Using Microsoft *Access* to Explore the OMEX II Database

Introduction

The OMEX II database contains a significant proportion of the project data set. The CD-ROM includes three copies of the database in different variants of the Microsoft JET format. Also included is a copy of the Microsoft *Access* run-time edition software, which may be freely distributed providing it is accompanied by an *Access* database application.

The database may be found in the DBJET directory on Disk 1 of the OMEX II CD-ROM package. There are three copies in different JET format variants, named OMEX2V7.MDB, OMEX297.MDB and OMEX200.MDB. The *Access* and JET variants are discussed more fully in the next section, but the following summary table of which file to use with which software may prove helpful.

Access 7.0	OMEX2V7.MDB
BODC Database Explorer	OMEX2V7.MDB
Run-time *Access* supplied	OMEX2V7.MDB
Access97	OMEX297.MDB
Access2000	OMEX200.MDB

Access Versions

At the time of writing, five versions of Microsoft *Access* have been released. The first of these, version 1.0, was current for a very short period of time and was not widely adopted. *Access* 2.0 was the main version used under *Windows* 3.n. Three versions, 7.0, *Access97* (sometimes called *Access* 8.0) and *Access2000*, have been available for *Windows95, Windows98, Windows2000 and NT* 4.n.

The database variants supplied on the OMEX II CD-ROM cover *Access* 7.0, *Access97* and *Access2000*. Previous BODC *Access* database releases have also covered *Access* 2.0 but this is now considered obsolete.

Each version of *Access* has its own version of the JET database format associated with it and each of these formats is different with only limited compatibility between versions. One would expect that earlier versions of the software would not be able to read later versions of the format. However, whilst the later versions of the software are able to read earlier versions of the format, their functionality is limited. It is possible to use objects, such as

forms, to retrieve data and to modify the data held in tables, but it is not possible to modify the design of objects or to create additional objects.

There is an additional compatibility problem for databases on CD-ROM. The later versions of *Access* insist on placing some information into earlier version database files the first time they are opened by the later version software. CD-ROM is a read-only medium. Consequently, when one tries to open a database on CD-ROM with a later version of *Access*, the result is a fatal error. It is therefore important to use the correct copy of the database to match the version of *Access* being used.

Some users may wish to copy the database from CD-ROM onto their hard disk. Modern hard disks have much faster seek times and higher data transfer rates than the fastest CD-ROM drive. This results in significantly improved performance when using the database.

Copies of the database on hard disk may be modified by the user. This may be desirable. For example, users might wish to add their own data into the database. However, a word of caution is required. The forms provided in the database have been built on the assumption that the database contents will be protected by the fact that CD-ROM is a read-only medium. Without this protection, it is possible to modify the data held in the database by simply typing something into a form field. This is surprisingly easy to do in *Access*. If you wish to guard against this possibility leave the 'read-only' attribute set for the database file (the default for files copied from CD-ROM).

Users with some experience of *Access* may find it helpful to set up additional queries, forms, reports or even *Visual Basic* applications (modules). These will allow output from the database to be tailored precisely to the user's requirements.

Access Objects

When a database is opened using *Access*, a dialog box is usually presented offering a choice of the objects available. However, the OMEX II database has been set up so that the only object that is visible is the switchboard form. To see the database in all its glory simply press the *F11* key.

There are six types of object: tables, queries, forms, reports, macros and modules. Users of the OMEX II database as supplied need only be concerned about the first three of these. The type of object required is selected by clicking one of the tabs on the top of the dialog. This presents a list of the available objects of that type. The desired object is simply highlighted by clicking on it and then opened by clicking on the 'Open' button. Alternatively, an object may be opened by double clicking on it.

Table objects, as their name suggests, are the actual tables of data that make up the relational database. Opening a table object provides a listing of the

table contents as a data grid. The tables may also be opened in design mode, which allows their detailed structure to be seen or, providing a full version of *Access* is installed, modified.

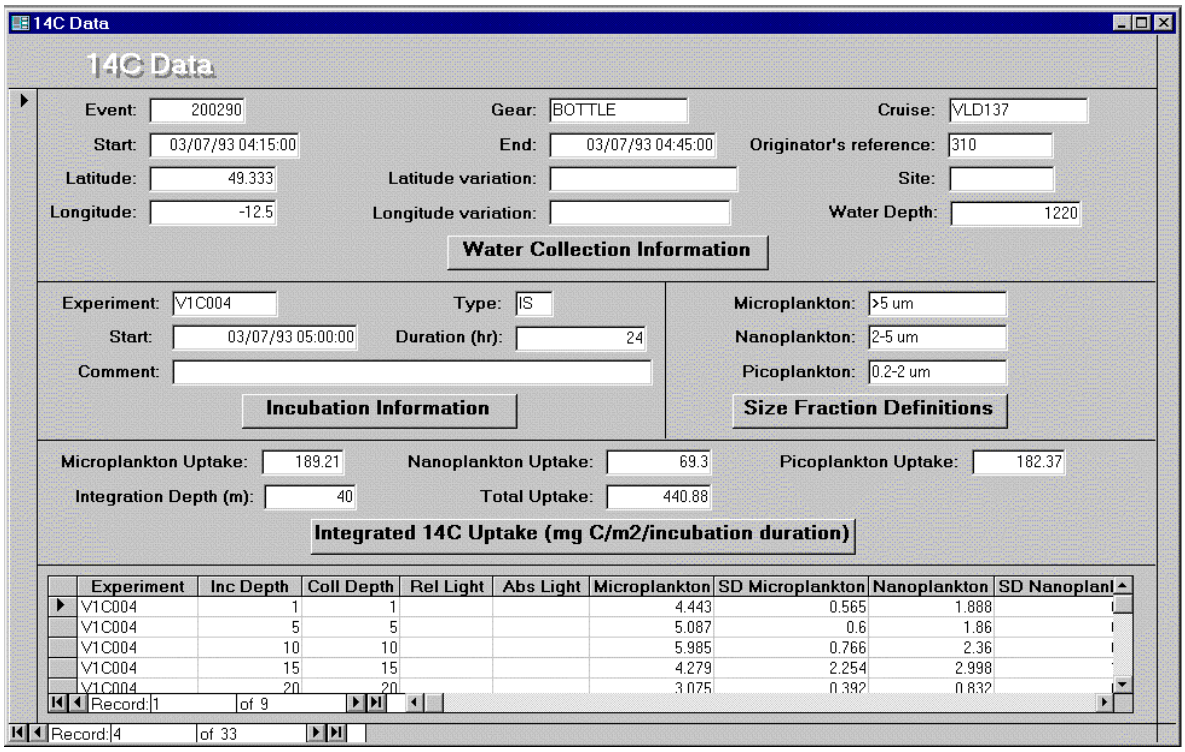
Within the data grid, the fields and records visible may be adjusted by using the horizontal and vertical scroll-bar controls. Any subset of rows and columns may be highlighted by dragging the mouse cursor over them and then transferred to other applications via the clipboard. However, as individual tables from a relational database are rarely self-contained, this may be of limited value.

Query objects, or 'views' as they are known in other relational database systems, may be thought of as additional table objects, which combine several base tables together, filter their contents or sort the data into a specified order. Their main reason for being is to underpin form objects. However, users may find them convenient for exporting some types of data into other applications. Like table objects, opening a query object causes its contents to be listed in the form of a data grid. Alternatively, the query may be opened in design mode, which provides a clear indication of the fields, source tables, sort keys and filter criteria that comprise the query. For those averse to GUI displays, the query may also be viewed as a native SQL statement.

The most important object type for those wishing to interrogate the database is the form object. These provide an interface to data from several tables combined in such a way that the user is provided with all the information necessary to make use of the data. Forms may be used to simply step through data records or display data as a data grid. However, when used in conjunction with *Access* ‘find’ and ‘filter’ tools, the forms provide a powerful and flexible mechanism for searching the database.

Using *Access* Forms

When an *Access* form is opened, the information is presented in what is termed 'form view' that looks like this:



The information presented is from a single database record, in this case a ^{14}C incubation experiment. The header information is displayed in discrete labelled boxes. In addition, there is a second embedded form that displays the data from child records (the individual sample bottles involved in the experiment) owned by the experiment record.

The record displayed may be changed using the 'video' control in the bottom left-hand corner of the window to step one record forwards, one record backwards or to jump directly to the first or last record.

For a form such as this to be of practical use there are two things that the user needs to be able to do. First, the user must be able to find the record or records of interest. Secondly, the user must be able to get the information from those records into another application such as a spreadsheet.

One method for locating the record or records of interest is to use the video controls to single step through the records until a record of interest is encountered. This is a perfectly satisfactory method for forms containing up to 100 records. However, many of the forms contain a significantly larger number of records and so more powerful tools are required.

The first tool provided by *Access* is the 'find' tool that may be found in the 'Edit' pull-down menu or invoked using the toolbar button with the binocular

icon. This tool can be used to search for a given string in either a single field (selected by clicking in the appropriate form box) or all fields. It is particularly useful for finding related records in other forms by searching for the BODC Event Number.

The most powerful tool provided by *Access* is the form filter. This allows the creation of a copy of the current form that contains a user-specified subset of records. The filter criteria are set up using the 'Edit filter/sort' option in the 'Records' menu, or as a button on the toolbar. Invoking the tool sets up a query design box displaying all the form fields into which the selection criteria are entered. The field names are dragged from the list provided. Sort criteria are simply selected from the list box.

However, the specification of the selection criteria is a little less straightforward. *Access* requires the information to include in the where clause of an SQL query and a rudimentary knowledge of SQL is definitely a help in giving it what it requires. However, the program is far from pedantic about syntax and usually accepts the obvious (e.g. =50, >50 or <50).

Tips. To select data values within a range use the syntax 'between value and value'. To match part of a string use the syntax 'like "wild-card string"'. The *Access* wild cards are '.' for a single character and '*' for any number of characters.

Access 7.0 and later include an alternative filtering interface termed 'Filter by Form'. In this, the filter criteria are entered directly into the boxes on a copy of the actual data form, rather than through a query design form. This is a much more convenient interface and is the preferred option of most *Access* users in BODC.

Once the filter criteria have been set up, the 'Apply filter/sort' option is invoked from the 'Records' menu or using the toolbar button. The result is a form of identical appearance to the original except for the number of records it contains.

Having identified a record of interest in an *Access* form, the next problem to address is how to get the information from that record into another application such as *Excel*. The best way of doing this by far is to copy the data over the Windows clipboard. To do this, click on 'Select Record' in the 'Edit' menu followed by 'Copy'. The record, including column names based on table field names, may then be pasted into *Excel*.

Providing the form doesn't contain a sub-form, that's all there is to it. However, most of the forms present in the OMEX II database include sub-forms. The data from these must be copied over in a separate operation. Select the sub-form data by clicking on the small box in its top left-hand corner. Then, simply click 'Copy' in the 'Edit' menu and paste the data into a convenient empty cell in the *Excel* spreadsheet.

This section has presented a very brief digest on how to use the forms included in the OMEX II database. *Access* is both intuitive to use and includes extensive documentation as on-line help. The more one learns about the program, the more one can get out of it. Use this section as a starting point but please do not regard it as an exclusive alternative to the documentation provided by Microsoft.

OMEX II Database Forms

Several forms have been created to allow the contents of individual tables or groups of tables to be conveniently viewed. These forms may be accessed through the 'Database Switchboard' form that appears when the database is opened. This contains a series of buttons. Pressing one of these opens another Switchboard page, which contains a further set of buttons. The bottom one of these returns to the main Switchboard page. The others open data forms.

The Database Switchboard offers the following pages:

Instrument Profile Data

Water Column Sample Data

Benthic Data

Production Data

Inventories and Parameter Dictionary

Currents

Please note that it is possible to use the Switchboard to open more than one form simultaneously. This is particularly useful for forms that display parameter codes (open the Parameter Dictionary to find out what they mean) or for the inter-comparison of different types of data.

Instrument Profile Data

CTD Data

This form provides access to all the CTD profiles held in the database. The header information (date/time, position, etc.) is in the upper section of the form. The profile itself is presented in the embedded form in the lower section. The profile for each measured parameter, including the independent variable (pressure) is listed sequentially. This is not the most convenient form for the data, but cross-tabulated queries including very large data tables proved problematical. Users may either transpose the data into a conventional grid by cutting and pasting in a spreadsheet or obtain data grids directly using the BODC Database Explorer software.

Filtering this form (e.g. by cruise or site) is advisable as it contains over 1000 records.

Instrument Profiles

This form provides access to the data collected by XBTs, profiling radiometers, and the FLY turbulence probe.

The form is divided into three sections. The upper left section provides information on when and where the profile was collected. The lower left section supplies information concerning the type of instrument used and the scientist who supplied the data.

The profile data themselves are displayed in the embedded form on the right-hand side. The profile for each measured parameter, including the independent variable (depth or pressure) is listed sequentially. This is not the most convenient form for the data, but it is all that is possible within the cross-tabulation rules imposed by *Access*. Users may either transpose the data into a conventional grid by cutting and pasting in a spreadsheet or obtain data grids directly with the BODC Database Explorer software.

Users interested in optics profiles should remember that some of the CTD profiles include downwelling and upwelling PAR scalar irradiance channels

Integrated Profile Data

This form displays ‘whole profile’ data derived from or related to series held in the Instrument Profiles table. It is designed to hold parameters such as K_d values derived from profiling radiometer data. In the case of OMEX II, it is

used to store the standard SeaWifs derived parameters, such as water leaving radiance values and surface chlorophyll, for the SeaWifs calibration casts collected during the 1998 Belgica cruise.

The form is divided into three sections. The upper left section provides information on when and where the profile was collected. The lower left section supplies information concerning the type of instrument used and the scientist who supplied the data. The parameter values are displayed in the embedded form on the right-hand side.

CTD Calibration Information

This form provides information on CTD calibrations derived and/or applied by BODC. It is important to realise that this table is provided for information only: all the calibrations listed have already been applied to the data.

The form displays a page for each parameter measured on each CTD cast. There are five sections to the form:

CTD Information	This section displays the basic metadata (date/time, position, water depth, etc.) for the CTD cast.
Parameter Information	This section displays the code and plain language name for the parameter together with details of the data originator.
Algorithm Information	This section describes the algorithm used, if any, to compute the parameter from other CTD data channels.
Calibration Information	This section provides further information on how the parameter was computed. There are three embedded forms that specify the input channels, the coefficients used and any plain language comments on the calibration.
Rig Geometry Information	This section displays the CTD/bottle rosette geometrical information used in the computation of water bottle depths from CTD pressure data.

Water Column Sample Data

Water Bottle Data

This form presents water sample data. Each record displayed represents a sample collection event. Consequently, all depths sampled by a CTD rosette are included as one record.

The embedded sub-form displays the profile for each parameter measured sequentially because of the cross-tabulation constraints imposed by *Access*. This is adequate for either a quick look at the data or if a small number of parameters are of interest. However, it is cumbersome if a significant number of parameters are required.

If a grid format with several parameters in each row is required, then use the BODC Database Explorer software to retrieve the data.

Integrated Bottle Data

This form displays column integrated data derived from water bottle profile data. In the case of OMEX II this table is used to display integrated primary production values derived from various types of artificial light incubation experiments. Note that the equivalent integrated values for natural light incubations are included in the Carbon Uptake form described below.

The form has a simple structure with station metadata parameters displayed in the upper portion and the integrated parameter values displayed as an embedded sub-form in the lower section of the form.

Zooplankton Net Data

This form presents the zooplankton abundance and biomass data, together with the results of experiments carried out on the net haul catches. The form displays one record for each catch. The form is in three sections.

The upper section contains station metadata fields such as date/time and position. The mid-section holds additional metadata fields that pertain to the net haul, such as mesh size and depths sampled. The embedded table at the bottom presents the profile for each parameter measured sequentially. This is adequate for either a quick look at the data or if a small number of parameters are of interest. However, it is cumbersome if a significant number of parameters are required.

If a grid format with several parameters in each row is required, then use the BODC Database Explorer software to retrieve the data.

CPR Phytoplankton Data

This form displays one record for each Continuous Plankton Recorder tow segment. A sub-form is included that displays the individual phytoplankton taxon abundances. Users are referred to the database table structure documentation for the definition of the abundance parameter used.

CPR Zooplankton Data

This form displays one record for each Continuous Plankton Recorder tow segment. A sub-form is included that displays the individual zooplankton taxon abundances. Users are referred to the database table structure documentation for the definition of the abundance parameter used.

Sediment Trap Data

This form displays sediment trap data and includes four sections. The top section provides trap deployment information common to all trap types. Below this are two sections, which display trap position information appropriate to moored traps on the left or drifting traps on the right.

The fourth section is a sub-form displaying each parameter profile time series sequentially. If a grid format with several parameters in each row is required, then use the Database Explorer software to retrieve the data.

Benthic Data

Core Profile Data

The Core Profile Data form gives access to a wide range of parameters measured as profiles along cores. The form is divided into three sections. The top section displays information on the coring time and position. The middle section provides additional information for the sub-core on which the profile was measured.

The bottom section is an embedded sub-form that lists the profiles for each parameter measured sequentially. If a grid format with several parameters in each row is required, then use the Database Explorer software to retrieve the data.

Whole Core Data

This form displays measurements made on a whole core, sub-core or grab sample. The form is in three sections. The top section displays information on the coring event. Below this is additional information on the individual sub-core used. The third section is a sub-form containing one record for each parameter measured.

The whole core data in the OMEX II database include relatively small numbers of parameters. Consequently, the form display provides adequate access to the data. However, if the parameters are required listed in grid format, the Database Explorer software may be used instead.

Benthic Macrofauna Data

This form presents benthic macrofauna abundance and biomass data on core samples, usually the top 15cm of the sediment. The main body of the form gives information on where and when the core was collected. The embedded sub-form provides a sequential display of taxon abundances.

Note that total macrofauna abundance profiles are included in the Core Profile data.

Benthic Meiofauna Data

This form provides access to profiles of benthic meiofauna abundance and nematode biomass. The main body of the form gives information on where

and when the core was collected. The embedded sub-form provides a sequential display of taxon abundance and biomass for each horizon sampled. Note that replicate core profiles are provided for most of the stations. These may be distinguished using the core reference numbers.

Benthic Megafauna Data

The Benthic Megafauna Data form displays biological data collected using an epibenthic sledge. The form is divided into three sections. The first section lists information about where and when the trawl took place. The second section presents information pertaining to the catch as a whole. The third section is an embedded sub-form that lists the abundance and biomass data by taxonomic group.

Production Data

Carbon Uptake

This form presents data from long duration (24 hour) ^{14}C incubation experiments, which may not be normalised to uptake per hour. Note that P:I parameters (alpha and P_{max}) and normalised (per hour) short-term uptake data are included as water bottle parameters and may be obtained using either the Water Bottle Data form or the Database Explorer software.

The form contains five sections. The top section displays information on where, when and how the water used in the experiment was collected. Below this are sections presenting information on the incubation conditions and the size fraction definitions. The fourth section displays column integrated uptake data. The fifth section is a sub-form that presents the uptake rates for the individual samples incubated during the experiment.

Nitrogen Uptake

This form presents data from long duration (24 hour) ^{15}N incubation experiments, which may not be normalised to uptake per hour. Note that normalised (per hour) uptake rates are included as water bottle parameters and may be obtained using either the Water Bottle Data form or the Database Explorer software.

The form contains five sections. The top section displays information on where, when and how the water used in the experiment was collected. Below this are sections presenting information on the incubation conditions and the size fraction definitions. The fourth section displays column integrated uptake data. The fifth section is a sub-form that presents the uptake rates of nitrate, ammonium and urea for the individual samples incubated during the experiment.

Phosphorus Uptake

This form presents data from ^{32}P and ^{33}P incubation experiments. The data included here are from long-term experiments where it is inappropriate to normalise the data in terms of uptake per hour or from P:I experiments that have not been parameterised in terms of alpha and P_{max}. Note that normalised (per hour) uptake data and P:I parameters are included as water bottle parameters and may be obtained using either the Water Bottle Data form or the Database Explorer software.

The form contains five sections. The top section displays information on where, when and how the water used in the experiment was collected. Below this are sections presenting information on the incubation conditions, including the phosphorous isotope used, and the size fraction definitions. The fourth section displays column integrated uptake data. The fifth section is a sub-form that presents the uptake rates for the individual samples incubated during the experiment.

Phosphate Uptake Kinetics

This form presents the results of experiments designed to show the effect of phosphate concentration on the rate of phosphate uptake by the particulate phase. The upper part of the form provides information on the origin of the water sample used for the experiment. The embedded sub-form below this displays the phosphate uptake rates against ambient phosphate concentration.

Inventories and Parameter Dictionary

Cruises

This form presents background information on all the cruises included in the database. These include Continuous Plankton Recorder tows and mooring deployment or recovery cruises in addition to the cruises dedicated to OMEX II.

Events

This form provides access to the information held on data events. The form is divided into four sections. The top two sections display information common to all events and their associated cruise. The lower sections provide information specific to point events or traverse events. Note that current practice is to include 'approximate' point event information for traverse events. This makes the latter visible to basic geographical searches.

Browsing this form can quickly provide an overview of the scope of the oceanographic measurements held in the database.

Note that most of the events in the database are underway ADCP profiles. If these are not of interest, applying an appropriate filter (Not like "ADCP*" in the Gear box) makes the data set much more manageable, reducing the record count by a factor of 5.

Moored Instruments

This form provides an index to the moored instrument records provided on the CD-ROM. The form is in two sections. The upper part provides information about the mooring, such as position and deployment/recovery times. The lower section supplies information about an individual instrument deployed on the rig.

The BODC Series Reference field is more useful than one might think as it is used as part of the data filename. Although the rules for deriving the full pathname from the Series Reference are not obvious, the relevant file may be easily found using wild cards with the *Windows Explorer* 'Find' tool.

Parameter Dictionary

This form provides the definitions of the BODC parameter codes, which are used extensively in the OMEX II database. Its primary purpose is to allow the meaning of a parameter code to be obtained. If using a form that includes a parameter code, it may be convenient to also have the Parameter Dictionary form open at the same time. Codes are most conveniently located by using the 'Find' tool on the ‘Parameter Code’ box.

Currents

Underway ADCP

This form presents the underway Acoustic Doppler Current Profiler data. At the top of the form there are three sections specifying the profile header information, the acquisition/processing methods and the calibrations applied.

The embedded form displays the profile datacycles, including signal amplitude, which provides semi-quantitative information on zooplankton biomass distribution.

Drifter Data

The Drifter Data form presents the track co-ordinates of drifting buoys. The main body of the form provides deployment information. The track co-ordinates are listed in the embedded sub-form. Surface temperatures from the buoy sensor are included.

Moored Instruments

This form provides an index to the moored instrument records provided on the CD-ROM. The form is in two sections. The upper part provides information about the mooring, such as position and deployment/recovery times. The lower section supplies information about an individual instrument deployed on the rig.

The BODC Series Reference field is more useful than one might think as it is used as part of the data filename. Although the rules for deriving the full pathname from the Series Reference are not obvious, the relevant file may be easily found using wild cards with the *Windows Explorer* 'Find' tool.