



Enhancing the utility of the Rapid Environmental Assessment database through the use of in situ and modelled data sets during Q316

Anthony W. Isenor

Defence R&D Canada – Atlantic

Technical Memorandum
DRDC Atlantic TM 2008-212
November 2008

This page intentionally left blank.

Enhancing the utility of the Rapid Environmental Assessment database through the use of in situ and modelled data sets during Q316

Anthony W. Isenor

Defence R&D Canada – Atlantic

Technical Memorandum

DRDC Atlantic TM 2008-212

November 2008

Principal Author

Original signed by Anthony W. Isenor

Anthony W. Isenor

Defence Scientist

Approved by

Original signed by Mark McIntyre

Mark McIntyre

AS/H Maritime Information Combat Systems

Approved for release by

Original signed by Ron Kuwahara for

Calvin Hyatt

Chair DRP

© Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defence, 2008

© Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, 2008

Abstract

The Rapid Environmental Assessment database (READB) is an assembly of ocean environmental data collected by DRDC Atlantic and its predecessors. The database was a result of a contract effort that combined temperature, sound speed, transmission loss, and ambient noise data into a single database management system. As well, the database contains bathymetry data and temperature-salinity climatology for the Scotian Shelf. The database is intended for use while ashore and also while at-sea. Trial Q316 was the maiden voyage of the database. During the trial, the database was used extensively by utilizing previously existing data within the database and by incorporating data collected during the trial. The database was used as a data source for other applications being used to assess the surrounding oceanographic environment. The trial also highlighted potential improvements in the coupling between the external data sources and the database. Based on the Q316 database investigation, numerous recommendations are made for the database, for the redesign of the database, and how external data sources can more easily be incorporated into the database.

Résumé

La base de données d'évaluation environnementale rapide (BDEER) est une série de données sur le milieu océanique recueillies par RDDC Atlantique et ses prédécesseurs. Elle a été créée grâce à des efforts déployés en vertu d'un marché pour combiner des données sur la température, la vitesse du son, l'affaiblissement de transmission et le bruit ambiant en un seul système de gestion de base de données. En outre, elle contient des données bathymétriques et des données climatologiques température-salinité pour le plateau néo-écossais. Elle a été conçue pour servir au large des côtes et en mer. L'essai Q316 a été marqué par sa première utilisation. Durant l'essai, on en a fait un usage exhaustif en utilisant les données qui s'y trouvaient déjà et en y intégrant les données recueillies au cours de l'essai. Elle a servi de source de données pour d'autres applications qui ont permis d'évaluer le milieu océanographique ambiant. L'essai a permis de faire ressortir des améliorations possibles au couplage entre les sources externes de données et la base. D'après l'examen qui en a été fait durant l'essai Q316, de nombreuses recommandations sont formulées au sujet de la base de données et de son réaménagement et de la façon d'y intégrer plus facilement les sources externes de données.

This page intentionally left blank.

Executive summary

Enhancing the utility of the Rapid Environmental Assessment database through the use of in situ and modelled data sets during Q316:

Anthony W. Isenor; DRDC Atlantic TM 2008-212; Defence R&D Canada – Atlantic; November 2008.

Introduction or background:

The Rapid Environmental Assessment database represents a compilation of data used in scientific and defence research conducted by DRDC Atlantic. The database represents the primary management system for data collected by DRDC Atlantic and contains much of the DRDC trial data including expendable bathythermograph profiles, nonacoustic environmental data such as air temperatures and humidity, ambient noise data, and transmission loss data.

The database (DB) was constructed under contract, with support provided by the Technology Evaluation for Rapid Environmental Assessment (REA) Applied Research Program. One of the functions of the DB is to provide data support to tactical decision aids. This trial tested the functionality of the READB.

Results:

The database was utilized in a quasi real-time mode by incorporating environmental data collected during a transmission loss experiment. The database ingested temperature data, sound speed data, and numerical model output, and provided these data to the internal network of CFAV Quest. Utilization by tactical decision aids resulted in sonar performance predictions in the form of range dependent transmission loss estimates. The investigation has produced recommendations on standardization of incoming data formats and the redesign of many database components.

Significance:

This at-sea evaluation of the REA database has provided valuable information on potential usage and design of the REA system. The investigation has shown that improvements are needed to the process of data transfer from external sources to the database, and also how the database can be restructured to be more accommodating of these data.

Future plans:

The READB is currently being redesigned to accommodate shortcomings of the current design. As well, efforts are underway to standardize many of the external data source formats into a common form. This involves the coordination of, and collaboration with, the many external data providers.

Sommaire

Enhancing the utility of the Rapid Environmental Assessment database through the use of in situ and modelled data sets during Q316:

Anthony W. Isenor; DRDC Atlantic TM 2008-212; R & D pour la défense Canada – Atlantique; Novembre 2008.

Introduction ou contexte:

La base de données d'évaluation environnementale rapide (BDEER) représente une compilation de données utilisées lors de recherches scientifiques et relatives à la défense menées par RDDC Atlantique. Elle constitue le principal système de gestion des données recueillies par RDDC Atlantique et contient une bonne partie des données d'essai de RDDC, y compris les profils bathythermographiques non récupérables, les données environnementales non acoustiques (comme l'humidité et la température de l'air), les données sur le bruit ambiant et les données sur l'affaiblissement de transmission.

Elle a été constituée en vertu d'un marché, le soutien étant fourni dans le cadre du programme de recherche appliquée sur l'évaluation de la technologie d'évaluation environnementale rapide (EER). Elle sert notamment à fournir un soutien sous forme de données à l'égard des aides aux décisions tactiques. L'essai a permis d'en vérifier la fonctionnalité.

Résultats:

La base de données a servi en mode de temps quasi-réel grâce à l'intégration des données environnementales recueillies lors d'une expérience sur l'affaiblissement de transmission. Elle a recueilli des données sur la température, des données sur la vitesse du son et des résultats de modèles numériques, puis elle a fourni ces données au réseau interne du NAFC Quest. L'utilisation de ses données par les aides aux décisions tactiques a permis de formuler des prédictions sur le rendement des sonars sous la forme d'estimations de l'affaiblissement de transmission en fonction de la distance. L'examen a permis de préparer des recommandations sur l'uniformisation des formats des données d'entrée et de réaménager de nombreux éléments qui s'y trouvent.

Portée :

Cette évaluation en mer de la BDEER a permis de recueillir des renseignements utiles sur la conception et l'usage possible du système EER. L'examen a montré qu'il faut apporter des améliorations au processus de transfert de données de sources externes à la base de données, et comment la base de données peut être restructurée pour mieux intégrer ces données.

Recherches futures :

La BDEER est en voie de réaménagement pour remédier aux lacunes de sa conception actuelle. En outre, des efforts sont déployés pour uniformiser de nombreux formats des sources externes de

données. Cette uniformisation suppose la coordination des nombreux fournisseurs de données externes et une collaboration avec eux.

This page intentionally left blank.

Table of contents

Abstract	i
Résumé	i
Executive summary	iii
Sommaire	iv
Table of contents	vii
List of figures	viii
List of tables	x
Acknowledgements	xi
1 Introduction.....	1
1.1 Outline.....	2
2 Connection to the READB	4
2.1 Granting permission for remote connection	4
2.2 RPC Setup	4
2.3 Constructing a data source connection point.....	5
3 At sea database access – Three use cases	9
3.1 Use case 1 - Existing data sources in the READB	9
3.2 Use case 2 - Ocean Feature Analysis reports as part of a GIS product	18
3.3 Use case 3 - Importing in situ and model data into the READB	20
3.3.1 Data sources	21
3.3.2 Importing in situ data	21
3.3.3 Importing numerical model data	23
4 The READB and large data sets	29
4.1 Issues on ingesting external data into a database.....	29
4.2 Options for storage of external data	30
4.2.1 All data in a DBMS.....	30
4.2.2 Portion of the data in the READB.....	31
4.2.3 Data external to a DBMS	31
4.3 When data should go into the READB.....	31
4.4 Three-way approach	32
5 Concluding remarks and recommendations.....	34
References	38
List of symbols/abbreviations/acronyms/initialisms	41
Distribution list.....	43

List of figures

Figure 1: The ODBC Data Source Administrator provides the ability to create a connection between a PC's applications and the data resource. The PostgreSQL system data source is used to create a specific PostgreSQL connection layer for the users of the PC.	6
Figure 2: The Create New Data Source identifies the driver used for the data source.	6
Figure 3: The ODBC driver setup requires information specific to the data resource. The computer needs to be identified from which the resource is served, and the name of the resource needs to be provided. The user details for the resource are also required. This is not the username and password for the user's account on the server. Rather, it is the username and password for user access to the database.	7
Figure 4: The SQL first ensures that a table named t_slice_1 does not exist. Next, a view is created which contains an id number, longitude, latitude, depth and temperature. These fields are filled based on content from the tsd_temperature table fields for point ID (pid), the x and y geometry (the_geom) positions, the depth value from position 3 of the array, and the temperature value from position 3 of the array; and only on the condition that the month value is 9 and the temperature point ID equals the grid point ID.	11
Figure 5: The Data Link properties provide information on which driver to use for the connection.	12
Figure 6: The Data Link Properties Connection tab provides specific information on which data source is being connected to the application.	13
Figure 7: A selection condition can be constructed within the ArcMap™ environment. The conditional statement shown here limits the "temp" field (i.e., temperature) to only those values greater than 0.	15
Figure 8: The ESRI® ArcMap™ Layer Properties Symbology tab allows the user to adjust the color scale used for the points.	16
Figure 9: The temperature climatology for 11.5 m depth, in September for the western Scotian Shelf. The color coding is as follows: blue: 12.3-15°C ; green: 15-16°C; yellow: 16-17°C; red: 17-18°C.	17
Figure 10: The ESRI® ArcMap™ software allows the user to import ASCII gridded data and create a raster representation. The MetOc OFA analysis was provided in netCDF, converted to ASCII and then imported using this ArcMap™ tool.	19
Figure 11: The imported OFA with a continental overlay (shown in brown) displayed in ArcMap™. The temperature range of the figure is 2°C (white region near Greenland) to 30°C in the Gulf Stream. The approximate operation area of Q316 is shown by the black box.	20
Figure 12: The XBT and XSV drop points (red dots) and the C-NOOFS / Mercator grid locations (blue dots). Grid locations are numbered according to the unique identifier assigned in the READB. The brown lines indicate bathymetry in	

intervals of 100 fathoms except around Brown’s Band which is 50 fathoms. The GMOOS Buoy “N” is also shown. The entire grid domain was loaded into the READB, but only ts profiles from the cells at the grid locations included in this figure were loaded into the READB. 25

Figure 13: Sound speed comparison from four different sources. Mercator model temperature profiles were used to compute the model sound speed profiles at four model locations. The in situ temperature data from the XBT and MCTD profiles were also used to compute sound speed. Finally, the XSV profiles of sound speed are included. The legend indicates the model grid location identifiers as assigned by the READB while in situ data is identified by a similar unique READB identifier. .. 27

Figure 14: A summary of the factors that influence the decisions of whether or not to store data within the READB or external to the READB..... 32

List of tables

Table 1: The table structure of two tables used to store the gridded Dalhousie temperature / salinity climatology. The <code>tsd_geopoints</code> table contains data which define the grid while the <code>tsd_temperature</code> table contains the temperature values at the previously defined grid cells.	10
Table 2: A comparison of the deepest model grid location and bathymetry estimated from digital chart CA176140 (corresponds to CHS paper Chart 4003). The next deeper model level is provided and whether or not this level could exist based on the chart bathymetry is also indicated. Ten of the 20 points indicate that the next model level should exist.	28

Acknowledgements

I would like to extend a special thank you to Adam Lundrigan at the Newfoundland Fisheries and Oceans office. Adam provided valuable advice and support on the netCDF numerical model output used during the Q316 trial.

This page intentionally left blank.

1 Introduction

Rapid Environmental Assessment (REA) refers to a military's ability to rapidly assess its surrounding environment in order to improve effectiveness of response [1]. In terms of the navy, this means a rapid assessment of the atmospheric and oceanographic conditions in the operation area for the goal of improving sensor and weapon performance. This environmental understanding is based in large part on data collection for those parameters important to the operation. For example, in antisubmarine warfare operations, the water column temperature profile is a critical factor in determining sound speed propagation. In the absence of collected data, historical data provide background information from which average conditions can be determined.

To be effective, both the collected and historical data need to be readily accessible. This typically means the data are in a system which properly manages the data and makes these data readily available for associated processing. For DRDC Atlantic, both historical data and data collected on individual trials have not traditionally been in a managed system. However, recent developments are placing these historical data in such a system for the purpose of researching activities, practices and concepts associated with REA.

Part of this DRDC research has resulted in the construction of the Rapid Environmental Assessment database (READB). The READB is a compilation of the environmental data collected by DRDC Atlantic. In fact, it is not an overstatement to suggest that the database now represents the primary management system for DRDC Atlantic historical environmental data. The database contains expendable bathythermograph (XBT) profiles, expendable sound velocity (XSV) profiles, nonacoustic environmental data such as air temperatures and humidity, ambient noise data, and transmission loss data [2, 3].

The database was constructed under contract [4], with funding provided by the Technology Evaluation for Rapid Environmental Assessment Applied Research Program (Project 11cq). During development, one of the envisaged functions of the READB was to provide data support to tactical decision aids (TDA). For anticipated sea trials the database was considered the logical location for the storage of collected data, effectively providing central data management functions and a singular access point for the TDAs data requirements.

The READB was constructed using the open source database management system known as PostgreSQL [5]. PostgreSQL is reported to have one of the more robust open source geographic information system (GIS) add-ons available, that being PostGIS. This combination of database management system and GIS has been used by numerous other groups facing similar environmental data issues to those faced by DRDC Atlantic [6-11].

The READB was used at sea for the first time on CFAV Quest trial Q316. The trial, conducted on 08-19 September 2008, had 14 primary objectives [12] including:

- the testing and evaluation of newly developed equipment;
- supporting interdepartmental collaborations;
- supporting departmental projects currently underway; and also

- to develop and test interfaces relating to the Environmental Model Manager (EMM) [13] and the Rapid Environmental Assessment database. This objective included external data sources such as forecast data from ocean circulation models (provided by Fisheries and Oceans Canada (DFO) and/or TRINITY Meteorological and Oceanographic (MetOc) Office Halifax), and/or in situ data from onboard instrumentation.

The Q316 trial objective dealing with interfaces with the READB made particular reference to the EMM. However, the EMM is simply an example of one application that could conceivably interact with and utilize the contents of the READB. In a more general sense, we are interested in the interaction between external applications and the READB.

1.1 Outline

This brief report documents the activities conducted on Q316 that exercised the capabilities of the READB. The report first describes how to connect a remote personal computer (PC) to the database. This is important as it provides a general connection approach which can be used by applications executing in the Microsoft Windows[®] environment on the remote PC. The connection also requires a subtle modification to the PostgreSQL database management system, which is also documented.

The report then provides three examples of utilizing the database. The examples include:

- using existing data within the database;
- using external data in combination with data from the database; and
- incorporating external data into the database for distribution and use.

The first example deals exclusively with data already existing in the DB. However, an additional manipulation of the data provides a more specific product thus exercising the capabilities of the DB.

The second example deals with the Ocean Feature Analysis (OFA) product produced by the Department of National Defence (DND) TRINITY MetOc office. The MetOc produces a map of surface ocean conditions for the North West Atlantic twice per week, based on accumulation of remote sensed imagery and in situ data. Incorporation of the OFA product into a GIS environment then allows the combination of this product with other data from within the READB.

The third example deals with external data, either in situ data collected during the trial or numerical model output. These data are read directly into the READB for use by other applications.

The report concludes with several recommendations. These recommendations are founded on the knowledge and experience gained during use of the READB on Q316. These recommendations

are directed at READB users, data providers, the modelling community, and DRDC Atlantic management and staff that intend to use the READB.

2 Connection to the READB

While ashore, the READB resides on a computer running Linux and managed by the Maritime Environmental Awareness group of DRDC Atlantic. During Q316, the database was deployed on a Dell Latitude D820 laptop running Windows XP[®]. This computer was known on the Quest local area network as “Polar”.

The connection between remote computers and the Polar laptop provides these remote PCs with the ability to access the READB. However, the connection process is not trivial. For the instructions that follow, we assume both Polar and the remote PC have successfully connected to the Quest local area network.

2.1 Granting permission for remote connection

To connect a remote PC (RPC) to the Polar machine running the READB, the PostgreSQL environment must first grant permission for the connection to be established. This is done via the PostgreSQL configuration file `pg_hba.conf` (i.e., a text configuration file for the database management system; DBMS). The file is contained in the `PostgreSQL/8.2/data` directory. The file may be accessed via the folder structure or via the PostgreSQL menus under the Windows Start Programs menu system. The Administrative account on Polar should be used to modify the `pg_hba.conf` file.

The `pg_hba.conf` file controls remote access to the PostgreSQL environment. Without proper permissions, the PostgreSQL system will not allow users from remote computers to access databases in the PostgreSQL environment. To allow access by the RPC, obtain the RPC IP number; assume this IP number is 123.123.12.123. Then load the `pg_hba.conf` file into a text editor and add the following line to the bottom of the file:

```
host all all 123.123.12.123/32 md5
```

This line instructs PostgreSQL to allow all users on the RPC (denoted by its IP address) access to all databases hosted by PostgreSQL. More detailed instructions for editing `pg_hpa.conf` can be found in the configuration file; however, the above line will serve to completely open the PostgreSQL environment to all users on the RPC.

2.2 RPC Setup

The RPC may also require additional software to access the READB. The access technique constructed here is based on open database connectivity (ODBC) protocol. For this method of connection, the RPC requires the PostgreSQL ODBC driver to be loaded. For this set of instructions we assume the RPC is running a version of the Microsoft Windows[®] environment.

The following steps should be followed to load the PostgreSQL ODBC drivers.

1. Acquire PostgreSQL ODBC driver. The Windows driver can be downloaded from the PostgreSQL site [14].
2. Unzip the driver file and double click the .msi file. Follow the install instructions.
3. Then double click the update.bat file to install any updates.
4. Reboot your computer.

2.3 Constructing a data source connection point

The ODBC driver allows applications on the RPC to access the READB via an ODBC connection point, known as an ODBC data source. The data source identifies all the connection parameters for the database, in this case the READB. The ODBC Data Source Administrator provides the ability to create this connection layer.

To construct the ODBC data source on the RPC, first access the Windows Control Panel followed by the Administrative Tools. Next, access the Data Sources (ODBC). A system level ODBC connection can be constructed, thereby allowing all users on the RPC to connect via the ODBC connection. Under the System DSN tab (see Figure 1), click the Add button.

The Create New Data Source window is then displayed (Figure 2). Here is where the specific ODBC driver is selected for accessing the resource. Select the PostgreSQL ANSI item and click Finish.

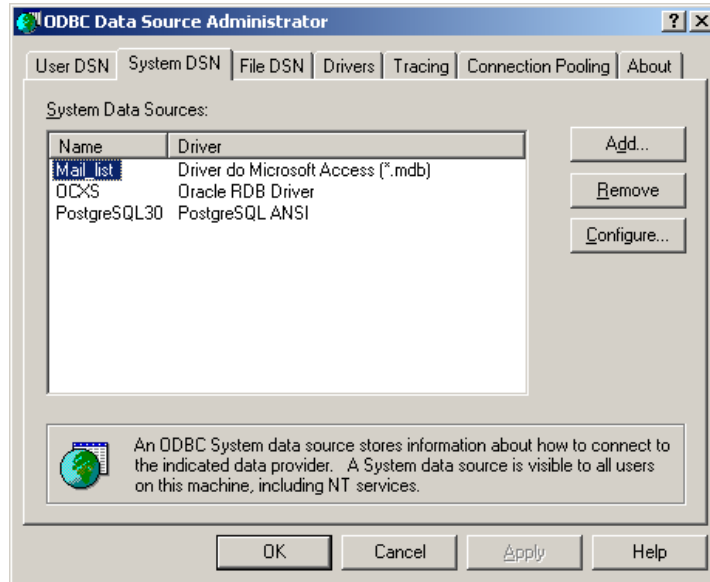


Figure 1: The ODBC Data Source Administrator provides the ability to create a connection between a PC's applications and the data resource. The PostgreSQL system data source is used to create a specific PostgreSQL connection layer for the users of the PC.

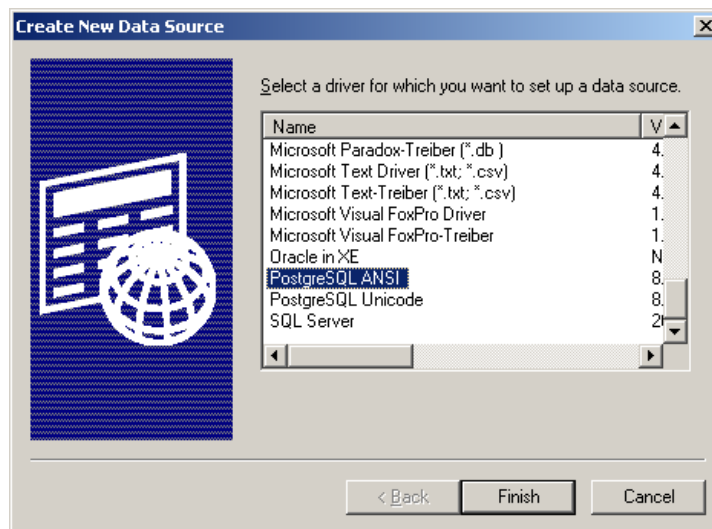


Figure 2: The Create New Data Source identifies the driver used for the data source.

Complete the ODBC connection entry window according to the following (see Figure 3):

- The Server IP address is the address of the PC hosting the PostgreSQL environment.
- The Database is the name of the READB that is being accessed. During Q316 we were accessing the READB version 3 beta, which was named “readbv3b”.
- SSL mode and Port should default to the values shown.
- The Username and Password corresponds to the REA database (i.e., the username and password for logging into the REA database itself). This is not the username and password for the remote user access to the computer hosting the PostgreSQL environment.

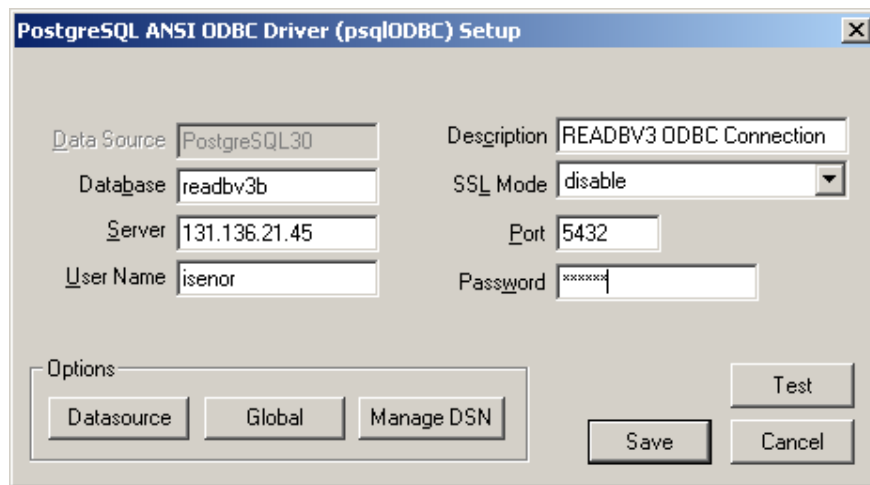


Figure 3: The ODBC driver setup requires information specific to the data resource. The computer needs to be identified from which the resource is served, and the name of the resource needs to be provided. The user details for the resource are also required. This is not the username and password for the user’s account on the server. Rather, it is the username and password for user access to the database.

Click on the Test button to test connectivity to the remote database. Assuming the test is successful, you have now created an ODBC data source which will act as a layer between the RPC applications and the READB.

This remote access provides some capabilities that are otherwise not available directly in the PostgreSQL environment in which the READB resides. For example, within the READB environment constructed as part of the contract development, the user can display data collection locations and some data via a browser interface (for this discussion the READB browser interface will be referred to as REA BI). The linking of the REA database tables to the REA BI requires programming expertise as well as knowledge of the PostGIS methods of encoding the geospatial positions. However, the remote access provides the user with the capability of creating remote

tables within another database management system (e.g., Microsoft Access[®]), and the linking of these tables to the existing tables within the READB.

Additionally, the ODBC connection allows applications on the RPC to access the READB while maintaining access to the RPC's local data resources. Local resources could be in the form of databases or specific to the RPC application. An example of one such application is ESRI[®] ArcMap[™] which allows the RPC user connection to local resources as well as the remote READB.

3 At sea database access – Three use cases

One objective of trial Q316 was to investigate the usage and utility of the READB in a sea going environment. In what follows, we consider three use cases that highlight different approaches to utilizing the DB while at sea. In particular, the use cases deal with:

- utilizing data currently stored within the READB,
- utilizing data external to the READB in combination with READB data, and
- importing in situ data into the READB for utilization by remote applications

The first use case utilizes data sources existing within the READB. This example demonstrates an enhancement to the functionality of the current READB system by allowing user control over the GIS interface. Although such control is not currently available in the REA BI, this should not be considered a limitation of the BI. The BI was intended to provide quick look access to the content of the DB. The BI was never intended to have the functionality of a full GIS package. This use case also highlights the benefits of having access to the capabilities of a full GIS while at sea.

The second use case utilizes a large spatial scale gridded data source external to the READB. Throughout Q316, the TRINITY MetOc office was providing Q316 with reports on the sea surface temperature (SST) conditions of the Northwest Atlantic. These reports were from the MetOc Ocean Feature Analysis (OFA) product. The OFAs are produced twice per week and represent a compilation of remote sensed imagery and in situ data over a three or four day period. These data are not part of the existing READB nor were the data added to the READB. The data are being accessed by the ESRI® ArcMap™ GIS independent of the READB, but are being combined with data from the READB using the ArcMap™ interface and the ODBC data source established previously.

The third use case will deal with external data that is imported directly into the READB and made available to all applications using the DB. This use case covers both in situ data collection and numerical model output. The use case covers the more traditional instrumentation platforms (e.g., conductivity-temperature-depth (CTD) vertical profiling instrument) used during Q316. The use case also deals with circulation model output. This use case demonstrates the utility of the READB as a shared real-time data repository for use by tactical decision aids.

3.1 Use case 1 - Existing data sources in the READB

This example application focuses on the ability of the user to manipulate the data within the READB, and produce an output product which is accessible via third party software. The product produced in this use case would not be accessible via the existing functionality provided by the REA BI.

The READB contains a wide assortment of data types. For this use case, we focus on the gridded temperature-salinity dataset. These data were a result of a 2001 Technology Investment Fund (TIF) [15] effort involving DRDC Atlantic and Dalhousie University. One output of the effort was a gridded climatology of monthly temperature and salinity (ts) over the Scotian Shelf and slope at 15 depth levels from 0 to 500m.

The READB stores the ts gridded data in three tables: `tsd_geopoints`, `tsd_temperature` and `tsd_salinity`. Table 1 indicates the table structure for the `tsd_geopoints` and `tsd_temperature` tables. Knowledge of the exact table structure is required to perform manipulations on the data. These manipulations will be carried out in the PostgreSQL database using Structured Query Language (SQL) [16].

Table 1: The table structure of two tables used to store the gridded Dalhousie temperature / salinity climatology. The `tsd_geopoints` table contains data which define the grid while the `tsd_temperature` table contains the temperature values at the previously defined grid cells.

Table name	Field name	Brief field description
tsd_geopoints	pid	A unique identifier for each point in the grid.
	depth	The depth array object. All 15 depths are present in the array.
	the_geom	The latitude and longitude stored in PostGIS geometry field.
tsd_temperature	pid	A unique identifier for each point in the grid.
	temp	The temperature values in a 15 element object array.
	month	The month for which these temperature values correspond.

Note from Table 1 that the temperature and depth (i.e., also salinity) values are stored in a PostgreSQL object; in this particular case a double precision array. This means that a single field in the table stores all 15 elements in the array. If we were to access these data via the ODBC connection (e.g., from Microsoft Access[®]), these objects would appear as simple text strings enclosed by braces such as `{}`. This is because the ODBC connection is not capable of understanding the complex object (i.e., the double precision array).

Manipulating such complex objects using SQL via an ODBC connection is quite difficult. Since the ODBC connection does not recognize the object, it considers the object as a text string. The ODBC connection would access the entire array in a form similar to `{2.3, 4.3, 5.3}` (here, only three data values are shown while in the actual table structure there would be 15 values). This means parsing of the string is required to identify the braces and the comma separated list of values. Parsing is not a standard function of SQL.

Thus, we will manipulate the ts data within the environment in which it was created, namely PostgreSQL. However, the product we produce will be accessible via external GIS software.

The manipulation will create a single layer of temperature data over the entire grid. The output of the selection will be stored in a PostgreSQL “view” table. A “view” represents a virtual table. The view table does not actually contain data, but rather it stores the SQL command that combines data from other tables. When the user selects the view, the SQL command is executed with the results made available to the user. Thus, view tables do not impose additional storage requirements on the DB, because a view only represents a particular combination of the data that is already contained in existing tables.

For the example, we create a view which represents a horizontal slice of the ts gridded data. We arbitrarily select a 11.5 m depth for the month of September (i.e., the month of the Q316 trial). The SQL which corresponds to this view selection is shown in Figure 4.

```
drop view t_slice_1;
create view t_slice_1 (id, lon, lat, dep, temp) AS (select
tsd_temperature.pid,  x(tsd_geopoints.the_geom),
y(tsd_geopoints.the_geom),  tsd_geopoints.depth[3],
tsd_temperature.temp[3] from tsd_temperature, tsd_geopoints where
tsd_temperature.month=9 and tsd_temperature.pid =
tsd_geopoints.pid);
```

Figure 4: The SQL first ensures that a table named t_slice_1 does not exist. Next, a view is created which contains an id number, longitude, latitude, depth and temperature. These fields are filled based on content from the tsd_temperature table fields for point ID (pid), the x and y geometry¹ (the_geom) positions, the depth value from position 3 of the array, and the temperature value from position 3 of the array; and only on the condition that the month value is 9 and the temperature point ID equals the grid point ID.

Once the view is created in PostgreSQL, it can be loaded into an application on the RPC. For this use case, we use the ESRI[®] ArcMap[™] application.

¹ The geometry is a PostGIS-specific terminology. In a PostGIS database, the latitude and longitude of a location are not stored in degrees North-South or East-West. Instead, the latitude-longitude position is encoded into what appears as a long text string (e.g., the position 45N, -63W is encoded as 0101000020AD1000000000000000804FC0000000000804640). This encoding plus the PostGIS functions allows more complicated queries. To convert back to latitude-longitude, the PostGIS provides the x and y functions (e.g., x(the_geom) activates the x function).

The view is accessed within ArcMap™ by first adding a data frame to the ArcMap™ display layer. The specific view can only be selected after the data source is added to ArcMap™ by doing the following:

- In ArcMap™, click Add Layer button or under the File menu, select Add Data.
- In the Look In window dropdown box, select Database Connections
- Pick Add OLE Database Connection
- Select the ODBC connections as shown in Figure 5. Click Next.
- Under the Connection tab (see Figure 6), select the ODBC connection point created in section 2.3. Enter the username and password for the READB.

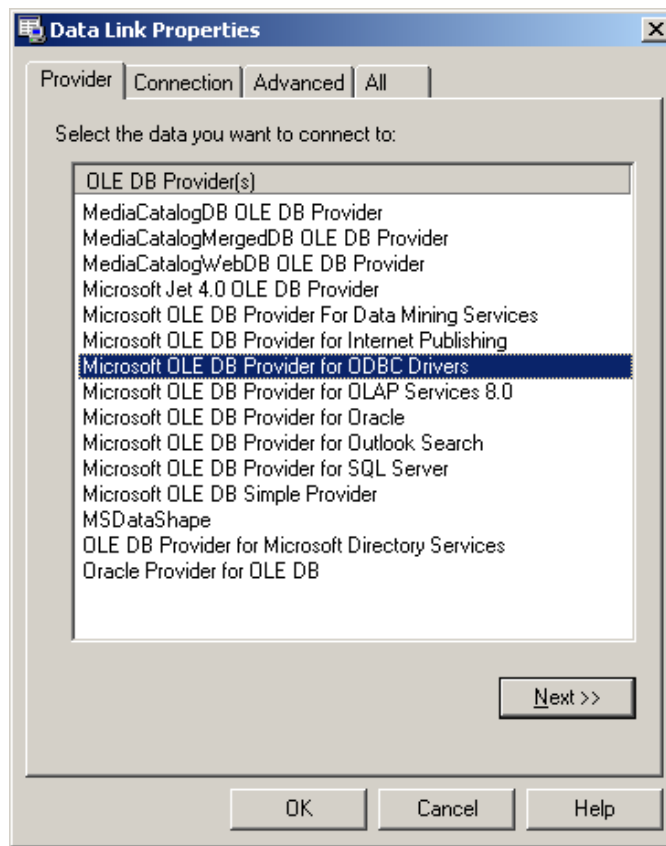


Figure 5: The Data Link properties provide information on which driver to use for the connection.

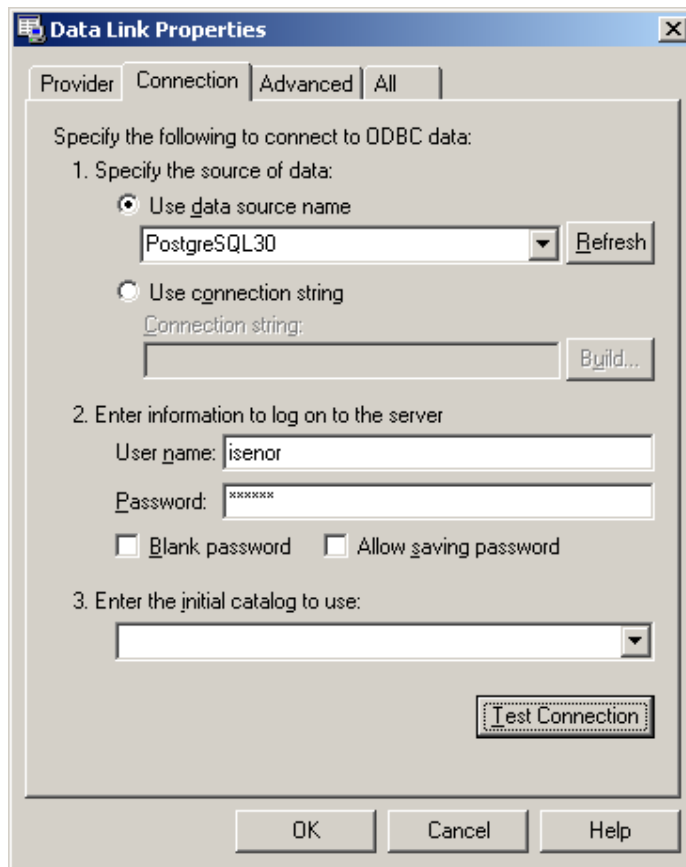


Figure 6: The Data Link Properties Connection tab provides specific information on which data source is being connected to the application.

- Select the newly created OLE Connection for the data source name.
- Enter the username and password for the data connection. Click OK.
- Select the tables or views you wish to use in the layer, in this case view “t_slice_1” is selected.
- In the ArcMap™ layer window panel, the additional layer should appear. Now instruct ArcMap™ which fields in the view represent positional coordinates. To do this, right click on the added layer (i.e., the view source) and select Display XY Data.

- Select the X and Y fields from the dropdown list boxes which correspond to the longitude and latitude values². Click the ok button.

The initial display of the data within the ESRI® environment shows a large region of grid points all with an identical color. As well, many points exist over Nova Scotia. Closer examination shows that these over-land values are in fact zero temperatures.

Additional criteria could have been placed on the creation of the view (see Figure 4) that would eliminate the zero temperature values. However, for this example it is instructive to utilize the ArcMap™ functionality to perform this sub-selection.

ArcMap™ provides the capability to perform queries on the ArcMap™ layer data. The capability is provided via a limited SQL statement; specifically that part of the statement that follows the WHERE in an SQL SELECT statement. In this example, we wish to omit the display of those temperature values at zero. To accomplish this within the ArcMap™ environment, we access the Properties window for the layer. The Definition Query tab provides access to a statement builder where we can create that part of the SQL text which follows the WHERE clause in “SELECT * from tablename WHERE”. Since we would like to eliminate the zero values over land, and since we recognize the fact that zero temperatures at 11.5 metres depth in September on the Scotian Shelf are not physically possible, we build a query to remove all zero temperatures based solely on the zero value. The query text is shown in Figure 7. The criteria simply states that only temperatures > 0 should be selected. The string “temp” refers to the field name in the view (see use of the field name temp in Figure 4).

² This is required because ArcMap is not naturally capable of understanding the geometry created by PostGIS.

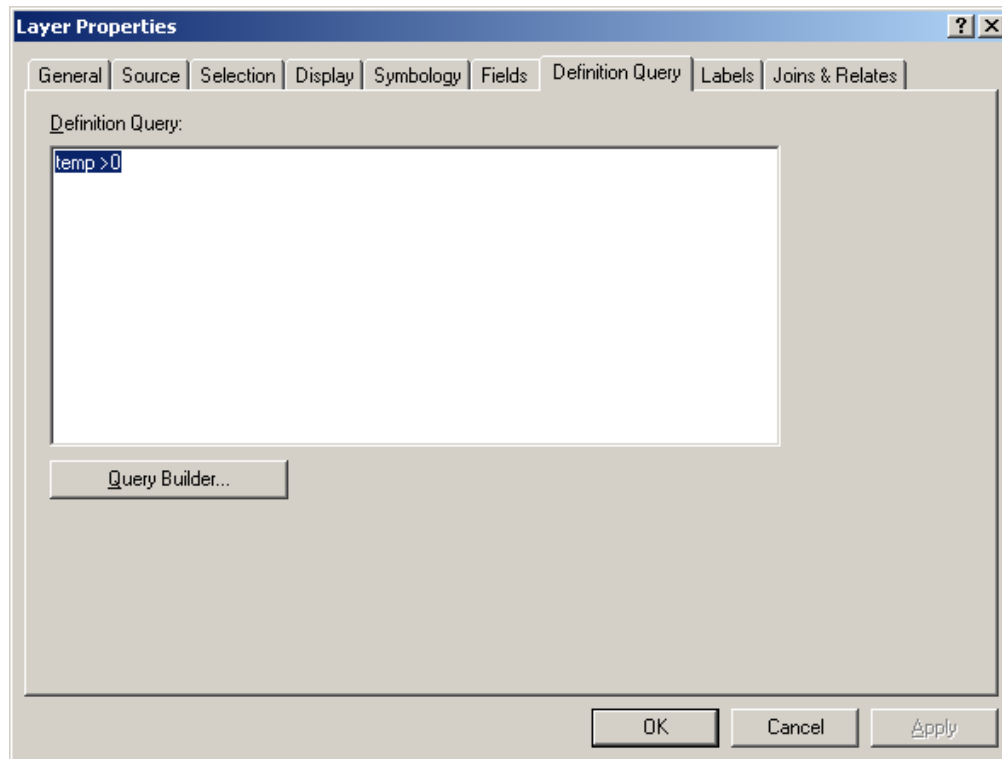


Figure 7: A selection condition can be constructed within the ArcMap™ environment. The conditional statement shown here limits the “temp” field (i.e., temperature) to only those values greater than 0.

The points which represent the ts grid values can also be color coded based on their temperature. This is possible through the Symbology tab on the properties window for the layer. The Symbology tab is shown in Figure 8.

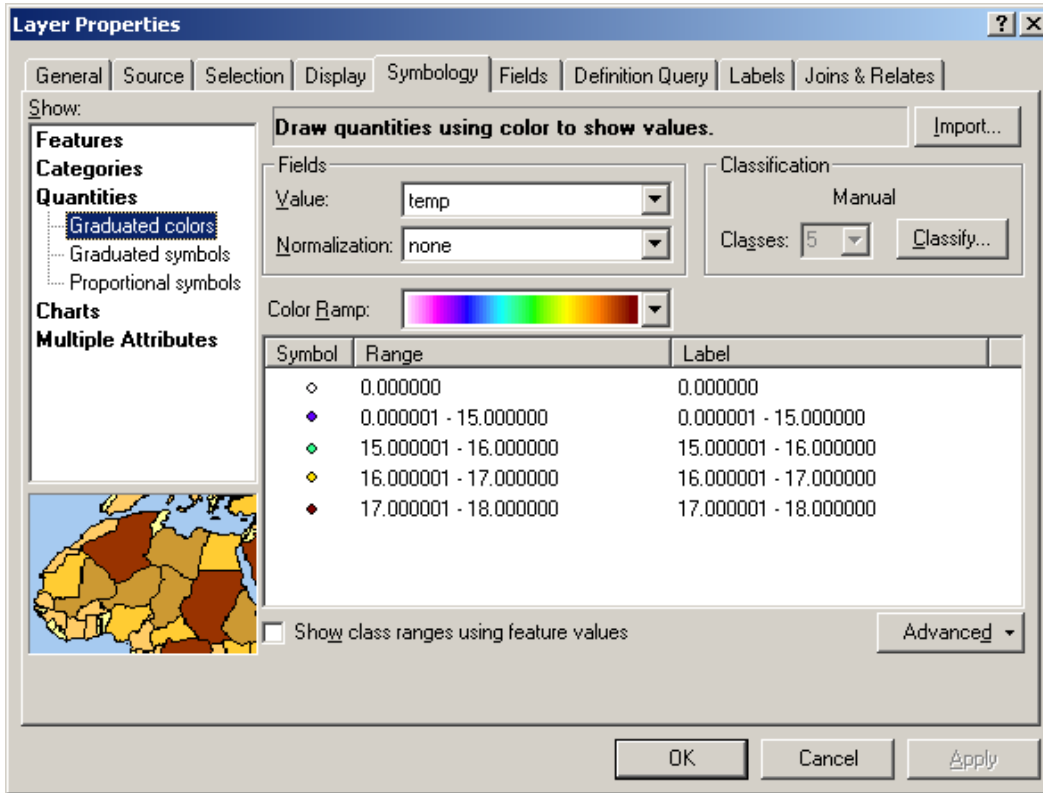


Figure 8: The ESRI® ArcMap™ Layer Properties Symbology tab allows the user to adjust the color scale used for the points.

Finally we add a continental coastline layer from data provided with the ArcMap™ installation. Zooming in to south western Nova Scotia (i.e., the general operating area for Q316), we have the resulting product as shown in Figure 9.

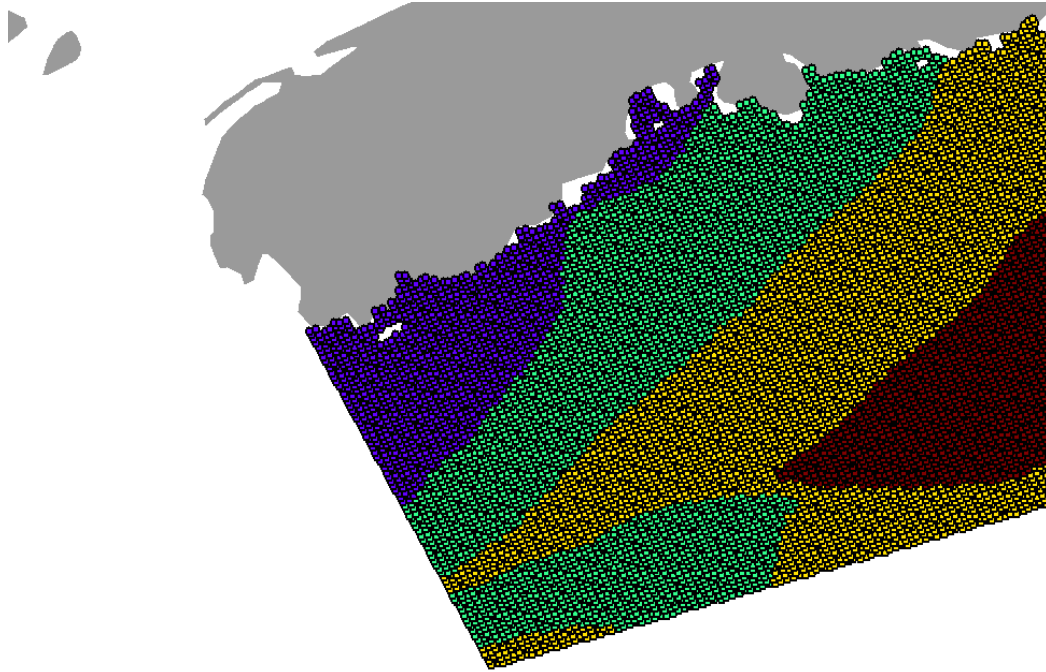


Figure 9: The temperature climatology for 11.5 m depth, in September for the western Scotian Shelf. The color coding is as follows: blue: 12.3-15°C ; green: 15-16°C; yellow: 16-17°C; red: 17-18°C.

Although the image produced in Figure 9 is useful, the image itself is not the most important aspect of the process. Two aspects of the process are more important. First, the DBMS provides us with the ability to access and manipulate the data using a standardized interaction method, that being SQL. The managed data source within the DBMS and the standardized access mean any user or application can access and utilize the data if they understand the database table structure and SQL. Second, the georeferencing provides the ability to include alternate data sources such as the coastline product included with the ArcMap™ GIS application. This means we can locate the data relative to other data products which others have produced, and present the results in familiar georeferenced map-like products.

3.2 Use case 2 - Ocean Feature Analysis reports as part of a GIS product

The second use case illustrates the incorporation of the MetOc OFA product into an external GIS application on the RPC. This capability is important because with the wealth of possible data sources, it's unlikely we will be able to incorporate all data sources into the READB. This example illustrates how an external source in network Common Data Form (netCDF) [17] can be read into the ESRI® ArcMap™ software and used in combination with other data sources from within the READB.

The MetOc office provides the Canadian Forces (CF) with a multitude of environmental data products covering the ocean, land and air. The OFA product is produced twice weekly from the DND TRINITY MetOc office. The product represents a compilation of the in situ and remotely sensed data accessible by the MetOc office. For Q316, MetOc provided the sea surface temperature (SST) and temperature at 150m depth in a latitude-longitude grid in netCDF.

Other products are also available from MetOc. For example, a coastal ocean circulation model developed by Dalhousie University [18, 19] (i.e., the model is commonly referred to as DalCoast) and running at the TRINITY MetOc office produces forecasts of temperature, salinity and current fields. Unfortunately, the model output is currently only available in graphical form or a modified ASCII-binary version of Over-the-horizon GOLD (OTH-GOLD) messages. These messages contain vertical profile data from the model, and can be read by the Allied Environmental Support System (AESS) system at the MetOc. However, the messages are a customized format and cannot be imported into other applications without development of customized readers. Due to these limitations, the Dalcoast output was not used during Q316.

High-end ArcMap™ products have the ability to read netCDF files directly. However, the product being used on Q316³ does not ship with this capability. However, this product does have the ability to import gridded ASCII files in a format specific to ArcMap™. Thus, using the MetOc OFA netCDF file we create an ASCII representation which can be ingested into the ArcMap™ system.

The netCDF file was read into Visual Information Solutions Interactive Data Language (IDL) [20] utilizing freeware code developed by Rupert and Morrison [21]. The code utilizes netCDF procedures within IDL to read the netCDF file and store the results in an IDL structure. Additional IDL code was developed during Q316 to export the structure data to the required ASCII formatted file necessary for ArcMap™ import.

ArcToolbox was then used to read the SST gridded ASCII data. The ArcToolBox Conversion Tools has an ASCII to Raster converter. The ASCII format, as specified in the ArcMap™ help system, has one serious limitation – the cell size must be equal size in latitude-longitude space.

The ASCII to Raster converter is executed from an ArcMap™ graphical user interface as shown in Figure 10. This results in ArcMap™ automatically creating a new layer using the raster image. Accessing the layer properties as described in the previous section, allows one to set the lower

³ ESRI® ArcMap™ 9.2 (build 1324)

temperature limit (i.e., set to 6°C) and change the color scale. The resulting image is shown in Figure 11.

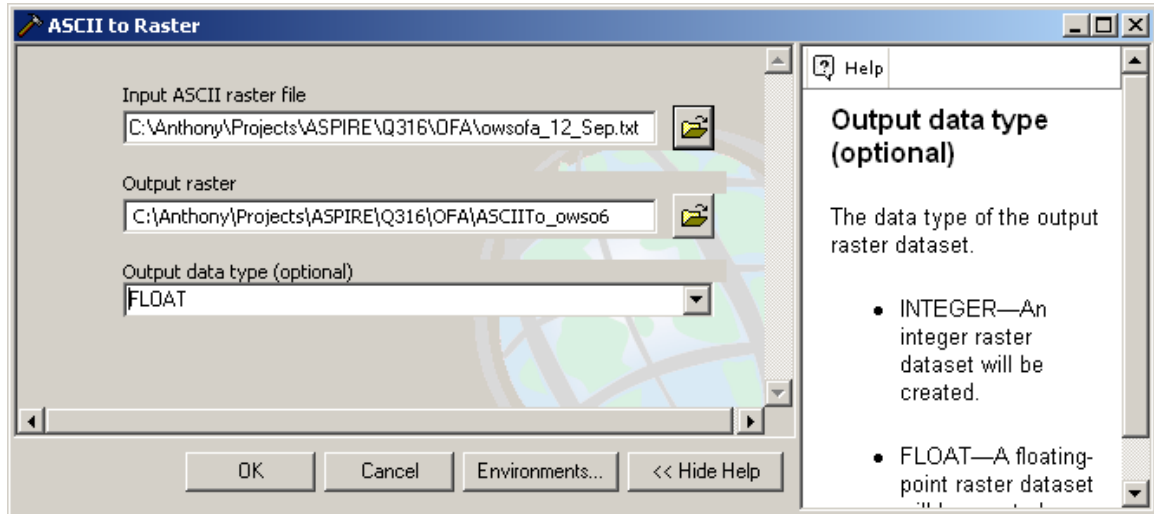


Figure 10: The ESRI® ArcMap™ software allows the user to import ASCII gridded data and create a raster representation. The MetOc OFA analysis was provided in netCDF, converted to ASCII and then imported using this ArcMap™ tool.

The Q316 trial divided its data collection activities between St. Margret’s Bay, Nova Scotia and the Brown’s Bank Northeast Channel area. St. Margret’s Bay is so close to land that the OFA is unlikely to provide much benefit. In fact, while examining the near-land OFA product, issues with the near-land OFA interpolation algorithm were detected. We speculate that the issue deals with the temperature smoothing algorithm that may include land points. We are therefore sceptical of the near-land points.

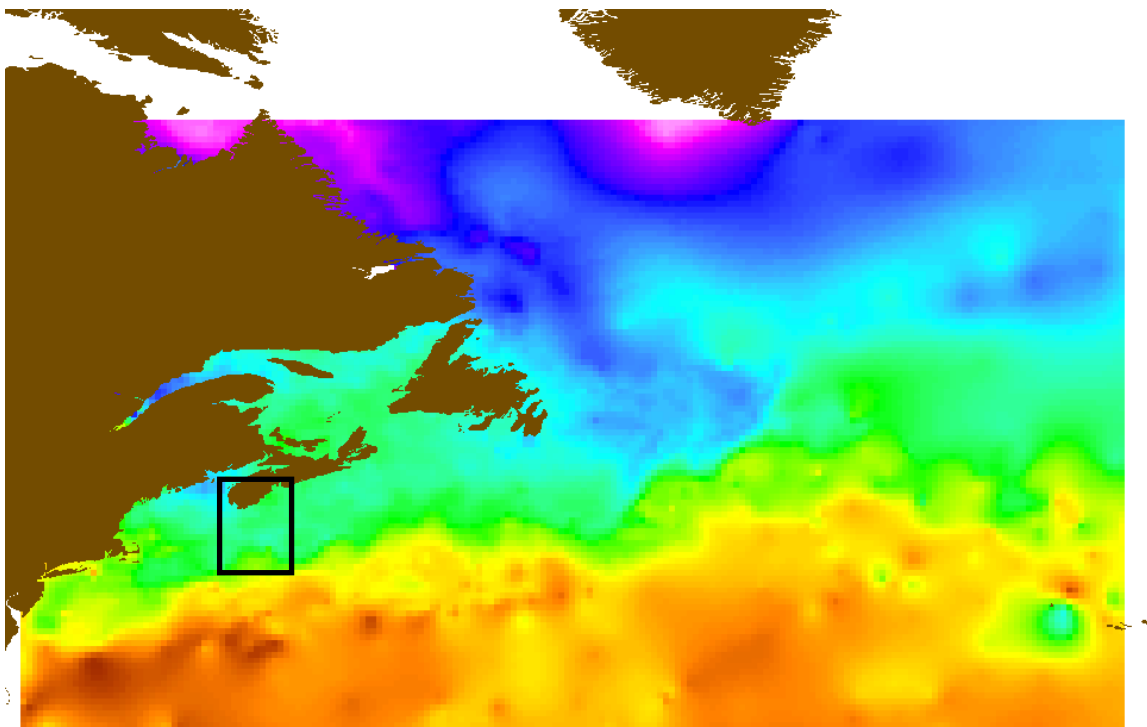


Figure 11: The imported OFA with a continental overlay (shown in brown) displayed in ArcMap™. The temperature range of the figure is 2°C (white region near Greenland) to 30°C in the Gulf Stream. The approximate operation area of Q316 is shown by the black box.

3.3 Use case 3 - Importing in situ and model data into the READB

The third use case describes the incorporation of data into the READB. Incorporating data directly into the DB has the advantage of providing a central location for user access to the data. This type of central access can support applications, models, or tactical decision aids that may be accessing and utilizing the data.

3.3.1 Data sources

The in situ oceanographic data collected during Q316 used the following instrumentation:

- expendable bathythermographs (XBT) collecting temperature profiles. These instruments record temperature as a function of time. Algorithms then compute depth and sound speed.
- expendable sound velocity (XSV) profiler which measures sound speed as a function of time. Algorithms then compute depth.
- conductivity-temperature-depth (CTD) profiles from the moving vessel profiler (MVP) system [22]. The MVP allows the automated deployment of the CTD while the ship is steaming. The CTD deployed from the MVP will be referred to as the MCTD.
- free fall cone penetrometer (FFCPT) [23] is a device which impacts the sea floor, measuring the deceleration upon impact and the pore pressure in the local sediment. Algorithms then compute grain size and sediment type information of the bottom. The FFCPT is also deployed from the MVP system.
- CTD data collected from a SLOCUM [24] glider. Due to personnel issues, the glider was not deployed as planned. As a result, these data are not considered in this report.
- acoustic data collected by hydrophones moored in the water column. These data are not considered in this report.

The physical oceanographic data (i.e., water temperature, conductivity, sound speed and pressure) were collected to support the transmission loss sensitivity study. One goal of this study is the development of a tool to estimate the range sensitivity of transmission loss to variability of the input parameters. This tool is known as the Portable Acoustic Sensitivity Transmission and Estimation Tool (PASTET) [25]. In general terms, the PASTET tool may be considered one example of a broad class of tools known as tactical decision aids (TDA)

Numerical model data provides another possible source of environmental input to the PASTET. Prior to the trial, we had established access to model output from the Canada-Newfoundland Operational Ocean Forecasting System (C-NOOFS) [26] model as supplied by DFO. As well, output from the French global numerical model Mercator [27] was also being supplied by the DFO. Both numerical model outputs provided temperature, salinity and current field forecasts in netCDF format.

3.3.2 Importing in situ data

The general procedures for importing the XBT, XSV and MCTD data were essentially the same regardless of the specific data source. For each source, the processing software supplied with the specific instrument produced output in ASCII form. IDL code developed during Q316 was then used to read the ASCII data. These data were then manipulated to form SQL INSERT statements which were then fed into the READB. Note that the SQL statements had to accommodate various foreign key relationships that exist in the DB.

There were other import limitations related to the tables within the READB. First, we had to accommodate the vertical profile data in the READB table named `xbt_profiles`. This approach expedited access by PASTET, since PASTET had previously been coded to use `xbt_profiles` as input. It was recognized that this approach violated the true intent of this DB table. As the table name implies, the table was designed to store temperature data from an XBT. However, during construction of the DB it was also used to store sound speed data computed from the XBT profile temperature data. As well, during construction it was again improperly used to store sound speed data from XSV instruments.

The import approach followed here continued to abuse the `xbt_profiles` table. Of course the XBT temperature data could appropriately be added to this table. However, we also added the XSV sound speed data and MCTD temperature data.

A second limitation was directly associated with the MCTD data. The CTD collects temperature, conductivity and pressure data. Calculations can provide salinity data. However, the `xbt_profiles` table is not capable of storing either conductivity or salinity. Fields could have been added to accommodate these parameters but in general, this is not considered to be an appropriate approach. Any additional fields would simply continue the abusive trend. A better approach would be to revise the design of the table itself. This issue is being addressed by the READB redesign effort.

A third limitation was the fact the MCTD measures pressure. The pressure data were being forced into the depth field of the `xbt_profiles` table, again misusing the field and table.

A fourth limitation is related to the use of the data by applications such as PASTET. At the moment, PASTET assumes the READB contains correct or valid data. This is clearly not the case for the historic data in the DB, as these data have been loaded in their non quality controlled state. For the in situ data, there are presently no procedures for the quality control (QC) of the data. However, it became clear that PASTET is very sensitive to the quality of the input data.

To avoid invalid data in the READB, a manual QC procedure was implemented on the import of the in situ data. The procedure was rudimentary at best, using visual examination of the numerical data to identify spiking resulting from surface entry and bottom impact. When bottom impact could not be identified, the sounding value at the time of launch was used as the approximate cut off level for the profile.

The FFCPT data was not actually imported into the READB. The import software that was being developed during Q316 was not completed before the end of the trial. As well, preliminary runs of PASTET indicate that the sensitivity due to bottom type was minimal. Thus, lower importance was placed on the bottom type data which resulted in lower priority for the development of FFCPT import software.

A second important factor weighting the decision to lower the priority of the FFCPT data import was the existing READB content related to bottom type. The PASTET currently uses available information in the READB to quantify the bottom type. However, the READB only stores a single text field describing the bottom type. The specific field being used has text content that characterises the bottom as SAND, GRAVEL, MUD, etc. PASTET reads the text description and converts this description to numeric parameters, which it then uses in calculations.

In somewhat the opposite scenario, the FFCPT produces output in numeric form that represents the measured accelerations and pressure field in the bottom sediment. The quickest solution for incorporating these data into the READB would be to convert the FFCPT numeric values into text content, which is consistent with the existing READB text content describing bottom types. The conversion of the numeric FFCPT data to text, and then from text to numeric using PASTET, is obviously a convoluted trail. It would be much more beneficial to examine the PASTET requirement as compared to the FFCPT output and determine the most appropriate data for storage. Such an investigation may benefit PASTET and the long-term FFCPT data archival.

Finally, it is recognized that the FFCPT data is a one-dimensional or vector data source [28]. This means the FFCPT creates a data set similar in form to an XBT, XVS or CTD instruments. It is important to recognize this similarity in data sets for the redesign effort.

3.3.3 Importing numerical model data

Q316 also had access to oceanographic numerical model output. DFO staff were providing the C-NOOFS and Mercator model output once per day, downloadable from a DFO ftp site. Both outputs were in netCDF format. Again using IDL code from Rupert and Morrison [21], the model output was read into IDL structures. Code developed during Q316 was then used to export the structures in ASCII SQL files for direct import into the READB.

NetCDF stores data in a gridded fashion, using named dimensions to describe the grid. NetCDF is self describing, which means the grid in the netCDF file is fully described by a header which precedes the data in the netCDF file. However, the IDL code places all the data into a structure of 1-dimensional arrays. This results in some difficulties when interpreting the data structures. In particular, the recombining of the multiple structures into vertical or horizontal dimensions (i.e., a profile or a horizontal map) is difficult to do with certainty.

Fortunately, DFO staff were responsive to requests for ASCII versions of single vertical profiles from the model. By providing a profile example, we were then able to check our developed IDL code which reconstituted the profiles from the netCDF grid structures. This check of our algorithms ensured we were properly decoding the grid into vertical profiles.

The C-NOOFS model is based on the same code base as Mercator. C-NOOFS also uses the same grid domain as Mercator. C-NOOFS is initialized using the Mercator model parameters of salinity, temperature, horizontal current field, and sea surface height. The open boundary conditions use the same Mercator parameters. However, C-NOOFS adds an hourly wind field at 10 m height [29] which is obtained from the Canadian Meteorological Centre (CMC). The C-NOOFS model at the time of Q316 covered a region of approximately 26°-86°N, 27°-92°W [30].

Unfortunately, the second week of Q316 corresponded to a planned upgrade to the C-NOOFS model. This meant we could not obtain C-NOOFS model output during the second week of the Q316 trial.

Since C-NOOFS was offline, it was decided to trial the PASTET using output from Mercator. Mercator is a global model that uses a $\frac{1}{4}$ longitudinal degree grid (at the equator) with 26 vertical levels in the upper 200m of the water column [31]. Mercator runs once per week, with the output representing a daily average of oceanographic conditions. Unfortunately, Mercator also went through a software release upgrade during the trial, moving from release PSY3V2R1 which had a total of 46 levels in the vertical, to model PSY3V2R2 which has 50 vertical levels. Fortunately this upgrade did not impact our use of the model output during Q316.

In the present form of the READB, to load model output we must first define the model grid within the READB. In the terminology for this report, we consider a *grid domain* to be composed of many *grid cells*, with each cell defined by a specific x-y-z position. A set of grid cells that make up a vertical column (i.e., occur at a single x-y position) are referred to as being at a single *grid location*. Thus, the grid location refers to the horizontal location of a vertical set of grid cells.

The Mercator upgrade necessitated that we load two grid domains in the READB; one domain defining the PSY3V2R1 46 level model and a second grid domain defining the PSY3V2R2 50 level model. Each grid location is given a unique grid location identifier. Note that the x-y positions of the grid locations did not change between releases.

Loading the Mercator ts data into the READB also required some inappropriate table usage. At present, the only gridded data within the DB is in the three tsd tables (see section 3.1). The first table contains a description of the horizontal and vertical grid. The remaining tables contain temperature and salinity data at the various depth levels for the grid. Each record in the temperature or salinity table also contains a field for the month that applies to the data.

These tables were constructed specifically to store the Dalhousie climatology [15]. To store model output from another model (e.g., the Mercator or CNOOF models) we were again faced with the decision to add additional specialized tables or to use existing tables inappropriately. Again, PASTET could already use the existing tsd tables and therefore we decided the most expeditious solution to be the inclusion of the model output into the tsd tables. Again, we recognize and acknowledge this as an abuse of the READB table structure.

The Mercator model domain is global. However, only a small region of this data set was required for the Q316 trial. Thus, a total of 20 grid locations were selected in the region of Brown's Bank. The temperature data from the cells at these locations was then placed in the READB. The grid locations, the Q316 in situ data collection sites and the Gulf of Maine Ocean Observing System (GMOOS) buoy "N" are shown in Figure 12.

The decision to use the tsd tables also meant we were inappropriately using fields within these tables. The tsd tables were structured based on the Dalhousie climatology, which is monthly. The model output from Mercator was daily. The tsd tables have no field for the day value because the Dalhousie Climatology did not require a day field. However, it became evident that storing the day of the month in the month field (i.e., having the month field contain numbers like 15, 16, 17 meaning September 15, etc.) was useful for PASTET. The benefit to this approach was that the existing PASTET code searched the data using a user defined spatial-temporal window. Thus, by selecting our area of operation and the real time period, PASTET could select and use only ts data from the Dalhousie climatology. The input to PASTET could also be quickly

changed to Mercator model data by selecting a time period of month=15, 16, 17, etc. thereby excluding Dalhousie data while including Mercator data. Functionally, this indicates a PASTET requirement to separate sources of modelled data.

Initial results from the PASTET application using input from the in situ and modelled data was promising. A report highlighting the use of PASTET during Q316 is being prepared by the contractor [25].

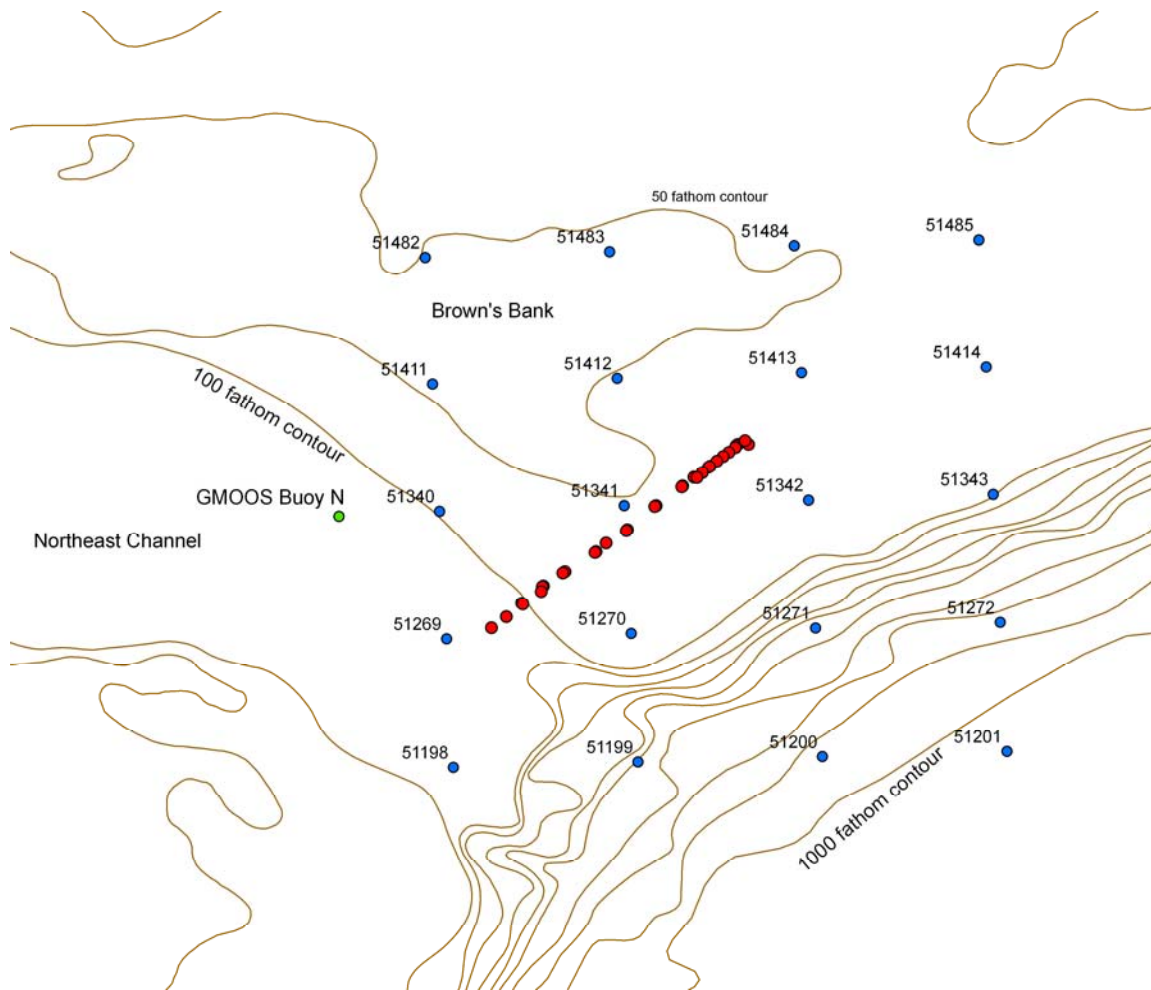


Figure 12: The XBT and XSV drop points (red dots) and the C-NOOFS / Mercator grid locations (blue dots). Grid locations are numbered according to the unique identifier assigned in the READB. The brown lines indicate bathymetry in intervals of 100 fathoms except around Brown's Bank which is 50 fathoms. The GMOOS Buoy "N" is also shown. The entire grid domain was loaded into the READB, but only ts profiles from the cells at the grid locations included in this figure were loaded into the READB.

A superficial comparison of the Q316 in situ data and the Mercator model output was also conducted (Figure 13). The comparison shows the computed sound speed profiles based on the model temperature profiles and an assumed salinity of 30 psu. The model output is from a model computation on September 10th, for September 17th. Note that this represents the 7th forecasted day in the model run. The sound speed profiles from in situ XBT temperature profiles and XSV profiles are also shown.

The first observation is regarding the general shape of the sound speed curves. The model is considered to be doing very well at representing the sound speed channel at about 50m, although overestimating the vertical extent (i.e., width in the vertical) of the sound channel. The bottom level of the model also appears to be shallow as compared to the in situ data. Based on Figure 12, the model location identified by READB ID=51269 is clearly located in > 182 m of water (i.e., > 100 fathoms of water) while the last active vertical level of the model is at 92 m.

An additional model grid location (i.e., READB ID=51198) was introduced to the comparison simply to determine the vertical extent of the sound channel. Sound speed from this grid location (see Figure 13) again indicates that the vertical extent of the sound channel is overestimated by the model. It also indicates that the total model depth is again being underestimated.

To briefly investigate the model depth issue, a comparison of model cells and bathymetry was conducted. The actual bottom depth used in the model is not available in the data provided. However, using the deepest active grid cell we can compare that depth to the bathymetric point value at the grid location, based on Canadian Hydrographic Service (CHS) electronic chart number CA176140. This electronic chart corresponds to CHS paper chart 4003.

Table 2 shows the comparison of bottom active model layer depth and the bathymetry at that point location. As well, the table indicates the depth of the next model layer and identifies if the model layer could exist at that point location. All the grid locations shown in Figure 12 are used in the comparison. The table indicates that half the locations could have an additional model layer. However, it needs to be reiterated that this comparison is using a single point bathymetric value from the CHS chart. No spatial averaging has been done on the bathymetric value over the horizontal extent of the grid location.

Model, XBT, XSV Comparison

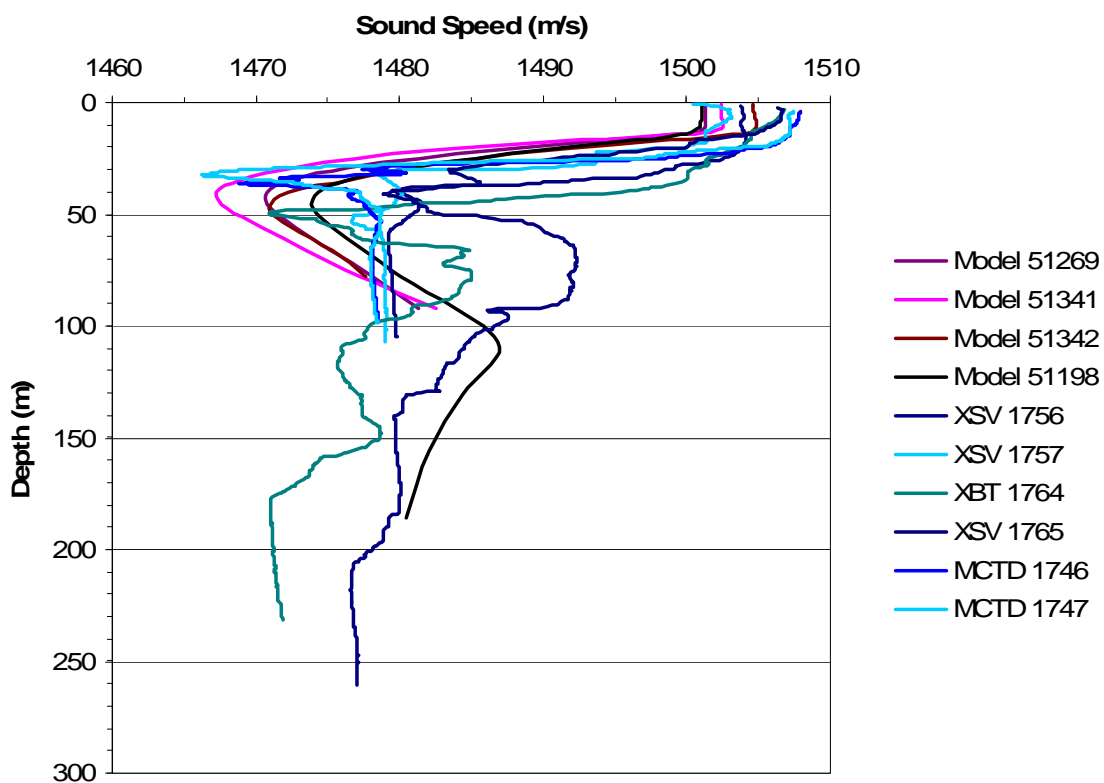


Figure 13: Sound speed comparison from four different sources. Mercator model temperature profiles from a forecast for September 17th were used to compute the model sound speed profiles at four model locations. The in situ temperature data from the XBT and MCTD profiles were also used to compute sound speed. Finally, the XSV profiles of sound speed are included. The legend indicates the model grid location identifiers as assigned by the READB while in situ data is identified by a similar unique READB identifier.

Table 2: A comparison of the deepest model grid location and bathymetry estimated from digital chart CA176140 (corresponds to CHS paper Chart 4003). The next deeper model level is provided and whether or not this level could exist based on the chart bathymetry is also indicated. Ten of the 20 points indicate that the next model level should exist.

READB Grid Location Number	Latitude (N)	Longitude (E)	Deepest Active Model Layer (m)	4003 Bathy- metry (m)	Next Model Level (m)	Model Level Should Exist
51198	41.9877	-65.7571	186	205	222	false
51199	41.9949	-65.5071	902	1098	1062	true
51200	42.0021	-65.2572	1245	1683	1452	true
51201	42.0093	-65.0072	1684	1866	1942	false
51269	42.1616	-65.7663	92	220	110	true
51270	42.169	-65.5163	92	146	110	true
51271	42.1763	-65.2664	131	1006	156	true
51272	42.1837	-65.0164	902	1500	1062	true
51340	42.3349	-65.7757	92	183	110	true
51341	42.3424	-65.5258	92	91	110	false
51342	42.3499	-65.2758	78	120	92	true
51343	42.3575	-65.0258	78	183	92	true
51411	42.5075	-65.7853	92	91	110	false
51412	42.5152	-65.5354	92	91	110	false
51413	42.5229	-65.2855	92	103	110	false
51414	42.5306	-65.0355	78	110	92	true
51482	42.6796	-65.7952	92	91	110	false
51483	42.6875	-65.5453	92	82	110	false
51484	42.6953	-65.2954	92	91	110	false
51485	42.7032	-65.0454	92	100	110	false

4 The READB and large data sets

The work conducted during Q316 also provided an opportunity to investigate options for dealing with external data sets. The issue here is whether or not external data should be incorporated into the READB.

A distinction is drawn between the REA system and the READB. The READB represents one part of a larger system i.e., the REA system. This larger system includes data collection, data processing, data management, and data delivery to clients or client applications. Unfortunately, the larger system does not actually exist as a single system. In other words, the larger system has been constructed from numerous individual pieces. As well, it is not obvious that a design plan exists for the larger system, but rather individual plans exist for the individual pieces. In this area, there is a need for a complete end-to-end plan of how the data should exist within the system, how the data are collected, processed, managed and utilized during REA trials and in general, during all DRDC Atlantic trials.

4.1 Issues on ingesting external data into a database

There are four basic issues related to the storage of external data in a database such as the READB.

The first issue relates to the size of the external data source. External data sources can be quite large (e.g., typically >10Gbytes). In some cases this is due to the vastness of the coverage area while in other cases it may be due to high temporal or spatial sampling of the data (e.g., sidescan sonar is an example of data with detailed spatial sampling). If the external source is large, then loading the source into the READB causes a similar increase in READB size. As the size of the READB increases, the portability of the DB decreases.

Second, storing external data in a database removes the ability to use the external software created for and packaged with the external data set. Often, data sets are delivered with software that provides access and utilization of the data set. Since the software specifically interacts with the data set as produced by the external source, it will not function on a copy of the data as it exists within a database. Thus, the development effort of the external organisation and functionality of the external software is lost.

Third, the updates of the external data source will necessitate the reload of the data into the database. This reload means the database administrator must purge the previous version of the data, and import the updated version.

Fourth, it is simply not a scalable solution to store all external data sets in a single database. Typical datasets that we require can easily exceed 10 Gbytes per data set. Importing these data into the DB is certainly possible, but there are performance and hardware issues that must be considered. As well, portability of the database is impacted by the size of the database.

In general terms, storing all external data sources within the READB may simply be an unrealistic goal. A new plan needs to be formulated for the storage of external data sets that are important to DRDC Atlantic.

4.2 Options for storage of external data

Following current procedures, when a data set is acquired from an external source for incorporation into the REA system the data are loaded directly into the READB. In the case of large volume data sets, this tends to bloat the database. As an example, bathymetry data currently makes up 20Gbytes of the READB (total READB size is 32Gbytes). This type of volume starts to impact the portability of the DB (e.g., TDA development requiring a standard alone READB).

In simplistic terms, the basic issue here is whether or not all data used in the REA system needs to be contained within the READB. In essence there are three available options for the data:

1. store all data in a DBMS; or
2. store a portion of the data from the external data source in the READB; or
3. do not store external data sources in a DBMS.

4.2.1 All data in a DBMS

Although we recognize that storing external sources within the READB introduces several problems, we also recognize that not all the external data need be stored in a single DB. This is because, as with any DBMS, PostgreSQL can have multiple databases defined within its environment. Thus, one option would be to store the data acquired from external sources in a separate DB, and link that DB to the READB.

In this scenario, the link could provide the means to move a sub-selection of data from the DB containing the external data to the READB. Thus, at the time of a trial the actual READB could be loaded with only parts of the external data from the second DB. Such a solution could be constructed using specialized software that forms the link between the databases.

Individual users would also have to utilize these linking functions. This is because PostgreSQL does not allow SQL commands to span multiple databases. In other words, a user cannot access both databases concurrently using SQL commands.

This solution does not address software that may be delivered with the external data or the update problem noted previously.

4.2.2 Portion of the data in the READB

An alternate solution is to place the data from the external sources in the READB, but isolate the tables that contain these data. This isolation would mean that the tables would not be relationally linked to the main tables that support the REA activity.

PostgreSQL does provide a mechanism to create new databases by sub-selecting tables from an existing database. This would be the mechanism to sub-select a portion of the database content for use while at sea.

Again, this solution does not address software that may be delivered with the external data or the update problem noted previously.

4.2.3 Data external to a DBMS

It is also possible to have developed code support functions that connect the DB to external data sources. In PostgreSQL, triggers on individual table operations (e.g., insert) can be constructed using JAVA code [32]. This JAVA code could reach out beyond the database, to the external data resource. This technique would also allow external software to be linked to the DB via the JAVA code thereby utilizing the external software while having the external data source appear as part of the REA system.

With this solution, there may be code maintenance issues if the external source formats were altered such as for an update or new release of the data set. However, if no format changes were made to the external data, incorporating an update would be a trivial replacement of the external source. External software could also be utilized in this approach by linking the external software to the developed JAVA code.

4.3 When data should go into the READB

With the diversity of issues, there remain some truisms from which we can base our solution. First, DRDC Atlantic remains the responsible custodian for the data it collects. Since we are responsible for those data it seems reasonable to expect that these data be managed in a proper system. At the moment, the only managed database system available is the READB. Thus, it appears a reasonable goal to have all environmental data collected by DRDC Atlantic directed towards the READB.

Second, we must recognize the value of the external data sets to DRDC trials. Some portion of these external data sets will be directly useful to trial participants. Thus, it is important for some or part of the external data to be accessible during trials.

Factors which influence when the data should or should not be in the database must also be considered. In some cases there may be temporal factors influencing the usefulness of the data.

In these cases, the data may be useful for a well defined time period, after which the data are not useful. An example would be forecasts from numerical models, where the usefulness of the previous forecast drops when a new forecast is issued. This is similar to a “best before” time which indicates when the data is of value (Figure 14).

Alternately, the data may be of interest to a specific trial but not of general interest to all trials. This may be related to the data directly supporting a TDA or application product that is useful to a specific trial, while not useful to all trials. In these cases, some of the data from the external source should likely be stored within the READB for a given period of time.

There are also cases when storage of the external data within the READB is not beneficial. If updates to the external data are frequent, it will not be practical to ingest and store the data in the READB. This situation applies to onboard processing systems which generate quasi real-time data products. In these cases, the update frequency may be seconds or minutes. If other systems use these products it may be more efficient to link the other system directly to the data product.

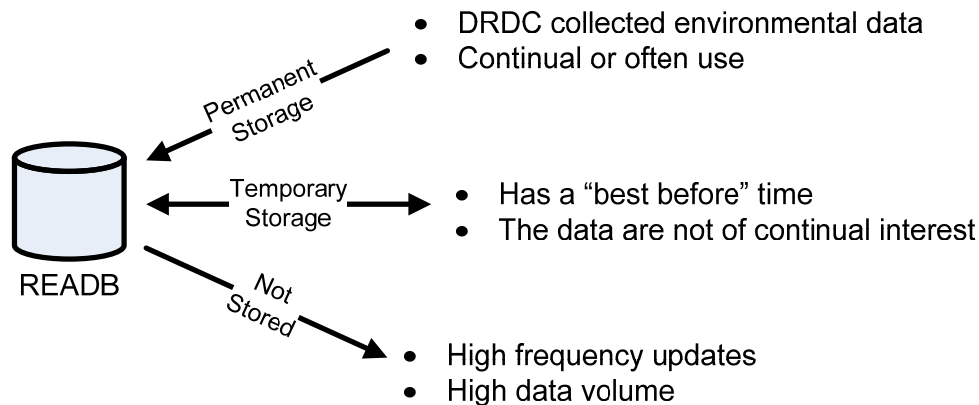


Figure 14: A summary of the factors that influence the decisions of whether or not to store data within the READB or external to the READB.

4.4 Three-way approach

It should be clear that the decision to store external data sets within or external to the READB is complicated. Numerous competing issues influence the situation and thus the decision. Even developing a single procedure for identifying which of the three options outlined in Figure 14 is best for a particular situation would be difficult. This is because individual users and individual trials will have different perspectives on what constitutes high frequency updates, high data volumes or the specific best before time.

The diversity in user requirements, trial requirements and external data sets means that any one of the three options posed in Figure 14 is a viable solution to the specific situation. Thus, it is likely that the best we can hope for is a case-by-case examination of which option provides the functionality we seek, without impacting maintenance or development issues of the database. Thus, the READB redesign effort must accommodate all three options outlined in Figure 14.

5 Concluding remarks and recommendations

This report was intended to provide information on Q316 activities that directly involved the READB. The report serves as general documentation for how to establish a connection from a remote PC to the DB and also what changes are required at the server PC to allow such a connection.

The report also outlined three use cases that involve specific interactions with the READB. The first use case dealt with sub-selecting data that already exists in the READB. This use case involved manipulations to the data within the PostgreSQL environment. The second use case involved the importing of the MetOc OFA product into a GIS environment. This environment also allowed access to the READB and thus the combining of the OFA product with data from the READB. Finally, the third use case dealt with the importing of in situ data and numerical model data into tables within the present READB structure. This use case provided an example of incorporating data for use by TDAs such as PASTET. However, the use case also pointed out that in order to meet our goals, we needed to inappropriately use several existing tables within the READB structure.

The use cases and Q316 activities involving the READB have provided valuable insights in three broad categories: the DB redesign activity; the use of READB data by external TDAs; and the importing of external numerical products into the READB. In particular, the following recommendations are made:

1. Numerical models supporting TDAs: The Mercator model output was found to be a very useful data source for input into PASTET. The agreement between the Mercator model output and in situ data was also encouraging. However, the modelled sound channel was noted to have a larger vertical extent as compared to the in situ data. Since the vertical resolution of the model at this level is about 10 m, the model is certainly capable of resolving such a feature. Since the model did not accurately represent the vertical extent of the sound channel, we are suspicious the temporal or horizontal resolution of the model output is influencing the vertical extent of the sound channel in the model results.

Finer horizontal grid scale or finer temporal model output may improve these results. The temporal resolution could be addressed by using the C-NOOFS model. C-NOOFS output can be generated every hour and this may show a sound channel that is not spread in the vertical. Both models use the same horizontal grid, so C-NOOFS cannot be used to investigate the horizontal resolution question. It is obvious that the present grid does not properly resolve the Northeast Channel, possibly because of bathymetry smoothing over the horizontal extent of a grid location. This problem could be reduced if a finer scale model were available. However, a static model domain will not meet the requirements of DRDC Atlantic or the CF in general. Any ocean model constructed to support CF operations needs to be highly portable. The concept of the “10-day model” has been proposed previously by the author and continues to be a suggestion directed to DND and as well to those involved in the Centre for Ocean Model Development and Application (COMDA) project. The 10-day model refers to the concept of having a fine scale ocean model operating within 10 days of defining any operational area anywhere in the world.

Recommendation 1: DRDC Atlantic should investigate the possibility of acquiring from DFO, output from finer horizontal and temporal scale models for future trials.

Recommendation 2: DRDC Atlantic should investigate the viability and usefulness of the 10-day model concept with DND personnel and COMDA representatives.

2. Automate the READB import process: PASTET is very sensitive to the quality of the input data. During Q316, the QC procedure for XBT, XSV and CTD data was based on visual inspection of the numeric profile data. This is both time consuming and highly dependent on the skill set of the person performing the QC.
 - a. The QC procedure needs to be automated from the collection point to the final import into the DB. The automation of the procedure should not be a major task. The XBT and XSV data acquisition is a routine operation on board Quest, and the collection procedures are at present highly automated. Extending the automation to include minor procedural changes (e.g., have the collection clock on GMT) and the automated trimming of the upper and lower portion of the profile would not be a difficult task.

Recommendation 3: DRDC Atlantic should develop and implement automated QC procedures for XBT and XSV data collected on board CFAV Quest.

- b. The above recommendation points to a more general requirement for environmental data collection on Quest sea trials. The present environmental data management procedures on Quest should be examined from an end-to-end perspective. This end-to-end process must consider all aspects of data collection, data management and the delivery of the data to clients.

Recommendation 4: DRDC Atlantic should develop an end-to-end management plan for data collected on CFAV Quest trials run by DRDC Atlantic.

3. MetOc Dalcoast output in netCDF: Dalcoast could provide valuable forecast information to a tool such as PASTET. However, MetOc must work toward providing the DalCoast output in netCDF format. MetOc needs the capability to sub-select a region of the model, and provide export of the model variables over user-defined vertical levels to netCDF. This would be an important step toward providing actual data products as opposed to visual products (i.e., an image of the data).

Recommendation 5: TRINITY MetOc should incorporate netCDF export capabilities into the DalCoast numerical model. The netCDF format should be compatible with other model netCDF formats such as Mercator and C-NOOFS output.

4. MetOc OFA:
 - a. Regarding interpolation, there appears to be a smoothing problem with the present OFA product when in the vicinity of land. This issue needs to be resolved in the present product and MetOc should ensure the same problem does not exist in the new OFA

product being developed under the Spaceborne Ocean Intelligence Network (SOIN) project.

Recommendation 6: TRINITY MetOc should investigate and correct the OFA smoothing problem which exists near land and ensure this problem does not exist in the SOIN version of the OFA product.

5. REA database redesign: There is currently a READB redesign effort underway at DRDC Atlantic. The activities of Q316 further highlighted the need for this redesign. Specifically, the following issues need to be addressed in the redesign.
 - a. Quality control procedures (such as described above) need to set quality flags on the individual data values. As well, TDA operators need the ability to set the flags on the data values, thus encouraging the TDA operator to be an active contributor to the quality control of the data set. This would also avoid the requirement for automated QC algorithms to be responsible for all QC issues.

Recommendation 7: DRDC Atlantic should ensure the READB redesign includes the ability to assign quality flags to all data collected and included in the DB.

- b. The redesign also needs to address very different data collection activities that collect very similar data. For example, the present DB has no means to include the FFCPT data in its initial form. The redesigned DB needs to consider FFCPT data but also needs to recognize the similarity between FFCPT data and more common data, such as XBT profiles. Such commonalities need to be recognized in the redesign effort.

Recommendation 8: DRDC Atlantic should ensure the DB redesign treats data with the same dimension in a similar way.

- c. The use of FFCPT data within the PASTET tool needs to be investigated. Presently, PASTET utilizes bottom types stored in the READB. However, PASTET converts these bottom types to numerical data values. The FFCPT may be capable of providing these data values directly, thus negating the need for the PASTET conversion.

Recommendation 9: DRDC Atlantic in collaboration with PASTET designers, should determine the form of the bottom type data used by PASTET and whether or not this form is directly available from the FFCPT. If the data are available in the appropriate data types, PASTET should be modified to use these data directly.

- d. The storage of the FFCPT data in the READB needs to be examined. Currently, FFCPT data are not in the READB. The redesign effort will incorporate these data into the database. This redesign must account for the characteristics of the data set, the general needs of TDAs, and also the specific and possibly modified (see Recommendation 9) needs of PASTET.

Recommendation 10: DRDC Atlantic should examine the FFCPT data in detail and incorporate these data into the redesigned database. The incorporation must take into

account the characteristics of the data set as well as data usage by TDAs such as PASTET.

- e. The redesign must also address the generalized model output issue. The MetOc OFA product required the grid data to be placed within an ASCII file for import to ArcMap™. The current redesign effort should allow gridded data to be imported directly into the DB without improperly filling fields. Again, this would allow TDA access to various model outputs via the shared database.

Recommendation 11: DRDC Atlantic should ensure the DB redesign treats gridded model output data in a similar way. In concept, this is the same recommendation as 8. Here, emphasis is placed on model data.

- f. The storage of external data sets in the REA system will require a three-avenue approach.

Recommendation 12: DRDC Atlantic should ensure that collected data are stored within the existing READB. Those data that are important for short term activities such (e.g., while at sea) should be stored on a temporary basis within the existing READB. Those data sets that are high volume or only occasionally useful, should be made accessible via links between the READB and the external data resource. The redesign effort should account for all three avenues.

6. Database support at sea: Going to sea with a READB or some equivalent will require trained personnel, especially in the early stages of development and use. This should not be a surprise, as any piece of specialized kit or gear that is used at sea typically is accompanied by trained personnel who are capable of maintaining and repairing the gear if it is damaged. The DB and associated import software will require a similar person.

Recommendation 13: DRDC Atlantic should encourage chief scientists who wish to take the READB to sea, to treat the DB as a specialised piece of kit, and include in their plans the need for trained personnel to accompany the READB.

7. End-to-end data plan: There needs to be an end-to-end plan developed for the management of at-sea environmental data. The plan needs to build on the existing infrastructure available both at-sea and ashore, effectively linking the processing streams into a single managed system. The plan needs to account for all aspects of data collection, quality control, visualization and utilization while at sea, and data delivery to ashore databases archives such as those maintained DND MetOc and national data centres.

Recommendation 14: DRDC Atlantic should develop an end-to-end data management plan for data collection activities on CFAV Quest and follow-on data management activities associated with these data.

References

- [1] Whitehouse, B.G., Vachon, P.W., Thomas, A.C., Quinn, R.J. and Renaud, W.M. (2006), Rapid Environmental Assessment (REA) of the Maritime Battlespace, *Canadian Military Journal*, Spring 2006, 66-68.
- [2] Young, C., Greene, C., Isenor, A. and Cole, A. (2006), Rapid Environmental Assessment Database Project, From the Sonobuoy Interoperability Conference, Berlin, Germany.
- [3] Isenor, A.W., Young, C., Osler, J., Renaud, W. and Reed, G. (2007), Rescuing data assets and allowing discovery using metadata, From the Ocean Biodiversity Informatics, Dartmouth, Nova Scotia.
- [4] Deveau, T.J. (2006), Rapid Environmental Assessment Database Project - Phase I, (DRDC Atlantic CR 2006-213) Defence R&D Canada – Atlantic.
- [5] PostgreSQL PostgreSQL: The worlds most advanced open source database (online), <http://www.postgresql.org/> (Access date: Oct. 3, 2008).
- [6] Giannecchini, S., Fabiani, A., Spina, F. and Grasso, R. (2008), Geospatial data fusion server architecture at NURC Fusion Centre, (NURC-FR-2008-002) NATO Undersea Research Centre, La Spezia, Italy.
- [7] Hodsdon, J., MacLeod, B., Briggs, D., Wolff, N. and Incze, L. (2007), The dynamic atlas of the gulf of maine: two approaches to web-based mapping and data distribution, From the Ocean Biodiversity Informatics Conference, Dartmouth, Nova Scotia.
- [8] Chic, O., Río, E.d. and García-Ladona, E. IVADO: an Oceanographic Data Analysis and Visualization Tool based on ION/IDL and Java using netCDF Interface, From the Oceans 2001.
- [9] Read, A.J., Halpin, P.N. and Crowder, L.B. (2003), A Digital Archive of Marine Mammal / Bird / Turtle Data for OBIS, (unpublished).
- [10] Zhu, W.H., Wetzler, M.A., Caress, D. and Denbo, D.W. The High-Resolution Multibeam Bathymetry Data Access System.
- [11] SEACOOS: SouthEast U.S. Atlantic Coastal Ocean Observing System SEACOOS Data Management and Visualization Cookbook (online), http://seacoos.org/documents/cookbook#PostgreSQL_noautolink_database_w (Access date: Oct. 3, 2008).
- [12] Osler, J.C. (2008), CFAV QUEST Cruise Plan – Q316.
- [13] Roger, W.A. and Webb, S. (2006), An Environment Modeling Manager for Tactical Decision Aid Support, From the Undersea Defence Technology, Hamburg, Germany.

- [14] PGFoundry Official PostgreSQL ODBC Driver (online), <http://gborg.postgresql.org/project/psqlodbc/genpage.php?downloads> (Access date: Oct. 3, 2008).
- [15] Bobanovic, J. and Thompson, K. (2001), Estimating three dimensional properties of the coastal ocean from remotely sensed data, Dalhousie University.
- [16] Digital Equipment Corporation (1992), Information Technology - Database Language SQL, ISO 9075:1992.
- [17] Unidata NetCDF (online), <http://www.unidata.ucar.edu/software/netcdf/> (Access date: Oct. 3, 2008).
- [18] Ohashi, K., Sheng, J., Thompson, K.R. and Ritchie, H.C. (2005), An Introduction to DalCoast3, a Shelf Circulation Forecast System (Version 1), Dalhousie University.
- [19] Thompson, K., Liu, Y. and Jones, C. (2007), Update on Dalcoast3, Dalhousie University.
- [20] ITT Visual Information Solutions, (online), <http://www.ittvis.com/Company.aspx> (Access date: October 2008).
- [21] Rupert, S. and Morrison, D. read_ncdf.pro (online), http://guvi.jhuapl.edu/data/tools/readers/read_ncdf.pro.txt (Access date: Oct. 3, 2008).
- [22] ODIM Moving Vessel Profiler (online), http://www.brooke-ocean.com/mvp_main.html (Access date: Oct. 3, 2008).
- [23] Furlong, A., Osler, J., Christian, H., Cunningham, D. and Pecknold, S. (2006), The Moving Vessel Profiler (MVP) - a Rapid Environmental Assessment Tool for the collection of water column profiles and sediment classification, (DRDC Atlantic SL 2006-273) Defence R&D Canada.
- [24] Webb Research Corporation SLOCUM Glider (online), <http://www.webbresearch.com/slocum.htm> (Access date: October 2008).
- [25] Giles, P., Kilistoff, S. and Brooke, G. (2008), Portable Acoustic Sensitivity Transmission and Estimation Tool, (in press) Defence R&D Canada.
- [26] Fisheries and Oceans Canada Canada-Newfoundland Operational Ocean Forecasting System (online), http://www.c-noofs.gc.ca/home_e.php (Access date: Oct. 3, 2008).
- [27] Mercator Ocean (2007), Mercator analyses and forecasts the ocean in real-time from surface to bottom (online), http://www.mercator-ocean.fr/html/mod_actu/public/welcome_en.php3 (Access date: October 2008).
- [28] Isenor, A.W., Keeley, J.R. and Linguanti, J. (2003), Developing an eXtensible Markup Language (XML) Application for DFO Marine Data Exchange via the Web, (DRDC Atlantic ECR 2003-025) Defence R&D Canada.

- [29] Davidson, J.M.F. and Ratsimandresy, A.W. (2007), Presentation of the MERSEA Canadian Operational Oceanography System C-NOOFS, *Mercator Quarterly Newsletter*, 24, 40-48.
- [30] DFO (2007), Canada-Newfoundland Operational Ocean Forecasting System, 1, 11.
- [31] Mercator Ocean 1/4 degree global configuration (ORCA025) (online), http://www.mercator-ocean.fr/html/science/piste_rouge/model/model_orca025_en.html (Access date: October 2008).
- [32] Douglas, K. and Douglas, S. (2006), PostgreSQL The Comprehensive guide to building, programming, and administering PostgreSQL databases, Second ed. Sams Publishing.

List of symbols/abbreviations/acronyms/initialisms

AESS	Allied Environmental Support System
BI	browser interface (refers to the REA BI)
CHS	Canadian Hydrographic Service
CNOOFS	Canada-Newfoundland Operational Ocean Forecasting System
CF	Canadian Forces
CMC	Canadian Meteorological Centre
COMDA	Centre for Ocean Model Development and Application
CTD	conductivity-temperature-depth
DB	database
DBMS	database management system
DFO	Fisheries and Oceans Canada
DRDC	Defence Research & Development Canada
DRDKIM	Director Research and Development Knowledge and Information Management
DRP	Document Review Panel
DND	Department of National Defence
EMM	Environmental Modelling Manager
FFCPT	Free Fall Cone Penetrometer
GIS	Geographic Information System
GMOOS	Gulf of Maine Ocean Observing System
IDL	Interactive Data Language
MetOc	Meteorological and Oceanographic Office
MVP	Moving Vessel Profiler
NATO	North Atlantic Treaty Organization
NURC	NATO Undersea Research Centre
ODBC	open database connectivity
OFA	Ocean Feature Analysis
OTH	Over-the-horizon
PASTET	Portable Acoustic Sensitivity Transmission and Estimation Tool
PC	Personal Computer

QC	quality control
R&D	Research & Development
REA	Rapid Environmental Assessment
READB	Rapid Environmental Assessment database
REA BI	REA browser interface
RPC	remote personal computer
S/H	Section Head
SOIN	Spaceborne Ocean Intelligence Network
SQL	structured query language
SST	sea surface temperature
TDA	tactical decision aid
TIF	Technology Investment Fund
ts	temperature and salinity
XBT	expendable bathythermographs
XSV	expendable sound velocity

Distribution list

Document No.: DRDC Atlantic TM 2008-212

LIST PART 1: Internal Distribution by Centre

- 2 DRDC ATLANTIC LIBRARY FILE COPIES
 - 3 DRDC ATLANTIC LIBRARY (SPARES)
 - 1 SECTION HEAD MICS
 - 1 TIM HAMMOND
 - 1 LIESA LAPINSKI
 - 1 ANDREW MACINNIS
 - 1 MARK MCINTYRE
 - 1 JOHN OSLER
 - 1 SEAN WEBB
 - 2 A. W. ISENER (1 CD COPY, 1 HARD COPY)
-
- 14 TOTAL LIST PART 1

LIST PART 2: External Distribution by DRDKIM

- 1 Library and Archives Canada
- 1 NDHQ/DRDKIM 2-2-5
- 1 Fraser Davidson
G10-117J, NAFC
80 East White Hills Rd.
St. John's, NL
Canada A1C 5X1
- 1 Peter Giles
GD Canada
31 Millbrook Avenue,
Dartmouth, NS
B2V 0A2
- 1 Adam Lundrigan
G10-117J, NAFC
80 East White Hills Rd.
St. John's, NL
Canada A1C 5X1

- 1 Debbie Powers
G10-117J, NAFC
80 East White Hills Rd.
St. John's, NL
Canada A1C 5X1
- 1 Senior Staff Officer MetOc
MetOc, Bldg D-201
CFB Halifax
Halifax, NS
B3K 5X5
- 1 Staff Officer Military Oceanography
MetOc, Bldg D-201
CFB Halifax
Halifax, NS
B3K 5X5
- 1 Dr. Brenda J. Topliss, CPhys, MInstP
DIRECTOR
Centre for Ocean Model Development for Applications (COMDA)
Fisheries and Oceans Canada
Bedford Institute of Oceanography, P.O. Box 1006
1 Challenger Drive, Dartmouth, N.S., CANADA, B2Y 4A2
- 1 Ulrich Suesser
ulrich.suesser@ec.gc.ca

10 TOTAL LIST PART 2

24 TOTAL COPIES REQUIRED

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence R&D Canada – Atlantic 9 Grove Street P.O. Box 1012 Dartmouth, Nova Scotia B2Y 3Z7		2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED	
3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) Enhancing the utility of the Rapid Environmental Assessment database through the use of in situ and modelled data sets during Q316:			
4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used) Isenor, A.W.			
5. DATE OF PUBLICATION (Month and year of publication of document.) November 2008	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 60	6b. NO. OF REFS (Total cited in document.) 32	
7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Technical Memorandum			
8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) Defence R&D Canada – Atlantic 9 Grove Street P.O. Box 1012 Dartmouth, Nova Scotia B2Y 3Z7			
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) Project 11ch, 11cn		9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)	
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC Atlantic TM 2008-212		10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) Unlimited			
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.) Unlimited			

13. **ABSTRACT** (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

The Rapid Environmental Assessment database (READB) is an assembly of ocean environmental data collected by DRDC Atlantic and its predecessors. The database was a result of a contract effort that combined temperature, sound speed, transmission loss, and ambient noise data into a single database management system. As well, the database contains bathymetry data and temperature-salinity climatology for the Scotian Shelf. The database is intended for use while ashore and also while at-sea. Trial Q316 was the maiden voyage of the database. During the trial, the database was used extensively by utilizing previously existing data within the database and by incorporating data collected during the trial. The database was used as a data source for other applications being used to assess the surrounding oceanographic environment. The trial also highlighted potential improvements in the coupling between the external data sources and the database. Based on the Q316 database investigation, numerous recommendations are made for the database, for the redesign of the database, and how external data sources can more easily be incorporated into the database.

La base de données d'évaluation environnementale rapide (BDEER) est une série de données sur le milieu océanique recueillies par RDDC Atlantique et ses prédécesseurs. Elle a été créée grâce à des efforts déployés en vertu d'un marché pour combiner des données sur la température, la vitesse du son, l'affaiblissement de transmission et le bruit ambiant en un seul système de gestion de base de données. En outre, elle contient des données bathymétriques et des données climatiques température-salinité pour le plateau néo-écossais. Elle a été conçue pour servir au large des côtes et en mer. L'essai Q316 a été marqué par sa première utilisation. Durant l'essai, on en a fait un usage exhaustif en utilisant les données qui s'y trouvaient déjà et en y intégrant les données recueillies au cours de l'essai. Elle a servi de source de données pour d'autres applications qui ont permis d'évaluer le milieu océanographique ambiant. L'essai a permis de faire ressortir des améliorations possibles au couplage entre les sources externes de données et la base. D'après l'examen qui en a été fait durant l'essai Q316, de nombreuses recommandations sont formulées au sujet de la base de données et de son réaménagement et de la façon d'y intégrer plus facilement les sources externes de données.

14. **KEYWORDS, DESCRIPTORS or IDENTIFIERS** (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

rapid environmental assessment; REA; REA database; REA DB; Postgre; PostgreSQL; PostGIS; tactical decision aid; TDA; netCDF; numerical model; oceanography;

This page intentionally left blank.

Defence R&D Canada

Canada's leader in defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale



www.drdc-rddc.gc.ca