.

In the future, MAI plans to prepare an AIM Application Outline (AAO) for each client prior to the execution of AIM runs, which spells out:

- How AIM will be run for their application, including the databases to be used, and their known limitations, such as: 1) what animals may need to be used as surrogates; 2) variances from actuality in the General Digitized Environmental Model (GDEM), bottom losses, etc.; 3) resolution of the models; etc.
- What statistical approaches will be applied, including possible conservative offsets, variances based on multiple propagation loss runs with different sound velocity profiles, etc.
- A clear and concise description of the regulatory thresholds to be applied to the output data and their basis. It may be appropriate to provide a copy of the AAO to the regulatory agency for comment before AIM runs are executed.

The completed AAO, as agreed to by all parties, will be made a part of the final AIM report.

The Panel's report contains several sets of comments and recommendations that will each be addressed below. Page numbers in parentheses refer to pages in the CIE Summary Report, and references at the end of quotes refer to panelist comments.

### **III. SUMMARY OF FINDINGS (p. 5)**

#### **a. Assess whether the AIM implementation is correct**

##### **Relevant Extracts: (p. 6)**

- “Each of the components of AIM need to be tested or “exercised” by running a set of simulations (a low frequency sonar vessel, a supertanker, and a seismic exploration vessel) 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 the Effects of Sound on the Marine Environment (ESME) workbench over a set of exercises that covers the full range for which AIM is designed to provide assessments (Getz 2006).”

##### **MAI response:**

- MAI has conducted hundreds of hours of testing on the components that make up AIM and have run hundreds of simulations that were analyzed to ensure the results were reasonable compared to expected results. Particular tests have been

run to simulate scenarios that could be compared to analytical solutions. While this is not possible for complicated scenarios, it is possible for basic scenarios and has given confidence that algorithms contained in AIM are correctly implemented. I should be understood that the application of AIM is intended as a time-based statistical estimate of the effect of a complex and typically large 3D underwater acoustic field on representative distributions of marine wildlife within that field. The accuracy of such results is based on statistical parameters, such as the mean and standard deviation of numbers of animals exposed to set thresholds- -not of a single individual animal but rather a collective result applied to a local population.

- Occasionally a run is incorrectly set up in AIM. MAI concurs that the operator must be careful and must fully document the set up. This set up and documentation must be complete and concise so that it can be easily analyzed for errors in the set up procedures. Further testing and documentation of code and updates thereto are addressed in several places below.
- MAI looks forward to comparing AIM with the ESME workbench, or any other comparable 3D underwater acoustic exposure model when they are available. It is important to note that ESME has not been established as being capable of producing the “right” answer. To date, only AIM has undergone an independent review. Therefore, any difference between AIM and ESME outputs would not necessarily indicate that one or the other is correct. An analysis would have to be performed to understand the differences.

**b. Assess the animal movement simulation within AIM (p. 6)**

**Relevant Extracts: (p. 7)**

- “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 animal 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).”

**MAI response:**

MAI concurs with the comments above; as additional information on marine mammal behavioral responses to external stimuli and environmental conditions becomes available, animats will be updated to include these behaviors. Aversion parameters to account for aggregative social, feeding, or predator avoidance behavior of individual animals are not currently implemented in AIM. Creation of a behavioral rule within AIM is possible; however, MAI might consider an alternative approach to account for effects such as animal aggregation in the context of impact analysis. If areas that are likely to create animal concentrations can be identified (e.g., Hudson Canyon), it would be possible to identify which animats were exposed within that aggregation area. Then, it would be possible to apply a multiplier to these modeled takes to account for the aggregation of animals within a specific location without increasing the animal density over the entire modeling area. Alternatively, a second set of animats could be placed in these areas of natural concentration to supplement the overall distribution. Thus, these areas of known concentration can be represented in the simulation. This approach has been used in recent simulations.

**c. Assess whether AIM meets the CREM guidelines for model development and evaluation (p. 7)**

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

The Panel agreed that this was generally true. Relevant extracts (p. 8):

- “Generally an excellent job has been done in applying the best science in the development of AIM (Porter 2006).”
- “The animats, 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 spatio-temporal 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.”

**MAI response:**

MAI concurs with the comments above; again, as additional data become available, behavioral models will be revised and updated as necessary.

**Has there been adequate peer review of AIM? (p. 8)**

“To a certain extent this question is answered by the existence of the CIE Panel review since NMFS required an “independent peer review of AIM” (see SOW, Appendix 2 of CIE Summary Report). 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 Panel 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 Active Sonar program. EISs are highly reviewed by both specialists employed by the applicant and by the Office of Protected Resources, of the National Marine Fisheries Service. EIS procedures include publication in the Federal Register and a lengthy 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 the EMSE workbench: when output differs by more than a percent or two, then both platforms need to be checked for likely sources of error...””

**MAI response:**

MAI generally concurs with the comments above and believes that reviews conducted to date, including the CIE Panel’s efforts, do constitute adequate peer review. Furthermore, the National Research Council’s report on Marine Mammal Populations and Ocean Noise: Determining When Noise Causes Biologically Significant Effects (NRC, 2005) states, “AIM has been used to predict the exposure of animals with different response patterns to sources with different modes of operation, monitoring, and mitigation; and it can help in selecting alternatives that minimize effects on marine mammals while maximizing operational effectiveness of the source.”

**Does the conceptual model have an adequate theoretical basis? (p. 9)**

“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 (p. 9):**

- “AIM is generally using the best available knowledge; however, the limitations of that knowledge are widely recognized (Porter 2006).”
- “Environmental databases, especially for the geo-acoustic 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).”

**MAI response:**

MAI concurs with the above comments; as additional behavioral data and environmental databases become available, they will be incorporated into AIM.

**Has the conceptual model been adequately implemented? (p. 9)**

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

**Relevant extracts (p. 9):**

- “... discussed many ways that AIM could be adapted ... to make the animal 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).”

**MAI response:**

MAI generally concurs with the above comments. As additional behavioral data and environmental databases become available, they will be incorporated into AIM. Because future daily variations of site-specific data are unknown, MAI is not convinced that including the full range of these past variations would be more accurate than using

monthly averages for the purpose of predicting future outcomes. MAI could investigate the NOAA sound velocity profile databases, which present monthly means AND standard deviations of those means. It would be feasible to produce several TL runs for areas within an AIM modeling site using the NOAA mean +/- the standard deviation (three SVPs). Perhaps three different locations within a modeling site, resulting in nine different TL runs, would suffice. The variability in the TL field could be described. These results could be included in the report of the AIM results. The variability in the TL field could also be used to investigate the possible effect upon impact predictions. For example, suppose there was a 3 dB variation (e.g., max-min) in the TL field based on variability in the SVP. Then, one might investigate the estimated number of marine mammal incidental takes based on take threshold +/- 3 dB. This would be a method for propagating the effect of SVP uncertainty to the take impact estimates. As with many of the statistical approaches to the analysis of AIM results, this approach can be accomplished without alterations to the AIM code base.

**Were adequate techniques and procedures used for model code verification? (p. 10)**

“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 Panelist’s reports had been written. However, the Panel 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 Terms of Reference #1, the panelists each had recommendations for further testing.”

**Relevant extracts (p. 10):**

- “There was limited information to answer this question. The CIE Panel could not and did not verify programming code. At the Review Meeting, 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 demonstrations 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).”

**MAI response:**

MAI agrees with the comments and will attempt to undertake a project to better document the development and verification of the AIM code. Thereby, more credibility will be afforded to AIM in the academic community in subsequent peer reviews. As for comparison of AIM results with ESME, please see MAI response to III.a. above.

**d. AIM model applications (p. 11)**

“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 transmission loss and a single simple sound source) or very complex (multiple species, multiple sound sources using PE or Bellhop transmission loss models, with animal aversions and attractions specified). The level of complexity needs to be tailored to the available data.”

**MAI response:**

MAI concurs with the above comments.

**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 animals 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).”

**MAI response:**

MAI concurs with the above comments.

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

“Two panelists answered this question with regard to AIM 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).”

**Relevant extracts (p. 12):**

- “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 behavior 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).”

**MAI response:**

Generally, sensitivity testing has been accomplished using MATLAB programs to analyze the output results as a function of variations in the input parameters. Certainly, more sensitivity analysis can be performed and it is agreed that for important policy decisions this should be done. MAI intends to expand the statistical analysis of future AIM runs consistent with the needs and desires of our customers.

**IV. PANEL RECOMMENDATIONS (p. 12)**

**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).”

**MAI response:**

MAI generally concurs with the comments and recommendations. It should be noted that all of the acoustic databases typically used by MAI and two of the acoustic propagation models (PE and BELLHOP) have been thoroughly tested and verified for correct implementation by the AIM code. These data were presented to the Panel as part of their preliminary briefings. Each recommendation above will be further implemented as time permits and resources are available.

**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; and
- Comparative analyses and addressing conditional (what if?) questions.”

**MAI response:**

MAI has performed similar analyses in the past and will again when appropriate, and with customer concurrence. In fact, since MAI’s clients are typically concerned with the number of Level A and B incidental takes (per the Marine Mammal Protection Act), the results most often dealt with are those of the tails of the received level (RL) distribution curves, not the averages.

“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; and
- Use the coherent option as the default in BELLHOP.”

**MAI response:**

MAI generally concurs with the above recommendations and will study and implement them, as appropriate, and with customer concurrence.

“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”.

**MAI response:**

Responses to the above comments can be found in Subsection V.a. below.

**c. Application of AIM (p. 13)**

“(Porter 2006): Use a burn-in period to stabilize animat behavior. In cases where the highest quality predictions are important, do calculations at multiple frequencies; and do sensitivity studies to provide error bars on model predictions.”

**MAI response:**

MAI concurs with the above recommendations, which will be implemented, as and where appropriate, and with customer concurrence.

“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; and
- 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.

**MAI response:**

MAI concurs with the above recommendations and will implement them, as and where appropriate, and with customer concurrence. It should be noted that already many of the data suggested by Getz (2006) are captured in “input” and “output” files. Included in the input data files are: 1) the run parameters and default settings for the acoustic propagation model, and 2) initial animat 3D positions. The output data files include underwater sound exposure (i.e., received level) data and can contain all of the 3D positions of all animats (if the simulation is configured to record these data).

## V. NOTES FROM INDIVIDUAL PANELIST'S REPORTS

### a. Dr. Thomas

P. 99: Dr. Thomas disagrees with the “blackfish” grouping, citing that incorporates species from four genera.

**MAI response:**

MAI proposes to review the data for these species and determine if there are sufficient data to split this group into smaller groups.

P. 99: Dr. Thomas states that the review document is limited to US waters, and suggests adding manatees to the animal review document.

**MAI response:**

Species have been added to the list as needed for individual projects and all past projects have included all species appropriate for the geographic location of interest. There is no systematic inclusion or exclusion of species.

P. 101: Dr. Thomas applauds the bootstrap-*t* testing of AIM output, but is unclear whether it is a part of AIM or not.

**MAI response:**

The bootstrap-*t* is a separate test that can be run on AIM outputs after the completion of the simulation. It is applicable to, but not part of, AIM.

P. 110: Dr. Thomas indicated that most hearing studies of marine mammals have used stimulus step sizes of 3 dB. She suggested that perhaps a +/- 3 dB bounding of received levels.

**MAI response:**

This could be accomplished in the impact predictions by a +/- 3 dB variation in take criterion thresholds.

P. 110: Dr. Thomas repeated the concern about point estimates and the lack of error bounds around AIM results.

**MAI response:** This is a valid concern; however, this is not a criticism of AIM itself, rather how it has been applied in the past. Note, however, that in the MMS reports, the uncertainty in animal density information was carried forward to provide error bounds around the take estimates.

Pp. 111-112: Dr. Thomas poses 14 suggestions/comments. “The patent description for AIM, which suggests 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?”

**MAI response:**

It would be more efficient to add the background ambient noise in post-processing of AIM output, but this is certainly possible.

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 user’s operations? This could be added on as a possible ambient noise masker.”

**MAI response:**

Receiver animals (e.g., whales) can already be programmed to be acoustic sources as well.

3. “Could the surface time be used to imitate haul-out periods for 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?”

**MAI response:**

Yes, surface time can last hours to imitate the haul-out behavior of pinnipeds.

4. “Can AIM model swim speed versus received level to estimate how quickly an animal could leave a received level to move to a tolerable, lower received level.”

**MAI response:**

AIM can export received sound level and speed of the animal at each time step in the model. This would allow the type of analysis proposed by Dr. Thomas to be conducted.

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 animals as being within the duct or below the duct?”

**MAI response:**

Yes, the user-specified sound velocity profile function can be used to add a surface duct to an existing sound velocity profile (SVP). The user can specify the depth (in meters) to ensure that the source is located within the duct, as suggested by Dr. Thomas.

6. “Movements of marine mammals in the water column are often affected by organisms of the Deep Scattering 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 mammal’s 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 scattering 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?”

**MAI response:**

Modeling a deep scattering layer (DSL) and incorporating it into AIM is possible. However, MAI suspects that it would take a large effort to incorporate it fully into all aspects of the model.

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.”

**MAI response:**

The TL matrices produced by the propagation models can already be used to estimate the zones of audibility of a source, given a specified ambient noise level and a detection threshold for the receiver.

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 animal, 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 animals that exceed that criterion, or the number subjected to TTS or PTS.”

**MAI response:**

Incorporation of the m-weighting can be done in post-processing of the AIM output without the need for any code modifications; or it can be accomplished by the judicious selection of source level (SL) when running the AIM problem.

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.”

**MAI response:**

MAI agrees that AIM is well suited for studies of masking effects of anthropogenic sounds on marine mammal populations. Post-processing of AIM output from the existing application can provide this type of analysis.

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, NIOSH (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. NIOSH 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.”

**MAI response:**

AIM can (and has) been run for extended periods of time, accumulating noise exposures over a period of days and even weeks. Post-processing of these data already allow for energy accumulation and Leq calculations.

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.”

**MAI response:**

Diurnal specificity of dive type would be an enhancement to the AIM software. This issue can currently be handled by creating daytime animats and night time animats that are run simultaneously. Then, the results from these animats can be selectively sampled. This approach does not explicitly address the transition zones between behavioral states (e.g., diurnal and nocturnal) but it does begin to address the issue.

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.”

**MAI response:**

Dr. Thomas raises the common concern about variability in SVP. In the responses to Dr. Getz's comments, an approach to address this concern is outlined below.

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."

**MAI response:**

This is an interesting proposed application of AIM. It would be necessary (and already possible) to program the animats to alter their behavior in response to the stimulus. However, MAI would argue that the amount of knowledge necessary to create such a simulation is also sufficient to produce the proposed data product; i.e., a plot of return of the animal to baseline behavior. This is because the responses to stimuli are deterministic in the model.

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 animal. This might help evaluate extra energy expenditures an animal incurs when exposed to noise; i.e., related to fitness."

**MAI response:**

AIM does not create the linearity index. The linearity index was measured from the AIM output using a separate MATLAB routine. Nevertheless, the AIM output can return the three-dimensional coordinates for each time step. The total distance traveled by the animat can be determined from the output of the existing AIM algorithms.

Dr. Thomas makes a list of four recommendations at the end of her report (p. 113). These include a recommendation to expand the suite of test projects that have already been created. Specifically, she recommends test projects for:

1. "A simulation(s) that test(s) for corner effects; i.e., an animat might get stuck in a right-angle boundary and continue to "bounce" off the bottom and the latitude/longitude boundary."

**MAI response:**

Corner effects have already been examined, in the internal robustness testing.

2. "A simulation, or set of simulations that test for "stickiness," where an animat gets stuck in a bathymetric feature."

**MAI response:**

The animat that got “stuck” during the Panel review was set to a 180 deg aversion to land. This was operator error, which demonstrates the importance of operator knowledge in order to properly run AIM. It is already known that animats should be set to 135 deg aversion to irregularly shaped features (e.g., land). The angle response, along with the variability of course heading, has been shown to be effective in preventing this “stickiness”. Inspection and analysis to detect this type of operator error is standard practice for AIM runs.

3. “Simulations to examine the effect of animat speed on sound exposure.”

**MAI response:**

Simulations were conducted to examine the relationship between animat speeds and sound exposures. These results were consistent with expected values and should have been presented to the Panel in more detail.

4. “Recommendation of testing the random animat placement routine using different sizes and shapes of initial areas, and testing the placement data.”

**MAI response:**

In the course of conducting each of the AIM test cases presented to the Panel, each test case used different area sizes and random animat placement. Analysis of the animat locations and movements were consistent with expected values and presented to the Panel in Dr. Frankel’s presentation.

Dr. Thomas also has a set of eight recommended tests for the animat movement engine.

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

**MAI response:**

Yes, AIM notifies the user of an incorrect input.

2. “If an animat is programmed to spend X percent of time in deep water and Y percent 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?”

**MAI response:**

Surface/resting behavior is specified separately from dive behavior, and more than one surface behavior can be specified (e.g., short respiratory sequence and extended resting behavior). However, AIM cannot (currently) associate specific dive behavior with specific surface behavior. Nevertheless, if a species has different surface behaviors, they can be incorporated into the simulation.

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

**MAI response:**

AIM does not have the ability to check the species parameters. The user does not select species from within AIM. Rather, the user has to input the values manually into the animat builder. There are no software checks here; it is incumbent upon the user to create the simulations properly.

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.”

**MAI response:**

Increasing surface time would reduce the probability of exposure, as the propagation models account for the pressure-release property of the surface. Currently, AIM is capable of modeling horizontal aversion, but not vertical aversion from a stimulus. Such a capability can be added to AIM.

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

**MAI response:**

The number of variables involved in predicting exposure levels precludes a simple response to this comment. Rather, it would need to be addressed on a case-by-case basis. Nevertheless, a comparison of TL predictions for a site with the dive behavior could indicate how a change in the number of dives would impact the exposure--specifically, comparison of received levels at the intermediate and bottom depths of the dive behavior. An increase in number of dives per time would result in more time being spent at intermediate depths and less at bottom depths.

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?”

**MAI response:**

Inquiry about animats that remain near the surface in a ducting environment—if the animats dive into the duct, or if it is a surface duct (e.g., arctic), and the principal sound source is in the duct, then animat exposure would be increased.

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?"

**MAI response:**

There are no interactions between/among species, so there is no effect of modeling separately vs. together. Animats are allowed to be co-located in space and time. Note that the range and depth resolution of the propagation models are typically larger than the body length of an animal. Therefore, if an animat should be displaced a few meters due to occasional co-location with another animat, it will have negligible effect upon exposure prediction.

8. "Rather than an animat being an individual animal, it could be a pod or subgroup of animals (lime mother/calf pair or a bachelor herd of sea lions). Some thought should be given to the question of whether 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."

**MAI response:**

Individual animats can be programmed to be groups of individuals as well as an individual animal. This effect has already been considered. For example, consider if animals were to travel in groups of ten, then the GROUP density should be 0.10 of the value of the INDIVIDUAL animal density. Therefore, the impact predictions will have the same mean value regardless of whether animats are treated as individuals or groups. However, the variation around that mean is likely to be different for groups vs. individuals.

Finally, Dr. Thomas states, "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."

**MAI response:**

MAI concurs with the above recommendation. Although it should be noted that this would be, in large part, solely an exercise of the propagation models and their implementation within AIM, MAI believes that this is the component of AIM with the least uncertainty, and other testing would have higher priority.

**b. Dr. Porter**

P. 54: "...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 Range-dependent Acoustic Model (RAM) option which is also available in the NSPE and generally used by the fleet in place of the narrow-angle PE.”

**MAI response:**

It would be straightforward to implement a refinement to address BAM's lack of surface reflection characterization. However, it would require modification of the AIM code base.

This criticism of the use of narrow-angle PE has been made previously (e.g., ITM 2002). RAM is now operational within AIM, but needs review and tuning of its implementation in order to be considered fully useful. Also, the AIM code base would need to be modified to achieve this goal.

P. 55: “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.”

**MAI response:**

Replacement of NSPE with RAM may be one approach to address this issue.

P. 58: Summary of BAM and NSPE criticism, and suggestion to use coherent mode in BELLHOP as the default. Additionally, Dr. Porter addressed the single frequency nature of the TL models and suggested multiple runs as a remedy.

**MAI response:**

Range averaging (Harrison and Harrison 1995) could be implemented in AIM, which would also provide a remedy without the need for additional TL runs.

P. 59: “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.

AIM testing focused on some simple metrics on histograms of the depth distributions of the animals. An example of one such test is a simulation where 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. 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.”

**MAI response:**

The initial distribution of animat depths does not reflect animat distribution during the AIM runs. One answer is to run the animats for a period of time prior to the modeling effort. This time period allows the depth distribution of animats to achieve their (statistical) stationary distribution for the remainder of the model run. The histograms produced for and reproduced in the report are a type of diagnostic that can be used to determine the duration of this burn-in period.

P. 62: Dr. Porter notes that errors in AIM were uncovered during the Panel testing. Those errors were corrected during the course of the review. However, Dr. Porter indicated that additional benchmarking is warranted.

**MAI response:** MAI concurs with the above comment and has established a set of test cases as part of the internal robustness testing. However, the Panel review showed that these test cases cannot cover all possible situations. MAI has plans to expand the number of test cases with the goal of near-constant code-checking, and upgrading the code when appropriate. This will accelerate the ability to test future versions of the AIM software and guard against code errors creeping into existing portions of the code. It should be noted that errors of the type noted during the Panel’s testing do not result in significant changes to the results in the final AIM product.

P. 62: “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.”

**MAI response:**

The proposals listed in the responses to Dr. Getz’s comments below are applicable to Dr. Porter’s comments above as well.

Dr. Getz

P. 27: “Recommend that 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.”

**MAI response:**

MAI will document the relevant entries in the .aim prj file (e.g., animal movement, TL model parameters, etc.). MAI believes it is less important, but potentially valuable, to document the display parameters (e.g., color maps). This approach can be accomplished without alterations to the AIM code base.

P. 27: “Recommend that a critical output should be reported minimally in terms of both averages and confidence intervals rather than just averages.”

**MAI response:**

MAI proposes investigating the NOAA sound velocity profile databases, which present monthly means AND standard deviations of those means. It would be feasible to produce several TL runs for areas within an AIM modeling site using the NOAA mean +/- the standard deviation (three SVPs). Perhaps three different locations within a modeling site, resulting in nine different TL runs, would suffice. The variability in the TL field could be described. These results could be included in the report of the AIM results. The variability in the TL field could also be used to investigate the possible effect upon impact predictions. For example, suppose there was a 3 dB variation (e.g., max-min) in the TL field based on variability in the SVP. Then, one might investigate the estimated number of marine mammal incidental takes based on take threshold +/- 3 dB. This would be a method for propagating the effect of SVP uncertainty to the take impact estimates. This approach can be accomplished without alterations to the AIM code base.

Furthermore, the variability in the animal population or density estimate can be carried forward along with the mean estimate.

P. 27: “Recommend that 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.”

**MAI response:**

MAI will consider applying for funding to design and carry out an appropriate sensitivity analysis. Ideally, this effort would include Dr. Getz as a co-investigator. The results of such an analysis could be submitted for publication in a peer-reviewed journal.

P. 27: “Recommend that 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.”

**MAI response:**

MAI concurs that within the text of the above recommendation, there are several notable comments.

P. 34: “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)...The best way to evaluate a black box is to have two black boxes independently developed under the same conceptual design and then see if they produce the same or very similar output.”

**MAI response:**

MAI disagrees with the characterization of AIM as a “black box.” Nevertheless, we would welcome the opportunity to compare AIM with ESME or any other comparable 4D acoustic propagation and animal movement model.

P. 36: “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.”

**MAI response:**

MAI contends Dr. Getz is correct in his evaluation of the animat movement engine, but incorrect in the intention of its application. The goal of the animat movement engine in AIM is not to replicate the details of any one animal, but rather to capture the aggregate movement of a group of animals. The fine details of individual animal movement are not important to the goal of impact prediction.

MAI may solicit funding for a comparative analysis between three dimensional animal movement data (e.g. WHOI d-tag data) and AIM animat movements. This has been done for two dimensions based on beaked whale time-depth-recorder (TDR) data. MAI could use the 3D DTAG data as a basis for a study on the fidelity of animal movement simulation by comparing actual animal data against animat simulations. Again, the results from such a study could be submitted for publication in a peer-reviewed journal.

P. 37: “At this point, 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).”

**MAI response:**

Creation of such a behavioral rule within AIM is possible. However, MAI might consider an alternative approach to account for effects such as animal aggregation in the context of impact analysis. If areas that are likely to create animal concentrations can be identified (e.g., Hudson Canyon), it would be possible to identify which animals were exposed within that aggregation area. Then, it would be possible to apply a multiplier to these modeled takes to account for the aggregation of animals within a specific location without increasing the animal density over the entire modeling area. Alternatively, a second set of animals could be placed in these areas of natural concentration to supplement the overall distribution. Thus, these areas of known concentration can be represented in the simulation. This approach has been used in recent simulations.

## VI. REFERENCES

- CIE Summary Report. 2006. Summary Report: Review of AIM, 25-27 September, Washington, D.C. Independent System for Peer Review: University of Miami. 134 p.
- Frankel, A.S. and K. Vigness-Riposa. 2006. Marine Animal Behavioral Analysis. Marine Acoustics, Inc. Tech. Memo. 63 p.
- Getz, W. 2006. Review of AIM 25-27 September, 2006. Washington, D.C. *For* University of Miami Independent System for Peer Review. xx p.
- Harrison and Harrison. 1995. A simple relationship between frequency and range averages for broadband sonar. *J. Acoust. Soc. Am.* 97: 1314-1317.
- Hursky, P., M.B. Porter, M. Siderius, and V.K. McDonald. 2004. High frequency (8-16 kHz) model-based source localization. *J. Acoust. Soc. Am.* 43: 3021-3032.
- ITM. 2002. Proceedings of the Twenty-First Annual Information Transfer Meeting, January 2002. US DOI, MMS, Gulf of Mexico OCS Region, MMS 2003-005.
- Ketten, D.R. 1994. Functional analysis of whale ears: adaptations for underwater hearing. *IEEE Proc. Und. Acoust.* 1: 264-270.
- Ketten, D.R. 1997. Structure and function of whale ears. *Bioacoustics Int. J. Anim. Sound. Rec.* 8: 103-105.
- McDonald, V.K., J.A. Rice, M.B. Porter, and P.A. Baxley. 1999. Performance measurements of a diverse collection of undersea, acoustic, communication signals. Proceedings of IEEE Oceans, 1999.
- National Institute for Occupational Safety and Health. 1998. Criteria for a recommended standard: occupational noise exposure. U.S. Department of Health and Human Services. NIOSH Publication No. 98-126.
- National Research Council. 2005. Marine Mammal Populations and Ocean Noise: Determining When Noise Causes Biologically Significant Effects. National Academies Press, Washington, D.C.
- Porter, M. 2006. Review of 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 AIM for determining noise exposure levels for marine mammals. *For* University of Miami Independent System for Peer Review. xx p.