BODC Project Database Structure

PARENT HOME

Introduction

A relational database is made up from tables. Each table contains one or more fields. Some of these fields, termed key fields, are contained in more than one table and provide the mechanism for linking tables, and hence the data they contain, together. If a key field occurs once in each of two tables, then a one to one relationship between the tables is established. If the key field occurs once in one table and many times in the other, then a 'one to many' relationship exists between the tables.

The BODC Community Research Project databases are relational databases that have been built for the specific task of storing project data in a way where it may easily be found when required. The end result is an effective, if not elegant, design.

Simple structures exist that match the data. These structures are extended in response to data sets supplied. However, as the databases have developed, certain patterns have been recognised in the data. In response to this, fully normalised structures (i.e. the type conventional database designers desire) have been developed. The advantage of these normalised structures is that providing the rules of their underlying data model are obeyed by the data, their scope may be expanded with no maintenance overheads. Their disadvantage is that significantly more work is required getting data in and getting data out.

We therefore have a situation in practice where simple and fully normalised structures exist side by side. Any spare resource is directed towards converting the simple structures into normalised structures, providing a clear advantage can be seen in doing the work. The structure of the database is therefore dynamic but it is supported by 'soft' documentation that can evolve in parallel.

Table Types

The database may be considered as containing five types of table. The database is built on an event-based data model. In other words, something has to happen to generate the data stored. The primary information in the database therefore has to describe what these events were, where they happened and when they happened.

This information is stored in the database primary index tables. In the following table definitions, the hot links to these tables are coloured red.

The data model assumes that the events are related to the data they generate by one to many relationships. These relationships are implemented in the database by one or more secondary index tables. These tables also provide storage for metadata that are specific to a single type of event. The hot links to these tables in the table definition index are coloured green.

The third type of table is the fully normalised data table. These may be regarded as stable, long term entities within the database. Because the structures are normalised, it is not possible to obtain the sort of cross-tabulated output most users require using simple SQL queries. Consequently, data access tools are provided by BODC. The hot links to these tables in the table definition index are coloured blue.

The normalised data tables are supported by a series of code tables, such as the parameter dictionary, that together may be termed the dictionary tables. The hot links to these tables in the table definition index are coloured dark yellow.

Finally, there are the simple-structured data tables. These may be effectively interrogated by simple SQL queries. However, they should be regarded as transient entities that may disappear from subsequent database releases. Obviously, if they do disappear, the data they contained will have been transferred to a fully normalised structure within the database. The hot links to these tables in the table definition index are coloured magenta (pale purple).

The Parameter Dictionary

The parameter dictionary is an essential feature of the normalised data storage tables. The identification of parameters is based on 8-byte codes. These have been designed using a hierarchical model. The first four bytes may be considered as the 'parent'. This provides information on the parameter at a low level of detail. This parent has one or more 'child' 4-byte codes. These subdivide the parent into more detailed information.

This relationship is exploited in different ways. For example, chemical parameters have a parent field identifying the basic parameter with the children identifying different measurement protocols. Thus 'CPHL' is chlorophyll-a, but CPHLHPP1 is chlorophyll-a measured by reverse-phase HPLC on an acetone extract from a GF/F filter. For biological species codes the parent specifies the genus and the children the species.

There are a large number (thousands) of parameters coded in the database. Finding out what a given code means is straightforward (!). A query matching on field CPMUSG of table ZUSG will provide the answer. However, specifying the parameter code for data retrievals requires some thought.

The secret lies in the use of wild cards which any database management system can incorporate into query searches. The recommended technique is to use table ZUPM to identify the parent code of the parameter you require. A wild card may then be set up to include as many of the child codes as required. One word of warning. Always check the meanings of all codes covered by a wild card as there are traps for the unwary. For example, the wild card CORG% covers both CORGCAP1 ("POC") and CORGCOD1 (DOC) which should not be merged into a single data set!

Documentation Structure

This document contains two main sections. These are:

Table Definitions: A description of the fields contained in each table of the database relevant to the PRIME project.
Linkage Definitions: Documentation that describes how the tables of the database are linked together through their key fields.

Database Table Definitions

This section provides a field level description of all the user-accessible tables in the OMEX and ODB databases which are relevant to the PRIME project. A complete description of ALL the database tables is given in the version of this document found on the OMEX I CD-ROM.

Table ARGOS

This table contains all the position information from the drifting buoys deployed as part of the PRIME cruise Lagrangian experiment.

Table BINCTD

This table contains the CTD profile data, averaged into either 1 db (casts shallower than 100 db) or 2 db bins for cruise data and 0.1db bins for mesocosm data.

Table BOTDATA

This table contains analytical data on water and air samples. A very wide range of parameters is stored here.

Table BOTTLE

This table provides an inventory of the water and air samples, collected by a variety of methods, held in the database. Note that the name BOTTLE is more of a historical relic than a description of current usage. Vital information, particularly the depth or height from which the sample was taken, is held in the table.

Table BOTYPINDX

This is a code table that supports table BOTTLE by defining the mnemonics used to identify bottle types.

Table CRSINDX

This table allows the user to identify which events were associated with which project by linking together the cruise fields from table EVENT. Metadata storage fields for each cruise are also provided.

Table CTDCAL

This table contains the calibration coefficients applied to the CTD profiles by the BODC processing system.

Table CTDINDX

This table provides an inventory of the CTD casts held in the database together with storage for CTD-specific information.

Table CTDTYP

This is a code table that supports table CTDINDX by defining the mnemonics used to identify CTD instrument types.

Table EVENT
An event is defined as any activity that results in the collection of data that are stored in the database. Table EVENT contains information on what the event was and where and when it occurred. It could therefore be considered as the most important table in the database and should certainly be involved at the start of any search for data.

Table G_CODE
G_CODE is a simple code table that defines the gear codes used in table EVENT.

Table FATACID
This table contains concentrations and d13c ratios of a number of different classes of fatty acids, measured on general particulates or doliolid samples.

Table MULTI
This table contains abundance, surface area and volume data for a huge number of different particle sizes studied as part of the mesocosm experiment.

Table N15DAT
This table contains ^{15}N uptake (new and regenerated production) data from long period (generally 24 hour) on-deck and in-situ incubations, including size-fractionated data.

Table N15HDR
This table provides an inventory of the ^{15}N uptake (new and regenerated production) uptake experiments held in the database. Vital metadata fields are included. The table also provides storage for column integrated data.

Table NETINDX
This table provides an inventory of the net haul samples, collected by a variety of methods, held in the database. Vital information, particularly the depth from which the sample was taken and the mesh size of the net, is held in the table.

Table NETDATA
This table contains analytical data on net haul samples. A wide range of data is stored here.

Table ORGCODE
Table ORGCODE is a code table defining the data originator codes used in the database.

Table SECCHI
This table contains secchi disk depths and weather and sea state observations made during a number of cruises to Ocean Weather Station India.

Table SSINDX
This table provides an inventory of the SeaSoar casts held in the database together with storage of SeaSoar-specific information.

Table WEATHER
This table contains details of weather observations made at Bergen airport for the duration of the mesocosm experiment.

Table ZUCT

This table is part of the parameter dictionary. Its function is to group the 4-byte 'parent' parameter codes into categories to enable them to be interrogated more easily.

Table ZUNT

This table is part of the parameter dictionary. It is a code table that defines the codes used to represent parameter units.

Table ZUPM

This table is part of the parameter dictionary. It contains the definitions of the 4-byte 'parent' parameter codes (i.e. the first four bytes of the parameter code).

Table ZUSG

This is the main table of the parameter dictionary, containing definitions of the full 8-byte parameter codes.

Table ARGOS
Field Definitions

BEN	NUMBER(6)	BODC event number.
DATIM	DATE	Date/Time (GMT) of position fix
LAT	NUMBER	Position of buoy, degrees positive North
LON	NUMBER	Position of buoy, degrees positive East

Table BINCTD
Field Definitions

BEN	NUMBER(6)	BODC event number.
PRESS	NUMBER(5,1)	Pressure (db).
TEMP	NUMBER(5,3)	Temperature (°C).
SALIN	NUMBER(5,3)	Practical salinity (PSU).
SIGMA	NUMBER(5,3)	Potential density anomaly (kg/m^3).
O2	NUMBER(4,1)	Dissolved oxygen (µM).
O2SAT	NUMBER(4,1)	Oxygen saturation (%).
CPHYL	NUMBER(4,2)	Chlorophyll (mg/m^3).
ATTEN	NUMBER(5,3)	Optical attenuance (per m).
DWIR	NUMBER(5,1)	Downwelling irradiance (µE/m^2/s).
UWIR	NUMBER(4,1)	Upwelling irradiance (µE/m^2/s).
POTEMP	NUMBER(5,3)	Potential temperature (°C).
POAT	NUMBER(5,3)	Potential attenuance (per m).

Notes
The pressure value signifies the midpoint of the bin. Thus, a pressure of 1.0 db signifies a bin extending from 0 db to 2 db (assuming a 2 db binning interval).

The density parameter computed is the potential density anomaly calculated at 0 db and is numerically equivalent to the parameter known as sigma-theta (computed by substituting potential temperature into the UNESCO SVAN function).

Oxygen saturation has been computed using the algorithm of Benson and Krause (1984).

Reference
Benson, B.B., Krause D. (1984). The concentration and isotopic fractionation of oxygen dissolved in fresh water and sea water in equilibrium with the atmosphere.
Limnol.Oceanogr., 29, 620-632.

Table BOTDATA
Field Definitions

IBTTLE	NUMBER(6)	BODC bottle/sample identifier.
CPCODE	CHAR(8)	Parameter code.
FPVAL	NUMBER	Parameter value.
CPFLAG	CHAR(1)	Parameter quality control flag.
IORGRF	NUMBER(6)	Originator's reference.
IDOCRF	NUMBER(8)	Document reference.
CILOAD	CHAR(6)	Record creation date (yymmdd).
TSGMOD	DATE	Last modification time stamp.

Notes
The primary key is formed from the three fields, IBTTLE, CPCODE and IORGRF. In other words, the table contains one row for each parameter measurement on each water or air sample by a given data originator. The parameter code consists of 8 bytes, which describe the parameter measured in some detail. The parameter code definitions are stored in the parameter dictionary (see the table names starting with 'Z').

The parameter flag field serves two purposes. First, it identifies parameter values identified as problems during quality control procedures. Different codes are used to differentiate between originator, BODC and user quality control. Secondly, it is used to identify samples where the measured parameter was either below detection limit or saturated the measuring apparatus. In these cases the data values are set to the detection limit (zero if no detection limit was specified) or the saturation value respectively. If no flag value has been assigned (signifying good data), the CPFLAG field is blank.

The flag values that may be encountered are:

K Uncertain/suspect value (source of quality control unknown).
L Uncertain/suspect value (data originator's quality control).
M Uncertain/suspect value (BODC quality control).
O Uncertain/suspect value (user quality control)
T Nearest value to bottle firing depth
< Below detection limit.
> In excess of stated value.

The 'T' flag is only found on records created for water bottle samples from CTD profile data. It means that no data were found at the bottle firing pressure. Instead, the nearest data value has been taken, providing this was within 2 db of the required pressure.

The originator's reference field allows the suppliers of individual data values to be identified. The objective when allocating these linkages is to provide a point of contact for users of the data to approach when initiating collaboration that will endure beyond the end of a project. Consequently, linkages have been assigned at the PI level and do not necessarily specify the individual who actually did the analysis.

Codes are used to eliminate potential problems with misspellings and the like. The codes used are documented in the table ORGCODE. Document references have not yet been implemented so the IDOCRF field is currently always null.

Table BOTTLE
Field Definitions

BEN	NUMBER(6)	BODC event number.
IBTTLE	NUMBER(6)	A unique identifier assigned by BODC to each sample.
MINP	NUMBER(5,1)	Minimum pressure for the sample (db).
MAXP	NUMBER(5,1)	Maximum pressure for the sample (db).
DEPTH	NUMBER(6,2)	Sampling depth (m).
BOTYP	CHAR(4)	Bottle/sample type identifier.
FLAG	CHAR(1)	Problem indicator flag.

Notes
Table BOTTLE was originally conceived for the management of water bottle data. However, as the BODC databases developed, it was realised that the table could be utilised for other data types. Data currently held include pumped air and water samples, stand-alone pump (SAP) samples, bucket samples and air bottle samples. The most important function of this table is to implement the 'one-to-many' relationship that may exist between samples and events. The table contains one row per sampling depth (multiple samples at a single depth are considered as one). Each record in EVENT can 'own' as many records as it likes in BOTTLE through the foreign key field BEN. Hence each EVENT can include many sampling depths.

The relationship between MINP, MAXP and DEPTH requires some explanation. MINP and MAXP only have relevance to bottles on a CTD rosette. In this case, bottle 'depths' are frequently logged as pressure ranges during CTD screening and loaded into BOTTLE. Subsequently, applying a pressure calibration to MINP and MAXP derives DEPTH (the distance from the surface to the midpoint of the bottle), correcting for CTD frame geometry, and applying the standard conversion from pressure to depth. In order to allow for pressure calibration drift, the minimum value is constrained at 0.5 m. The fields MINP and MAXP provide a direct linkage between BOTTLE and the CTD data, which is why they are retained. For other sample types, DEPTH is assigned a value from reports or logs and MINP and MAXP are left null. Note that air samples have negative depths to indicate height above sea level.

The BOTYP field specifies how the sample was collected. For water bottle data, this field identifies the type of water bottle used. The codes are defined in table BOTYPINDX. A number of other codes are used for other sampling methods:

The FLAG field is used to indicate known problems. The coding convention used is:

B	Filter burst (SAP samples)
L	Contamination through leakage suspected
M	No sample obtained
O	Bottles fired in incorrect order

The 'O' flag is used to flag stations where there was obvious confusion from the sample data set about which bottle was fired at which depth. These problems have been resolved during data load, but the flag is included to remind users that there may be problems with data from that station obtained outside the database.

Table BOTYPINDX
Field Definitions

BOTYP	CHAR(4)	Bottle type code
DESCR	CHAR(20)	Plain language description

Notes
This table explains the mnemonics used in table BOTTLE for different bottle types.

Table CRSINDX
Field Definitions

CRUISE	CHAR(8)	BODC cruise mnemonic,
PROJECT	CHAR(12)	Mnemonic of the project with which the cruise was associated.
PSO	CHAR(20)	Cruise chief scientist.
COUNTRY	CHAR(20)	Country responsible for organizing the cruise.
TBEGNS	DATE	Date the cruise sailed.
TENDS	DATE	Date the cruise docked.
LOCATION	CHAR(80)	Plain language description of the area studied.
COMM	CHAR(60)	Plain language comment field.

Notes
This table allows events associated with a particular project to be identified as well as providing limited background information on cruises.

Table CTDCAL
Field Definitions

BEN	NUMBER(6)	BODC event number.
FPCOR	NUMBER(5,2)	Pressure correction (db).
FCSLOP	NUMBER(7,4)	Chlorophyll calibration slope.
FCIRR	NUMBER(8,7)	Chlorophyll calibration irradiance term.
FCCEPT	NUMBER(7,4)	Chlorophyll calibration intercept.
FUSLOP	NUMBER(7,4)	Upwelling irradiance calibration slope.
FUCEPT	NUMBER(7,4)	Upwelling irradiance calibration intercept.
FDSLOP	NUMBER(7,4)	Downwelling irradiance calibration slope.
FDCEPT	NUMBER(7,4)	Downwelling irradiance calibration intercept.
FBASOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the base of the water bottle (m).
FTOPOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the top of the water bottle (m).
FTEMOF	NUMBER(5,2)	Distance between the CTD pressure sensor and the reversing thermometer (m).
FSSLOP	NUMBER(7,5)	Salinity calibration slope.
FSCEPT	NUMBER(7,5)	Salinity calibration intercept.
FTSLOP	NUMBER(7,5)	Temperature calibration slope.
FTCEPT	NUMBER(7,5)	Temperature calibration intercept.
FOSLOP	NUMBER(6,3)	Oxygen calibration slope.
FOCEPT	NUMBER(6,3)	Oxygen calibration intercept.

Notes
This table contains one row per CTD cast and therefore allows each CTD to have a separate calibration. However, in most cases calibrations have been set up on a cruise by cruise basis. Each calibration and its method of determination are now discussed.

Rig Geometry
The fields FBASOF, FTOPOF and FTEMOF contain the information required to compute the true water bottle depth from the CTD pressure channel. FTEMOF is used when extracting calibration temperatures. The water bottle depth for a given CTD pressure reading (calibrated and converted to depth) is given by:

BOTTLE DEPTH = CTD DEPTH - (FBASOF + ((FTOPOF - FBASOF)/2.0))

This equation assumes bottle depth to be defined as the depth to the midpoint of the water bottle. The depth of the reversing thermometer is obtained by simply subtracting FTEMOF from the calibrated and converted CTD pressure reading.

The values used in these fields were obtained from actual measurements of the CTD rigs.

Pressure
The pressure correction, FPCOR, is a simple offset, which is added to the uncalibrated CTD pressure. It is derived by consideration of data logged when the CTD was obviously out of the water.

Temperature

The temperature calibration has two components, FTSLOP and FTCEPT, which are applied to the uncalibrated temperature using the equation:

TCAL = TRAW * FTSLOP + FTCEPT

The temperature calibration is derived by comparison of the CTD temperature channel with calibrated digital reversing thermometer data for a specified cruise. A mean offset is computed after rejection of suspect reversing thermometer readings and stations where the reversing thermometers were fired on a temperature gradient.

In most cases, the accuracy of CTD resistance thermometers exceeds that of the digital reversing thermometers in common use. Consequently, the calibration coefficients are set to 1 and zero unless a problem is suspected with the CTD calibration.

Salinity

The salinity calibration is identical in form to the temperature calibration and has been derived in a similar manner using water bottle samples assayed on a bench salinometer.

Chlorophyll

The chlorophyll concentration (in mg chlorophyll-a/m^3) may be obtained from the fluorometer voltage using the following equation:

CHLOROPHYLL = EXP (VOLTS * FCSLOP + FCIRR * DWIR + FCCEPT)

The chlorophyll calibrations were set up by multiple regression of fluorometer voltages and downwelling irradiance (at the water bottle firing depths) against the log of the associated extracted chlorophyll measurements. The calibration was done on quality controlled data.

It should be noted that, on many cruises, the FCIRR term is zero because no quantitative relationship between voltage, light and chlorophyll could be established.

Oxygen

The oxygen calibration is of the form:

OXCAL = OXRAW * FOSLOP + FOCEPT

and is derived by regression of CTD channel values at the bottle firing depths against Winkler titration results.

Irradiance

The raw irradiance (upwelling and downwelling) data are held as voltages. These are calibrated using the equation:

IRRADIANCE = EXP (VOLTS*SLOPE + INTERCEPT)

This returns calibrated values in units of $\mu W/cm^2$

For the PML 2-pi PAR meters currently used on CTDs on the NERC ships, an empirical calibration factor (0.0375) has been determined to convert these data into $\mu E/m^2/s$. The calibration is only valid over the range -1.5V to +1V.

Table CTDINDX
Field Definitions

BEN	NUMBER(6)	BODC event number.
TBEGNC	DATE	Date/time of the start of the downcast.
TENDC	DATE	Date/time of the end of the downcast.
MAXP	NUMBER(5,1)	Maximum pressure in the downcast (db).
FMAXP	CHAR(1)	Set to 'C' if the pressure calibration held in table CTDCAL has been applied to MAXP. Otherwise left null.
EXTCO	NUMBER(5,3)	Downwelling irradiance extinction coefficient.
MLD	NUMBER(3,1)	Mixed layer depth (m).
EZD	NUMBER(4,1)	Depth to the base of the euphotic zone (m).
TYPE	CHAR(3)	Type code of the CTD used (i.e. NB3 for Neil Brown Mk. III).

Notes

The downcast start and end times have been derived from the CTD data time channel and may be used to regenerate that channel if required.

The fields EXTCO, MLD and EZD were set up for the BOFS programme. In practice, it has been found that providing universally acceptable algorithms for their computation is an impossible task. Consequently, current practice is to leave the fields null unless agreed values are provided by the scientific community.

The CTD type codes are defined in table CTDTYP.

Table CTDTYP
Field Definitions

TYPE	CHAR(3)	CTD type code mnemonic.
DESCR	CHAR(30)	Plain language definition of the mnemonic.

CTD type code definitions.

Table EVENT
Field Definitions

BEN	NUMBER(6)	BODC event number. A unique numerical identifier assigned each event.
OID	CHAR(12)	What the event was known as during the cruise(originator's identifier)
GCODE	CHAR(8)	Code used to specify the gear pertaining to the event.
TBEGNS	DATE	Event start date/time (GMT).
TENDS	DATE	Event end date/time (GMT).
LAT	NUMBER(7,5)	Average latitude for event (°+ve North).
LON	NUMBER(7,5)	Average longitude for deployment (°+ve East).
VARLAT	NUMBER(7,5)	Maximum deviation of latitude from mean during station.
VARLON	NUMBER(7,5)	Maximum deviation of longitude from mean during station.
WDEPTH	NUMBER(5,1)	Average bathymetric depth for the event (m).
LATS	NUMBER(7,5)	Latitude at time TBEGNS (°+ve North).
LONS	NUMBER(7,5)	Longitude at time TBEGNS (°+ve East).
LATE	NUMBER(7,5)	Latitude at time TENDS (°+ve North).
LONE	NUMBER(7,5)	Longitude at time TENDS (°+ve East).
CRUISE	CHAR(8)	Cruise mnemonic.
SITE	CHAR(12)	Fixed station name.

Notes

This table has been built from the best available information from cruise reports, log sheets and information accompanying data. Automatically logged navigation has been used to match times and positions wherever possible.

There are two types of event, point events and traverse events. Point events may be considered as those events that effectively happen at a fixed position. Their positions are specified by the fields LAT, LON, VARLAT and VARLON with the other four position fields left null. Traverse events, such as tows and trawls, involve the ship steaming a significant distance. In this case, the start and end positions are stored in LATS, LONS, LATE and LONE. Note that some point events have data entered into the point event position fields to allow them to be handled as very low resolution points as required. Water depths are only included for point events. Note that the PRIME Mesocosm experiment was really neither a point event nor a traverse event and consequently has no position specified in EVENT.

Wherever possible, the fields LAT and LON are derived by averaging the data from the ship's navigation log over the event duration. VARLAT and VARLON are the maximum deviation of the data set from the mean. If VARLAT and VARLON are null then the data in LAT and LON have been take from logs or reports. Obviously, the average of the ship's positions are not used for moorings. If VARLAT and VARLON are set then the information has been derived from the difference of the recorded positions on deployment and recovery.

The BODC event number (BEN) is a concept introduced to overcome the problem that it is impossible to guarantee that the identifiers assigned during the cruise will be unique within database incorporating many cruises. It is a very important field because it is used within the database as a 'primary key' which by definition must be unique. Data elsewhere in the database, resulting from a specified event, will either be labelled directly, or via a linkage record to its BEN.

OID, the originator's identifier, is the label that was assigned to the event during the cruise, mesocosm experiment or Ocean Weather Station India campaigns.

Event start and end times have been specified to bracket the event. Thus, for a CTD cast, the time span is from the instrument leaving the deck until its return. Some events are regarded as instantaneous, for example non-toxic samples. In these cases, the end times are set null. Wherever possible, cores are regarded as instantaneous events at the time when the corer reached the bottom.

The gear codes are mnemonics used to describe the data collection activity or the equipment used. The codes have been chosen to convey as much meaning as possible, but a plain language description of each code is provided in table G_CODE.

The cruise identifiers are made up from a ship code concatenated with the cruise identifier. For example, 'DI' is used for Discovery and 'CD' for Charles Darwin and a typical cruise would be labelled DI182 or CD46. There are a number of situations in the PRIME project where the concept of a cruise is not relevant, but the sampling details are put into table EVENT for convenience. In these cases, the ‘cruise’ has been artificially defined and the cruise identifier set accordingly. For the mesocosm experiment, the cruise identifier has been set as ‘MESOCOSM’. Similarly, for the historical data relating to the Ocean Weather Station India, there are many observations where the actual cruise has not been specified. In these cases, the cruise identifier has been set as ‘OWSINDIA’.

Table FATACID

Field Definitions

BEN	NUMBER(6)	BODC event identifier
SOURCE	CHAR2(2)	Source of material
D13C	CHAR2(2)	Concentration/δ^{13}C discriminator
POLAR	NUMBER	Polar lipids
PIGMENT	NUMBER	Pigments
STEROL	NUMBER	Sterols
UNKNOWN	NUMBER	Unknown grouping
FFA	NUMBER	Free fatty acids
TAG	NUMBER	Triacylglycerols
STEST	NUMBER	Sterol esters
C14X00	NUMBER	C14:00 fatty acid
C15X00	NUMBER	C15:00 fatty acid
C16X00	NUMBER	C16:00 fatty acid
C16X1NY7	NUMBER	C16:1,N-7 fatty acid
C16X2NY3	NUMBER	C16:2,N-3 fatty acid
C17X00	NUMBER	C17:00 fatty acid
C16X3NY3	NUMBER	C16:3,N-3 fatty acid
C16X4NY1	NUMBER	C16:4,N-1 fatty acid
C18X00	NUMBER	C18:00 fatty acid
C18X1NY9	NUMBER	C18:1,N-9 fatty acid
C18X1NY7	NUMBER	C18:1,N-7 fatty acid
C18X2NY6	NUMBER	C18:2,N-6 fatty acid
C18X3NY6	NUMBER	C18:3,N-6 fatty acid
C18X3NY3	NUMBER	C18:3,N-3 fatty acid
C18X4NY3	NUMBER	C18:4,N-3 fatty acid
C18X5NY3	NUMBER	C18:5,N-3 fatty acid
C20X00	NUMBER	C20:00 fatty acid
C20X1NY9	NUMBER	C20:1,N-9 fatty acid
C20X4NY6	NUMBER	C20:4,N-6 fatty acid
C20X4NY3	NUMBER	C20:4,N-3 fatty acid
C20X5NY3	NUMBER	C20:5,N-3 fatty acid
C22X00	NUMBER	C22:00 fatty acid
C22X5NY6	NUMBER	C22:5,N-6 fatty acid
C22X6NY3	NUMBER	C22:6,N-3 fatty acid
SFA	NUMBER	Saturated fatty acids
MUFA	NUMBER	Monounsaturated fatty acids
PUFA	NUMBER	Polyunsaturated fatty acids

Notes

The concentration or δ^{13}C discriminator indicates whether the data values held in the record are fatty acid concentrations in nanomoles/litre or ^{13}C enrichments in per mil.

Table G_CODE
Field Definitions

GCODE	CHAR(8)	Standardised gear code.
DESCR	CHAR(60)	Plain language description of the gear described by GCODE.

Table MULTI
Field Definitions

IBTTLE	NUMBER(6)	BODC bottle reference number
CHANID	NUMBER	Multi-sizer channel reference number
MINDIAM	NUMBER	Minimum particle diameter included in the channel (μm)
MAXDIAM	NUMBER	Maximum particle diameter included in the channel (μm)
ABUND	NUMBER	Number of particles (per l)
SA	NUMBER	Total particle surface area (μm^2 per l)
VOL	NUMBER	Total particle volume (μm^3 per l)

Notes
The source of the water may be identified through IBTTLE. The table contains one record for each multi-sizer channel.

Table N15DAT
Field Definitions

EXPREF	CHAR(6)	BODC experiment reference.
IBTTLE	NUMBER(6)	BODC bottle reference number.
DEPTH	NUMBER(4,1)	Depth (or depth equivalent) at which the sample was incubated.
TNO3	NUMBER(4,2)	Nitrate uptake (µM/day).
TNH4	NUMBER(4,2)	Ammonia uptake (µM/day).

Notes
The experiment reference provides a linkage between the metadata held in table N15HDR and the individual uptake measurements held in N15DAT.

The source (position and depth) of the incubated water may be identified through IBTTLE. Note that IBTTLE will not be unique for every record in cases where a common water sample was incubated at several depths.

Note that the units are quoted in terms of uptake per day. This is a loose definition. Strictly speaking, the uptake is quoted over the period of the incubation duration. Normally this is approximately 24 hours but users are advised to check the duration in the appropriate field of N15HDR.

Table N15HDR
Field Definitions

EXPREF	CHAR(6)	BODC experiment reference.
TYPE	CHAR(2)	Experiment type code. (OD for on deck experiments, IS for in-situ experiments)
BENCOL	NUMBER(6)	BODC event number of the water collection event.
BEN	NUMBER(6)	BODC event number assigned to the incubation..
SDATE	DATE	Date and time of the start of the incubation.
INCDUR	NUMBER(3,1)	Incubation duration (hours).
COMM	CHAR(30)	Plain language comment field.
DEPINT	NUMBER(4,1)	Depth over which the integrated production was calculated.
INTNO3	NUMBER(6,2)	Integrated nitrate uptake (µmol/m^2/day)
INTNH4	NUMBER(6,2)	Integrated ammonia uptake (µmol/m^2/day)
MICDEF	CHAR(8)	Microplankton definition.
NANDEF	CHAR(8)	Nanoplankton definition.
PICDEF	CHAR(8)	Picoplankton definition.

Notes

Fields BENCOL and BEN require some explanation as the presence of two BODC event numbers in a single table may at first sight seem confusing. BENCOL specifies where the water used in the production experiment came from. In some ways it is superfluous because the same information may be derived from the IBTTLE field in N15DAT. However, it is included to simplify the task of linking integrated production data held in table N15HDR to the place and time to which they relate.

BEN is a reference given to some production experiments. This invariably relates to *in situ* experiments where a rig has been cast adrift from the ship. On-deck incubations have never been considered as events. The reason for this is more historical than logical: the event entries are drawn up from ship’s logs and whilst a rig being deployed has often (but not always) merited a log entry, the placing of samples in an on-deck incubator has not.

Integrated production data are only included if they were computed and supplied by the data originator. They are not routinely determined by BODC.

Table NETDATA
Field Definitions

INET	NUMBER(6)	BODC net sample identifier.
CPCODE	CHAR(8)	Parameter code.
FPVAL	NUMBER	Parameter value.
CPFLAG	CHAR(1)	Parameter quality control flag.
IORGRF	NUMBER(6)	Originator's reference.
IDOCRF	NUMBER(8)	Document reference.
CILOAD	CHAR(6)	Record creation date (yymmdd).
TSGMOD	DATE	Last modification time stamp.

Notes

The primary key is formed from the three fields, INET, CPCODE and IORGRF. In other words, the table contains one row for each parameter measurement on each net sample by a given data originator. The parameter code consists of 8 bytes, which describe the parameter measured in some detail. The parameter code definitions are stored in the parameter dictionary (see the table names starting with 'Z').

The parameter flag field serves two purposes. First, it identifies parameter values identified as problems during quality control procedures. Different codes are used to differentiate between originator, BODC and user quality control. Secondly, it is used to identify samples where the measured parameter was either below detection limit or saturated the measuring apparatus. In these cases the data values are set to the detection limit (zero if no detection limit was specified) or the saturation value respectively. If no flag value has been assigned (signifying good data), the CPFLAG field is blank.

The flag values that may be encountered are:

K	Uncertain/suspect value (source of quality control unknown).
L	Uncertain/suspect value (data originator's quality control).
M	Uncertain/suspect value (BODC quality control).
O	Uncertain/suspect value (user quality control)
<	Below detection limit.
>	In excess of stated value.

The originator's reference field allows the suppliers of individual data values to be identified. The objective when allocating these linkages is to provide a point of contact for users of the data to approach when initiating collaboration that will endure beyond the end of a project. Consequently, linkages have been assigned at the PI level and do not necessarily specify the individual who actually did the analysis. Codes are used to eliminate potential problems with misspellings and the like. The codes used are documented in the table ORGCODE.

Document references have not yet been implemented so the IDOCRF field is currently always set to null.

Table NETINDX
Field Definitions

BEN	NUMBER(6)	BODC event reference number
INET	NUMBER(6)	BODC net reference number
MINDEP	NUMBER(5,1)	Minimum depth of net haul
MAXDEP	NUMBER(5,1)	Maximum depth of net haul
GCODE	CHAR(8)	Gear code
MESH	NUMBER(4)	Mesh size of the net in microns
LATS	NUMBER(7,4)	Start latitude of tow segment (°N)
LONS	NUMBER(7,4)	Start longitude of tow segment (°E)
LATE	NUMBER(7,4)	End latitude of tow segment (°N)
LONE	NUMBER(7,4)	End longitude of tow segment (°E)

Notes
The position fields are included to allow multiple-sample net tows to be treated as a single event.

Table ORGCODE
Field Definitions

IORGRF	NUMBER(6)	Originator's reference code.
CORGNM	CHAR(20)	Originator's name.
CORGO	CHAR(40)	Originator's organisation.

Table SECCHI
Field Definitions

BEN	NUMBER(6)	BODC event number of the observation
CLOUD	NUMBER	Degree of cloud cover in eighths.
SUNALT	NUMBER	Altitude of the sun in degrees
SEASTATE	NUMBER	Sea state according to the Beaufort Scale
SECDEPTH	NUMBER	Maximum depth (metres) at which Secchi disk visible
SEACOLOUR	NUMBER	Observed colour of the sea
COMMENTS	CHAR(60)	Comments on weather/sea conditions
WI	NUMBER	Estimated wind speed in knots
SURFTEMP	NUMBER	Sea surface temperature (°C).
SHIPSPD	NUMBER	Ship's speed in knots

Notes
It should be noted that the Secchi disk depths were originally noted as feet and have subsequently been converted to metres.

Table SSINDX
Field Definitions

BEN	NUMBER(6)	BODC event number of the observation
BMPNTR	NUMBER(5)	Pointer to binary merge file
MINP	NUMBER(4,1)	Minimum pressure (db) of SeaSoar profile
MAXP	NUMBER(4,1)	Maximum pressure (db) of SeaSoar profile

Notes
The binary merge pointer refers to a cycle number in the underway data file. This is useful for calibration/ comparison purposes. Note that the actual profiles are given in table BINCTD.

Table WEATHER
Field Definitions

BBEGNS	DATE	Date of measurement
TEMP00	NUMBER	Temperature recorded at midnight (Celsius)
TEMP06	NUMBER	Temperature recorded at 06:00 (Celsius)
TEMP12	NUMBER	Temperature recorded at noon (Celsius)
TEMP18	NUMBER	Temperature recorded at 18:00 (Celsius)
MAXTEMP	NUMBER	Maximum temperature recorded (Celsius)
MINTEMP	NUMBER	Minimum temperature recorded (Celsius)
AVGTEMP	NUMBER	Average temperature recorded (Celsius)
SKY00	NUMBER	Extent of sky obscured at midnight (eighths)
SKY06	NUMBER	Extent of sky obscured at 06:00 (eighths)
SKY12	NUMBER	Extent of sky obscured at noon (eighths)
SKY18	NUMBER	Extent of sky obscured at 18:00 (eighths)
PRECIP06	NUMBER	Precipitation recorded at 06:00 (mm)
PRECIP18	NUMBER	Precipitation recorded at 18:00 (mm)
PRECIPALL	NUMBER	Total precipitation (mm)
WIND00	NUMBER	Wind speed (knots) recorded at midnight
WIND06	NUMBER	Wind speed (knots) recorded at 06:00
WIND12	NUMBER	Wind speed (knots) recorded at noon
WIND18	NUMBER	Wind speed (knots) recorded at 18:00

Notes
These measurements were obtained from Bergen airport to provide a meteorological context for the mesocosm experiment.

Table ZUCT

Field Definitions

CCTREF	CHAR(4)	Category code.
CCTFUL	CHAR(40)	Category description in plain language.
CILOAD	CHAR(6)	Date of record creation (yymmdd).
TCTMOD	DATE	Record modification time stamp.

Notes

The category codes are designed to group parameters into logical subgroups according to general operational practices. However, there will inevitably be parameters that could be fitted into more than one category depending upon one's point of view. This should be borne in mind when searching the dictionary. Always check out all possible categories.

Table ZUNT
Field Definitions

CPUREF	CHAR(4)	Unit code.
CPUABB	CHAR(10)	Abbreviated unit description.
CPUFUL	CHAR(40)	Full unit description.
CILOAD	CHAR(6)	Date record was created (yymmdd).
TPUMOD	DATE	Last modification time stamp.

Table ZUPM
Field Definitions

CPMCAT	CHAR(4)	Category code.
CPMREF	CHAR(4)	4-byte code for the parameter name.
CPMABB	CHAR(20)	Abbreviated parameter name.
CPMFUL	CHAR(80)	Full parameter name.
CPMUNT	CHAR(4)	Parameter storage unit code.
FABSNT	NUMBER	Absent data value.
FPMINM	NUMBER	Minimum value for parameter.
FPMAXM	NUMBER	Maximum value for parameter.
CINVER	CHAR(1)	Plot inversion flag.
CILOAD	CHAR(6)	Date of record creation (yymmdd).
TPMMOD	DATE	Date and time of last modification.

Notes
Most of the fields in this table are of more interest to BODC personnel than to database users. The exceptions are CPMCAT, CPMREF, CPMFUL and CPMUNT.

The category code (CPMCAT) provides the linkage to table ZUCT and hence identifies which generalised parameter descriptions belong to which category. CPMFUL contains the parameter description in plain language and provides the hook by which users can recognise just what is meant by a particular code.

The field CPMUNT specifies the units in which the parameter is stored in the database. This is present as a code (to prevent problems arising from differing descriptions being given to the same unit e.g. degrees, deg. and the like) which may be translated using table ZUNT.

Table ZUSG

Field Definitions

CPMREF	CHAR(4)	Parameter name code (bytes 1-4)
CSGREF	CHAR(2)	Parameter subgroup code (bytes 5-6).
CDSREF	CHAR(2)	Parameter discriminator code (bytes 7-8).
CPMUSG	CHAR(8)	Full 8-byte parameter code.
IPMBEF	NUMBER(1)	Number of digits before the decimal point.
IPMAFT	NUMBER(1)	Number of digits after the decimal point.
CSGABB	CHAR(20)	Abbreviated parameter code description.
CSGFUL	CHAR(100)	Full parameter code description.
CSGMTH	CHAR(100)	Methodology description.
ISGREF	NUMBER(8)	Narrative document reference.
CILOAD	CHAR(6)	Date record was created (yymmdd).
TSGMOD	DATE	Record modification time stamp.

Notes

The complete parameter code (CPMUSG) is constructed by concatenation of the parameter name, parameter subgroup and parameter discriminator codes.

The fields IPMBEF and IPMAFT are included to allow software to format data sensibly. Note that the data covered by the parameter codes are stored to a precision of some 16 decimal places. IPMAFT indicates how many of these have significance.

The meaning of a given code is specified in plain language by the fields CSGFUL and CSGMTH. These fields are designed to give a user-friendly reference to the full parameter code. If they don’t, please let us know. All the details which make the parameter unique (including filtration details where appropriate) are included.

The ISGREF field allows a linkage point for data documentation. It is designed to allow general methodology description documents to be linked to a parameter code. This on-line documentation is not currently implemented and the field is set null.

Database Linkage Definitions

The tables in this section of the document show the linkages that exist between the database tables. The linkages chains run along the rows of the table and always start with table EVENT. The type of linkage is shown by bolding the text. A linkage from normal text to bold text is a 'one to many' relationship. Links from normal text to normal text or bold text to bold text are 'one to one' relationships.

Drifting Buoy Data

EVENT	ARGOS
BEN	**BEN**

CTD Data

EVENT	CTDINDX	CTDCAL	CTDTYP	BINCTD
BEN	BEN	BEN	BEN	**BEN**
	TYPE		TYPE	

Water Sample Data (Fully Normalised)

EVENT	BOTTLE	BOTDATA	ZUSG	ORGCODE
BEN	**BEN**			
	IBTTLE	**IBTTLE**		
		IORGRF		IORGRF
		CPCODE	CPMUSG	

Net Haul Sample Data (Fully Normalised)

EVENT	BOTTLE	BOTDATA	ZUSG	ORGCODE
BEN	**BEN**			
	INET	**INET**		
		IORGRF		IORGRF
		CPCODE	CPMUSG	

N15 Production Data

EVENT	BOTTLE	N15DAT	N15HDR
BEN	**BEN**		
	IBTTLE	**IBTTLE**	
		EXPREF	EXPREF
BEN			BENCOL

Parameter Dictionary

ZUCT	ZUPM	ZUNT	ZUSG
CCTREF	**CPMCAT**		
	CPMUNT	CPMUNT	
	CPMREF		**CPMREF**

Primary Index

CRSINDX	EVENT	G_CODE
	BEN	
	GCODE	GCODE
CRUISE	**CRUISE**	

Fatty Acid Data

FATACID	EVENT	G_CODE
BEN	BEN	
	GCODE	GCODE
	CRUISE	

Multi-sizer Data

EVENT	BOTTLE	MULTI
BEN	**BEN**	
	IBTTLE	**IBTTLE**

Secchi Disk Data

SECCHI	EVENT	G_CODE
BEN	BEN	
	GCODE	GCODE
	CRUISE	

SeaSoar Data

EVENT	SSINDX	BINCTD
BEN	BEN	**BEN**