Analysis of Marine Mammal Vocalization in the MARLANT Operations Areas

Matthew Coffin
Heather Anderson
Joe Hood
Akoostix Inc.

Charlotte Dunn
Bahamas Marine Mammal Research Organisation

Prepared by:
Akoostix Inc.
10 Akerley Blvd, Suite 12
Dartmouth, NS B3B 1J4

Contractor's Document Number: AI-CR-2013-008
Contract Project Manager: Joe Hood, 902-404-7464

PWGSC Contract Number: W7707-4501017273
Technical Authority: James A. Theriault, Defence Scientist
Contractor's Publication Date: March 2013

Contract Report
DRDC-RDDC-2017-C142
June 2017

Principal Author

Original signed by Matthew Coffin

Matthew Coffin

Research Analyst

Approved by

Original signed by Robert A. Stuart

Robert A. Stuart

Head, Technology Demonstration

Approved for release by

Original signed by Calvin V. Hyatt

Calvin V. Hyatt

Chair, Document Review Panel

Abstract

Marine mammals can be adversely affected by anthropogenic sound sources, so mitigation efforts are regularly implemented to reduce the risk to mammals during active-sonar activities. Mitigation efforts can be more effective if prior knowledge of the marine-mammal population in the area is known. This contractor report summarizes Akoostix's effort to develop a database of marine-mammal activity in the waters around Nova Scotia, and includes analysis of existing data sets and identification of potential sources of more data that have either already been collected or could be collected in the future. It was found that the readily available data sets have very sparse coverage, and much more data is required to complete a comprehensive database for this area. A database structure has been designed and the existing observations have been inserted into the database. Notes on these experiments are provided herein, while additional detail is available in Software Tools for Analysis and Research (STAR)-formatted trial archives.

This page intentionally left blank.

This page intentionally left blank.

Table of contents

List of figures

List of tables

Acknowledgements

Akoostix would like to acknowledge the efforts of the Defence Research and Development Canada (DRDC) and the crew of Canadian Forces Auxiliary Vessel (CFAV) Quest. The data analyzed under this trial would not have been gathered without their efforts, care, and diligence. We would also like to acknowledge the efforts of Hilary Moors-Murphy in analysing the Q338 data set under a previous effort.

This page intentionally left blank.

1 Introduction

This contractor report documents work performed under contract W7707-4501017273 for Project Authority (PA), Jim Theriault. The work was performed between February 2013 and March 2013.

Anthropogenic underwater acoustic signals, which are often used during scientific experiments and military exercises, can have adverse effects on marine mammals. These can range from permanent injury to behavioural changes of large groups of marine mammals. For this reason, efforts are regularly put in place to mitigate the risk posed to marine mammals. These mitigation efforts are greatly improved if there is prior knowledge of mammal activity, such as the types of species that may be present in an area, when they might be present, and the characteristics of the acoustic signals they generate. The objectives of this project were therefore aimed at improving our knowledge of marine mammal activities in the Maritime Forces Atlantic (MARLANT) Operational Areas (MARLOAS). Specific objectives included development a database for the cataloguing of marine-mammal vocalizations and observations, the analysis of existing data to populate the catalogue, and the identification of potential sources of marine-mammal data in the community. These were achieved by performing the following:

- A Microsoft Access database was developed, in consultation with the PM, which is used to store data related to marine mammal observations, and provide a user interface allowing users to query the database by time and geographic location.
- Five existing datasets were analysed for marine-mammal activity. Several vocalization types were identified and recorded in the database. Observations of each type of vocalization, including the date, time, and location of the observation, were recorded in the database.
- Potential sources of acoustic recordings containing marine mammal data in both the public and private sectors were identified. Contact was made with various persons of interest to gauge their potential willingness to provide data and collaborate on the future expansion of the work described in this report.

Software was also developed to provide a streaming command-line interface and graphical user interface to the DIFAR demultiplexor.

This report provides a detailed description of the work performed for each of the tasks listed in the Statement of Requirements (SOR). The report is organized as follows:

- Section 2 describes potential sources of existing and/or future data sets that could be used for the detection of marine mammals, and includes a description of five existing data sets that were analysed and used to populate the database.
- Section 3 describes the data analysis technique, including the software tools used, to identify and classify the vocalizations inserted into the database.
- Section 4 describes the structure of the database, including how the how the user interface is configured and how the data is organized inside the database.
- Section 5 outlines recommendations for future work.
- Section 6 summarizes the results of the project with conclusions.
- Annex A provides configuration management (CM) information to help users understand which version of the STAR software was used for this work.

- Annex B provides a brief description of the software used to support this project.
- Annex C describes software development that was performed under this contract.
- Annex D provides a summary of the MARLOAS inserted into the database.
- Annex E provides a list of vocalization types and sub-types that are used to classify marine mammal sounds.
- Annex F describes a data quality rating scale.
- Annex G describes a vocalization classification confidence scale.
- Annex H describes the individual components of the Microsoft Access database.

2 Concept

Our understanding of marine-mammal data activities is largely dependent the quantity and quality of available data from which information can be harvested. Unfortunately, collecting large quantities of marine-mammal data is a significant challenge for a number of reasons. Because of the global distribution of marine mammals, a comprehensive data set would ideally include data from all the world's oceans. Measurements would also have to cover considerable time spans, on the order of years or even decades, to provide reliable information about patterns and/or variations in marine-mammal activities from year to year. While continuous global measurements are not currently feasible due to enormous technical and budgetary considerations, there are numerous groups and agencies around the world that regularly collect marine-mammal data in specific oceanic regions. A practical approach to attaining the best coverage in space and time, then, is to mine data collected from these organizations.

While the scope of this contract wasn't intended to be global in nature, the concept of compiling data from various sources was implemented to assess available data sets collected in the waters off of Nova Scotia. This section describes potential data sources, and includes an analysis of some existing data sets that have been compiled. It is hoped that the initial effort described in this report can be expanded to include more sources of data resulting in a more comprehensive database.

The remainder of this section describes potential data sources and the data sets used to create the initial database, delivered under this contract.

2.1 Potential data sources

During this contract, several organizations were contacted in an effort to assess the current availability of marine-mammal data and potential for further data gathering. Input was received from individuals at each organization regarding their respective data gathering methods and ongoing/future plans, which are summarized below.

There are several other organizations that would likely have valuable input in this area. Further investigation into the types of data currently being gathered, plans for future data gathering and analysis, and interest in collaborating on an international marine-mammal vocalization database would be beneficial.

2.1.1 Government agencies

2.1.1.1 Department of Fisheries and Oceans Canada (DFO)

DFO has a mandate to develop knowledge of marine-mammal life processes and habitat usage. As such, the three eastern DFO Regions (Maritimes, Newfoundland and Labrador, and Quebec) have undertaken a variety of efforts related to acoustic detection of marine mammals (Pacific region has also been actively involved in marine-mammal acoustics). Individual scientists have developed and used technologies to collect acoustic observation data on specific species. For example, Newfoundland and Labrador Region has been developing an on-line catalogue of marine-mammal vocalizations (though many are not geo-referenced).

DFO has been and is expected to continue to work cooperatively with DND. Where possible, they should be engaged in any future acoustic marine-mammal vocalization-data collection efforts. The Maritimes and Newfoundland & Labrador Regions in particular are interested in marine-mammal acoustic data collection and/or analysis (including studies related to assessing habitat usage). Future collaboration may provide an efficient means for addressing knowledge gaps, and through combined effort enable more advanced acoustic research with greater scope and resources.

2.1.1.2 Acoustic Data Analysis Centre (ADAC)

ADAC is responsible for post-mission analysis and archival of all underwater acoustic data recorded by the Canadian Armed Forces during trials, exercises, and operations. These data are primarily recorded and processed by operational systems, which may not be optimized for the detection of all species of marine mammals.

ADAC could be requested to investigate acoustic detections with the same marine-mammal detection parameters as used elsewhere in this study. They could post-process data using flexible systems such as their TruView Rapid Processing and Analysis (RPA) System that may be adapted to enhance detection performance over operational processors.

2.1.1.3 Defence Research and Development Canada (DRDC) - Atlantic

DRDC Atlantic has a sixty-year history of acoustic data collection in the western Atlantic. Over the years, the research ship CFAV QUEST has been at sea for 100 to 180 days per year. This is expected to continue for the foreseeable future. Of the data collection trials, acoustic experimentation has only been included in a portion of the efforts.

Unfortunately, data-collection trials have not generally included objectives involving marine-mammal detection. The acoustic sensors may not have been appropriate for the purpose. Any marine-mammal observations should be treated as opportunistic in that it will be difficult, but not impossible, to determine the effort expended in the data collection.

2.1.2 Universities

2.1.2.1 Dalhousie University and the University of St. Andrews

University researchers (specifically Hal Whitehead's Lab at Dalhousie University in Halifax and the University of St. Andrews Sea Mammal Research Unit [SMRU] in Scotland) have a history of conducting passive-acoustic studies focused on cetaceans and measuring acoustic vocalizations in eastern Canadian waters. In many cases, these studies focus on particular species in particular areas, so the geographic range may be narrow. It may be advisable to inform these researchers of knowledge gaps of interest to DND and consider collaborations to address these knowledge gaps.

2.1.3 Oil and Gas Producers

2.1.3.1 Canada-Nova Scotia Offshore Petroleum Board

The Canada-Nova Scotia Offshore Petroleum Board is responsible for the regulation of petroleum activities in the Nova Scotia offshore Area. It's counterpart, Canada-Newfoundland Offshore Petroleum Board has parallel responsibility in the Newfoundland & Labrador offshore area.

Both organizations expect that offshore oil and gas exploration adhere to the behaviour outlined in "Statement of Canadian Practice with respect to the Mitigation of Seismic Sound in the Marine Environment." The statement of practice includes passive acoustic monitoring for marine mammals as an option, but does not mandate the use. In some instances, where the environmental assessment identifies specific risks to marine mammals, a passive acoustic monitoring effort may be required.

It is less likely that useful data will be available through O&G, than other identified sources of marine mammal observation data.

2.1.4 Industry

2.1.4.1 Akoostix

Akoostix is a small Nova Scotian company that focuses on developing software and processes for surveillance, including that of marine mammals. They are currently exploring opportunities to conduct acoustic research trials that will gather data for developing new marine-mammal surveillance algorithms. They plan to focus on regions of interest to DND such as the Emerald Basin. Akoostix is open to collaboration with DND to extend the utility of these data sets for additional purposes, including those covered under this contract.

2.2 Data sets

Five data sets were analysed for marine-mammal vocalizations, the results of which were inserted into the database described in Section 4. A summary of the datasets is provided in Table 1 and a map showing the time and location for each data set is shown in Figure 1.

The majority of the marine-mammal observations were extracted from the Q325 and Q338 trials, as these were the two largest of all the available data sets. No positive observations were noted in the Q347 or Glider 2012 data sets, and these data were polluted by active sonar and self-noise, respectively. Several types of vocalizations were detected in the data sets, with the most common types including blue and fin whale moans, pilot-whale and delphinid whistles, and sperm- and beaked-whale clicks, among others.

Table 1: List of data sets analysed including the cetacean species that were detected.

DATA SET	DATE	GENERAL LOCATION	RECORDING DEVICES	DETECTED SPECIES

DATA SET	DATE	GENERAL LOCATION	RECORDING DEVICES	DETECTED SPECIES
DRDC trial - Q307	Jul-Aug 2007	43° 55' N, 63° W	Sonobuoy (CO mode only)	Delphinid whistles
DRDC trial - Q325	Oct-Nov 2009	43° 45' N, 62° 45' W	Sonobuoy, SHARP	Blue whale moans, fin whale moans, possible sei whale moans, beaked whale clicks, delphinid whistles
DRDC trial - Q338	May 2011	Marine Protected Area Whale Sanctuary The Gully 44° N, 59° W	Passive acoustic reusable buoy (PARB), sonobuoy, towed array	Delphinid whistles, sperm whale clicks, bottle nose whale clicks, other clicks
DRDC trial - Glider 2012	Feb-Mar 2012	44° 23' N, 63° 20' W	Slocum Glider, sonobuoy, PARB, SHARP	None
DRDC trial - Q347	Oct 2012	43° 45' N, 62° 50' W	Sonobuoy	None

Figure 1:Map showing the positions and dates of the data sets analysed for marine mammal vocalizations.

2.2.1 Spatial and temporal coverage of the existing data sets

The breadth of the data sets is quite limited in both space and time. Spatial coverage exists for only two general areas. Trials Q307, Q325, and Q347 all occurred near Emerald Basin, while the Q338 data covers the area near The Gully. On a broad temporal scale, the existing data has very coarse resolution with gaps of up to two years between data sets. Furthermore, although the collection of data sets spans five years,

the only month that has a repeated measurement is October, for which data exists from Q325 and Q347. Also, there are several months of the year for which there are no data at all.

It is interesting to note that this overlapping month of October related to Q325 and Q347 resulted in much different observations. In one case there were many species observed while none were observed for the other trial. The reason for this difference is not known, but highlights the importance of repeated observation to assess normal patterns. Also, extended observation may also be prudent to identify transient species that may be missed for short trials.

On a smaller temporal time scale, the existing data only provides 'snapshots' of information as the recorded data is limited to, at most, several hours per data set. This is due to the nature of the experiments for which data were collected, resource limitations, and instrument limitations. For example, sonobuoys have a limited operational life ranging from two to eight hours, and towed array data is limited to operational and budget constraints of the ship towing the array.

Furthermore, with the exception of Q338, the measurement of marine-mammal signals was not the focus of the sea trials analysed under this contract, and therefore the sensors were not optimally configured for this purpose. For example, many of the sonobuoys had sampling rates that were too low to measure high-frequency clicks. Other issues such as active sonar and system noise dominating the data further reduced the amount of useful data and complicated the analysis.

If marine-mammal measurements are to be given priority during future experiments, these issues will have to be considered during the trial-planning phase. The result might be extra effort required in data collection through extended measurement periods, or the use of instruments that provide a combination of longer duration measurements and optimal measurement parameters for marine-mammal detection. In any case, taking measurements dedicated to marine-mammal detection will require some sort of modification to the measurement techniques used for the data sets described in this section.

3 Data processing and formatting

Other challenges associated with marine-mammal data are associated with the complexity of the data and variability in the analysis tools and techniques. Marine mammals generate many different types of vocalizations, and although they can be generalized into specific categories and even species, it is often unclear which category a given signal falls into. Knowledge of marine mammal vocalization types and variability is limited, as even researchers focused on this area can be unsure of vocalization classification. A measured signal may fit the criteria for several different, or perhaps no, classifications. Vocalizations from different species may have similar acoustic properties, making them difficult to distinguish. For these reasons, classification often includes a subjective assessment that is based on experience and level of expertise of the analyst.

In addition to subjective differences, the way in which the data is processed can have a dramatic impact on what the analyst is able to see. Parameters pertaining to the FFT size, averaging, filtering, and normalizing, among others, all affect how information is presented to the analyst. For example, certain combinations of processing parameters can be ideal for visualizing a particular type of vocalization, but at the same time can cause other types of vocalizations to become undetectable.

Because of these issues, a well-defined procedure was implemented when analysing the sample data sets. This procedure ensured that the existing data were processed consistently, and provides a guideline that other analysts can follow for other data sets.

3.1 Data analysis methodology

All data were analysed using the Omni Passive Display (OPD) software. OPD was used to create spectrograms for the data, revealing features such as vocalizations and active transmissions. The analysis was generally performed in two passes to isolate the low- and high-frequency vocalizations, respectively. Low-frequency vocalizations include blue and fin whale moans that have dominant frequencies below 100 Hz and durations on the order of seconds. Whistles and clicks, on the other hand, have dominant frequencies ranging from 2-50 kHz. Whistles tend to have durations on the order of seconds, while clicks tend to have durations of fractions of a second.

To view the low-frequency vocalizations, the data were decimated to remove high-frequency components, leaving only the lowest frequencies, typically 0-200Hz. The FFT parameters were adjusted to give a high frequency-resolution of 0.15 Hz, and a time resolution of 0.65 seconds. For high-frequency vocalizations, decimation was disabled, resulting in processing of the full bandwidth data. The FFT parameters were set such that the resulting frequency and time resolutions were typically 12-50 Hz and 0.002-0.02 seconds, respectively.

OPD has a built-in ability to save the current display to an image file and to extract sub-sets of the data to save to a .wav audio file. Both of these features were employed to create sample data to support the vocalization types listed in the database. Whenever possible, the image and audio files for a specific vocalization were created from the same segment of data. Also, attempts were made to ensure that other types of vocalizations did not contaminate the image and audio files for a specific vocalization, although this was not always possible.

3.2 Track information

With the exception of the Q338 data, complete track information of the recording devices was not included with the data sets. Therefore, most of the observations that were inserted into the database have a single position associated with them, corresponding to the deployment site of the device. Track information from PARB buoy and towed-array data for Q338 was available and was included in the observation list in the database. To reduce the number of data points, the positional information for these devices was down-sampled to one data point per hour.

3.3 Digitization of MARLOAS

The MARLANT breaks the waters surrounding Atlantic Canada into operating areas, and the coordinates of these operating areas were inserted into the database. The boundaries of the operating areas were transcribed from [4] and a map [5] obtained from the DND Hydrographical Services Office.

For several areas, the coordinates entered in the database are not the same as coordinates specified by MARLANT. This is because these areas have non-rectangular shapes and/or are bounded by land on one or more sides, while the simple MS Access database will only function properly if the areas are defined by a bounding rectangle as described in Section 4.2.3. For this reason, the boundaries of some of the areas were adjusted to form a rectangle, resulting in some areas overlapping with land, adjacent areas, or both. The adjusted coordinates of the areas are provided in Annex C.

4 Observation database

A database was created using Microsoft Access to contain marine-mammal observation data. The database is intended to act as a comprehensive resource providing information about which species are present in a given area, and when they are present. The database is expandable and could potentially be used by multiple organizations as a common place to pool data from various sources. It is intended that the database will be continually populated as more data is collected and analysed. The data described in Section 2.2 has already been inserted into the database, but much more data are required before it can be used as a standalone reference.

The database may also be used in other ways such as to determine vocalizations related to a specific species or to support classification of unknown vocalizations. These options could prove very useful for training, data analysis, and specification/configuration of monitoring systems.

The database consists of a user interface that allows the user to submit queries and extract information from the database, input records into the database, or modify existing records in the database. The database also has internal structures containing the data, which, in general, are not visible to the user. This section gives an overview of both the user interface and internal structures.

4.1 User interface

The database has three operational modes, allowing the user to:

- Query existing records
- Modify existing records
- Add new data to the database

Only one of these three modes can be invoked at a time.

In query mode, the user enters criteria (for a time range and a geographic area) to return a list of observations of species that meet that criteria. While each query is similar in output, the user has the ability to choose the type of entry format for the geographical area of interest. The user must specify a date range and a geographical area of interest (or use the defaults which will output all available records) to search for observations of species. The geographical area can be specified either by entering the bounding coordinates of the area of interest, or by selecting a pre-defined MARLOA. The geographical search accommodates the specification of the bounding coordinates in degrees, minutes, and seconds, or simply decimal degrees. The date range can be selected from drop-down calendars for simplicity. Query results are returned in the form of a document, which displays observations of species that fall inside the bounds of the query criteria. The search results also show the link between specific observations and the data set for which the data was collected. Note that if blank records are returned, this indicates that an attempted observation was made but no species were observed.

In modify mode, the user has the ability to view the existing records and make changes and corrections if necessary. The user may want to do this if new information about existing records has become available. For example, a vocalization in the database may have been classified as unknown or incorrectly assigned to the wrong species based on limited knowledge at the time of the analysis. Further additions to the

database, or perhaps re-analysis of the existing data by another analyst may result in a different classification being applied to the observations, in which case the database would have to be modified.

In add mode, the user can insert new observations, vocalizations, operational areas, data set definitions, etc. into the database. This would be the case, for example, at the end of a trial or field campaign when there are new data to be analysed and inserted into the database. Or, an analyst may want to insert the boundaries of some defined operating area so that it can be linked to records in the database. This mode would facilitate that need.

4.2 Database backend

The internal structure of the database, where the data is stored, has two primary components: vocalizations and observations. Smaller components are used to contain ancillary data: data sets, species, and operating areas.

4.2.1 Vocalizations

A vocalization is defined as a sound or group of sounds that have similar acoustic properties and are created by a specific species or group of species. The physical properties of a vocalization are defined in the database by its:

- Main vocal type (as outlined in Annex E)
- Sub vocal type (as outlined in Annex E)
- Minimum dominant frequency, in Hz
- Maximum dominant frequency, in Hz
- Minimum inter-click interval or inter-call Interval (ICI), in seconds
- Maximum ICI, in seconds
- Minimum duration of the vocalization, in seconds
- Maximum duration of the vocalization, in seconds

Vocalizations can be attributed to a species (classification) if known, and a reference to support the classification, such as a journal article, can be included with the vocalization. The database supports entry of a free-text description of the vocalization to add context to the vocalization. In general, a vocalization should refer to a sound generated by single species, although this may not be possible if sounds from multiple species have similar acoustic properties, and distinguishing between them is not possible. In this case, giving the vocalization a more generic classification may be necessary. For example, whistles from several species of delphinids may not be distinguishable from each other, so they could all be classified together as “delphinid whistles”.

Due to the variability of marine-mammal sounds, vocalization parameters should not be viewed as rigid definitions of a specific sound into which all signals must fall, but rather a general summary that is subject to variability and change. It is acceptable, and in fact ideal, for multiple vocalization examples for the same type of sound to exist if there are notable differences between the acoustic parameters. This will allow for a qualitative assessment of the variability associated with a certain type of sound. For example, if an analyst detects a sperm whale click with an ICI outside of the range specified in the existing

vocalization for sperm whale clicks, then a new vocalization could be created with the new ICI value. Conversely, simply attributing the detection to the existing vocalization will cause loss of information, as the detection would be linked to a vocalization that doesn't accurately describe it.

Processing parameters, such as frequency resolution, time resolution, and normalization algorithm that the analyst used to visualize or listen to the vocalization can also be specified. A sample image showing the spectrogram, and sample audio file of the vocalization can also be linked to the vocalization to aid future users of the database in identifying the vocalization and comparing it to acoustic signals found in other data sets. In its current form, the image and audio file are not linked directly to the database, and just the names of the files are listed.

4.2.2 Observations

An observation is defined as any measurement that could potentially, but not necessarily, contain signals form marine mammals. Types of observations include acoustic recordings, photographs, or records of mammal sightings. Properties that define an observation, as defined in the database, include:

- Start date and time
- End date and time
- Location of the measurement, given by a latitude and longitude, or list of latitudes and longitudes if the data were collected using a moving sensor over some period of time.
- Mode (acoustic, visual, habitat model, etc.)
- Minimum detectable frequency of the sensor (for acoustic recordings)
- Maximum detectable frequency of the sensor (for acoustic recordings)
- Type of sensor used (i.e. sonobuoy, towed array, camera, etc.)
- Data quality (as outlined in Annex F)
- Whether or not the observation contains any detection (True or False). It is useful to have an observation with no detections to indicate that marine-mammal encounters are less likely.
- Link to the data set that contains the observation (see Section 4.2.3)

If a specific vocalization is present in an observation, that vocalization can be linked to the observation. In this case, the analyst can also add a confidence level as to how certain they are that the signal present in the data corresponds to the selected vocalization. A guide for selecting confidence levels is outlined in Annex G. If there is a signal in the data for which no matching vocalization exists, a new vocalization can be created. Only one vocalization can be linked to an observation entry. If there are multiple vocalization types in a single measurement, multiple observations can be listed, each linking to a different vocalization.

4.2.3 Ancillary data types

Other data types defined in the database include:

- **Data set**: The details of sea trials, field campaigns, or other organized experiments can be inserted into the database as a data set. A data set is comprised of a start date, end date, description, and a reference providing more information and a possible point of contact for data access. Observations

are linked to specific data sets to provide context to the observations, and allow the user to pursue further data analysis if desired.

- **Operational area**: The bounding coordinates of defined regions such as military operational areas, marine protected zones or shipping lanes can be inserted into the database as operational areas. The database uses this information to determine if an observation occurred within a defined operational area, and provides that information to the user. An operational area is defined by a bounding rectangle that is specified by four points: minimum longitude, minimum latitude, maximum longitude and maximum latitude. At this time only rectangular operational areas can be inserted into the database due to limitations with queries and MS Access.
- **Species**: Specific details about each marine-mammal species can be stored as a species table in the database. Vocalizations must be linked to a specific species listed in the database (if the species is not known, the vocalization should be linked to an "Unknown" or "Unidentified" species). Species are defined by:
  - Order
  - Suborder
  - Family
  - Species name
  - Common name
  - Vernacular
  - Description
  - Reference to literature

5 Recommendations

A number of recommendations resulted that should be implemented to improve on this approach:

- It is clear that the set of existing data is not dense enough (temporally and geographically) to gain a thorough understanding of marine-mammal activities and patterns, and that measurement not designed to observe marine-mammals are often not adequate for this purpose. If future measurements are going to be made, they should be coordinated such that the resulting data density is optimized and that the collection parameters are optimal for marine-mammal detection.
- The database currently is not linked to sonogram images or audio files, as this would make the database too large to be portable. A solution that allows these data to be linked to the database while maintaining portability is desirable.
- Operational areas can only be defined as rectangles in the current database. This is restrictive in that it doesn't allow for accurate representation of features such as coastline or areas that are bounded by polygons. This reduces the effectiveness of defining operational areas and can potentially cause confusion if the boundaries of known operational areas have to be intentionally entered into the database incorrectly to satisfy the rectangle requirement.

6 Summary and conclusions

The primary object of this contract was to analyse five acoustic data sets collected from 2007 to 2012 of the coast of Nova Scotia, with the purpose of extracting marine mammal signals. The analysis involved classifying detected vocalizations according to their acoustic properties, and noting the date, time, and position of observations of specific vocalizations. Observations of numerous vocalization types were noted, including blue and fin whale moans, delphinid whistles, and high frequency clicks from sperm and beaked whales.

The results of the analysis are stored in a Microsoft Access database, the design and development of which was a second objective of this contract.

The sparseness of the existing data sets underscores the need to have more data inserted into the database and the need to execute measurements on a more consistent basis. For this reason, several members of the scientific community were approached with the intent of indentifying existing data sets and/or the possibility or desire to collect more data. It is hoped that by pooling together data and resources, the database created under this contract will grow to become a useful tool, both as a data archive and a point of reference, for those interested in marine mammal activities.

References

[1] McInnis, J. and Ryan, G. (2010), OPD User Manual Version 1.4 (AI MA 2008-001), Akoostix Inc., Dartmouth, Nova Scotia.

[2] Flogeras, D. (2010), Active LAND Buoy Operator Manual Version 1.0 (AI MA 2010-009), Akoostix Inc., Dartmouth, Nova Scotia.

[3] Hood, J. (2011), Introduction to the Software Tools for Analysis and Research (STAR): Tools and Processes Version 1.0 (AI MA 2011-007), Akoostix Inc., Dartmouth, Nova Scotia.

[4] MARLANT (2011), Advanced MARLANTLORD, eff 1 Jul 2011, issued under COS Auth, Halifax, Nova Scotia

[5] Department of National Defence (2011), MARLANT Operating Areas (MARLOAS) [Map]. Department of National Defence, Halifax, Nova Scotia., Canada, Hydrographic Services Office, 2011.

This page intentionally left blank.

Annex A Configuration management

The final software deliverable for this contract was provided on the STAR release CD, which was generated and delivered to James Theriault, DRDC's lead scientist for STAR, on 28 Mar 2013. The release coincides with delivery for several STAR DISO call-ups. This section of the document describes the content of that CD.

A.1 STAR branch and release information

Each logical grouping of software modules has been independently versioned on the CD that is also versioned. The current STAR release version is 6.6.12 and contains the following:

- OPD 2.6.1
- ACDC 2.1.8
- SPPACS 1.3.1
- Analysis Tools 6.14.0 (STAR-IDL)

The 6.6.12 release CD was generated in conjunction with other DISO call-ups. Installation instructions are located in the root directory on the release CD.

A.1.1 STAR software documentation

Some manuals, API documentation, and other design documents are provided with the 6.6.4 software release CD. In a standard STAR distribution they can be found by opening *the /usr/local/atools/star-6.6.12/documentation.html* file in a standard web browser. This page contains links to several sets of documentation including:

- Software revision history
- Software API documentation including IDLDoc for the analysis tools (STAR-IDL) and DOxygen generated documents for OPD, ACDC, and SPPACs
- The STAR user manual[1]
- STAR quick reference guides
- STAR-IDL application user manuals
- STAR application user manuals
- Useful third-party Documentation

[1] This manual has not been updated for some time and is in the process of being superseded by newer documentation included on this CD.

A.2 Issue summary

The issue summary in Table 2 shows the current state of known defects for all of the software release candidates listed in Section A.1 as of 28 Mar 2013.

The distribution of issues is indicative of the maturity of the software. Though maturing, much of this software is composed of various evolutions of an iterative design, especially command line SPPACS applications and STAR-IDL components. This software would benefit from general design improvements and refactoring. There are no active blocker issues but there are several critical issues. These are obscure or infrequent bugs that were discovered during current work, but budget or schedule has been insufficient to address them yet. Critical issues are issues that still allow the operator to perform their function but could cause erroneous results or loss of data in those instances. These bugs should be fixed in the near future. Only Blocker issues do not have a work-around and need to be addressed before a contact can be completed successfully.

Table 2: Issue Summary (Severity vs. Status) for all software on STAR release 6.6.12

	OPEN	REOPENED	RESOLVED	CLOSED
Blocker	0	0	0	38
Critical	11	1	0	76
Major	124	4	1	229
Minor	43	2	0	28
Trivial	7	0	0	3
Undecided	0	0	0	5

Table 3 summarizes the critical issues that remain open, but only for software relevant to this contract. None of the critical issues had any effect on the success of this contract. Resolution of these issues may increase efficiency during the execution of future call-ups or contracts.

Table 3: Known Critical issues for OPD, SPPACS and STAR-IDL.

MODULE	ISSUE ID	SUMMARY	DESCRIPTION
OPD	OPDY-77	OPD hangs when validating data sources	OPD tries to verify the data stream by reading as much data as required to determine the format and sensors. This stalls the system until it receives this information. An array server can be accepting connections but not sending data, and OPD connects to the Northern Watch array server automatically as soon as a valid IP and port are given. OPD even saves the IP and port to save users from retyping them each time. If the server stops feeding data on one port (but still allows connections) and starts another, and the user restarts OPD, OPD tries to connect to the previously saved port and hangs. Workaround: 1. Port being saved is in the registry settings. It can be changed manually before resetting OPD or 2. The user can start the array server on the previously saved port. This will free OPD from the hang. Permanent fix (tentative): 1. Prevent OPD from validating data sources automatically by using some kind of connect button. 2. Have a timeout for validating sources since we still have the issue that a socket connection can be made but the data is not flowing.
OPD	OPDY-160	Crash with extreme processing parameters	It is still possible to crash OPD when memory limits are exceeded. The minimum amount of buffer size required to set up an extreme processing stream may still exceeds a PC's memory resources.
OPD	OPDY-174	Bad combination of zero padding/overlap in the processing parameters dialog can cause a hang	If you choose something like 8192 FFT size, 8191 zero pads, and anything but a 0% overlap, you end up with (8192-8191)*(1-overlap) which for any overlap but 0% will cause an integer round to 0. This is an extreme case and is not likely to be used.
OPD	OPDY-244	OPD crashes when stopping processing after EADAQ stops	OPD will crash if the user tries to process EADAQ when EADAQ is not recording.
OPD	OPDY-245	OPD crashes if EADAQ changes state during processing and OPD hasn't stopped processing	OPD will hang if the user stops processing and EADAQ resets. The display will freeze and the user is unable to stop processing.

MODULE	ISSUE ID	SUMMARY	DESCRIPTION
OPD	OPDY-275	Crash on deleting data	A crash was experienced while deleting old data. The user was processing data when the user commanded the deletion, but it appears that processing and deletion were completed at the same time, as the data being processed was also deleted, which may only occur once processing is completed. An OPD crash resulted. Effort to reproduce the crash has been unsuccessful, reinforcing that the user may have gotten the timing exactly right, deleting the data at the exact time that processing completed. See comments for system dump.
SPPACS	AKSP-72	*sp_median_nrmf* is referencing a null pointer occasionally in win32	*sp_median_nrmf* causes a crash (in OPD) by dereferencing a null pointer. This seems to only happen on win32, but may just be hidden in Unix environments (Windows has a history of being more strict, especially in debug mode). Before the crash (during construction of the module) warnings are printed to *stderr,* "Can't find remove point" An examination of the code revealed that there is a comment saying "if we got here, there is a bug".
STAR-IDL	AKOT-163	Failure with Ping to Source Mapping - Tracking Feature Extraction	A failure was experienced when attempting to map a specific ping to a source when running the tracking feature-extraction application. The following error message appeared: "Could not find active source for blast. Ignoring blast." Workaround (verified): Remove all but the desired ping from the NAD ping file.
STAR-IDL	AKOT-153	Animation with selected tracks - bad behaviour	When a user selects tracks in ITAC then starts animation, strange tracks were observed. If the user turns off selections, the animation behaves normally again.

MODULE	ISSUE ID	SUMMARY	DESCRIPTION
STAR-IDL	AKOT-150	Crash / Hang in STAR++	A hang was experienced using STAR++ on OSX (for the Q320 tracking3 data set, while a crash resulted for TMAST02 data on Linux. The user was executing a Monte-Carlo simulation for the OSX runs and had this feature both off and on for Linux. The settings were generally: - 2 detections, 10% threshold - 2 Clusters, 3km radius - El-El, El-Brg, El-Hy crossings - Monte-Carlo simulation with all error types and cluster results - Search time adjustment of 0.5 seconds On TMAST02 data the user had designated data on channels 2,3,4,5. For the Q320 tracking3 data set a threshold of 550 was used.
STAR-IDL	AKOT-107	Ownership of overlays and contained data	There is a general problem with ownership of overlays and contained data by the tactical plot. The tactical plot can be closed and reopened several times during an application's lifetime, so if it destroys all data that it contains it will be lost to the application and cannot be used on subsequent instantiation of the tactical plot, or usage by other modules (i.e. tracker). It may be reasonable just to own the overlay itself and not the contained data, but then ownership needs to be assigned to something. Another option may be to notify the tactical plot when it owns an overlay. Currently image overlays are owned by the tactical database, because it would be more effort to create a data container in the database and then force the creation of an overlay after the fact. This may need to change depending on how this issue is resolved. If two objects are created (image container and image overlay) then caution must be exercised when data is passed between them to avoid expensive data copies for large images.

Annex B Software Tools

This section provides background information necessary to understand the role that DRDC software played in this contract. This flexible reusable software enabled the project to make better use of the available budget, advancing the research more than would be otherwise possible. Their relationship to the project is described below while a high-level description of the tool itself is provided in the subsections below:

- SPPACS was used to basic manipulation of the raw data prior to analysis. For sonobuoy data, useful channels were extracted using the *sp_extract*, and the *sp_ph* module was used to modify the headers of some .wav files that had no or incorrect time stamps.
- STAR-IDL was used to extract Quest and PARB buoy track information from the NADAS and NMEA data files that accompanied the Q338 data set.
- OPD was used for the main analysis of the data. Vocalizations were detected by displaying the data on OPD's gram display using appropriate processing parameters, and excerpts of data were extracted using the .wav extractor tool and the save current view tool.

B.1 Signal Processing Packages (SPPACS)

SPPACS is a group of software programs that are written in the C/C++ programming languages, with each application providing a specific processing or utility function. They are designed to run on Linux and OSX based PCs and typically work with Defence Research Establishment Atlantic (DREA) formatted data files (DAT), though format converters are also contained in the suite. SPPACS has slowly evolved to its present day state.

The SPPACS software suite consists of two types of software. One type is runtime executables. These applications have proven to be very useful in simplifying data management and sonar processing tasks by providing a set of tools from which to build the necessary, often much customized, processing streams. These streams can be run from the command line or assembled into scripts to perform batch-processing tasks allowing large amounts of data to be automatically and incrementally processed.

The second form of the software is a group of library functions that can be used by other programs to efficiently perform standard tasks. These library functions are extensively used by the runtime software, but can also used for other applications, such as OPD. There are several types of libraries of which three are most commonly used in SPPACS:

- Utility (e.g. math, geo, filesystem, …) libraries that consist of utility routines for performing tasks, such as header manipulation, geospatial data representation, and command line parsing.
- Signal Processing (e.g. splib) libraries that contain modules for low-level signal-processing. A new SPPACS module typically consists of one or more SPLIB modules linked together with an SPPACS user interface.
- Sonar Processing (e.g. sonlib) libraries that contain modules consisting of several SPLIB modules linked internally to create a complex sonar module, such as passive processing.

B.1.1 Background and Design Information

More generic and reusable software was created by separating the library code above from SPPACS. These modules are independent of the data header format, time-stamping method, etc., and are suitable for integration in real-time processing systems. The libraries can be built to run on a number of UNIX, OSX or Microsoft Windows platforms and on less common processors such as the ARM core and Texas Instruments (TI) DSP. Once successfully ported, the CMAKE build environment supports subsequent builds with a command line option.

The C and C++ elements of the libraries are intentionally separated to ensure that the core capability, found primarily in the C modules, can be readily ported to systems that don't support the more complex language features employed in the C++ version of the libraries. For the most part, the C++ layer consists of a wrapper on the C layer that provides a more generic method of instantiating, connecting and running modules. This is provided by inheritance that is, in part, the adoption of a common interface from a base class allowing parts of the system to interact with a module without knowing the details of the module. Connection of SPPACS applications using UNIX pipes provides similar functionality at the application layer.

SPPACS is also supported by a number of libraries, such as the Fastest Fourier Transform in the West (FFTW), helping to ensure that the SPPACS software runs as efficiently as possible, while providing a significant reduction in coding effort. These dependencies, and the associated licenses, are tracked for those projects that require knowledge of intellectual property (IP).

B.2 STAR-IDL

The STAR-IDL[2] tools were developed to support general research and analysis objectives at DRDC Atlantic. The actual software goes hand-in-hand with an analysis process that is intended to help formalize a reliable and consistent research and analysis methodology [3]. The primary objectives of the STAR-IDL tools are:

- Provide scientific grade analysis tools that allow for efficient, detailed quantitative and qualitative analysis of a data set.
- Provide scientific grade algorithm prototyping and refinement tools that can be used to quickly realize a variety of algorithm options, validate the basis of the algorithm, and determine the best approach to use for system prototypes.
- Support synergy between DRDC groups and the Department of National Defence (DND) by providing a common software base for analysis. This synergy encourages inter-group communication and simplifies user training, analysis process development, documentation, and data portability.
- Support cost and analysis efficiency by providing software reuse and common tools and data formats. Examples of efficiency would be using the output of analysis from one group to feed the inputs of another, or using common software components to lower development cost of several custom analysis tools.

[2] The STAR-IDL tools were formerly referred to as the Software Tools for Analysis and Research (STAR). The STAR Software Suite has now come to mean the greater tool set, including OPD, ACDC, SPPACS, etc.

Most STAR-IDL components are currently implemented using Interactive Data Language (IDL), though the design is not restricted to IDL. For example, localization algorithms contained in C++ libraries are accessed from IDL.

Applications in the STAR-IDL tools are built using a combination of reusable and custom components that meet the requirements of each application. The layered design and common components allow for rapid and logical development of new capabilities. Though currently focused on two main areas - sonar data processing and analysis, and target localization, tracking, and multi-sensor data fusion - the tools are capable of expanding to meet other analysis and research requirements.

B.3 Omni-Passive Display (OPD)

OPD is a standalone signal processing application designed to run on UNIX, OSX, and Microsoft Windows platforms. It can be used to quickly produce sonogram, energy-time integration (ETI), amplitude-line integration (ALI), and time-series output from DREA digital acoustic tape (.DAT/.DAT32) files, wave files, sound card, EADAQ, Rapidly Deployable System (RDS), and Northern Watch. The following functions summarize its capability (detailed information can be found in the OPD User Manual [1]):

- A user can quickly set up the desired signal processing by loading in a preset configuration from storage, or by simply defining the desired frequency and time resolution. A more sophisticated user can define a wide range of parameters, including Fast Fourier Transform (FFT) size, zero padding, overlap, quantization range, decimation, sonogram compression and much more.
- OPD provides an optional beamformer and is capable of processing complex heterodyned time-series data.
- Annotations can be added to the data.
  - The user can assign a category (or classification) to the annotation from a list of presets as well as provide free-form text to associate with the annotation.
  - Previously generated annotations are displayed on screen when processing data associated with the annotation.
  - The annotation format is compatible with STAR-IDL and ACDC.
- Each processing result is stored in memory and can be selected for viewing and analysis. Analysis tools include a crosshair cursor for time-frequency measurements.
- The entire sonogram can be saved to an image file to capture the output for reports, etc.
- A WAV extraction tool allows the operator to define a region within a sonogram and clip the raw data associated with the selected bounds into a wave file.
- Operational measurement tools such as harmonic, banding, periodic-event and Doppler cursors can be used to analyze advanced features in data and learn tactical information about potential targets.

Annex C Software development

DIFAR data consists of three acoustic data channels, which are multiplexed into a single data channel for transmission via a Radio Frequency (RF) channel. A DIFAR demultiplexing application was developed to separate these data channels again, generating three output data channels for each input data channel when provided with appropriate data.

The core software consists of DRDC-supplied code that performs the processing, but the code were improved to make it suitable for cross-platform use (MS Windows, OSX, Linux) and generalized integration into signal processing applications by wrapping the code to create a STAR-compatible Signal Processing Library (SPLIB) module. The SPLIB module was then integrated into an SPPACS application to allow it to be connected to other modules for efficient batch processing and into a Qt-based application for simplified point-and-click use via a graphical user interface. A user manual was created to accompany the Qt-based application that is provided on the STAR release CD, while standard command-line help is available for the SPPACS application.

As the core DRDC DIFAR demultiplexing software contains valuable intellectual property (IP), the source code is not delivered with the STAR distribution, nor is it contained in the main Akoostix repository that holds the rest of STAR. Instead, it is maintained in a protected repository and binary-only builds are provided with the STAR distribution.

Annex D MARLANT Operating Areas (MARLOAS)

Table 4: Modified MARLANT Operating Areas (MAROLOAS).
This table shows the boundaries of each area, shown in the "Area Name" column. The boundaries are defined by the minimum longitude, minimum latitude, maximum longitude, and maximum latitude. The "Modified" column indicates whether or not the area boundaries have been modified from the official MARLOAS definition. The "Overlap" column indicates which areas overlap the area defined in that particular row.

AREA NAME	MIN LONG	MIN LAT	MAX LONG	MAX LAT	MODIFIED	OVERLAP
ALPHA	-63.667	44.32	-63	44.47	yes	
BRAVO	-64.083	44.32	-63.67	44.7	yes	C1
CHARLIE 1	-64.8	44	-64	44.6	yes	B
CHARLIE 2	-65	43.5	-64	44	yes[3]	
CHARLIE 3	-65	43	-64	43.5	no	HALIFAX
ECHO 1	-63	44	-62	44.98	yes	
DELTA 1	-64	44.17	-63.75	44.32	no	
DELTA 2	-63.75	44.17	-63.5	44.32	no	
DELTA 3	-63.75	44	-63.5	44.17	no	
DELTA 4	-64	44	-63.75	44.17	no	
ECHO 2	-63.5	44	-63	44.32	no	
FOXTROT 1	-66.9	44	-64.8	45.6	yes[4]	
FOXTROT 2	-66.9	43	-66	44	no	
FOXTROT 3	-66	43	-65	43.85	yes	HALIFAX
FOXTROT 4	-66	42	-65	43	no	HALIFAX
FOXTROT 5	-66.9	42	-66	43	no	
GOLF 1	-64	43.5	-63.5	44	no	
GOLF 2	-63.5	43.5	-63	44	no	
GOLF 3	-63.5	43	-63	43.5	no	
GOLF 4	-64	43	-63.5	43.5	no	
HOTEL 1	-63	43.5	-62.5	44	no	
HOTEL 2	-62.5	43.5	-62	44	no	
HOTEL 3	-62.5	43	-62	43.5	no	
HOTEL 4	-63	43	-62.5	43.5	no	
HALIFAX	-66	42.78	-64	44.32	yes	C3,F3,F4,L1[5]

[3] Southwestern corner of C2 moved to F3.
[4] Much of the Minas Basin and Chignecto Basin have been excluded from the area to avoid overlap with areas A, B, C1, D1-4, and E2.

AREA NAME	MIN LONG	MIN LAT	MAX LONG	MAX LAT	MODIFIED	OVERLAP
INDIA	-62	44	-61	45.33	yes	
JULIET	-61	44	-60	45.9	yes[6]	P,Q1
LIMA 1	-65	42	-64	43	no	HALIFAX
LIMA 2	-65	41	-64	42	no	
LIMA 3	-65	40	-64	41	no	
MIKE 1	-64	42	-63	43	no	M1H
MIKE 2	-64	41	-63	42	no	
MIKE 3	-64	40	-63	41	no	
MIKE 1 HOTBOX	-64	42	-63	42.5	no	M1[7]
NOVEMBER 1	-63	42	-62	43	no	
NOVEMBER 2	-63	41	-62	42	no	
NOVEMBER 3	-63	40	-62	41	no	
PAPA	-68	45.6	-58	49.2	yes	J,Q1,Q2
QUEBEC 1	-60.7	45.28	-55.5	48	yes	J,P,Q2,S1
QUEBEC 2	-60	44.75	-55.5	46	yes	P,Q1,Q3
QUEBEC 3	-60	44	-55.5	45.27	yes	Q2
SIERRA 1	-55.5	46	-50	48.66	yes	Q1,S2,S3
SIERRA 2	-55.5	45.27	-50	46.8	yes	S1,S3
SIERRA 3	-55.5	44	-50	46.17	yes	S1,S2
MARLAOS whole	-68	40	-50	49.2	NA[8]	

[5] Overlap present in MARLAOS definition, and is not due to adjustment of boundaries.
[6] Chedabucto Bay and Strait of Canso have been excluded from the area to avoid overlap with area I
[7] Overlap present in MARLAOS definition.
[8] Area defined by the extremities of all sub-areas. Includes regions that are not officially part of MARLOAS.

Annex E Vocalization types and sub types

VOCALIZATION TYPE	AVAILABLE SUB-TYPES
Call	Call AM call Constant call Harmonic call
Click	Click Click bursts Click codas Click trains Deep-water clicks Echolocation Clicks Fast click sequence Neonate clicks Rapid clicks Single clicks or slow click sequence Slow clicks Surface/slow clicks Usual clicks Various clicks
Chirps	Chirps AM Down-sweeps FM Up-sweeps
Song	Moans Social Song Tonal moans
Whistles	Harmonic FM whistles Whistle Whistles/rapid narrowband pulses

VOCALIZATION TYPE	AVAILABLE SUB-TYPES
Other	Bark Creaks Doublets "A" and "B" Grate Growls Grunts Gunshot Harmonic grunt Horn blasts Low-freq narrowband Noisy vocalizations Rasp Ratchet Screams Shrieks Slaps Slow creaks Squawks Squeals Thump trains Thumps in pairs Tones Underwater blows Yelp
Pulse	Pulse Discrete pulses Modulated Pulse Pulsed sounds Pulse bursts Pulsed moans Pulsed tones Pulse trains Pulsive Ragged pulse

Annex F Data quality scale

DATA QUALITY SCALE NUMBER	DESCRIPTION
1	Poor quality or degraded signal that cannot be clearly distinguished even when filtered (e.g. bad clipping or RF interference)
2	Fair quality with aural characteristics that are hard to hear without filtering
3	Average quality with audible aural characteristics
4	Good quality with easily identified aural characteristics
5	Excellent quality with very clear aural characteristics

Annex G Vocalization classification confidence scale

CONFIDENCE SCALE NUMBER	DESCRIPTION
1	Low confidence – The classification is based on limited information and analyst judgment.
2	Medium confidence – The classification is based on characteristics that are known to be related to the selected classification from direct experience, formal training, or an intelligence guide.
3	High confidence – The classification is based on multiple, well-known or generally accepted cues in the larger data segment (i.e. sequence of events surrounding the signal). Cues could be referenced from formal training (lesson plans) or intelligence guides.
4	Certain (Cert) – There must be known, direct confirmation of the signal's origin (e.g. whales of the selected species were sighted at the same time that a sound was identified to come from the same location).

Annex H Observation database objects

H.1 Object descriptions

OBJECT TYPE	NAME	DESCRIPTION
Table	tblConfidenceLevel	Lookup table; provides a numbered scale for describing the analyst's level of confidence in a species/vocalization identification.
	tblDataQuality	Lookup table; provides a numbered scale for describing the quality of data from which observations are made.
	tblDataSet	Stores information about a data set as a whole, such as name and dates that data was gathered.
	tblDataStatus	Lookup table; provides a list of states that are associated with a data set (e.g., analysed, unanalysed, etc.)
	tblIUCNstatus	Lookup table; provides a list of states associated with a species as designated by the International Union for Conservation of Nature
	tblMainVocalType	Lookup table; provides a list of vocalization types.
	tblObservation	Stores information about specific observations within a data set, such as the time, location, sensor frequency, and confidence in the detection.
	tblObservationMode	Lookup table; provides a list of modes for observations (how the observation was detected).
	tblObservationPoint	Stores individual points (time and location) where an observation occurred (multiple related observation points possible for a single species observation)
	tblOpArea	List of defined operations areas (simplified as a 'box', with four bounding points).
	tblOpAreaPrefix	List of operations area categories (or prefixes for the individual operations areas, e.g. MARLOA).
	tblReference	List of reference sources for vocalization and species descriptions (can include literature, web sources, personal communications, etc.).
	tblSARAstatus	Lookup table; provides a list of states associated with a species as designated by the Species at Risk Act.
	tblSensorType	Lookup table; provides a list of sensor types that are used to detect a species.
	tblSpecies	Lookup table; provides information about a species, such as Latin and common names, description, and SARA/IUCN status.
	tblSubVocalType	Lookup table; provides a list of vocalization subcategories.
	tblVocalization	Stores information about known vocalizations as related to a species; includes characteristics of the vocalization, processing parameters, example image and sound files, and references to literature or other sources.
Form	frmConfidenceLevel	Pop-up form which displays the confidence scale with descriptions; based on tblConfidenceLevel.

OBJECT TYPE	NAME	DESCRIPTION
	frmDataSet	Main form based on tblDataSet which is used to enter details about a data set. Contains the subforms for entering data points and observations (linked via DataSetID).
	frmNavigation	Main navigation form which provides buttons to access forms for user input; automatically opened when database is opened.
	frmOpArea	Provided for adding additional operations areas (default datasheet view). Based on tblOpArea.
	frmQueryCriteria	Provides fields for entering query parameters and buttons which generate reports based on the entered criteria. Default values present all recorded observations in the database.
	frmSpecies	Provided for adding (or modifying) species in form view (as opposed to the table method). Based on tblSpecies.
	frmVocalization	Used for adding (or modifying) known vocalizations. Based on tblVocalization.
	sfrObservation	Subform based on tblObservation used to enter multiple observations associated with a single data set.
	sfrObservationPoint	Subform based on tblObservationPoint used to enter multiple observation points associated with a single species observation.
Query	qryConfidenceLevelDescription	Used to display the description for a selected confidence level on the observation subform within frmDataSet.
	qryDataQualityDescription	Used to display the description for a selected level of data quality on the observation subform within frmDataSet.
	qryFilterVocalizationsOnSpecies	Used to display vocalizations for the combo box on the observation subform within frmDataSet.
	qryGeographicRange	Displays all records of observations based on a time range and latitude/longitude range in degrees (decimal).
	qryGeographicRange2	Displays all records of observations based on a time range and latitude/longitude range in degrees (minutes, seconds, direction).
	qryOpAreaObservations	Displays all records of observations based on a time range and selected operations area.
	qryTimeRange	Displays all records based solely on a time range.
Report	rptGeographicRange	Visual output for qryGeographicRange.
	rptGeographicRange2	Visual output for qryGeographicRange2.
	rptOpAreaObservations	Visual output for qryOpAreaObservations.
	rptTimeRange	Visual output for qryTimeRange.

H.2 Field descriptions

FIELD	TYPE	DESCRIPTION	VALIDATION
tblConfidenceLevel			
ConfidenceID	AutoNumber	Primary key; numbered identifier.	

FIELD	TYPE	DESCRIPTION	VALIDATION
Label	Text	Defines (with text) the level of confidence associated with the primary key values one through four.	
Description	Text	Explanation of the criteria that should be met to use the associated confidence level.	
tblDataQuality			
DataQualityID	Number	Primary key; numbered identifier.	
Label	Text	Defines (text) the quality of the data associated with primary key values one through five.	
Description	Text	Explanation of the criteria that should be met to use the associated data quality level.	
tblDataSet			
DataSetID	AutoNumber	Primary key; numbered identifier.	
DataSetName	Text	Descriptive label given to the data set; should generally be unique.	Required
DataSetDescription	Text	General information about the data set.	
StartDate	Date/Time	Beginning date of the data set.	Required; format: mm/dd/yyyy
EndDate	Date/Time	End date of the data set.	Required; format: mm/dd/yyyy. Must occur on or after StartDate.
DataStatusID	Number	Foreign key; stores the value from lookup table *tblDataStatus*.	
Reference	Text	Reference to the location or owner of the data.	
tblDataStatus			
DataStatusID	AutoNumber	Primary key; numbered identifier.	
Status	Text	Defines the state the data is in (e.g., analysed or unanalysed)	
tblIUCNstatus			
IUCNstatusID	AutoNumber	Primary key; numbered identifier.	
Status	Text	Assessment of a species according to the IUCN.	
Description	Memo	Detailed explanation of the associated IUCN status (what it means).	
tblMainVocalType			
MainVocalTypeID	AutoNumber	Primary key; numbered identifier.	
MainVocalType	Text	Name of the vocalization.	
tblObservation			
ObservationID	AutoNumber	Primary key; numbered identifier.	
DataSetID	Number	Foreign key; relates the observation(s) to a data set (multiple observations to a single data set).	

FIELD	TYPE	DESCRIPTION	VALIDATION
StartDate	Date/Time	Beginning date of the observation.	Required; format: mm/dd/yyyy
StartTime	Date/Time	Beginning time of the observation	format: hh:mm (24 h)
EndDate	Date/Time	End date of the observation (form dd/mm/yyyy).	Required; format: mm/dd/yyyy. Must occur on or after StartDate.
EndTime	Date/Time	End time of the observation	format: hh:mm (24 h)
ObservationModeID	Number	Foreign key; stores the value from lookup table *tblObservationMode*.	
MinFrequency	Number	Minimum frequency of the sensor (hertz).	Required; >= 0
MaxFrequency	Number	Maximum frequency of the sensor (hertz).	Required; >= 0
SensorTypeID	Number	Foreign key; stores the value from lookup table *tblSensorType*.	
DataQualityID	Number	Foreign key; stores the value from lookup table *tblDataQuality*.	
SpeciesID	Number	Foreign key; stores the value from lookup table *tblSpecies*.	
VocalizationID	Number	Foreign key; stores the value from lookup table *tblVocalization*	
ConfidenceID	Number	Foreign key; stores the value from lookup table *tblConfidenceLevel*.	
DetectionsFound	Yes/No	Checkbox (of value true or false); true if detections are found during this observation attempt and false if no detections found during this observation attempt.	Default value is true
tblObservationMode			
ObservationMode	AutoNumber	Primary key; numbered identifier.	
Mode	Text	Defines an observation method or mode (e.g., Acoustic or Visual).	
tblObservationPoint			
ObservationPointID	AutoNumber	Primary key; numbered identifier.	
ObservationID	Number	Foreign key; relates the observation point to a specific observation in *tblObservation*.	
StartDate	Date/Time	Start date of the observation point.	Required; format: mm/dd/yyyy
StartTime	Date/Time	Start time of the observation point.	Required; format: hh:mm (24h h)
Latitude	Number	Latitude of the observation point in degrees.	Required; Between -90 and 90.
Longitude	Number	Longitude of the observation point in degrees.	Required; Between -180 and 180.
tblOpArea			

FIELD	TYPE	DESCRIPTION	VALIDATION
OpAreaID	AutoNumber	Primary key; numbered identifier.	
OpAreaPrefixID	Number	Foreign key; stores the value from lookup table *tblOpAreaPrefix*.	
OpAreaName	Text	Defines (text) the operation area (area of interest).	
MinLat	Number	Minimum latitude of the bounding box in degrees.	Required; between -90 and 90
MaxLat	Number	Maximum latitude of the bounding box in degrees.	Required; between -90 and 90 and greater than MinLat.
MinLong	Number	Minimum longitude of the bounding box in degrees.	Required; between -180 and 180.
MaxLong	Number	Maximum longitude of the bounding box in degrees.	Required; between -180 and 180 and greater than MinLong.
tblOpAreaPrefix			
OpAreaPrefixID	AutoNumber	Primary key; numbered identifier.	
OpAreaPrefixName	Text	Name of the operations area category.	
tblReference			
ReferenceID	AutoNumber	Primary key; numbered identifier.	
ReferenceLabel	Text	Short label (similar to the in-text portion of a bibliography item).	
ReferenceFull	Text	Full source listing.	
Link	Hyperlink	Location of the reference (link to website or offline publication).	
tblSARAstatus			
SARAstatusID	AutoNumber	Primary key; numbered identifier.	
Status	Text	Assessment of a species according to SARA.	
Description	Memo	Detailed explanation of the associated SARA status (what it means).	
tblSensorType			
SensorTypeID	AutoNumber	Primary key; numbered identifier.	
SensorType	Text	Defines the type of sensor used to gather data.	
tblSpecies			
SpeciesID	AutoNumber	Primary key; numbered identifier.	
Order	Text	Scientific classification.	
Suborder	Text	Scientific classification.	
Family	Text	Scientific classification.	
SpeciesName	Text	Latin name of the species (should be unique).	Required.
CommonName	Text	Most commonly known name of the species.	

FIELD	TYPE	DESCRIPTION	VALIDATION
Vernacular	Text	Other commonly known or used name(s).	
Description	Memo	General description of the species.	
ReferenceID	Number	Foreign key; stores the value from lookup table *tblReference*.	
SARAstatusID	Number	Foreign key; stores the value from lookup table *tblSARAstatus*.	
SARAstatusDate	Date/Time	The most recent update of the species' SARA status.	Format: dd/mm/yyyy
IUCNstatusID	Number	Foreign key; stores the value from lookup table *tblIUCNstatus*.	
IUCNstatusDate	Date/Time	The most recent update of the species' IUCN status.	Format: dd/mm/yyyy
tblSubVocalType			
SubVocalTypeID	AutoNumber	Primary key; numbered identifier.	
SubVocalType	Text	Name of a secondary vocalization.	
MainVocalTypeID	Number	Foreign key; links the secondary vocalization (*SubVocalTypeID*) to a main vocalization category from *tblMainVocalType*.	
tblVocalization			
VocalizationID	AutoNumber	Primary key; numbered identifier.	
SpeciesID	Number	Foreign key; links a species from *tblSpecies* to this *VocalizationID*	Required
MainVocalTypeID	Number	Foreign key; primary vocalization category (stores value from *tblMainVocalType*).	
SubVocalTypeID	Number	Foreign key; secondary vocalization category (stores value from *tblSubVocalType*).	
Description	Text	General description of the vocalization.	
MinDomFreq	Number	Minimum dominant frequency (hertz).	>0
MaxDomFreq	Number	Maximum dominant frequency (hertz).	>0
MinICI	Number	Minimum inter-click interval.	>0
MaxICI	Number	Maximum inter-click interval.	>0
MinDuration	Number	Minimum duration of the vocalization (seconds).	>0
MaxDuration	Number	Maximum duration of the vocalization (seconds).	>0
FreqResolution	Number	Frequency resolution (hertz) of processing.	>0
TimeResolution	Number	Time resolution (seconds) of processing.	>0
NormalizationType	Text	Type of normalization used for processing.	
ImgFile	Hyperlink	Link to a sample image.	
WavFile	Hyperlink	Link to a sample .wav file.	
Reference	Hyperlink	Link to information source - literature, sea trial etc.	

This page intentionally left blank.

List of symbols/abbreviations/acronyms/initialisms

DND	Department of National Defence
DRDC	Defence Research & Development Canada