Proceedings of the Institute of Acoustics

THE SEA ANIMAL NOISE DATABASE SYSTEM (SANDS)

N J Felgate *, L J Lloyd

ASW Data Analysis Centre (ADAC), Farnborough, Hampshire.
(* For part of the development, the author was employed by EASAMS Ltd., Camberley, Surrey.)

1. INTRODUCTION

1.1 The Sea Animal Noise Database System (SANDS) is a database and display facility for data relating to marine mammals. It has been developed within the ASW Data Analysis Centre (ADAC), part of the Defence Evaluation and Research Agency of the United Kingdom's Ministry of Defence. It is a research tool that has been developed primarily to provide a reference set of biological noises for the training of sonar operators. It contains multi-media information on a number of species of marine mammal, including sounds, movies, stills, diagrams and text. Facilities are provided to search for specific species and examples on various criteria.

2. BACKGROUND

2.1 In anti-submarine warfare (ASW), sound is the foremost mechanism for the detection of potential threats. To detect a potential enemy underwater the sonar operator must be familiar not only with the acoustic signatures of the threats, but also with the many and varied noises from non-threat sources in the ocean. Some of the most prolific of these sources are biological in nature; in particular, marine mammals are a main source of noise in open ocean.

2.2 The potential of biological noises to create false alarms is illustrated by the research undertaken in the United States in the sixties. Operators reported detecting a source at 22 Hz which lasted for several hours, even days. This signal consisted of a sequence of pulses, each one second in duration, at a near constant frequency and repetition rate. Because of its stability it was initially thought by the defence community to be man-made, possibly emanating from Russian nuclear submarines. Much research was conducted to determine its origin and eventually it was established that it was actually the call of lone, male fin whales.

2.3 Historically, sonar operators have been given little formal training on biological sources and the noises that they produce. What training is given is provided at sea, while "on the job", by experienced sonar operators. This has lead to sonar operators devising their own terminology for the sounds that are detected; for instance, the noises of fin whale have become known as the "Jezzie Monster", those from sperm whale are known as "Carpenter Fish", and those from the grey bearded seal are "ice noise". This is considered unsatisfactory for a number of reasons:

- It ignores a large body of research undertaken by the scientific community which can provide additional collateral information on each species, such as pattern of behaviour, locations and seasonal distributions.

- As the political climate has changed, ASW platforms have been required to deploy away from their traditional areas of operation, into regions in which operators may be unfamiliar with the acoustic environment.

- There is little scope to provide prior training to operators as to the expected characteristics of these sources as seen on new sonar systems.

2.4 In order to overcome these problems, ADAC embarked in 1994 on a three year programme of work to set up a database of underwater animal sounds which could be used to improve the training of sonar operators. The aim was to establish the essential features of the database, and provide a reasonably comprehensive database for a limited number of animals and an initial start on the database of a larger number of animals. At the start of this project a symposium was organised to elicit suggestions as to the structure of the database, to identify mammals of particular interest, and to identify possible sources of data [1].

3. THE DATABASE

3.1 SANDS has been designed to store the following items of data:

- Examples of the sounds made by each species for aural comparison (stored in the form of digitised time-series) together with associated textual information on the data collection process.

- Text, maps and sketches providing socio-ecological information with regard to each species, together with historical, bibliographical, and explanatory textual data.

- Still pictures and movie sequences of each species for identification purposes with supporting information on the collection arrangements.

- Collated facts on each of the species in a list format allowing subsets of species to be identified by various physical features, distribution, acoustic characteristics, and taxonomy.

3.2 In 1994, at the inception of the project, a Macintosh QUADRA 840AV was chosen as the computer platform for the development of SANDS; at the time it was considered that the audio-visual capabilities of PC platforms were not sufficiently advanced for this use. An

Proceedings of the Institute of Acoustics

THE SEA ANIMAL NOISE DATABASE SYSTEM

external hard disk is used to store the textual and diagrammatic information, while the multimedia examples are stored on a number of CDs in a jukebox system.

3.3 The FoxPro database application is used to store the textual and diagram data in a number of tables. The examples are stored as individual files in formats applicable to the particular type of media e.g. the Audio Interchange File Format (AIFF) for sounds, the Joint Photographic Experts Group (JPEG) format for Stills and the Apple standard QuickTime format for movies. These formats were chosen as being accessible on other computing platforms. In particular AIFF was chosen for the storage of sounds since it allowed for multiple channel data and allowed for storage of data with a high sample-rate (some are over 80 kHz).

4. FEATURES OF THE DATABASE

4.1 The database has been developed to be highly structured and use is made of relational dependencies in order to reduce memory requirements, to help in data consistency and to assist in identifying where data is stored (so simplifying maintenance).

4.2 The socio-ecological text has been compiled for each species on a number of subjects. The items of text are stored separately enabling quick access to a particular subject of interest. The subjects are:

Vocalisation habits	Social behaviour	Size
Physical description	Identification	Possible confusions
Habitat	Population	Distribution
Motion	Diet	Feeding
Migration	Reproduction	Maturation
Mortality	Common names	Synonyms

4.3 For each species this text is held at three levels of detail:

- A one -line summary statement on each subject.

- Reasonably detailed statements for each subject, aimed at giving the intelligent laymen a comprehensive overview of the subject.

- Highly detailed statements for each subject identifying the source material and providing the arguments used to derive the overview statements from this material. This is aimed at those wishing to research the subject and to allow future maintenance of the overview statements in the light of fresh information.

4,4 Illustrative diagrams have also been included on the subjects of distribution, migration and physical description.

4.5 Provision has been made to store a large quantity of background data alongside the examples. There are a number of reasons for this:

- It enables the user to search for examples with certain characteristics.

- It enables the cause of problems with certain data to be researched and the affected data easily identified, especially necessary since the data has been collated from a number of sources, employing various equipment.

- It allows the investigation of hypotheses with regard to the correlation between characteristics derived from examples and environmental conditions, perhaps obviating the need for special, costly data collection trials.

- It helps establish the copyright owner of the data by identifying the organisations and people involved in the data collection.

4.6 The sound, stills and movies examples are linked to the textual data in the FoxPro tables by storing the path name to the example file (disk, directories and filename) within the tables. This allows the examples to be accessed in an integrated manner with the textual information and diagrams. There is provision in the database to store within the FoxPro tables a preview of each example (for sounds a low resolution spectrogram, for stills a low resolution version of the image, and movies a low resolution representative frame). This feature overcomes the possible delay in accessing the example (especially when it is necessary to change a disk in the autochanger) and overcomes memory restrictions to allow a number of examples to be compared simultaneously in the identification of a particular one of interest.

4.7 It has been recognised that often items of data are not available, or may be incomplete or inexact. Provision has been made for representing such fuzziness in the storage of numbers, dates, times and logicals in the database by specifying formats to represent these as character strings. Other items are restricted to a particular set of options, thus allowing meaningful search to be made on particular characteristics. For instance, sets of sound descriptors and sea areas have been compiled. It is usual in database design to devise codes for such restricted items and thus reduce storage requirements. However much of the data in the database has been entered by hand, so it was decided to use meaningful names which provide some redundancy in case of error (with codes a simple one character error can change the meaning radically and the error cannot easily be detected). Explanations and/or diagrams are stored alongside the list of options in the database to aid the user.

THE SEA ANIMAL NOISE DATABASE SYSTEM

5. THE SEARCH AND DISPLAY FACILITIES

5.1 ADAC has developed an application which allows the various items of multimedia data stored within the database to be selected and viewed in an integrated manner. This application has been developed using FoxPro's command language. It uses pull-down menus to allow the user to control the program flow, modal dialogue boxes to allow the user to select search criteria, display windows for the presentation of the selected text or diagrams and modal information boxes for the display of background data relating to these items.

5.2 The system allows different items of data to be displayed within their own display window. The same type of data for separate species may also be displayed in separate windows. Buttons on these windows provide access to background data displayed in information boxes relating to the main item. The system also provides a summary display window in which all the summary socio-ecological information for a particular species can be viewed at once. This window has buttons attached to the summary statements to enable the user to launch the display windows for the corresponding overview text and diagrams directly.

5.3 There is a considerable amount of data within SANDS already. To try to keep the user from becoming overwhelmed, the selection process has been designed to be highly structured. At the very top level of the selection process data is accessed species by species, the data for only one particular species being accessible at a time. The user can also form subsets of the species, and of the examples based on various search criteria. These search criteria can be concatenated allowing complex questions to be formed, such as "what species are to be found in the Red Sea with a falcate fin and produce whistles?". Specific species or examples can either be selected from these subsets either in a browse manner (using 'next' and 'previous' controls) or directly by means of dialogue boxes listing features of the species or examples (e.g. names, codes, descriptions).

5.4 Rather than develop specific software for the display of the sound, still and movie examples, use has been made of the proprietary software that is available. Such software has been integrated with the main system by means of AppleScript, the Macintosh's batch command language; essentially FoxPro runs a particular AppleScript code which requests the Macintosh Finder application (essentially the Operating System) to open the file using the chosen viewer application.

6. DATA SOURCES

6.1 While ADAC has access to some sonar recordings featuring vocalisations of the various species, it was decided to include in SANDS only sound examples obtained from researchers. There are a number of reasons for this; these include:

- The recordings made by UK ASW units invariably have the sounds of the platform in them, something of which the authorities are justifiably sensitive; using researcher collected data means that the system is totally unclassified.

- In order to give improved detection ranges, ASW platforms often have beamformed arrays. While the full compliment of arrays gives wideband coverage, each sensor only covers a narrow frequency range, and the time-series from them is usually distorted in some way in consequence of the beamforming. The researcher's examples are usually wideband and do not have any spectral shaping.

- Researchers usually have visual sighting of the particular animals making the noise, and also have the experience to make positive identification as to species; often sonar operators have neither.

- Researchers collect data in many areas into which UK ASW assets rarely go.

6.2 Most of the acoustic samples have been provided under contract by the Woods Hole Oceanographic Institution (WHOI), U.S.A from their archives, and the Wildlife Conservation Research Unit (WCRU), Oxford University, U.K. from recent collections. In some instances, the sound cuts have been provided to WHOI and WCRU by other research organisations, including Tethys and the International Fund for Animal Welfare.

6.3 Margaret Klinowska, the University of Cambridge, U.K., under contract collated the socio-ecological information for five of the species, later edited by WCRU to ensure consistency of style.

6.4 WCRU have also supplied, under contract, the socio-ecological information for the remaining 25 species, the distribution maps, the still photographs and the movie sequences. Many of the stills and movie sequences have been supplied by organisations such as the International Fund for Animal Welfare, the Sea Watch Foundation and the BBC Library Services, on the basis that the main use is for education and training purposes. In addition WCRU have provided a range of statistics for the time-frequency characteristics of the 'tonal' vocalisations of most of the species [2].

6.5 The diagrams of the species have been drawn by one of the ADAC team, Dr. L.Lloyd, under advisement of WCRU.

Proceedings of the Institute of Acoustics

THE SEA ANIMAL NOISE DATABASE SYSTEM

7. DATA

7.1 With the resources available during the past three years it has not been possible to cover all the vocal species. After consultation with WHOI and WCRU , emphasis has been placed on thirty species chosen as being the most vocal of the mammals found in the northern hemisphere. In all, data has been collated on 37 species as shown in the following tables:

Baleen Whales

Species		Sounds	Stills	Movies	Text
Blue whale	Balaenoptera musculus	√	√	√	√
Bowhead whale	Balaena mysticetus	√			√
Bryde's whale	Balaenoptera edeni				√
Fin Whale	Balaenoptera physalus	√	√	√	√
Minke whale	Balaenoptera acutorostrata	√	√	√	√
Sei whale	Balaenoptera borealis	√			√
Humpback whale	Megaptera novaeangliae	√	√	√	√
Northern right whale	Eubalaena glacialis	√	√	√	√
Southern right whale	Eubalaena australis	√			

Toothed Whales

Species		Sounds	Stills	Movies	Text
Sperm whale	Physeter macrocephalus	√	√	√	√
Narwhal	Monodon monceros	√	√	√	√
Pilot whale (long-finned)	Globicephala melas	√	√	√	√
Pilot whale (short-finned)	Globicephala macrorhynchus	√	√	√	√
Killer whale	Orcinus orca	√	√	√	√
False killer whale	Pseudorca crassidens	√	√	√	√
White whale	Delphinapterus leucas	√	√	√	√
Melon-headed whale	Peponocephala electra	√			√
Northern bottlenose whale	Hyperoodon ampullatus		√	√	√

THE SEA ANIMAL NOISE DATABASE SYSTEM

Pinnipeds

Species		Sounds	Stills	Movies	Text
Bearded seal	Erignathus barbutus	√			√
Grey seal	Halichoerus grypus	√	√		√
Harp seal	Phoca groenlandica	√	√	√	√
Harbour seal	Phoca vitulina	√			
Walrus	Odobenus rosmarus	√	√	√	√

Dolphins

Species		Sounds	Stills	Movies	Text
Bottlenose dolphin	Tursiops truncatus	√	√	√	√
Common dolphin	Delphinus delphis	√	√	√	√
Risso's dolphin	Grampus griseus	√	√	√	√
Striped dolphin	Stenella coeruleoalba	√	√	√	√
Atlantic spotted dolphin	Stenella frontalis	√	√	√	
Pan-tropical spotted dolphin	Stenella attenuata	√	√	√	√
Clymene dolphin	Stenella clymene	√			
Rough toothed dolphin	Steno bredanensis	√	√	√	√
Harbour porpoise	Phocoena phocoena	√	√	√	√
White-sided dolphin	Lagenorhynchus acutus	√	√	√	√
Spinner dolphin	Stenella longirostris	√	√	√	
Dusky dolphin	Lagenorhynchus obscurus	√			
Fraser's dolphin	Lagenodelphis hosei	√			
White-beaked dolphin	Lagenorhynchus albirostris	√	√	√	√

7.2 It is noted that the number of examples contained within the database varies greatly with species. For instance, there is only one sound example of sei whale in the database, while there are over 450 examples of humpback whale.

8. COMMENTS ON DATA AVAILABILITY

8.1 It was originally intended to include within the database detailed information as to the total population of a species and its distribution throughout the world, including any changes

Proceedings of the Institute of Acoustics

with season as a result of migratory behaviour. However it soon became apparent that such detailed information is not readily available. Sometimes there are detailed population estimates for a species in a particular area, but rarely any reliable collated population estimates for a species in all oceans throughout the world. Moreover the distribution of the species tends to be reported in qualitative rather than in quantitative terms, and the extremes of the species' range are often all that is available.

8.2 In the course of compiling the socio-ecological text and search characteristics with regard to the species' vocalisation habits, a survey of the open literature was undertaken. It was apparent from this study that the various researchers use an inconsistent set of subjective descriptors for animal vocalisations and are also inconsistent in the time-frequency statistics that they report [2]. Moreover they sometimes do not make a distinction between obviously different vocalisation types (even sometimes between clicks and whistles). Thus it has been very difficult to make any correspondence between the vocalisations reported by different researchers, and consequently, between reported vocalisations and the sound examples held in SANDS.

8.3 In collating the examples it has been noted that there is a marked absence of any systematic recording of collateral information relating to the circumstances surrounding the actual recording (e.g. weather conditions, animal behaviour, group size) and relating to the arrangements of the recording (e.g. the equipment used, the person making the recording, the vessel involved). As indicated earlier, this information is important for quality assurance purposes, for establishing the copyright of the examples, and for hypothesis testing.

9. FUTURE PLANS

9.1 Currently a reduced set version of the system is in production for the Macintosh platform This system will have the full functionality of the larger system but will only contain that number of examples that will adequately fit on one CD-ROM (approximately 650 MB). It will contain both typical and some atypical examples. It is intended to supply the application software on a second CD-ROM for installation onto the hard disk of the target platform.

9.2 It is also intended to devise a HTML version of the system containing the reduced data set. While this is unlikely to have the full search capabilities of the main application, it will have the advantage of being accessible on any computing platform with Internet viewing capabilities.

9.3 ADAC is seeking to expand the data set contained in the system, both by obtaining a full data coverage for the current set of species and by expanding this species set.

10. ACKNOWLEDGEMENTS

11. REFERENCES

[1] L J LLOYD (Ed.), ' Proceedings of the Underwater Animal Sounds Database Seminar 31 January - 1 February 1994', Defence Research Agency, Farnborough, Hants, Report No. 5/94, 1994.
[2] J N MATTHEWS, L E RENDELL, A GILL & J C D GORDON, 'Marine Mammal Vocalisations and the Sea Animal Noise Database System (SANDS)', Wildlife Conservation Research Unit, Dept. of Zoology, University of Oxford, 1997.