Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

Guide to the National Invertebrate Database (NID)

Cynthia Davies [1] & John Davy-Bowker [2]

[1] CEH Wallingford,
Maclean Building,
Benson Lane,
Crowmarsh Gifford,
Wallingford,
Oxfordshire, OX10 8BB
Tel: 01491 838800

[2] The Freshwater Biological Association,
East Stoke,
Wareham,
Dorset, BH20 6BB, UK

September 2009

Contents

1 Design and Structure of the NID

1.1 Introduction

The National Invertebrate Database is designed to store, retrieve and manipulate biological, environmental and chemical data collected for a variety of projects from sampling on rivers throughout the UK. The database is designed around the River Communities 'RIVPACS-style' invertebrate sampling methodology so the tables and columns reflect the types of data normally collected. The database was originally a fully-relational structure designed for an Oracle platform, which was then modified to run on Microsoft Access 97 with the data entry system on Microsoft Access 2.0 (Symes 1997). Subsequently, the structure was extensively redesigned by John Davy-Bowker (November 2006) and now operates in Microsoft Access 2002. This Manual describes the functioning of the current system.

The principal entities are Rivers, Sites, project site visits (termed Project_sites), Samples and invertebrate Taxa. The attribute data consist of geographical and physical attributes of sites as well as biological and chemical attributes of samples collected at each site. The physical site variables fall into two categories, time-invariant and time-variant data. Time-invariant data are collected only once during a site project visit in any year and are attributes of the entity Project_site. Time-variant data are measured during each sampling event for any one project and so are attributes of the entity Sample. The full database is populated from a number of different projects, some of which span several years.

The operation of the National Invertebrate Database is split into four sub-systems as follows:

- Data storage
- Data entry
- Frequently-asked question Queries and Reports
- Project Querying and Reporting

For security, the data are stored in a Microsoft Access 2002 database on a network drive which has restricted access. For practical purposes data input is carried out in separate Data Entry databases for each project. Once validated by the project leader, the new data are loaded into the main database by the database manager. After each update or addition of new data to the storage (Back-end) database, the data are copied to a series of satellite databases (Figure 1 The National Invertebrate Database collection): a user-friendly, form-driven Front end database application based on a series of Frequently Asked Questions, Data Entry databases, and project specific databases for more complex querying and reporting. The Front-end, Back-end and Project specific databases have identical core table structures. The Data Entry systems were designed originally with the same structure but are being modified to only include the data necessary for data input (August 2009). The input databases are archived following uploading of the new data.

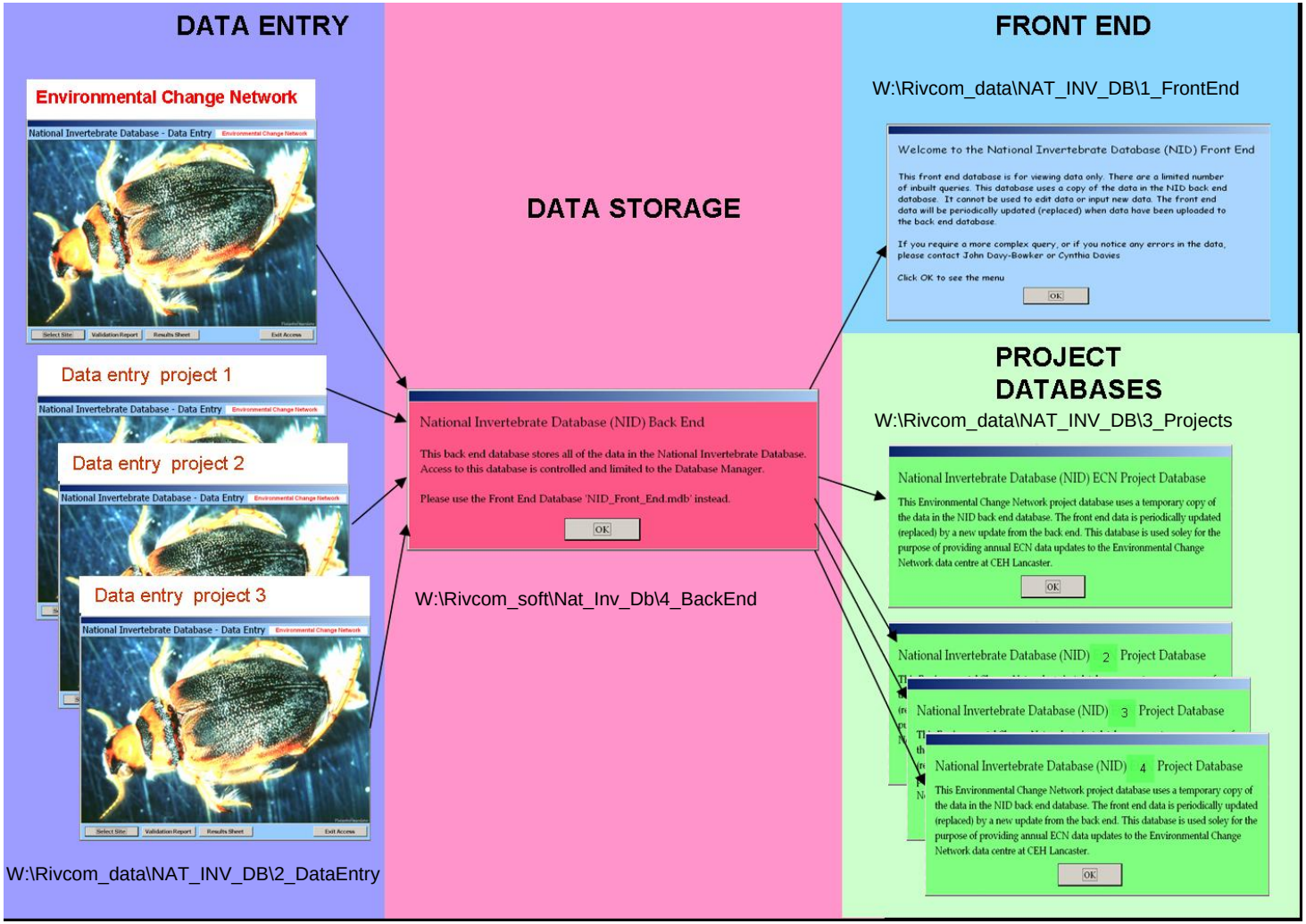


Figure 1 The National Invertebrate Database collection

1.2 Description of the data

A technical description of the tables and columns in the database is given in Appendix 3. The relationships between the tables are shown in Appendix 4, Entity Relationship Diagram of the NID.

1.2.1 Rivers

An hierarchical numbering system was devised so that sites can be located on a uniquely numbered watercourse (Symes, 1997). The hierarchical numbering allows all tributaries of any given watercourse to be derived from its number. A full description of the river numbering is given in Appendix 6
The table Dta_river stores a unique 6 digit river identifier (RIVER_ID), the hierarchical river number (comprising 8 fields) and a name for each watercourse. The table Dta_River also includes the facility to store the number and name of the relevant regional River Authority: Environment Agency (EA) region in England and Wales; former River Purification Board (RPB) region in Scotland; Island Authority (ISL) region which includes the outer Islands in Scotland and Northern Ireland; and Water Resource Region (IRE) for the Republic of Ireland[1]

1.2.2 Sites

Each sampling site is located on only one river, but a river may have many sampling sites. The sites are numbered (Site_ID) but this number is not unique so the combination of River_ID and Site_ID is needed to provide a unique site reference in table Dta_Site. The table holds time invariant attributes of sites including: site name (usually the nearest village or named land feature in the map); easting and northing (Ordnance Survey GB National Grid Reference in digits[2]); discharge category; distance from source (km); altitude a.s.l. (m); slope (m/km); Strahler stream order; catchment area; Country and County names. The table also has an attribute called SiteCode which in the earlier Oracle version of the database was the combination of River_ID and site_ID but the field is currently used by project leaders to retain project-specific site codes. Historically, sites were sequentially numbered along a river system regardless of the project; this system rapidly broke down as there were too few available numbers to retain the sequence. Sites may be numbered differently for each project, regardless of whether there is an existing site at the same location for a different project. This is because site lengths may be different and area samples may be different. The project leader should determine whether to use existing site codes or to generate a new site number. If possible, sites within a project are numbered in an upstream direction. Sites may be overlaid in a GIS using the eastings and northings to identify overlapping sample sites.

[1] A watercourse may lie in more than one authority region and some parts of the river system may lie in a different region than the primary river. The information is not given for many of the rivers. Historically, these data were held in a separate table, but the structure has been partially de-normalised to facilitate querying

[2] Although the database can store eastings and northings up to 1m resolution, many grid references were only recorded at 100m resolution and the remaining spaces are padded with zeros. There is no flag for the level of resolution of the record.

1.2.3 Projects

A project is a coherent study linked by a common set of objectives. Most projects are commissioned studies but some may be funded from NERC Science budget. Project names and an internal 6-character Project_ID code are stored within the table Lkp_Projects. The table also includes the CEH/IFE/FBA project code, the name of the project leader(s) and details of the Intellectual Property rights to that dataset. The start and end year of data collection under that Project_ID are also included. The UKEDI or Information Gateway or other catalogue reference number is provided to enable cross-reference to further metadata in CEH data catalogues.

1.2.4 Project Sites

A site may be visited a number of times for different projects or for the same project in different years. In table Dta_Project_Site, each project_site visit is identified by the combination of the site identifier and the project name and project year. The time-dependent physical variables measured for each site are stored as attributes and include: survey date[3]; maximum, minimum and mean channel width; depth category; bank stability; dominant bankside vegetation types; shading category; dominant land-use on the banks; and anthropogenic influences on the survey area.

1.2.5 Samples

Samples are collected during project_site visits and the details of the specific sampling occasion are held in table Dta_Sample. The site may be visited more than once in a project_year (e.g. spring, summer and autumn samples) and one or more samples may be taken at each visit. Each Project_Site record can therefore have more than one sample associated with it, but each sample can only be associated with one project_site visit. Samples are numbered (Sample_ID) but this number is not unique so the combination of the Site_ID, Sample_ID, and season number are needed to obtain a unique row reference[4]. The time-variant physical site characteristics that are measured when and where the samples are collected are stored as attributes of each sample. These include: water width; water depth at ¼, ½ and ¾ of the full water width; surface velocity category; clarity; composition of the substrate as percentages of rock pavement, boulders/cobbles, pebbles/gravel, sand and silt/clay; dominant particle size; macrophyte cover and presence of detritus. Other attributes include subsidiary number (used for storing sample code from earlier versions of the database or any other sample numbers of use to a given project) and codes for the person or organisation which collected the sample, who identified it and the level to which it was identified.

[3] This is the date the channel variables were measured. The date may be different for different samples in the same year.

[4] The table Dta_Sample does not use Project and Project_year from the table Dta_Project_site as part of the primary key. This forces restrictions on the data: the sample number must be unique for any given river-site-season combination regardless of the project. It also means that for an existing site the next available sample number for that site and season must be identified before samples are numbered from any new sampling trip to that site.

1.2.6 Species data

Individual taxa in the samples are recorded using an 8-character code known as the Furse Code (Furse et al. (Version 4.1)). The code was developed from the Maitland coding system (Maitland, P S. 1977) and enables a taxonomic list to be produced according to a standard identification level. The full list of the hierarchical codes is held in the look-up table Lkp_Taxa.

The presence of a taxon in a sample can be recorded according to a number of different abundance systems. These systems are defined in the look-up table Lkp_Abundance System. Each abundance system is uniquely identified by an Abundance_System_ID of 2 – 5 characters. Methods of storing abundance include IFE Log Abundance, Absolute Abundance, and Presence/Absence (i.e. no abundance measure recorded). It is possible to store data for a particular sample using any combination of abundance systems.

A sample may contain zero or many taxa and each taxon may occur in zero or many samples. Also the same sample may be processed at more than one hierarchical level using different abundance systems. For example, taxa may be identified and enumerated at the family level (e.g. IFE Log Abundance or absolute abundance) while a particular species of a family may be recorded as being Present with no abundance measure. Table Dta_Taxa allows for this double entry system. Each record is uniquely identified by the combination of the sample identifiers from the sample table (Dta_Sample), the Species_code from the table Lkp_Taxa, and the abundance-system_id from the table Lkp_Abundance System. Included in the record is an attribute for the abundance recorded (for records stored under the Presence/Absence abundance system, this attribute is left blank) and whether the sample was identified to species or family level. As it is not always possible to identify individuals further than to family or genus level even when the remainder of the sample has been recorded to species, the attribute Type enables all taxa in a sample to be included in the analysis of a sample record at a particular taxonomic level regardless of the final level of identification. This is useful where the double entry system has been employed.

The procedures and level of identification for analysis of samples are defined by the project and should be documented in the Project Data Management Plan.

1.2.7 The taxonomic table

There are six taxonomic levels encoded within the 8 digit Furse Species Code. However the actual names of the levels vary within the first three levels, so generic names have been given to these parts of the code. The code breaks down as follows:

Digit	Name of column	Taxonomic names represented
1-2	Major_Gp_Name	Phylum, Order, Class
3	Level2_Name	Superclass, Class, Sub-Class, Super-family, Family
4	Level3_Name	Family, Sub-family, Tribe
5-6 Gen	us_Name	Genus
7 Su	bgenus_Name	Sub-Genus[5]
8 Species	_Name	Species

Each taxon has a Scientific name, common name and other attributes associated with it. Taxa recorded at the species level also have the attribute species authority (including date).

In earlier versions of the National Invertebrate Database, each taxon level was modelled as a separate entity within the database. However, in the current version, all levels are held within a single table Lkp_Taxa. Appendix 5 Examples of Records in the Taxonomic Table shows an extract from the table Lkp_Taxa and a list of the names for each of the selected codes.

1.2.8 BMWP Families

In order to facilitate the calculation of BMWP scores for samples, the table Lkp_Taxa_List_BMWP_Families stores the taxon codes, family names and scores for the families nominated by the Biological Monitoring Working Party (Hawkes (1997) [6,7]

The table Lkp_Crossref provides a means of linking the main taxa table (Lkp_Taxa) to Lkp_Taxa_List_BMWP_Families so that all taxonomic records in the database can be standardised to the BMWP list of families.

1.2.9 AWIC Families

In order to facilitate the calculation of AWIC scores for samples, the table Lkp_Taxa_List_AWIC_F stores the taxon codes, family names and scores for the list of families which contribute to the Acid Waters Indicator Community (Davy_Bowker et al (2005)). [7]

[5] The Level2 (3rd digit) and sub-genus (7th digit) levels are complicated by the fact that they are not used by all species codes to represent these levels. For example, if a family is not broken down into sub-families or tribes, then both the 3rd and 4th digit may be used to number that family. Similarly, if there are no sub-genera within a genus then both the 7th and 8th digits may be used for the species number, in which case there will be no sub-genus name and the first 7 digits of the full taxon code will not be significant on their own (and the attribute Subgenus_name will be left blank). Where the 3rd or 7th digits are not being used to identify a particular taxonomic level (e.g. sub-family or sub-genus), then the 4th or 8th character positions may not be zero if the 3rd or 7th positions are not zero themselves.

[6] Origin and development of the Biological Monitoring Working Party Score System. Wat Res. vol 32 no 3. Table 2 -1980 (nb, Clambidae, Curculionidae & Chrysomelidae were dropped).

[7] Separate tables are used for the BMWP and AWIC families rather than storing the scores as attributes within the taxonomic table, firstly because the list of families is a small subset of the family codes and secondly because the BMWP list also includes some 'artificial' taxa created to accommodate changes in current taxonomy for example where taxa have been split since the BMWP system was devised. There is not a 1:1 mapping between taxa in the main taxonomic table and the BMWP list

The table Lkp_Crossref provides a means of linking the main taxa table (Lkp_Taxa) to Lkp_Taxa_List_AWIC_F so that all taxonomic records in the database can be standardised to the AWIC list of families.

1.2.10 Chemical Data

Physical and chemical determinands measured from water samples taken from sites are held in the table Dta_Chemistry. The table has the attributes of date of sampling, time of sampling, determinand code and absolute measurement for that sample. The combination of the determinand identifier and the identifier of the site (in table Dta_Site) with which it is associated uniquely identifies each record. The table also includes an attribute for the sample code used by the originator organisation where samples were analysed outwith CEH. Each determinand has a unique number identifier. The numbers, names and units of measurement for each determinand are held in the table Lkp_Determinand. Chemistry data are no longer collected as part of the NID but are available for certain rivers and sites and are held elsewhere in CEH; data in the table relate to samples collected between 1974 and 1999.

1.2.11 Project Metadata

Metadata for each batch of data for each project are held in the table Mtd_Metadata. The table has the attributes of project identifier, project manager, identification level and abundance systems used as well as information about the progress of data entry and validation. (Project level metadata are held in the table Lkp_Projects)

2 User Interfaces via Microsoft Access

2.1 Data Entry System

The following section provides a description of the preparation required prior to data entry and an outline of the procedures to be followed during data input. It also provides a description of the functionality of the Data Entry System.

2.1.1 Preparation of Raw Data

The most time consuming aspect of data entry process is that of data preparation, in particular sample coding. This is because the database requires unique identifiers for each river, site, project_site and sample. It is therefore important to ensure that there is no duplication of codes.

The first stage is to complete a list of sites for which there are samples. The form 'SITE LIST FOR DATA ENTRY' should be used to compile the list. (Appendix 7)

When coding samples there are several situations that may occur with respect to sample coding:

1) A new sample for an existing project is taken at a site which has been previously sampled (existing project, existing river, existing site, new sample)
2) The sample is taken on a river which is coded in the database but which currently has no site for the current project (existing project, existing river, new site, new sample)
3) The sample is taken for an existing project on a river that is not yet in the database (existing project, new river, new site, new sample)
4) The sample has been taken for a new project on a river which is coded in the database (existing river, new project, new site, new sample)
5) The sample has been taken for a new project and a new river (new project, new river, new site, new sample)

The following procedure aims to address each of these situations.

a) *Allocate data to a Project*
The list of existing projects can be obtained by selecting 'See Project metadata' in the Front End FAQ database. If data are to be entered under a new Project it is necessary to provide a Project Name and a unique identifier which uses 6 characters. The Project code, PI and IPR details must also be recorded for inclusion in the table Lkp_Projects . In addition, projects should be fully catalogued in a recognised data catalogue (currently the CEH Information Gateway, http://gateway.ceh.ac.uk, previously UKEDI) and the catalogue number recorded. The database manager should add the information to the main database.

b) *Identify river and site codes for an existing Project.*
Record the Project_ID on the 'SITE LIST FOR DATA ENTRY' form. A list of existing sites on the river should be extracted from the database. Using the Front End FAQ database, select the option, 'Sites and river codes for project' on the

main menu (Figure 2). The codes can be viewed or downloaded to Excel and entered onto the data preparation forms. Possible new sites can then be identified for numbering.
Samples are numbered using the next available number for the combination of River_ID, Site_ID. The Season_ID forms part of the unique key and each sample must therefore have a unique combination of River_ID, Site_ID, Sample_ID and Season_ID.

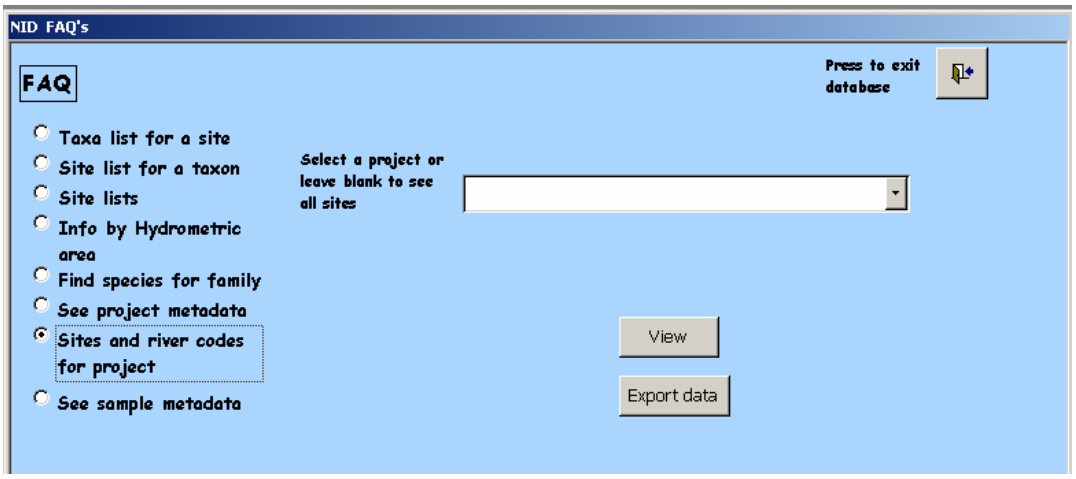


Figure 2

The Database Manager must add the new River and Site details to the Back-End database before creating the Project Data Entry System. Further information about the batch of samples can be added to the table Mtd_Metadata when data entry commences.

c) *Allocate a new site code for an existing river.*
If the site does not already exist in the database, a new site number must be allocated. Sites may be numbered differently for each project, regardless of whether there is an existing site at the same location for a different project. A list of river_ID's and codes for a project can be obtained from the Front End FAQ database. Using the database, select the option, Sites and river codes for project' on the main menu (Figure 2). A list of all existing sites in the database can be seen by selecting the option 'Site lists' and leaving the project selection box blank. The codes can be viewed or downloaded to Excel. The new sites can then be coded taking into account existing sites on that river. The project leader should determine whether to use existing site codes or to generate a new site number.

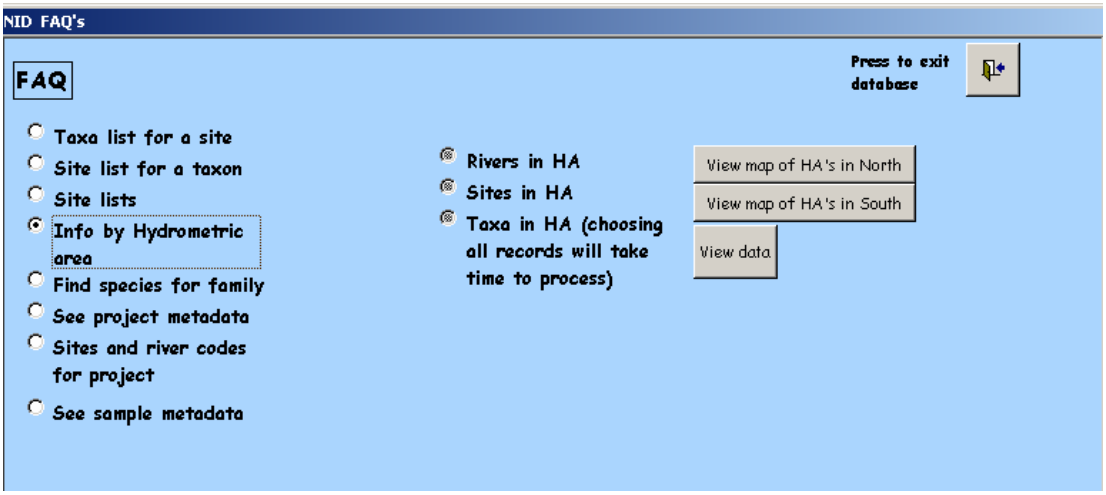


Figure 3

To code new sites, locate the site on an OS 1:50,000 map using the National Grid Reference. Locate the river name; if the site is on an un-named river, then make a note of the river into which it flows. If the site is on an existing river, note the River_ID. Locate existing sites numbers on the watercourse and insert the new site number in the appropriate place. If possible, sites are numbered in an upstream direction. If there are no available unallocated numbers to slot into the sequence, then the next available site number for that river can be used. The combination of River_ID and Site_ID **must** be unique.

Hint: The Intelligent River Network (IRN) can be used to either locate the appropriate map number or to view the location on-line on Streetmap.co.uk. To find the appropriate map number open the IRN and select the Useful Tools Enter the appropriate National Grid Reference (Figure 4).

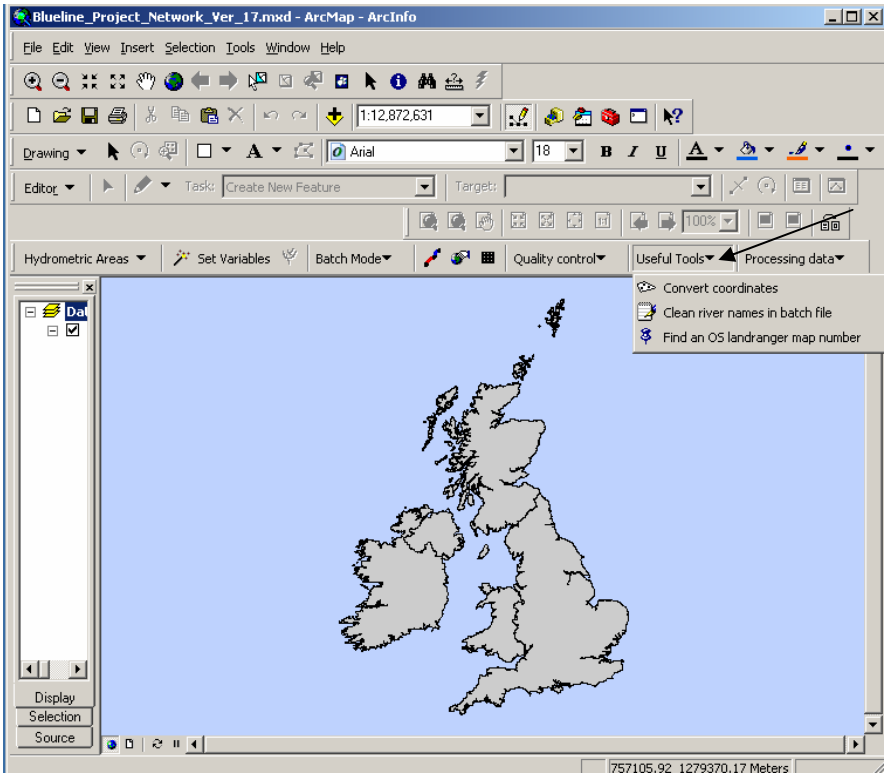


Figure 4

The location can be viewed on-line by selecting Go to Web Site (Figure 5).

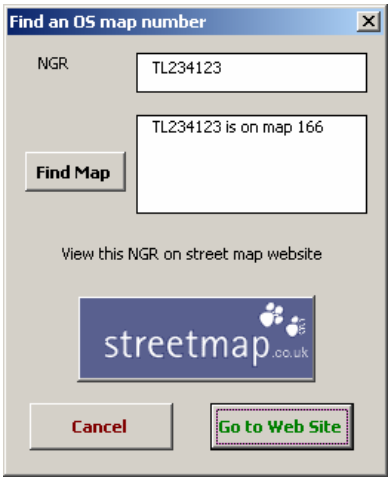


Figure 5

If required, the IRN can be used to generate the attributes such as Distance from source, Altitude etc. (Distance from source generated by the IRN should be checked by generating a scatter plot of Distance from Source against Discharge Category. The relationship should be tight, and can flag up where points have snapped to the wrong point on the system or there is an error in the line work in the IRN.)

d) *Allocate a site code on a new river.*
If a river is not included in the master list, it should be numbered according to the information in Appendix 6. A paper form has been designed in order to assist with river numbering - 'DETAILS OF RIVER NUMBERING' - as shown in Appendix 7. The database manager can then add the new rivers to the database.

The form can be filled in as follows:-

i) Find out which Hydrometric Area the river flows through (see map in Appendix 8) or use IRN and select Hydrometric Areas, Load and display by hydrometric area and Load by NGR (Figure 6). The appropriate hydrometric area is loaded and the number should be transferred to the form.

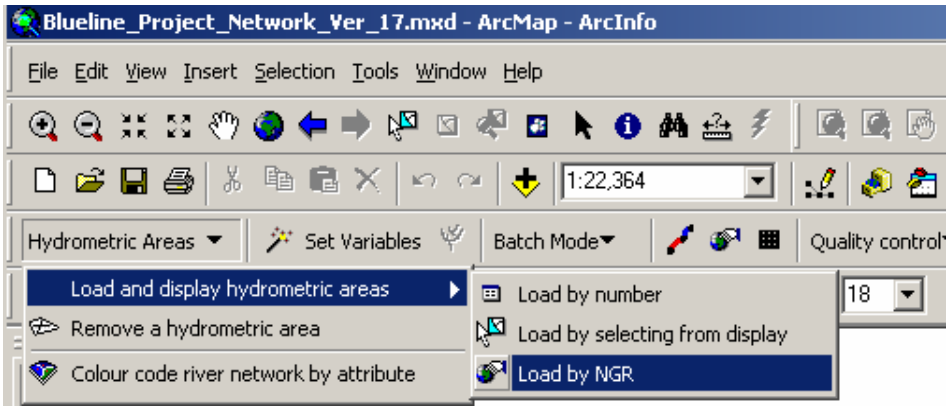


Figure 6

ii) Using the Front End database FAQ 'Info by Hydrometric Area' and 'Rivers in HA' export the query showing rivers in the hydrometric area ordered by their full code (see Table 1. Appendix 6)
iii) Using this spreadsheet, find the maximum River_ID used in the database for the hydrometric area in question
iv) Allocate River_ID by using the maximum River_ID for the appropriate Hydrometric Area and increment by 1 for each new river within that hydrometric area.
v) Highlight all "Primary Rivers" (those which have an outlet to the sea).
vi) Number the new rivers according to the procedure in Appendix 6.
vii) Once validated, the database manager adds the new rivers and codes to the main database.
viii) Number the new sites as in c) above.

d) *The sample has been taken for a new project and a new river*
If data are to be entered under a new project it is necessary to provide a Project description and a unique identifier for the project of up to 6 characters. The information is added to the table Lkp_Projects.

2.1.2 Creation of Data Entry Databases

In the Back-end database, the Database Manager (DM) adds the new project to Lkp_PROJECTS table if necessary then enters any new River and Site details into the main tables Dta_River and Dta_Site. The full site list for the project is appended to the table Dta_SITE_Data_Entry. The Project Leader must decide what level of taxonomic identification is appropriate for the project (species level abundance, dual entry etc.) and if required, a new entry should be made to the Look-up table, Lkp_ID_LEVEL. Any new staff name and initials should be added to table Lkp_STAFF.

The DM creates a copy of an existing Data Entry database (NB data entry systems are then modified for each individual project).

The four data entry tables are emptied in sequence (Dta_Taxa_Data_Entry, Dta_SAMPLE_Data_Entry, Dta_PROJECT_SITE_Data_Entry and Dta_SITE_Data_Entry) then tables Dta_RIVER, Dta_Site_Data_Entry and all lookup tables are refreshed from the main back-end database using the macro _mcr_UPDATE_TABLES. The macro runs the code mdl_Refresh_tables (Appendix 10)

The DM may add in details of new samples if this appropriate for a specific project.

The forms and reports are then modified to meet the requirements of the project.

Forms:

i) Frm_001_Main_Menu: Alter form caption and header label to indicate the project
ii) Frm_003_Select_Project_Site: Amend the default value in the PROJECT text box to the appropriate project.

iii) Frm_004_Select_Sample: Amend text help in labels for Sampling method and time, ID level and Subsidiary number (the project leader will have decided on the relevant coding in advance).

iv) Frm_005_Select_Abundance_Type: Amend code in combo box Fam_Sp on After Update event so that values are appropriate (especially where 2 different types are needed for family and species level enumeration). For example:

v)

```
Private Sub Fam_Sp_AfterUpdate()
If Fam_Sp.Value = "F" Then
Selected_Abundance_Type.Value = "Count"
Combo20.Value = "F"
Else
If Fam_Sp.Value = "s" Then
Selected_Abundance_Type.Value = "Count"
Combo20.Value = "S"
End If
End If
btnSurvey.SetFocus
End Sub
```

On combo box Selected_Abundance_Type alter Row Source sql to select desired abundance system from table Lkp_ABUNDANCE SYSTEM (or set appropriate default value if only one option).

Reports:

i) Rpt_Results: Amend title and column headers

ii) Rpt_Validation: Amend title and column headers

A checklist of actions required is provided in Appendix 9

2.1.3 The Main Menu Form

The Data entry clerk can then begin data entry. When the data entry database is opened the user is presented with the main form (Figure 7):

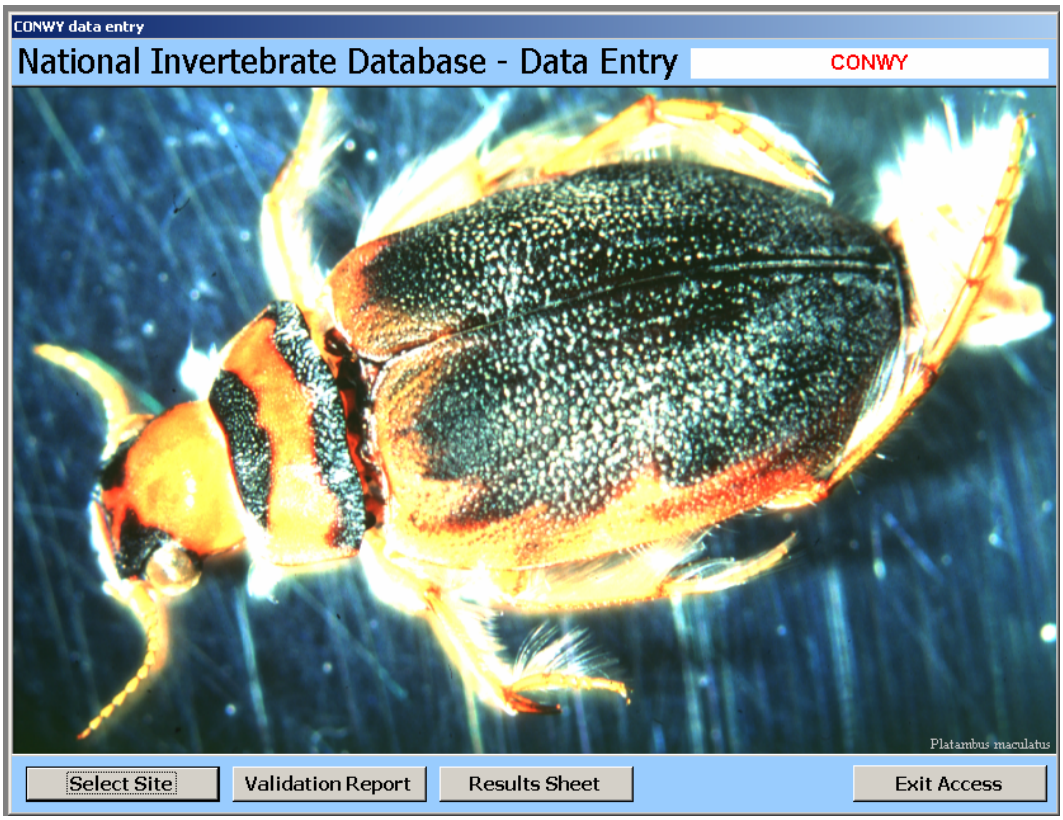


Figure 7

Click on Select Site to open form Frm_002_Select_Site (Figure 8)

2.1.4 Site Details Entry Form

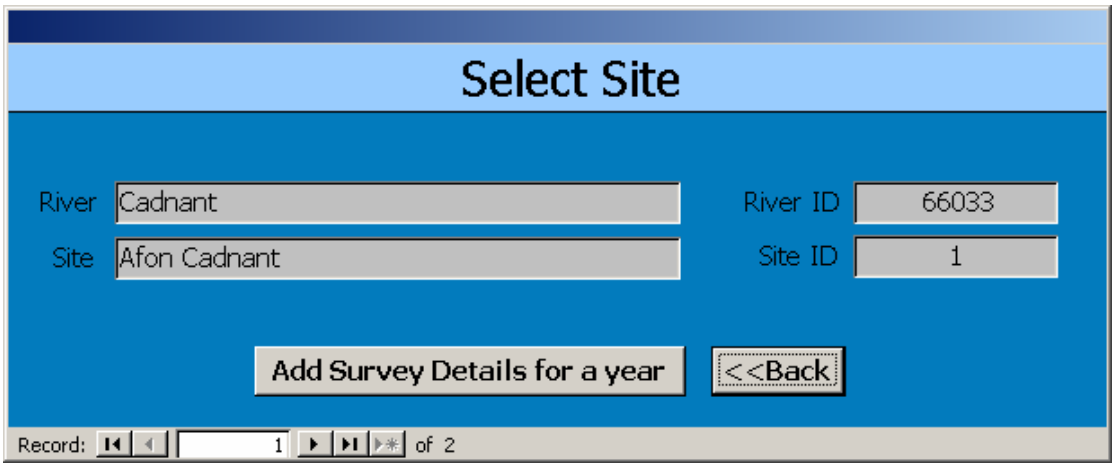


Figure 8

The form is based on Qry_Underlying_Frm_002_Select_Site. Only sites previously entered into the data entry site table can be selected.

```
SELECT Dta_SITE_Data_Entry.RIVER_ID, Dta_SITE_Data_Entry.SITE_ID, Dta_RIVER.RIVER_NAME,
Dta_SITE_Data_Entry.SITENAME, Dta_SITE_Data_Entry.EASTING, Dta_SITE_Data_Entry.NORTHING
FROM Dta_SITE_Data_Entry INNER JOIN Dta_RIVER ON Dta_SITE_Data_Entry.RIVER_ID =
Dta_RIVER.RIVER_ID
ORDER BY Dta_RIVER.RIVER_NAME, Dta_SITE_Data_Entry.SITENAME
WITH OWNERACCESS OPTION;
```

Next click on Add Survey Details to add survey details for that site for a given year to open Frm_003_Select_Project_Site (Figure 9)

2.1.5 Project Site Details Entry Form

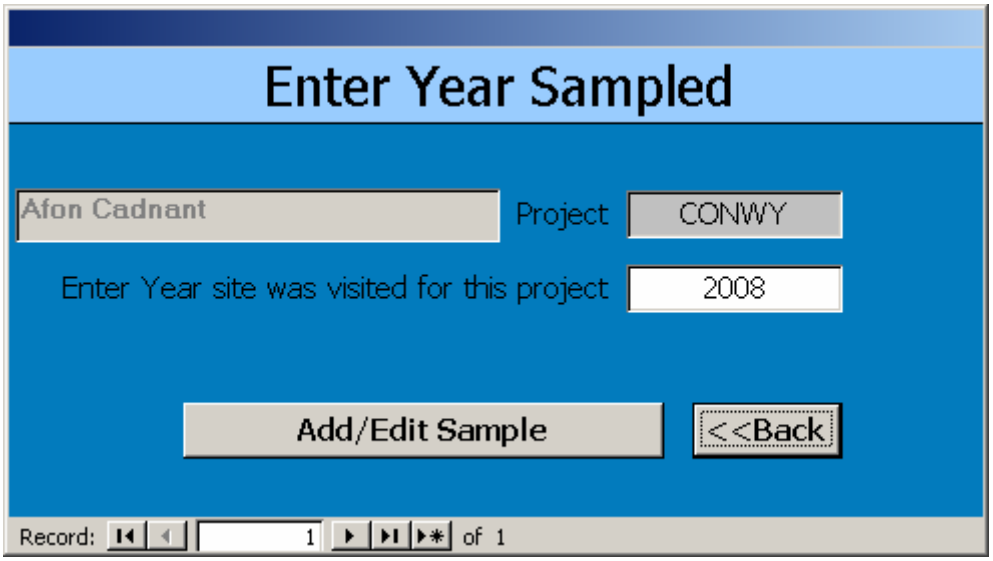


Figure 9

The form is based on Qry_Underlying_Frm_003_Select_Project_Site.

```
SELECT Dta_PROJECT_SITE_Data_Entry.RIVER_ID, Dta_PROJECT_SITE_Data_Entry.SITE_ID,
Dta_PROJECT_SITE_Data_Entry.PROJECT, Dta_PROJECT_SITE_Data_Entry.PROJECT_YR
FROM Dta_PROJECT_SITE_Data_Entry
WHERE (((Dta_PROJECT_SITE_Data_Entry.RIVER_ID)=[Forms]![Frm_002_Select_Site]![RIVER_ID]) AND
((Dta_PROJECT_SITE_Data_Entry.SITE_ID)=[Forms]![Frm_002_Select_Site]![SITE_ID]))
WITH OWNERACCESS OPTION;
```

Enter year samples were taken. One entry is required for each year the site was sampled (the data are stored in the table Dta_Project_Site_Data_Entry). **If coming back to data entry, the user must ensure the correct site and year and sample are selected!**

Click on Add/Edit Sample to open Frm_004_Select_Sample and begin data entry (Figure 10)

2.1.6 Sample Details Entry Form

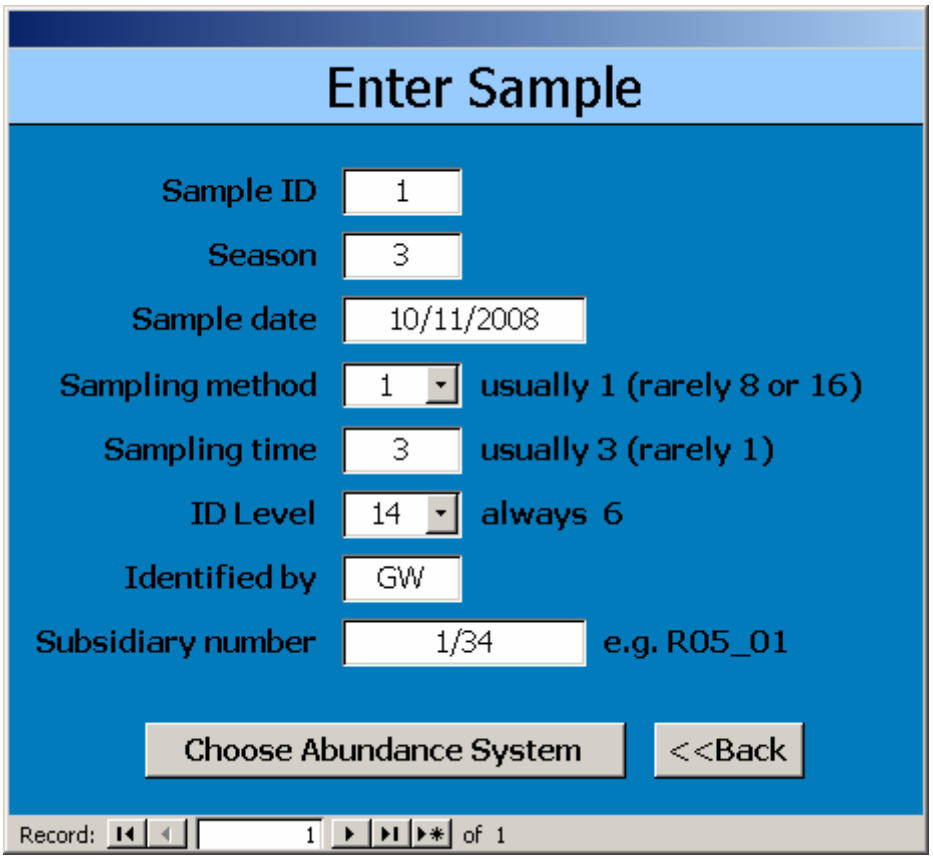


Figure 10

The form is based on Qry_Underlying_Frm_004_Select_Sample.

```
SELECT Dta_SAMPLE_Data_Entry.RIVER_ID, Dta_SAMPLE_Data_Entry.SITE_ID,
Dta_SAMPLE_Data_Entry.PROJECT, Dta_SAMPLE_Data_Entry.PROJECT_YR,
Dta_SAMPLE_Data_Entry.SAMPLE_ID, Dta_SAMPLE_Data_Entry.SEASON_ID,
Dta_SAMPLE_Data_Entry.SAMPLE_DATE, Dta_SAMPLE_Data_Entry.METHOD,
Dta_SAMPLE_Data_Entry.SAMPLING_TIME, Dta_SAMPLE_Data_Entry.SUBSID_NO,
Dta_SAMPLE_Data_Entry.IDENTIFIED_BY, Dta_SAMPLE_Data_Entry.ID_LEVEL
FROM Dta_SAMPLE_Data_Entry
WHERE (((Dta_SAMPLE_Data_Entry.RIVER_ID)=[Forms]![Frm_003_Select_Project_Site]![RIVER_ID]) AND
((Dta_SAMPLE_Data_Entry.SITE_ID)=[Forms]![Frm_003_Select_Project_Site]![SITE_ID]) AND
((Dta_SAMPLE_Data_Entry.PROJECT)=[Forms]![Frm_003_Select_Project_Site]![PROJECT]) AND
((Dta_SAMPLE_Data_Entry.PROJECT_YR)=[Forms]![Frm_003_Select_Project_Site]![PROJECT_YR]))
WITH OWNERACCESS OPTION;
```

Sample details can then be entered. It is possible to have more than one sample for a site in a given year. A new record must be created for each sample. The form will have been modified to give specific guidance on codes used for the particular project.

Samples are numbered according to previously entered numbers in order to avoid a clash (see Section 1.2.5). Samples should be numbered incrementally within a given project. Sample numbering should be pre-determined by the project leader together with the Data Manager. (NB Any sample numbers on the lab sheet should be entered into the field Subsidiary number - see below.)

Enter the correct Season (as in RIVPACS protocol: 1 = Spring (March-May), 2 = Summer (June – August), 3 = Autumn (September –November). There is no Winter

code. If samples were taken in winter, December and January are treated as autumn and February as spring. The only valid values are 1, 2 and 3.)
The data should be entered in the format DD/MM/YYYY.

Sampling method is a drop-down list. For example, Kick/Sweep = RIVPACS sampling method – 3 minute kick plus optional 1 min search.

Enter Sampling time in minutes (for RIVPACS method = 3; use 1 if less than 1 minute).

Select the correct ID level from the pick list. **This must have been agreed** before any processing starts and the correct version of the lab sheet must be used.

The initials of the member of staff carrying out the identification should be input.

The Subsid number field is used as designated by the project. This may be the number on the lab sheet. For example for the Conwy project it would be Site number plus chemical site number or for ECN, the number would be in the format R05_01.

An alternative Sample Details Data Entry form has been created to allow the entry of additional Sample variables from the field sample sheet.

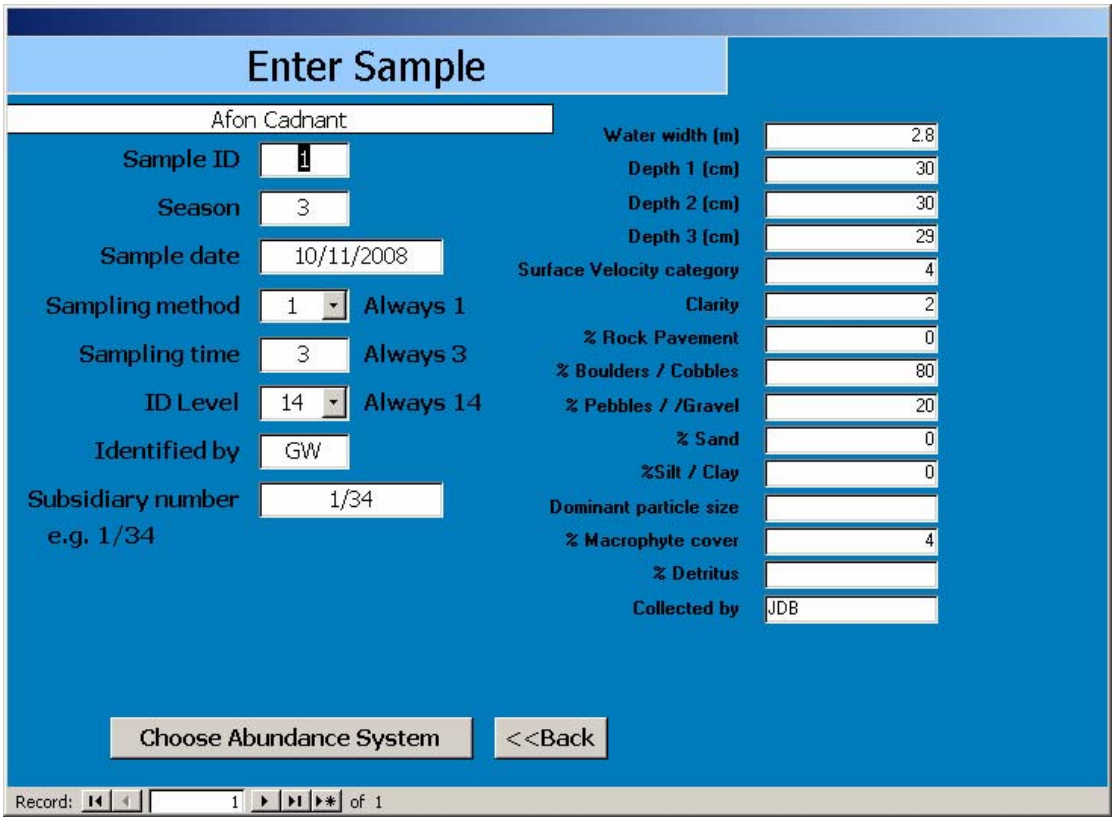


Figure 11

This form is based on Qry_Underlying_Frm_004i_Select_Sample

```
SELECT Dta_SAMPLE_Data_Entry.RIVER_ID, Dta_SAMPLE_Data_Entry.SITE_ID,
Dta_SAMPLE_Data_Entry.PROJECT, Dta_SAMPLE_Data_Entry.PROJECT_YR, Dta_SITE_Data_Entry.SITENAME,
Dta_SAMPLE_Data_Entry.SAMPLE_ID, Dta_SAMPLE_Data_Entry.SEASON_ID,
Dta_SAMPLE_Data_Entry.SAMPLE_DATE, Dta_SAMPLE_Data_Entry.METHOD,
Dta_SAMPLE_Data_Entry.SAMPLING_TIME, Dta_SAMPLE_Data_Entry.SUBSID_NO,
Dta_SAMPLE_Data_Entry.IDENTIFIED_BY, Dta_SAMPLE_Data_Entry.ID_LEVEL,
Dta_SAMPLE_Data_Entry.WATER_WIDTH, Dta_SAMPLE_Data_Entry.DEPTH1,
Dta_SAMPLE_Data_Entry.DEPTH2, Dta_SAMPLE_Data_Entry.DEPTH3,
Dta_SAMPLE_Data_Entry.SURF_VEL_CAT, Dta_SAMPLE_Data_Entry.CLARITY,
Dta_SAMPLE_Data_Entry.ROCK_PAVEMT, Dta_SAMPLE_Data_Entry.BOULDER_COBBLE,
Dta_SAMPLE_Data_Entry.PEBBLE_GRAVEL, Dta_SAMPLE_Data_Entry.SAND,
Dta_SAMPLE_Data_Entry.SILT_CLAY, Dta_SAMPLE_Data_Entry.DOM_PART_SIZE,
Dta_SAMPLE_Data_Entry.MACROPH_COVER, Dta_SAMPLE_Data_Entry.DETRITUS,
Dta_SAMPLE_Data_Entry.COLLECTED_BY
FROM Dta_SITE_Data_Entry INNER JOIN (Dta_PROJECT_SITE_Data_Entry INNER JOIN
Dta_SAMPLE_Data_Entry ON (Dta_PROJECT_SITE_Data_Entry.PROJECT_YR =
Dta_SAMPLE_Data_Entry.PROJECT_YR) AND (Dta_PROJECT_SITE_Data_Entry.PROJECT =
Dta_SAMPLE_Data_Entry.PROJECT) AND (Dta_PROJECT_SITE_Data_Entry.SITE_ID =
Dta_SAMPLE_Data_Entry.SITE_ID) AND (Dta_PROJECT_SITE_Data_Entry.RIVER_ID =
Dta_SAMPLE_Data_Entry.RIVER_ID)) ON (Dta_SITE_Data_Entry.SITE_ID =
Dta_PROJECT_SITE_Data_Entry.SITE_ID) AND (Dta_SITE_Data_Entry.RIVER_ID =
Dta_PROJECT_SITE_Data_Entry.RIVER_ID)
WHERE (((Dta_SAMPLE_Data_Entry.RIVER_ID)=[Forms]![Frm_003_Select_Project_Site]![RIVER_ID]) AND
((Dta_SAMPLE_Data_Entry.SITE_ID)=[Forms]![Frm_003_Select_Project_Site]![SITE_ID]) AND
((Dta_SAMPLE_Data_Entry.PROJECT)=[Forms]![Frm_003_Select_Project_Site]![PROJECT]) AND
((Dta_SAMPLE_Data_Entry.PROJECT_YR)=[Forms]![Frm_003_Select_Project_Site]![PROJECT_YR]) AND
((Dta_SITE_Data_Entry.SITENAME)=[Forms]![Frm_003_Select_Project_Site]![SITENAME]))
WITH OWNERACCESS OPTION;
```

Click on Choose Abundance System to open Frm_005_Select_Abundance_Type (Figure 12).

2.1.7 Abundance System Details Entry Form

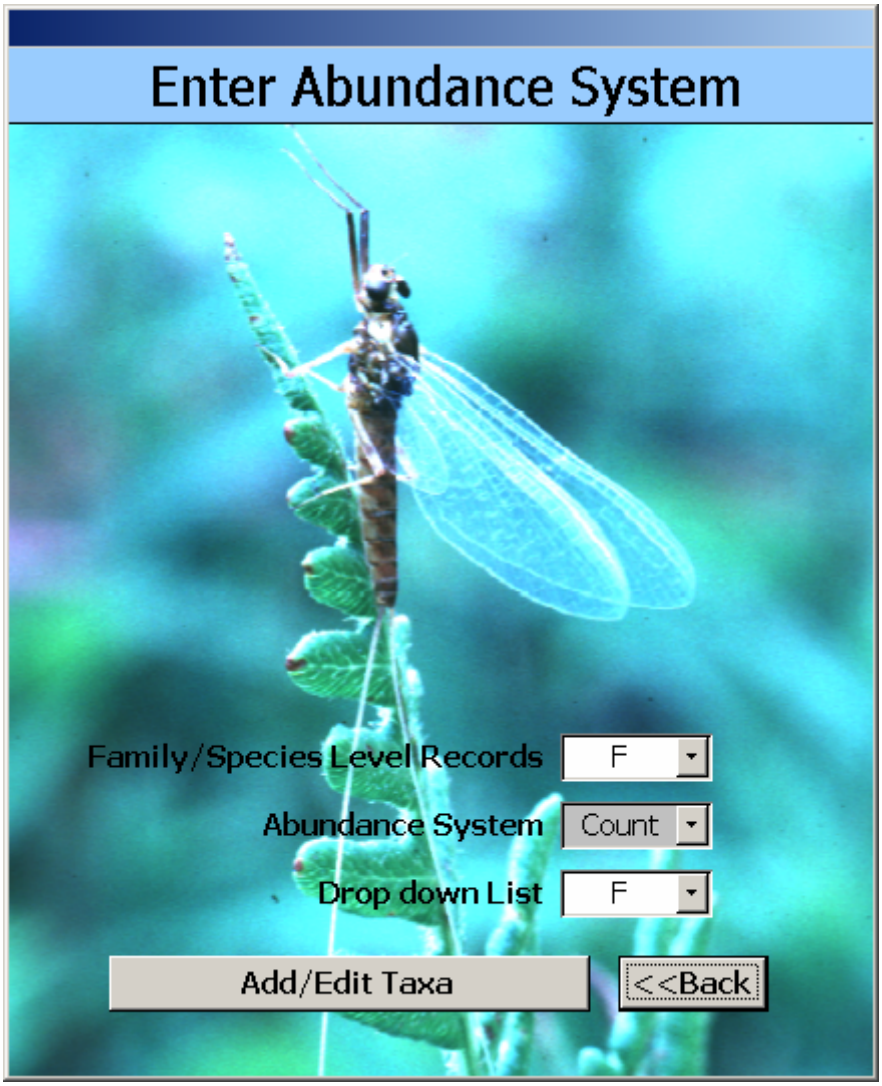


Figure 12

This form is not based on a query. It is used to set values to the defaults. The user must select either family or species level records before entering the data as some projects are analysed at both levels with different abundance measures (e.g. Count at Family level but Presence/Absence for species records). (NB Where it is not possible to identify a particular specimen further than to family in a Species level data set, it is still possible to enter the family name.)

Abundance System is based on an sql statement selecting appropriate data from Lkp_Abundance_System. Only values pre-selected for the project will be available.

Click on Add/Edit Taxa to begin entry of taxonomic data. This opens Frm_006_Select_Taxa (Figure 13).

2.1.8 Taxonomic Data Entry Form

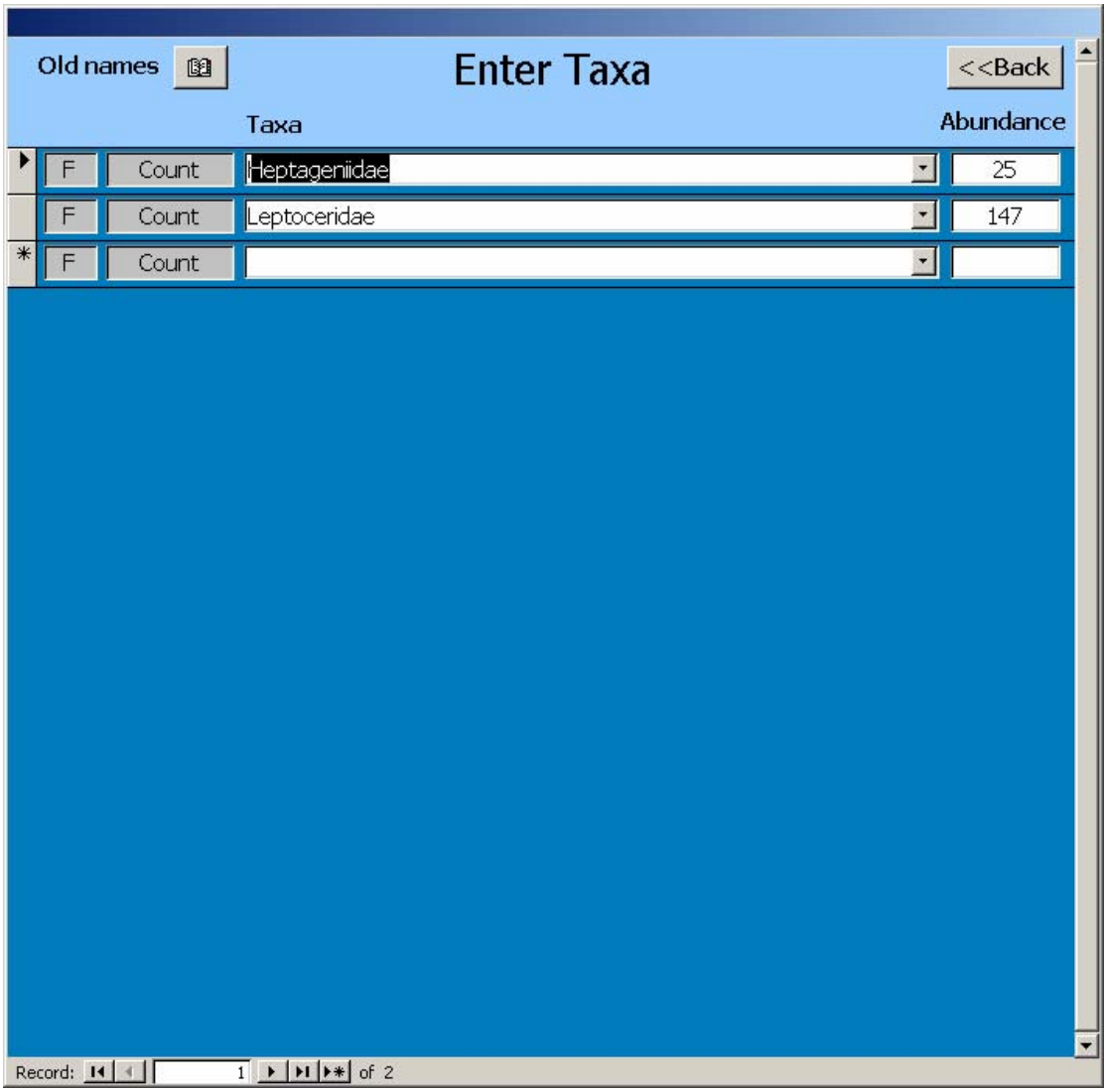


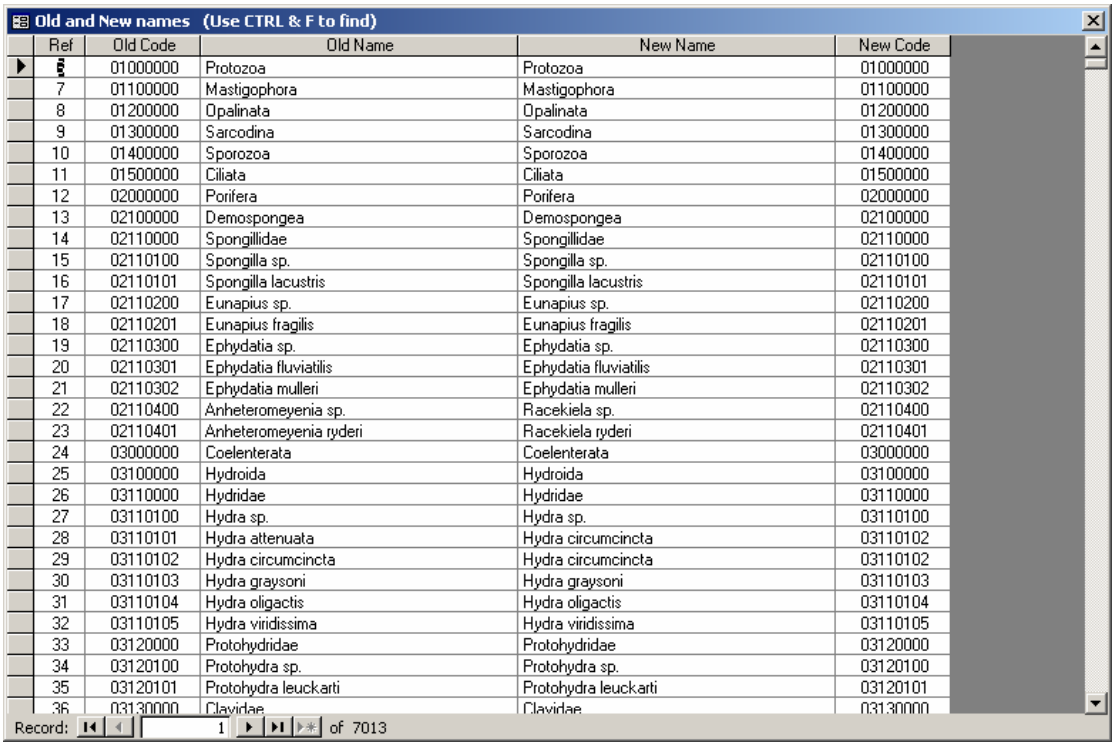
Figure 13

This form is based on Qry_Underlying_Frm_006_Select_Taxa.

```
SELECT DISTINCTROW Dta_Taxa_Data_Entry.RIVER_ID, Dta_Taxa_Data_Entry.SITE_ID,
Dta_Taxa_Data_Entry.SAMPLE_ID, Dta_Taxa_Data_Entry.SEASON_ID,
Dta_Taxa_Data_Entry.SPECIES_CODE, Dta_Taxa_Data_Entry.TYPE,
Dta_Taxa_Data_Entry.ABUNDANCE_SYSTEM_ID, Dta_Taxa_Data_Entry.ABUNDANCE
FROM Dta_Taxa_Data_Entry
WHERE (((Dta_Taxa_Data_Entry.RIVER_ID)=[Forms]![Frm_004_Select_Sample]![RIVER_ID]) AND
((Dta_Taxa_Data_Entry.SITE_ID)=[Forms]![Frm_004_Select_Sample]![SITE_ID]) AND
((Dta_Taxa_Data_Entry.SAMPLE_ID)=[Forms]![Frm_004_Select_Sample]![SAMPLE_ID]) AND
((Dta_Taxa_Data_Entry.SEASON_ID)=[Forms]![Frm_004_Select_Sample]![SEASON_ID]) AND
((Dta_Taxa_Data_Entry.TYPE)=[Forms]![Frm_005_Select_Abundance_Type]![Fam_Sp]) AND
((Dta_Taxa_Data_Entry.ABUNDANCE_SYSTEM_ID)=[Forms]![Frm_005_Select_Abundance_Type]![Selected_A
bundance_Type]))
ORDER BY Dta_Taxa_Data_Entry.RIVER_ID, Dta_Taxa_Data_Entry.SITE_ID,
Dta_Taxa_Data_Entry.SAMPLE_ID, Dta_Taxa_Data_Entry.SEASON_ID,
Dta_Taxa_Data_Entry.SPECIES_CODE, Dta_Taxa_Data_Entry.TYPE,
Dta_Taxa_Data_Entry.ABUNDANCE_SYSTEM_ID;
```

The correct names should be selected from the drop down list (you can jump to the name required by typing the first few letters of the name) and the abundance measure entered. There is a button available to look up the correct new name of a taxon if an old name has been recorded on the lab sheet. (This opens

Frm_007_Taxa_Old_Name_LookUp based on the table Lkp_Compare_Old_and_New_Names. Figure 14)



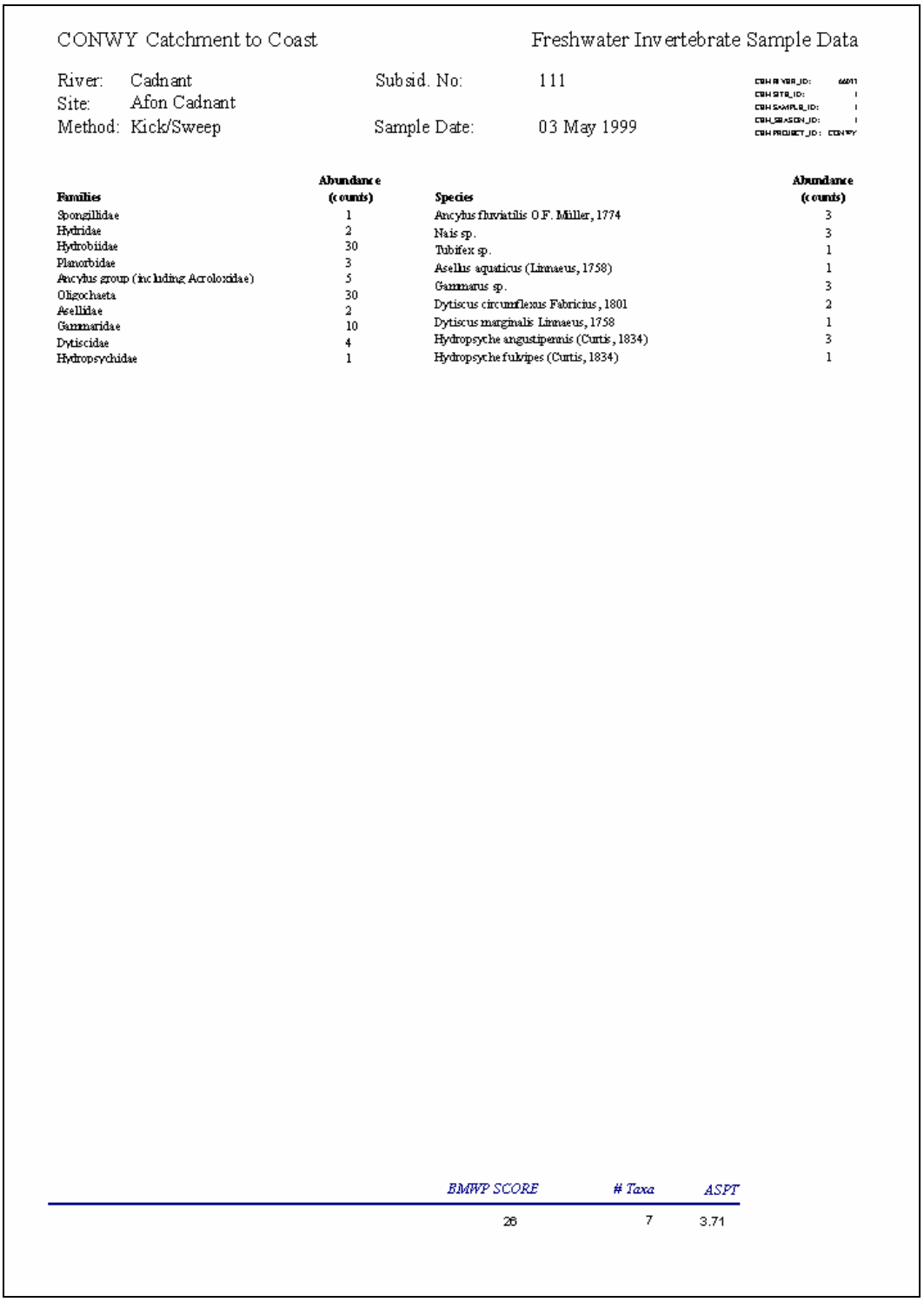
Ref	Old Code	Old Name	New Name	New Code
6	01000000	Protozoa	Protozoa	01000000
7	01100000	Mastigophora	Mastigophora	01100000
8	01200000	Opalinata	Opalinata	01200000
9	01300000	Sarcodina	Sarcodina	01300000
10	01400000	Sporozoa	Sporozoa	01400000
11	01500000	Ciliata	Ciliata	01500000
12	02000000	Porifera	Porifera	02000000
13	02100000	Demospongea	Demospongea	02100000
14	02110000	Spongillidae	Spongillidae	02110000
15	02110100	Spongilla sp.	Spongilla sp.	02110100
16	02110101	Spongilla lacustris	Spongilla lacustris	02110101
17	02110200	Eunapius sp.	Eunapius sp.	02110200
18	02110201	Eunapius fragilis	Eunapius fragilis	02110201
19	02110300	Ephydatia sp.	Ephydatia sp.	02110300
20	02110301	Ephydatia fluviatilis	Ephydatia fluviatilis	02110301
21	02110302	Ephydatia mulleri	Ephydatia mulleri	02110302
22	02110400	Anheteromeyenia sp.	Racekiela sp.	02110400
23	02110401	Anheteromeyenia ryderi	Racekiela ryderi	02110401
24	03000000	Coelenterata	Coelenterata	03000000
25	03100000	Hydroida	Hydroida	03100000
26	03110000	Hydridae	Hydridae	03110000
27	03110100	Hydra sp.	Hydra sp.	03110100
28	03110101	Hydra attenuata	Hydra circumcincta	03110102
29	03110102	Hydra circumcincta	Hydra circumcincta	03110102
30	03110103	Hydra graysoni	Hydra graysoni	03110103
31	03110104	Hydra oligactis	Hydra oligactis	03110104
32	03110105	Hydra viridissima	Hydra viridissima	03110105
33	03120000	Protohydridae	Protohydridae	03120000
34	03120100	Protohydra sp.	Protohydra sp.	03120100
35	03120101	Protohydra leuckarti	Protohydra leuckarti	03120101
36	03130000	Clavidae	Clavidae	03130000

Figure 14

Once data entry is complete for a given sample, press the back button to select another site or another sample.

The Main Menu (Figure 2) gives the options to run validation reports. These must always be checked against the lab sheets to ensure data entered is correct. Edits should be made and the validation run again until no further errors are discovered. Edits are made using the same forms as for data entry. Care must be taken to select the correct site, sample and identification level.

Final results sheets (Figure 15) can be printed off if required using the button on the Main Menu (Figure 2). BMWP scores are calculated on the results sheets using the queries BMWPStage1-5.
(These queries select all taxonomic records at both species and family level and relate them to the relevant BMWP family. The scores are then calculated based on the unique list of BMWP families in each sample.)


CONWY Catchment to Coast

Freshwater Invertebrate Sample Data

River: Cadnant

Site: Afon Cadnant

Method: Kick/Sweep

Subsid. No: 111

Sample Date: 03 May 1999

CEH_RIVER_ID: 66011
CEH_SITE_ID: 1
CEH_SAMPLE_ID: 1
CEH_SEASON_ID: 1
CEH_PROJECT_ID: CONWY

Families	Abundance (counts)
Spongillidae	1
Hydridae	2
Hydrobiidae	30
Planorbidae	3
Ancylus group (including Acroloxidae)	5
Oligochaeta	30
Asellidae	2
Gammaridae	10
Dytiscidae	4
Hydropsychidae	1

Species	Abundance (counts)
Ancylus fluviatilis O.F. Müller, 1774	3
Nais sp.	3
Tubifex sp.	1
Asellus aquaticus (Linnaeus, 1758)	1
Gammarus sp.	3
Dytiscus circumflexus Fabricius, 1801	2
Dytiscus marginalis Linnaeus, 1758	1
Hydropsyche angustipennis (Curtis, 1834)	3
Hydropsyche fulvipes (Curtis, 1834)	1

BMWP SCORE	*# Taxa*	*ASPT*
26	7	3.71

Figure 15

2.1.9 Uploading New Data

Once data entry and validation are completed, the Database Manager refreshes the links to Dta_Taxa_Data_Entry, Dta_SAMPLE_Data_Entry and Dta_PROJECT_SITE_Data_Entry in the relevant Data Entry database and runs the macro Mcr_Append_Dta_Data_Entry to append data from the Data Entry tables to the back end tables.

The Database Manager should also update the table Mtd_METADATA with details of progress with data entry.

2.2 Frequently-Asked Questions (Front-end) Database

The FAQ database was designed to offer users unfamiliar with MS Access the opportunity to extract data for frequently asked queries from the NID without having to ask the database manager. Staff were asked what information they would like to have included and the forms were designed to meet those needs. Additional queries can be added to the form by the database manager should the need arise. More complex queries and project specific analysis should be carried out in a project database using a copy of the main database.

The database opens with a Welcome message (Figure 16).

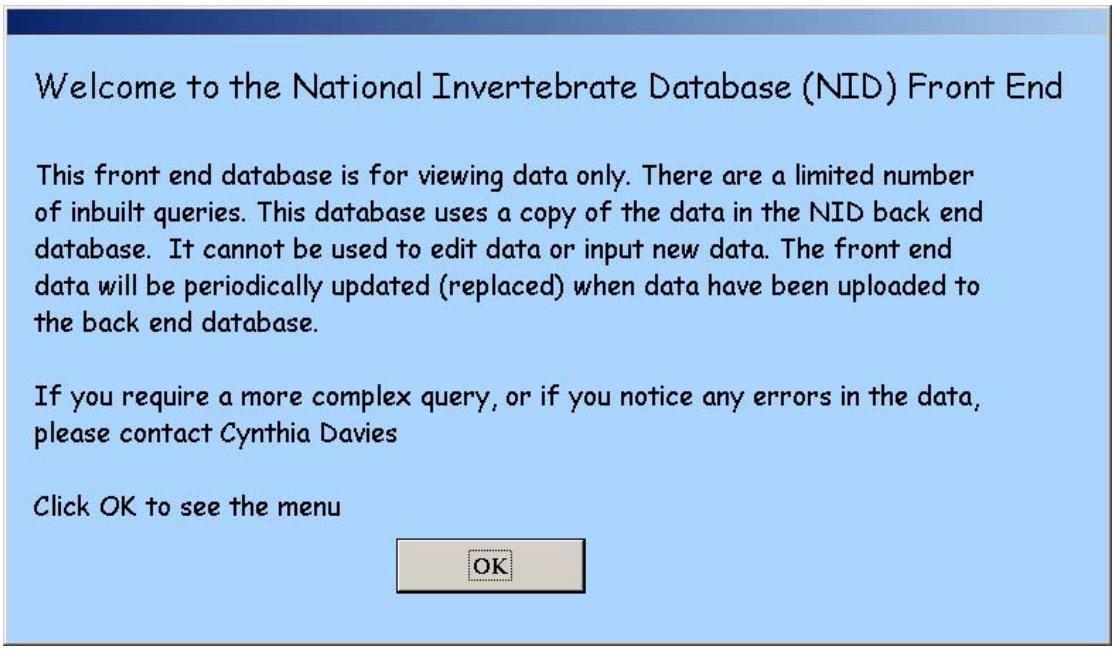


Figure 16

The main menu (Figure 17) offers the user a choice of pre-defined queries. Data can be exported to Excel by using the Export buttons provided or by 'copy and paste'.

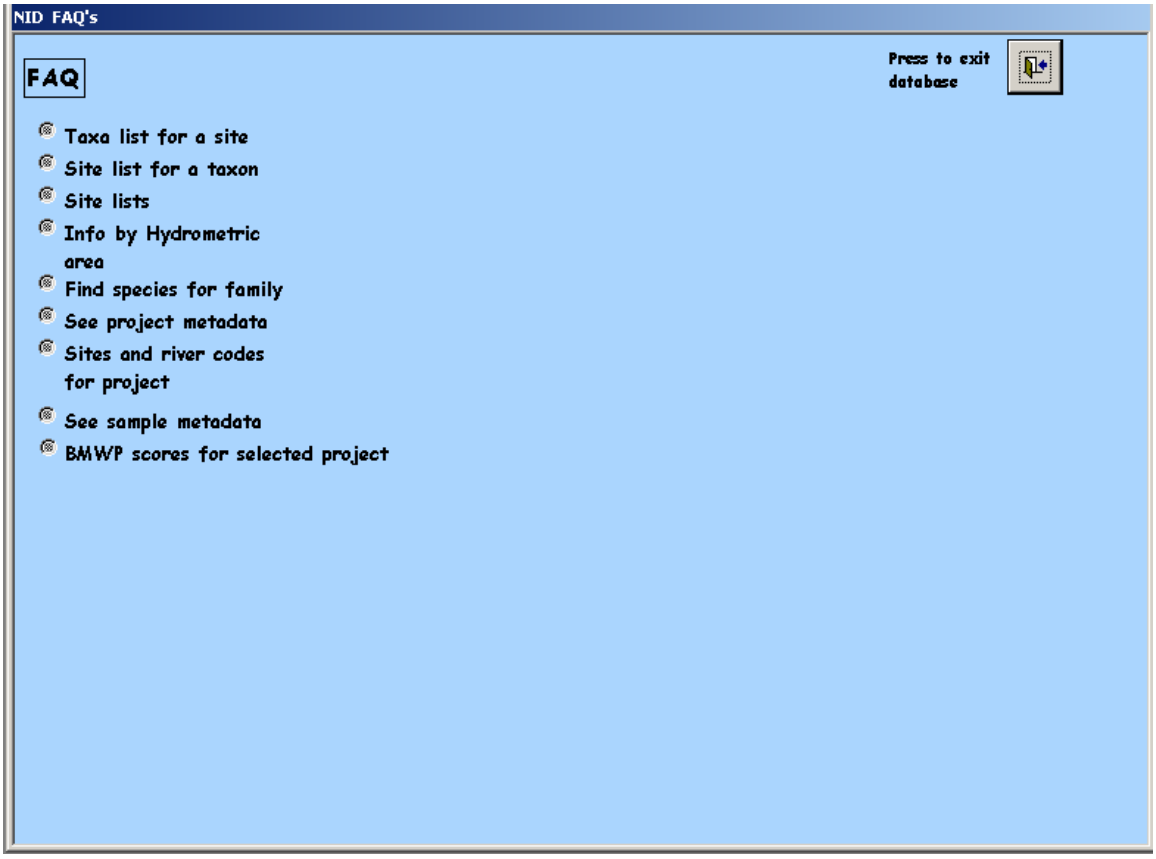


Figure 17

Currently the user is able to select from the following queries.

i) Taxa list for a site: the user can select the site required from either the whole site list or by typing in part of a site name to narrow the options. The results query includes river and site names, the project and project year, project specific site code, date of the sample, the RIVPACS season, the name and code of the taxon, abundance measure and the initials of the person who identified the organism.

ii) Site list for a taxon: the user can select the taxon required by either name or code. Drop down boxes are provided (HINT: type the first few characters of a name or code). The results query includes site details (river and site name, eastings and northings, the project and project year, date, RIVPACS season, abundance measure and the initials of the person who identified the organism.

iii) Site lists: the full list of sites or sites used in a selected project can be viewed. Results include River name and ID, all fields from the table Dta_Site, plus all details from table Dta_Project_Site ().

iv) Information by Hydrometric area: Data for all rivers or for a selected hydrometric area (all fields from table Dta_River); data for all sites or sites for a selected hydrometric area; or taxa by hydrometric are or for a selected hydrometric area. Users can view a map of hydrometric areas to identify the required HA number.

v) Find species for family: component taxa names and codes for a selected family (type the first few letters of the family required and press Return).

vi) See project metadata: Results include the project codes used in the NID and in RMS, the project name and project leader, CEH catalogue number (currently UKEDI), start and end dates for the project and information about IPR. In addition the identification level used for each year of the project is available.

vii) Sites and river codes for project: Results include all fields from Dta_River and Site name, ID and sites code plus eastings and northings.

viii) See sample metadata: the form shows details of data entered for batches of data for each project year (see Figure 18)

ix) BMWP scores for a selected project or all projects

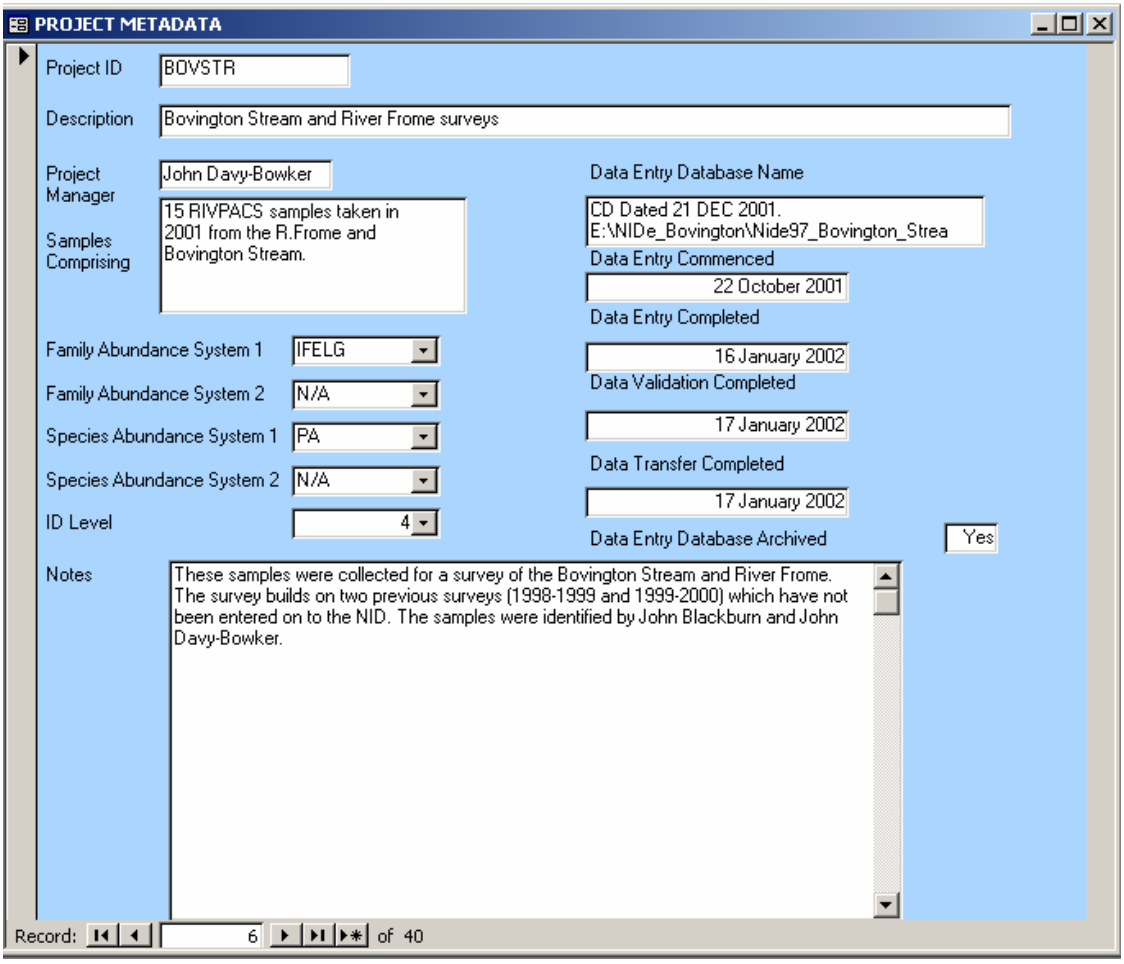


Figure 18

2.3 Project Databases

These database are not form driven. They use a copy of the NID back end and are available for experienced Access users to design their own queries. The Database Manager should make a copy of the Back-end database on request. The module mdl_Refresh_Front_End_Data (Appendix 11) can be run by the Database Manager whenever data are added or amended in the Back-end database to update the project databases with all current data from the Back-end database.

3 Procedure used for the Taxonomic Update

The old NID (Nat97.mdb) contained data up to 23/11/06. The taxonomy used was that in the Coded Checklist circulated as Version 1.1 (dated 1-Aug-1989) compiled by Mike Furse (Centre for Ecology and Hydrology), Ian McDonald (Thames Water Authority) and Bob Abel (Department of the Environment) (Furse et al.)

A decision was taken to update taxonomic data in the NID in November 2006 at the same time as the new NID database structure (as documented in this report) was implemented. Existing data were transferred to the new structure and the taxonomic update was applied to bring the data into line with the Coded Checklist produced by Mike Furse (Biolist_01-09-06.xls) (Furse et al.).

The following procedure was used:

- A spreadsheet was produced cross-referencing the names and codes used in Nat97 with the names and codes in the coded checklist (Biolist_01-09-06.xls); this included the artificial groups needed by old and new data.

- The Spreadsheet was imported to a separate database (Logic_Checks.mdb) database as table Read_Across_v22.

- Validation checks were carried out on the Checklist names and codes, for example to check there were no duplicates in the new codes; that all taxa within a family had the same family code; any given genus only maps to single family etc. The Checklist was amended as necessary before it was published on the Internet (http://www.ceh.ac.uk/data/furse_checklist/furse_animal_freshwater.html).

- The final version of the cross-referencing spreadsheet (v27) was imported to the Logic_Checks database.

- The columns Original_species_code and Species_Code (new code) were selected from ReadAcross (version 27) as table Code_Convertor.

- A new Taxon list (table Taxa_List) was created from the right hand side of Read_Across_v22 (new names and codes etc) where the names were valid (i.e. Key not like 1 & 7).

- The taxon list from Nat97.mdb was cross-checked against the new names using concatenated names to ensure the links were correct.

- When all validation checks were successful the new taxonomic lookup table (Taxa_List) was copied to new NID databases and input databases as the table Lkp_TAXA.

- The taxonomic data table (Dta_TAXA) in the new database structure includes a column to distinguish family and species level records. Data from the ABUNDANCE table in Nat97.mdb were imported to the new structure (table Dta_taxa) and the new column (Type) was updated manually. The code conversion needed to be applied to this table.

- The original data in Nat97.mdb contained samples where 2 taxa had been originally identified as separate taxa, but which in the new checklist mapped to the same new taxon. As these samples were both Presence Absence records there were no issues regarding counts.

- Species_codes in the table Dta_TAXA were updated using the table Code_Converter.

- This resulted in 3 duplicate rows where 2 old names mapped to the same new name. The duplicates were deleted.

- The updated table Dta_TAXA was copied to the new NID databases.

Although there have been amendments to the Coded Checklist subsequent to this date, there have been no major revisions and the Taxonomic information in the NID has not been revised since 2006.

4 The Future

The data in the NID need to be made inter-operable and accessible to other staff across CEH. The current structure in MS Access is not the most appropriate method of distribution and access. In addition, having multiple copies of the data in different databases could lead to inconsistencies. It requires a database manager to carry out several operations to ensure the databases are synchronised. As data result from different projects, records are 'owned' by different funding organisations so there needs to be some method of restricting access to parts of the database.

In order to maintain security of the source data and control access rights it is proposed that a modified database structure is implemented in Oracle. Instead of multiple copies of the data, access can be controlled with read-only access to either the tables or pre-defined views (queries) for most users. The data can be queried using SQL, through Oracle Forms or by linking tables to an Access front end. Alternatively, a web based application could be built to provide access to the source data.

The current database structure does not allow an audit trail of species taxonomy. No record is retained of the name under which a taxon was originally recorded when a taxonomic upgrade is applied. This issue needs to be addressed. In addition, the Furse Code system requires a considerable input of time and high level expertise to maintain in line with current taxonomy. It is proposed that an alternative externally-maintained coding system is sourced (eg National Biodiversity Network or Natural History Museum codes) and the data are linked to these codes and an audit trail of changes is instituted. The proposed revision in the way taxa are coded within the NID will improve inter-operability of data within CEH and with external bodies such as the Environment Agency.

Species distribution data held in the database are of potential interest to the CEH Biological Records Centre. A protocol and mechanism for access to the relevant data need to be agreed.

Currently a large amount of time has to be invested to set up new codes for rivers and sites before new data can be entered. The functionality of the Intelligent River Network for should be investigated for allocating codes in a more consistent manner.

Data entry should continue to be via separate Data Entry tables and the data only transferred to the main tables once fully validated. Forms should be used to ensure that data are entered into the correct tables, using the correct codes.

APPENDIXES

Appendix 1. References

Furse, M.T, McDonald, I & Abel, R. *A Coded Checklist of Animals Occurring in Fresh Water in the British Isles* (http://www.ceh.ac.uk/data/furse_checklist/furse_animal_freshwater.html).

Maitland, P S (1977). *A coded checklist of animals occurring in fresh water in the British Isles*, Edinburgh: Institute of Terrestrial Ecology, 76pp.

Symes, K.L. (1997). Guide to the National Invertebrate Database: Querying and Reporting System. Data Entry System. IFE Report T04053Z2/2

Appendix 2. Definition of terms used in the Manual

ASPT
The total BMWP score is divided by the number of taxa contributing to the score, to give the Average Score Per Taxon (ASPT). The ASPT is independent of sample size and perhaps less influenced by season than the BMWP score. It provides an additional, more consistent index of water quality.

Attribute
An Entity has properties that are associated with it or that describe it. These properties are attributes. Such as a river has the attribute River Name associated with it, a project has the attribute Project Code.

AWIC
Acid Waters Indicator Community index (AWIC) is a biotic index for assessing the impact of acidity on streams and rivers in England and Wales based on the sensitivity of macroinvertebrate organisms to acid water stress. The AWIC index distinguishes acid sites by the absence of any of a large number of acid sensitive families.

BMWP
Biological Monitoring Working Party (BMWP) protocol, which is the principal system of measuring water quality using freshwater macroinvertebrates used in the UK. Animals are given a biotic score based on their sensitivity to organic pollution. The total score for a sample can then be calculated; the higher the score the better the water quality in terms of organic pollution.

Determinand
A constituent or property of the water that is determined, or estimated, in a sample for example, chloride, turbidity, pH.

Entity
An Entity is an object of interest that exists and can be distinguished from other objects, such as a macroinvertebrate, a river or a project.

Form
A Form is an interface between the data and the user. For example, Forms enable users to enter valid data into the correct tables, look up information such as correct species codes to enter or to carry out complex pre-defined operations on the stored data in a user-friendly manner.

GIS
Geographic information system. Automated systems for the capture, storage, retrieval, analysis, and display of spatial data.

Hydrometric area
Hydrometric areas are groupings of catchments for monitoring and reporting purposes. They are either integral river catchments having one or more outlets to the sea or tidal estuary or, for convenience, they may include several contiguous river catchments having topographical similarity with separate tidal outlets.

Macro
A series of keyboard and mouse actions (program commands or instructions) which are stored in a file as a single computer instruction which can be recalled when necessary. Macros are helpful when you perform a task often. Macros can be attached to command buttons on forms. In Access 2002, any macro called Autoexec will run as soon as the database is opened.

Metadata
Data that describes the structure, organisation and/or location of data. Metadata is commonly called 'data about data'.

Module
Access 2002 database operations can be programmed using Visual Basic for Applications language (VBA 6.5). A module is a block of programming code consisting of one or more procedures stored in a single location.

Query
A Query is a question about the data stored in the Tables. Common queries are those which select certain attributes from a table or a number of linked tables. Criteria can be specified to select only the required data.

Relationship
A relationship is an association between Entities, for example a Sample site is situated on a river and a River may have one or more sampling sites along its length.

Report
A Report is a method of presenting data as information in a repeatable format based on the data currently held in the database. Additional text can be pre-set in the Report as required. Reports can be viewed on screen, printed out or exported to word processing packages.

RIVPACS
River Invertebrate Prediction and Classification System. The RIVPACS Software provides an output consisting of faunal lists Observed by sampling and Expected (generated by the software) from measured environmental variables, Observed and Expected biotic index values and associated Environmental Quality Indices (EQI) values. The EQI values can be banded into quality classes to aid interpretation of the data across the entire region or country.

SQL
Structured Query Language, a language for processing digital data.

Table
A Table is a 'container' for data about a particular subject or entity.

Taxon

A taxon (plural: taxa) is a group of (one or more) organisms, which a taxonomist adjudges to be a unit within biological classification. Examples include: species, genera, families, orders, etc.

Appendix 3. Description of Database Tables and Columns

Primary key fields are emboldened.

Table: Dta_CHEMISTRY: Table of chemical data

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
SITE_ID	Integer	2	CEH Site Identification Number
SAMPLE_DATE	Date/Time	8	Sample Date
SAMPLE_TIME	Date/Time	8	Sample Time
DETERMINAND	Text	255	CEH Determinand Code (includes details of units)
MEASUREMENT	Text	12	Chemical Measurement
CHEM_CODE	Text	50	Originator Organisation Sample Code (usually an Environment Agency Sample Code)

Table: Dta_PROJECT_SITE: Table of sites sampled for a particular project in a particular year

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
SITE_ID	Integer	2	CEH Site Identification Number
PROJECT	Text	6	CEH Project Identification Number
PROJECT_YR	Integer	2	CEH Project Year (groups samples taken for a given project in a given year)
SURVEY_DATE	Date/Time	8	Survey Date
MAX_WIDTH	Double	8	Maximum River Width (m)
MIN_WIDTH	Double	8	Minimum River Width (m)
MEAN_WIDTH	Double	8	Mean River Width (m)
DEPTH_CAT	Integer	2	Depth Category (Depth over 50% of the survey area. 1 = <25cm, 2= <50cm, 3 = <75cm, 4 = <1m, 5 = <2m, 6 = >2m)
BANK_STAB	Integer	2	Bank Stability. Score each bank separately: 0 for stable, 1 for eroding. Bank stability = sum of score
TREES	Integer	2	Extent of Trees. Score each bank separately: 0 trees absent, 1 trees present. Extent = sum of score
BUSHES	Integer	2	Extent of Bushes. Score each bank separately: 0 absent, 1 present. Extent = sum of score
REED_RUSHES	Integer	2	Extent of Reeds and Rushes. Score each bank separately: 0 absent, 1 present. Extent = sum of score
LOW_PLANTS	Integer	2	Extent of Low Plants. Score each bank separately: 0 absent, 1 present. Extent = sum of score
OTHER_SPEC	Integer	2	Extent of Other Plants. Score each bank separately: 0 absent, 1 present. Extent = sum of score
SHADING	Integer	2	Extent of Shading. Score each bank separately: 0 Nil/low shade, 1 Moderate shade, 2 Great shade. Extent = sum of score
URBAN	Integer	2	Extent of Urban Land. Score each bank separately: 0 absent, 1 present. Extent = sum of score
ARABLE	Integer	2	Extent of Arable Land. Score each bank separately: 0 absent, 1 present. Extent = sum of score
GRASSLAND	Integer	2	Extent of Grassland. Score each bank separately: 0 absent, 1 present. Extent = sum of score
HEATH_MOOR	Integer	2	Extent of Heath Moorland. Score each bank separately: 0 absent, 1 present. Extent = sum of score

DECID_WOOD	Integer	2	Extent of Deciduous Woodland. Score each bank separately: 0 absent, 1 present. Extent = sum of score
CONIF_WOOD	Integer	2	Extent of Coniferous Woodland. Score each bank separately: 0 absent, 1 present. Extent = sum of score
WEED_CUTTING	Integer	2	Extent of Weed Cutting. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
DREDGING	Integer	2	Extent of Dredging. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
BANK_MAINT	Integer	2	Extent of Bank Maintenance. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
CHANNEL_STRG	Integer	2	Extent of Channel Straightening. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
ORG_POLLN	Integer	2	Extent of Organic Pollution. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
INORG_POLLN	Integer	2	Extent of Inorganic Pollution. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
PHYS_POLLN	Integer	2	Extent of Physical Pollution. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
BRIDGE	Integer	2	Extent of Bridges. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
WEIR	Integer	2	Extent of Weirs. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s
OTHER_INFRAST	Integer	2	Extent of Other Infrastructure. 0= No influence, 1 = influence u/s, co-incident or not stated, 2 influence d/s

Table: Dta_RIVER: Table of rivers together with their names & hydrometric codings

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
RIVER_NAME	Text	60	CEH River Name
HYDROM_AREA	Integer	2	Hydrometric area
PRIMARY_RIVER	Integer	2	CEH Primary river number
PRIMARY_TRIB	Integer	2	CEH Primary tributary number
SECOND_TRIB	Integer	2	CEH Secondary tributary number
TERT_TRIB	Integer	2	CEH Tertiary tributary number
QUATERN_TRIB	Integer	2	CEH Quaternary tributary number
QUIN_TRIB	Integer	2	CEH Quinary tributary number
OTHER_TRIB	Integer	2	CEH Other tributary number
AUTHORITY_ID	Integer	2	River Authority Identification Number
AUTHORITY_NAME	Text	30	River Authority Name
AUTHORITY_ORG	Text	3	River Authority Organisation/Type

Table: Dta_SAMPLE: Table of invertebrate samples with associated project, date & environmental variables etc

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
SITE_ID	Integer	2	CEH Site Identification Number
SAMPLE_ID	Integer	2	CEH Sample Identification Number
SEASON_ID	Integer	2	CEH Season Identification Number, where: 1(Spring) = 1st March to 31st May; 2(Summer) = 1st June to 31st Aug; 3(Autumn) = 1st Sept to 31st Nov (although some samples are given Season ID numbers outside these date limits)
PROJECT	Text	7	CEH Project Identification code
PROJECT_YR	Integer	2	CEH Project Year (groups samples taken for a given project in a given year)

SAMPLE_DATE	Date/Time	8	Date Sampled
WATER_WIDTH	Double	8	Average Water Width (m)
DEPTH1	Double	8	River Depth Measurement 1 (cm)
DEPTH2	Double	8	River Depth Measurement 2 (cm)
DEPTH3	Double	8	River Depth Measurement 3 (cm)
SURF_VEL_CAT	Integer	2	Water Surface Velocity Category, where 1 = <=10m/s; 2 = >10 to 25 m/s; 3 = >25 to 50 m/s; 4 = >50 to 100 m/s; 5 = >100 m/s
CLARITY	Integer	2	Water Clarity (1 = Clear & bright, 2 = Cloudy, 3 = Turbid)
SAMPLING_TIME	Integer	2	Sampling Time (values 1-8 indicate minutes; value 15 indicates seconds)
METHOD	Integer	2	Sampling Method (cross-ref Lkp_Sampling_Method)
ROCK_PAVEMT	Integer	2	Surface Cover of Rock Pavement (this is independent of the other 4 substratum categories)
BOULDER_COBBLE	Integer	2	% Surface Cover of Boulders/Cobbles (excluding Rock Pavement)
PEBBLE_GRAVEL	Integer	2	% Surface Cover of Pebbles/Gravel (excluding Rock Pavement)
SAND	Integer	2	% Surface Cover of Sand (excluding Rock Pavement)
SILT_CLAY	Integer	2	% Surface Cover of Silt/Clay (excluding Rock Pavement)
DOM_PART_SIZE	Double	8	Dominant Particle Size (Phi Scale)
MACROPH_COVER	Integer	2	% Surface Cover of Instream Macrophytes
DETRITUS	Integer	2	Extent of Detritus. 0 absent, 1 present. Extent = sum of score
SUBSID_NO	Text	100	CEH Subsidiary Sample Number (project/dataset specific identification number with meaning to the project manager)
COLLECTED_BY	Text	5	Person who Collected the Sample
IDENTIFIED_BY	Text	5	Person who Identified the Sample
ID_LEVEL	Integer	2	Identification Level of the Sample
RIVPACS SAMPLE?	Yes/No	1	Identifies if this sample has ever been used in a version of RIVPACS

Table: Dta_SITE: Table of sites with associated National Grid References & environmental variables etc

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
SITE_ID	Integer	2	CEH Site Identification Number
SITENAME	Text	125	CEH Site Name
EASTING	Text	50	Great Britain National Grid Reference Easting
NORTHING	Text	50	Great Britain National Grid Reference Northing
DISCHARGE	Integer	2	Discharge Category (categories of mean annual discharge in cubic metres per second, where 1 = <0.31; 2 = 0.31 to 0.62; 3 = 0.62 to 1.25; 4 = 1.25 to 2.5; 5 = 2.5 to 5.0; 6 = 5.0 to 10.0; 7 = 10.0 to 20.0; 8 = 20.0 to 40.0; 9 = 40.0 to 80; 10 = >80.0)
DIST_F_SOURCE	Double	8	Distance from source (km)
ALTITUDE	Long Integer	4	Altitude above sea level (m)
SLOPE	Double	8	Slope (m/km)
STRM_ORDER	Integer	2	Stream Order (Strahler)
CATCHMT_AREA	Double	8	Catchment Area (km2)
SITECODE	Text	50	Old CEH (IFE) Site Code
TWINSPAN_GRP	Text	4	RIVPACS TWINSPAN Classification Group (as used in RIVPACS)
RIVPACS_SAMPLE	Yes/No	1	Identifies if this site has ever been used in a version of RIVPACS
COUNTRY	Text	50	Country
COUNTY	Text	50	County

Appendix 3

Table: Dta_TAXA: Table of individual invertebrate records as species codes

Name	Type	Size	Description
RIVER_ID	Long Integer	4	CEH River Identification Number
SITE_ID	Integer	2	CEH Site Identification Number
SAMPLE_ID	Integer	2	CEH Sample Identification Number
SEASON_ID	Integer	2	CEH Season Identification Number
SPECIES_CODE	Text	8	CEH Species Code (Furse Code)
ABUNDANCE_SYSTEM_ID	Text	5	CEH Abundance System Identification Number
TYPE	Text	50	Identifies if the taxon record is a Family (F) or Species (S) level record
ABUNDANCE	Text	50	Abundance (may be counts, log10 categories etc - see Abundance System ID to interpret Abundances)

Table: Lkp_ABUNDANCE SYSTEM: Look Up table of abundance systems

Name	Type	Size	Description
ABUNDANCE_SYSTEM_ID	Text	5	CEH Abundance System Identification Number
ABUNDANCE_SYSTEM_NAME	Text	255	Description of Abundance System used

Table: Lkp_Compare_Old_and_New_Names: LOOKUP old to new names/codes (used in data entry forms) (Updated August 09

Name	Type	Size	Description
Ref	Long Integer	4	
Old_code	Text	255	Old code
Old_name	Text	4	Old name
Int_name	Text	255	Intermediate name
Int_code	Text	255	Intermediate code
Current_code	Text	255	Current code (updated August 09)
Current_name	Text	255	Current name (updated August 09)

Table: Lkp_CROSSREF: Cross reference from Lkp_Taxa to BMWP and AWIC family codes

Name	Type	Size	Description
SPECIES_CODE	Text	8	CEH Species Code (Furse Code)
BMWP_F_CODE	Text	50	Equivalent BMWP Family code
AWIC_F_CODE	Text	50	Equivalent AWIC Family code
TAX_RICH	Text	50	Code to Convert to for taxon richness calculation

Table: Lkp_DETERMINAND: Look Up table of chemical determinand codes and descriptions

Name	Type	Size	Description
DATA_ORIGIN	Text	255	Describes where the data was obtained from (the Originator)
DETERMINAND	Text	255	CEH Determinand Code
DET_CODE	Text	4	Originator Determinand Code
DET_NAME	Text	255	Originator Determinand Name
DET_DESCRIPTION	Text	255	Originator Determinand Description
DET_UNITS	Text	255	Determinand Units
PREFERRED_NAME	Text	255	CEH Preferred Determinand Name (standardises Determinand Names for data from different Originators)
PREFERRED_DET	Text	255	CEH Preferred Determinand Code (standardises Determinand Codes for data from different Originators

Table: Lkp_ID_LEVEL: Look Up table of identification level descriptions

Name	Type	Size	Description
ID_LEVEL_ID	Integer	2	Identification Level of the Sample
ID_LEVEL_DESCRIPTION	Text	255	Identification Level Description

Appendix 3

Table: Lkp_PROJECTS: Look Up table of project descriptions

Name	Type	Size	Description
Project ID	Text	255	CEH Project Identification Number
Project description	Text	255	Brief Description of the Project
Project Code	Text	100	FBA/IFE/CEH Project Code(s)
Project Leader(s)	Text	100	Project Leader at the time the work was done
Catalogue_Number	Text	100	United Kingdom Environmental Information Data Index number or CEH Information Gateway URI
Sample_Start_Year	Long Integer	4	Year of the First Sample
Sample_End_Year	Long Integer	4	Year of the Last Sample
IPR	Text	255	Intellectual Property Rights information

Table: Lkp_SAMPLING_METHOD: Look Up table of sampling methods

Name	Type	Size	Description
SAMPLE_METHOD_ID	Integer	2	Unique reference code
SAMPLE_METHOD_DESCRIPTION	Text	50	Description of code

Table: Lkp_STAFF: Look Up table of staff

Name	Type	Size	Description
STAFF_ID	Text	50	Unique id (usually person's initials)
STAFF_NAME	Text	100	Name (May include staff from other organisations, students etc)

Table: Lkp_TAXA: Look Up table of invertebrate taxon names

Name	Type	Size	Description
SPECIES_CODE	Text	8	CEH Species Code (Furse Code)
MAJOR_GP_NAME	Text	25	Major Taxonomic Group Name
LEVEL2_NAME	Text	35	Level 2 Taxonomic Group Name
LEVEL3_NAME	Text	100	Level 3 Taxonomic Group Name
GENUS_NAME	Text	35	Genus name
SUBGENUS_NAME	Text	35	Subgenus Name
SPECIES_NAME	Text	35	Species Name
AUTHORITY	Text	50	Authority
COMMON_NAME	Text	35	Common Name
Red_Data_Book	Text	50	British Red Data Book designation; last bulk update = XX/XXX/XXXX; source = XXXX
Nationally_Notable	Text	50	Nationaly Notable; last bulk update = XX/XXX/XXXX; source = XXXX

Table: Lkp_TAXA_LIST_AWIC_F: Look Up table of AWIC (family) taxa and AWIC (family) scores

Name	Type	Size	Description
SPECIES_CODE	Text	8	Taxon code in Lkp_taxa
FAMILY_NAME	Text	100	Name in Lkp_Taxa
AWIC_F_SCORE	Integer	2	AWIC family level score
PUB_Name	Text	100	Published name (Davy_Bowker et al (2005))

Table: Lkp_TAXA_LIST_BMWP_FAMILIES: Look Up table of BMWP taxa and BMWP taxon scores

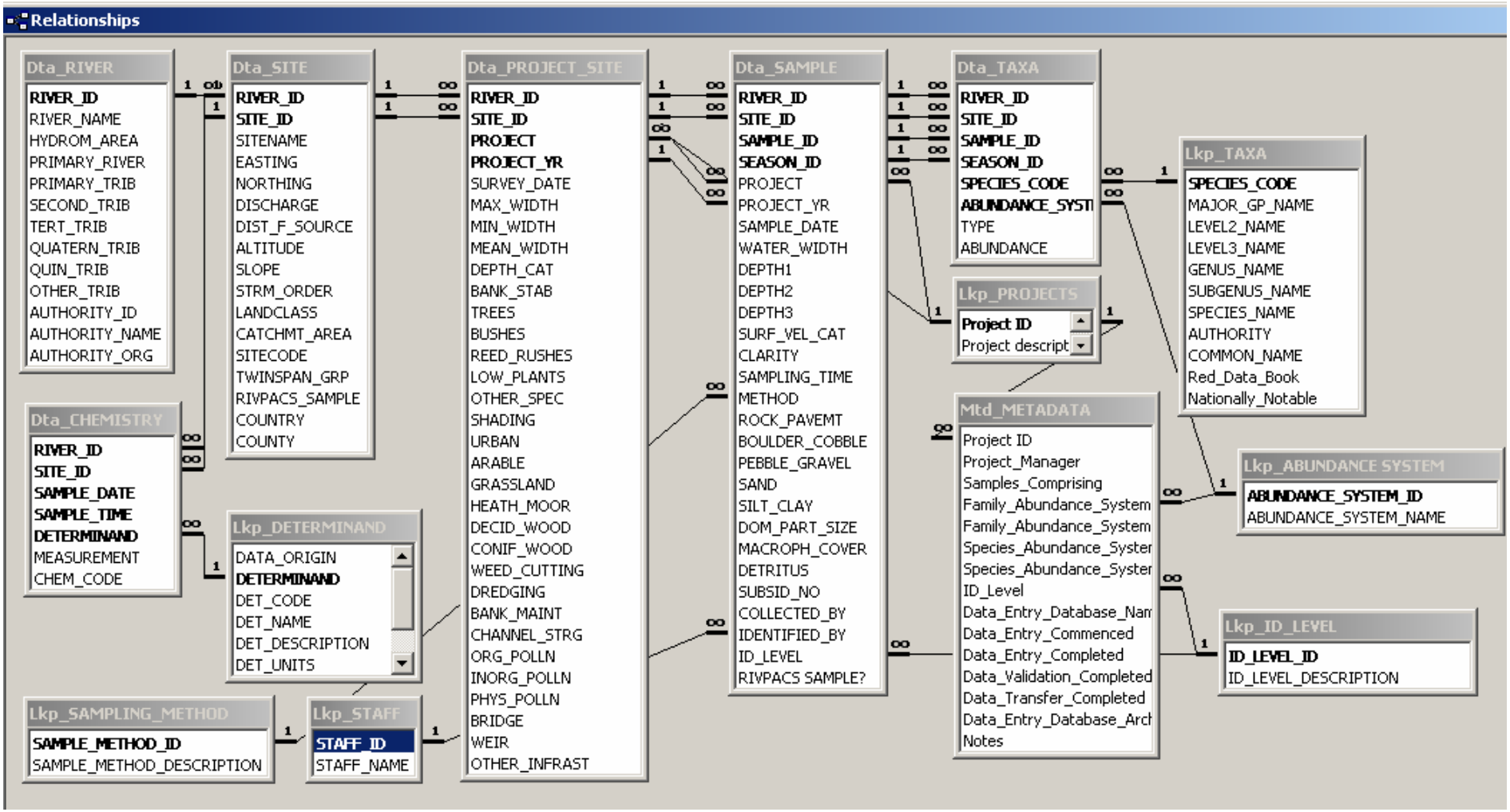
Name	Type	Size	Description
SPECIES_CODE	Text	8	Code in Lkp_Taxa
FAMILY_NAME	Text	100	Name in in Lkp_Taxa
BMWP_SCORE	Integer	2	BMWP score
Pub_Name	Text	100	Name used in Hawkes (1997) Technical note, Origin and development of the Biological Monitoring Working Party Score System. Wat Res. vol 32 no 3. Table 2 -1980 (nb, Clambidae, Curculionidae & Chrysomelidae were dropped)

Appendix 3

Table: Mtd_METADATA: Metadata table for projects

Name	Type	Size	Description
Project ID	Text	255	CEH Project Identification Number
Project_Manager	Text	20	Full name of in house project manager
Samples_Comprising	Text	255	Description of the batch of samples e.g. all samples we have done up to the year 2000
Family_Abundance_System_1	Text	5	CEH Abundance System used for 'families' eg IFELG, or ECNLG. First system of entry.
Family_Abundance_System_2	Text	50	CEH Abundance System used for 'families' eg IFELG, or ECNLG. Second system of entry.
Species_Abundance_System_1	Text	5	CEH Abundance System used for 'species' eg PA, IFECO. First system of entry.
Species_Abundance_System_2	Text	50	CEH Abundance System used for 'species' eg PA, IFECO. Second system of entry.
ID_Level	Integer	2	Identification Level of the Samples
Data_Entry_Database_Name	Text	255	FULL name of data entry database eg Nide97_Env_Change_Network_DATA_ENTRY .mdb
Data_Entry_Commenced	Date/Time	8	Date initial data entry commenced
Data_Entry_Completed	Date/Time	8	Date initial data entry completed
Data_Validation_Completed	Date/Time	8	Date data fully validated and corrected so that ready for transfer to main database
Data_Transfer_Completed	Date/Time	8	Date data transferred from data entry system to main database
Data_Entry_Database_Archived	Yes/No	1	Has the data entry database been archived to CD or other permanent storage device
Notes	Memo	-	Other notes

Appendix 4. Entity Relationship Diagram of the NID



Appendix 5. Examples of Records in the Taxonomic Table

Extract from Table Lkp_TAXA

SPECIES_CODE	MAJOR_GP_NAME	LEVEL2_NAME	LEVEL3_NAME	GENUS_NAME	SUBGENUS_NAME	SPECIES_NAME	AUTHORITY	COMMON_NAME
32000000	Branchiura							Fish lice
32010000	Branchiura		Argulidae					Fish lice
32010100	Branchiura		Argulidae	Argulus				Fish lice
32010101	Branchiura		Argulidae	Argulus		coregoni	Thorell, 1865	Fish lice
48000000	Trichoptera							Caddisflies
48320000	Trichoptera	Limnephiloidea	Brachycentridae					Caddisflies
48320100	Trichoptera	Limnephiloidea	Brachycentridae	Brachycentrus				Caddisflies
48320101	Trichoptera	Limnephiloidea	Brachycentridae	Brachycentrus		subnubilus	Curtis, 1834	Caddisflies
50000000	Diptera							True-flies
50100000	Diptera	Tipuloidea						
50110000	Diptera	Tipuloidea	Tipulidae					True-flies
50110300	Diptera	Tipuloidea	Tipulidae	Tipula				True-flies
50110320	Diptera	Tipuloidea	Tipulidae	Tipula	Schummelia			True-flies
50110321	Diptera	Tipuloidea	Tipulidae	Tipula	Schummelia	variicornis	Schummel, 1833	True-flies
50400000	Diptera	Chironomidae						True-flies
50420000	Diptera	Chironomidae	Tanypodinae					True-flies
50420500	Diptera	Chironomidae	Tanypodinae	Procladius				True-flies
50420510	Diptera	Chironomidae	Tanypodinae	Procladius	Holotanypus			True-flies
50420511	Diptera	Chironomidae	Tanypodinae	Procladius	Holotanypus	choreus	(Meigen, 1804)	True-flies

The names at the different levels can be combined to produce the name of the taxon for each code. The table below shows the taxon name for the extract shown above.

Appendix 5

Taxon name

Code	Name with sub genera	Name
32010000	Argulidae	Argulidae
32010100	Argulus sp.	Argulus sp.
32010101	Argulus coregoni Thorell, 1865	Argulus coregoni Thorell, 1865
48320000	Brachycentridae	Brachycentridae
48320100	Brachycentrus sp.	Brachycentrus sp.
48320101	Brachycentrus subnubilus Curtis, 1834	Brachycentrus subnubilus Curtis, 1834
50100000	Tipuloidea	Tipuloidea
50110000	Tipulidae	Tipulidae
50110300	Tipula sp.	Tipula sp.
50110320	Tipula (Schummelia) sp.	Tipula sp.
50110321	Tipula (Schummelia) variicornis Schummel, 1833	Tipula variicornis Schummel, 1833
50400000	Chironomidae	Chironomidae
50420000	Tanypodinae	Tanypodinae
50420500	Procladius sp.	Procladius sp.
50420510	Procladius (Holotanypus) sp.	Procladius sp.
50420511	Procladius (Holotanypus) choreus (Meigen, 1804)	Procladius choreus (Meigen, 1804)
32000000	Branchiura	Branchiura
48000000	Trichoptera	Trichoptera
50000000	Diptera	Diptera

Appendix 6. CEH River Numbering System

1 Description of the River Number

The river numbering system is hierarchical with the Hydrometric Area (HA) numbers at the first level. There are 108 Hydrometric Areas covering Great Britain, including major islands (e.g. the Isle of Wight). A further 39 Areas cover the whole of Ireland. All Irish HAs are numbered from 200 to differentiate them from the British HA's[8]. After the HA number there are seven levels; each level representing progressively smaller levels of tributaries from the main river.

For example,

Level	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
	27	60	15	85	20	0	0	0	Dove

Level	1 -	hydrometric area
	2 -	primary river (flows into the sea)
	3 -	primary tributary (tributary of primary river)
	4 -	secondary tributary (tributary of primary tributary)
	5 -	tertiary tributary (tributary of secondary tributary)
	6 -	quaternary tributary (etc…)
7	-	quinary tributary
	8 -	other tributaries (beyond quinary)

When the database resided on an Oracle platform, Levels 2,3 and 4 were 3 digit numbers to allow for a maximum of 999 tributaries at these levels; the remainder were two digits allowing up to 99 tributaries at these levels. Although in the current MS Access version the fields are all Integer and allow up to 9999 tributaries at any level, only the correct number of digits should be used at any level.

Within each Hydrometric Area the numbering proceeds in a clockwise direction around the coast. At level 2, the numbering proceeds in an upstream direction starting from the mouth of the river and at subsequent levels from the confluence of the tributary with the watercourse at the next lower level. No two watercourses may be identically numbered across all levels (including Hydrometric Area). Each watercourse must have a number at all levels, even if this number is 0.

2 River numbers in the National Invertebrate database

All biological sampling sites that are entered into the National Invertebrate Database (NID) must be located on a numbered watercourse. Not all the watercourses in the country have been numbered with this system. An initial attempt at numbering the largest rivers and their tributaries in each HA was made using the NRA River Quality Survey maps in England and Wales and the River Purification Maps in Scotland, from which all rivers with a discharge

[8] (NOTE: The numbers do not correlate with the current Irish Environmental Protection Agency numbering system; The EPA uses 40 areas for the whole of Ireland, numbered from 01 – 40.)

greater than the smallest class (shown as thick lines on the maps) were numbered. These watercourses were numbered at intervals of 5 or 10 to allow for the numerous small streams left in between. The initial numbered network was then expanded and updated as each sampling sites in the database was located on its watercourse, using the 1:50,000 maps to number and name the smallest streams. The numbering of watercourses even from the 1:50,000 maps still has intervals to allow for even smaller streams which do not appear on maps at this scale. It should be noted that it has not been possible to retain the sequential nature of the numbering in the higher levels (4-6) as additional tributaries have been added.

In addition, a single unique number is required to identify each river within the table Dta_River and within other related tables in the database. The numbers themselves have no intrinsic meaning but are purely sequential; however, by convention, the appropriate hydrometric area number is used as the first digits of the River_ID. When a new watercourse is added to the database, the new River-ID is allocated by incrementing the largest River_ID used so far in the HA.

Example from Hydrometric Area 44 of how the numbering is applied
Table 1:

River name	HA	Prim. River	1^y Trib	2^y Trib	3^y Trib	4^y Trib	5^y Trib	Other trib	River ID
Frome	**44**	**20**	0	0	0	0	0	0	44005
Luckford Lake	**44**	**20**	10	0	0	0	0	0	44017
Unnamed	**44**	**20**	10	5	0	0	0	0	44020
Unnamed	**44**	**20**	10	7	0	0	0	0	44018
Unnamed	**44**	**20**	10	7	5	0	0	0	44019
Holy Stream	**44**	**20**	15	0	0	0	0	0	44006
Win	**44**	**20**	17	0	0	0	0	0	44021
Wool Stream	**44**	**20**	20	0	0	0	0	0	44007
Bovington Stream	**44**	**20**	25	0	0	0	0	0	44022
Tadnoll Brook	**44**	**20**	30	0	0	0	0	0	44008
South Winterborne	**44**	**20**	45	0	0	0	0	0	44014
Cerne	**44**	**20**	55	0	0	0	0	0	44023
Sydling Water	**44**	**20**	60	0	0	0	0	0	44015
Hooke	**44**	**20**	65	0	0	0	0	0	44016
Rampisham Brook	**44**	**20**	75	0	0	0	0	0	44024
Wraxhall	**44**	**20**	75	5	0	0	0	0	44025

Appendix 6

.

Primary river
Frome Code: HA **44, 20**,0,0,0,0,0,0 (Columns for tributaries (Primary to Other) are all 0)
Primary tributaries of the Frome
Luc kford Lake: Code: **44,20,**10,0,0,0,0,0
Tributaries of Luckford Lake
Un named Code: **44,20,**10,5,0,0,0,0
Un named Code: **44,20,**10,7,0,0,0,0
Tributary of 2nd Unnamed
Unnamed Code: **44,20,**10,7,5,0,0,0
Ho ly Stream Code: **44,20,**15,0,0,0,0,0
W in Code: **44,20**,17,0,0,0,0,0
Wo ol Stream Code: **44,20,**20,0,0,0,0,0
Bovington Stream Code: **44,20,**25,0,0,0,0,0
Tadnoll Brook Code: **44,20,**30,0,0,0,0,0
South Winterborne Code: **44,20,**45,0,0,0,0,0
Cerne Code: **44,20,**55,0,0,0,0,0
Sydling Water Code: **44,20,**60,0,0,0,0,0
Hooke Code: **44,2**0,65,0,0,0,0,0
Rampisham Brook Code: **44,20,**75,0,0,0,0,0
Tributary of Rampisham Brook
Wraxh all Code: **44,20,**75,5,0,0,0,0

Appendix 6

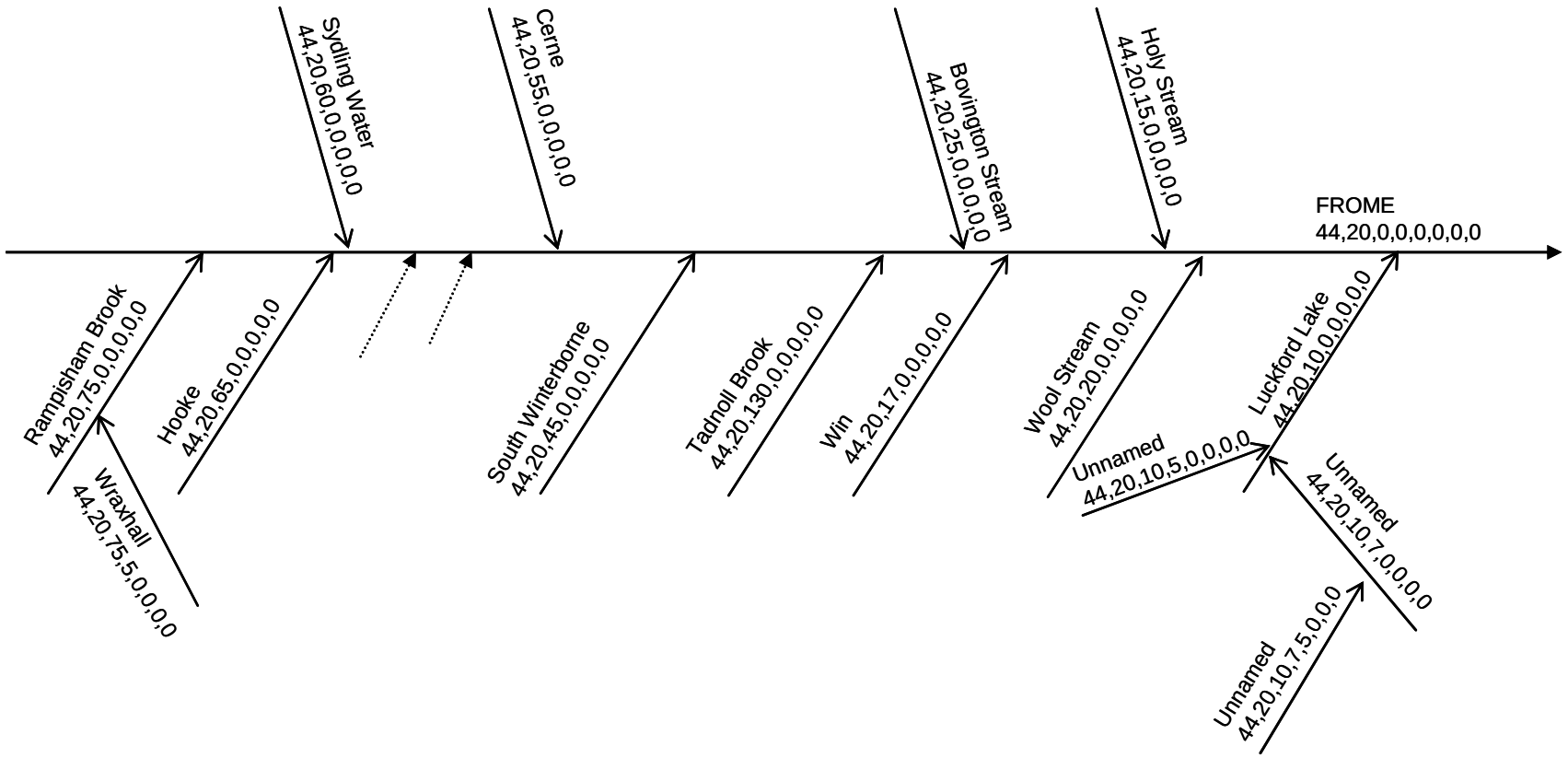


Figure 19 is a schematic representation of the numbering system. Dotted arrows indicate as yet un-numbered tributaries

3 River names

The naming protocol of the River Quality Maps was adopted for consistency and to reduce unnecessary storage in the database. This means the word *River* has been assumed for all rivers that are named *River* <*name*> on the maps so that just the <*name*> is entered into the database. This applies in Wales where the word **Afon** is used to mean **river**. Some rivers appear on the maps as <*name*> *River*; in this case the distinction has been maintained and the word *River* is entered into the database. All other descriptive terms have been kept, e.g. Stream, Brook, Nant, Burn, Water, Pill and Allt. Where a watercourse does not have a name, it is named *Unnamed* in the database. It is possible that a watercourse that has been numbered just from the River Quality map appears as *Unnamed* although when found on the 1:50,000 map it does have a name. In this case, the known a should replace *Unnamed*.

4. Numbering new watercourses

If a site that is being added to the database is not located in a numbered watercourse, then that watercourse must first be numbered. This should be done using the following procedure:

a) Find out the HA that the site falls into.
b) Locate the site on its watercourse from its grid reference on the appropriate 1:50,000 map.
c) Follow the watercourse downstream until it flows into a named watercourse that appears on the complete list if numbered watercourses for the HA.
d) Locate the numbered watercourses up and downstream of the new watercourse.
e) Number the new watercourse appropriately, taking care to ensure that the numbers are correct at all levels and do not duplicate an existing river number.

Appendix 7. Data Preparation Forms

SITE LIST FOR DATA ENTRY

Name of project (for which the site numbering was completed):

Date of completion:

River name	River_ID	Site Name	Site_ID	NGR	Country	Country

Date of entry to NID:

Entered By:

Appendix 7

DETAILS OF RIVER NUMBERING FOR ADDITION TO THE NATIONAL INVERTEBRATE DATABASE

Name of project (for which the river numbering was completed):

Date of completion (for river numbering)	***Hydrometric Area:***	***Maximum River_ID noted for HA:***

Rivers to be added (PTO for additional space)

RIVER_ID	HA	Primary River	Primary Trib	Secondary Trib	Tertiary Trib	Quarternary Trib	Quinary trib	Other Trib	River Name

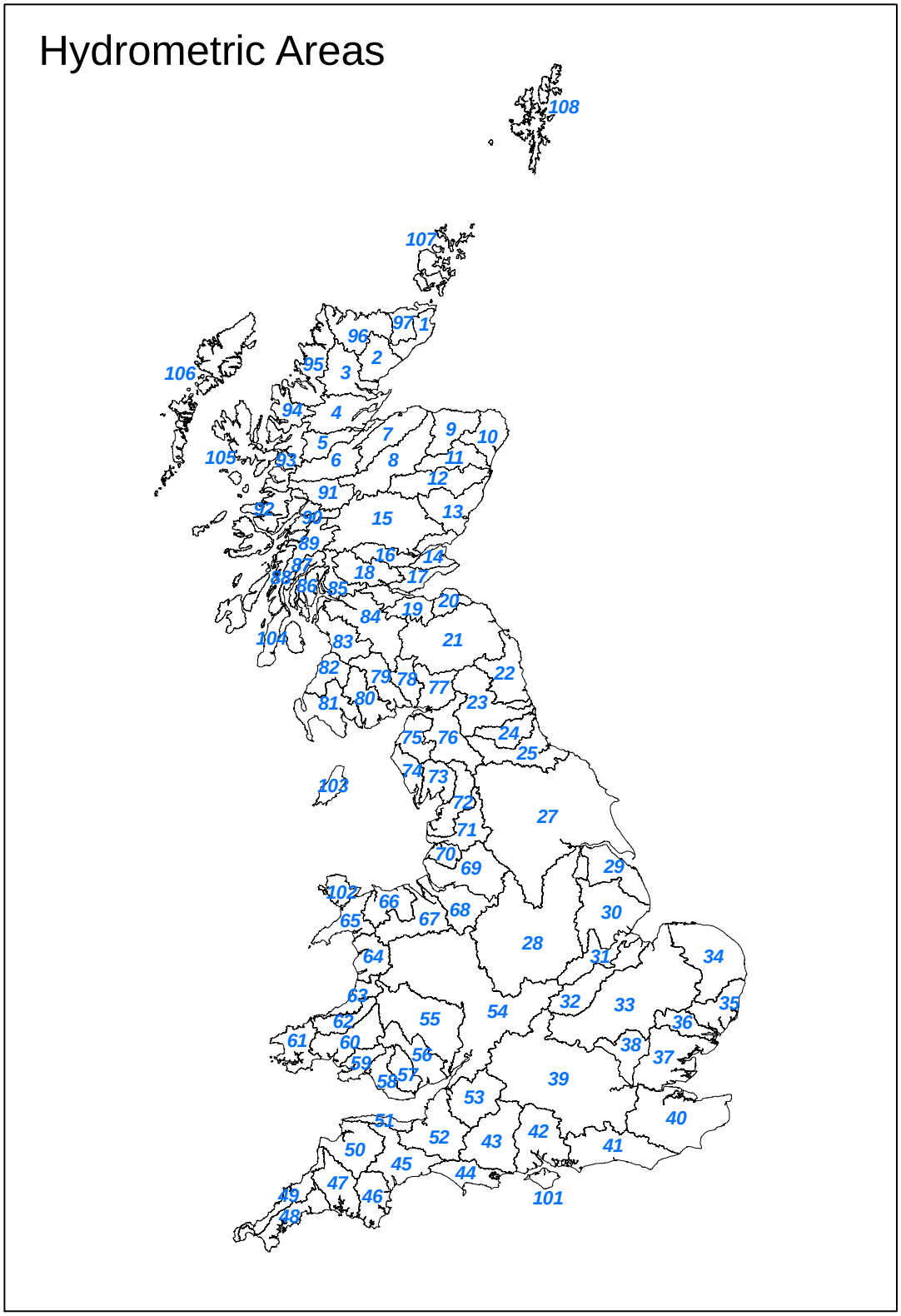
Appendix 7

RIVER_ID	HA	Primary River	Primary Trib	Secondary Trib	Tertiary Trib	Quarternary Trib	Quinary trib	Other Trib	River Name

Date of entry to NID:

Entered By:

Appendix 8. Hydrometric Area Map



Appendix 9. Checklist for Data Manager

On start of new project:
In NID_BackEnd.mdb

- Add new Rivers (Dta_River)
- Add new Sites (Dta_Site)
- Add all sites to Dta_SITE_Data_Entry
- Add Project details (Lkp_Project)
- Add new Sampling method (Lkp_Sampling_method)
- Add new identification level (Lkp_ID_Level)
- Add Sample progress metadata (Mtd_Metadata)
- Add new staff (Lkp_Staff)

In Data Entry System

- Refresh data in Data_entry system
- Identify next sample number for project
- Add new sample numbers if appropriate (or this can be done during data entry)
- Modify forms to add default values or explanatory text
    - Frm_001_Main_Menu: Alter form caption and header label to indicate the project
    - Frm_003_Select_Project_Site: Amend the default value in the PROJECT text box to the appropriate project.
    - Frm_004_Select_Sample: Amend text help in labels for Sampling method and time, ID level and set default if appropriate.
    - Frm_004_Select_Sample: Amend text help in label for Subsidiary number (the project leader will have decided on the relevant coding in advance).
    - Frm_005_Select_Abundance_Type: Amend code in combo box Fam_Sp on After Update event so that values are appropriate (especially where 2 different types are needed for family and species level enumeration).
    - Frm_005_Select_Abundance_Type: On combo box Selected_Abundance_Type alter Row Source sql to select desired abundance system from table Lkp_ABUNDANCE SYSTEM (or set appropriate default value if only one option).
- Modify reports
    - Rpt_Results: Amend title and column headers
    - Rpt_Validation: Amend title and column headers

On completion of data entry:
In NID_BackEnd.mdb

- Update Sample progress metadata (Dta_Metadata)
- Append new data to Back-end tables

In other databases: Refresh data in NID_FrontEnd.mdb and any other satellite databases as required.

Appendix 10. Data Entry Refresh Tables Module

Workflow in Refresh Tables module for data entry system

Delete existing data in Data Entry Database	Transfer equivalent tables from Back-end database with 'Imported---' at start of table name	Append data from Imported tables to empty table structure in Data Entry Database	Delete 'Imported---' tables
DoCmd.OpenQuery "Qry_Del_TaxaDE", acViewNormal, acEdit 'Deletes all in Dta_Taxa_Data_Entry			
DoCmd.OpenQuery "Qry_Del_SampleDE", acViewNormal, acEdit 'Deletes all in Dta_SAMPLE_Data_Entry			
DoCmd.OpenQuery "Qry_Del_ProjSiteDE", acViewNormal, acEdit 'Deletes all in Dta_PROJECT_SITE_Data_Entry			
DoCmd.OpenQuery "Qry_Del_SiteDE", acViewNormal, acEdit 'Deletes all in Dta_SITE_Data_Entry	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_SITE_Data_Entry", "IMPORTED---Dta_SITE_Data_Entry", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_Site_data_entry", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_Site_data_entry"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_Taxa_List_BMWP_Families", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_Taxa_List_BMWP_Families", "IMPORTED---Lkp_Taxa_List_BMWP_Families", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_Taxa_List_BMWP_Families", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_Taxa_List_BMWP_Families"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_STAFF", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_STAFF", "IMPORTED---Lkp_STAFF", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_STAFF", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_STAFF"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_SAMPLING_METHOD", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_SAMPLING_METHOD", "IMPORTED---Lkp_SAMPLING_METHOD", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_SAMPLING_METHOD", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_SAMPLING_METHOD"

Appendix 10

Delete existing data in Data Entry Database	Transfer equivalent tables from Back-end database with 'Imported---' at start of table name	Append data from Imported tables to empty table structure in Data Entry Database	Delete 'Imported---' tables
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_PROJECTS", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_PROJECTS", "IMPORTED---Lkp_PROJECTS", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_PROJECTS", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_PROJECTS"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_ABUNDANCE SYSTEM", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_ABUNDANCE SYSTEM", "IMPORTED---Lkp_ABUNDANCE SYSTEM", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_ABUNDANCE SYSTEM", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_ABUNDANCE SYSTEM"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_ID_LEVEL", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_ID_LEVEL", "IMPORTED---Lkp_ID_LEVEL", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_ID_LEVEL", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_ID_LEVEL"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_cfOldNew_Names", acViewNormal, acEdit 'Deletes all in Lkp_Compare_Old_and_New_Names	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_Compare_Old_and_New_Names", "IMPORTED---Lkp_Compare_Old_and_New_Names", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_cfOldNewNames", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_Compare_Old_and_New_Names"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_CROSSREF", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_CROSSREF", "IMPORTED---Lkp_CROSSREF", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_CROSSREF", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_CROSSREF"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_TAXA", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_TAXA", "IMPORTED---Lkp_TAXA", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_TAXA", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_TAXA"
DoCmd.OpenQuery "DELETE_ALL_IN---Mtd_METADATA", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Mtd_METADATA", "IMPORTED---Mtd_METADATA", False	DoCmd.OpenQuery "APPEND_ALL_IN---Mtd_METADATA", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Mtd_METADATA"
DoCmd.OpenQuery "DELETE_ALL_IN---Dta_RIVER", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_RIVER", "IMPORTED---Dta_RIVER", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_RIVER", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_RIVER"

Appendix 11. Refresh Data Module for all Front-end Databases

Workflow in Mdl_Refresh_front_end_Data for all Front_end databases

Delete existing data in Front end Database(s)	Transfer equivalent tables from Back-end database with 'Imported---' at start of table name	Append data from Imported tables to empty table structure in Front-end Database	Delete 'Imported---' tables
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_STAFF", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_STAFF", "IMPORTED---Lkp_STAFF", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_STAFF", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_STAFF"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_Taxa_List_BMWP_Families", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_Taxa_List_BMWP_Families", "IMPORTED---Lkp_Taxa_List_BMWP_Families", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_Taxa_List_BMWP_Families", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_Taxa_List_BMWP_Families"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_TAXA_LIST_AWIC_F", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_TAXA_LIST_AWIC_F", "IMPORTED---Lkp_TAXA_LIST_AWIC_F", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_TAXA_LIST_AWIC_F", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_TAXA_LIST_AWIC_F"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_CROSSREF", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_CROSSREF", "IMPORTED---Lkp_CROSSREF", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_CROSSREF", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_CROSSREF"
DoCmd.OpenQuery "DELETE_ALL_IN---Mtd_METADATA", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Mtd_METADATA", "IMPORTED---Mtd_METADATA", False	DoCmd.OpenQuery "APPEND_ALL_IN---Mtd_METADATA", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Mtd_METADATA"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_SAMPLING_METHOD", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_SAMPLING_METHOD", "IMPORTED---Lkp_SAMPLING_METHOD", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_SAMPLING_METHOD", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_SAMPLING_METHOD"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_PROJECTS", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_PROJECTS", "IMPORTED---Lkp_PROJECTS", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_PROJECTS", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_PROJECTS"

Appendix 11

Delete existing data in Front end Database(s)	Transfer equivalent tables from Back-end database with 'Imported---' at start of table name	Append data from Imported tables to empty table structure in Front-end Database	Delete 'Imported---' tables
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_ABUNDANCE SYSTEM", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_ABUNDANCE SYSTEM", "IMPORTED---Lkp_ABUNDANCE SYSTEM", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_ABUNDANCE SYSTEM", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_ABUNDANCE SYSTEM"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_DETERMINAND", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_DETERMINAND", "IMPORTED---Lkp_DETERMINAND", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_DETERMINAND", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_DETERMINAND"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_cfOldNew_Names", acViewNormal, acEdit 'Deletes all in Lkp_Compare_Old_and_New_Names	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_Compare_Old_and_New_Names", "IMPORTED---Lkp_Compare_Old_and_New_Names", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_cfOldNewNames", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_Compare_Old_and_New_Names"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_TAXA", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_TAXA", "IMPORTED---Lkp_TAXA", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_TAXA", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_TAXA"
DoCmd.OpenQuery "DELETE_ALL_IN---Lkp_ID_LEVEL", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Lkp_ID_LEVEL", "IMPORTED---Lkp_ID_LEVEL", False	DoCmd.OpenQuery "APPEND_ALL_IN---Lkp_ID_LEVEL", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Lkp_ID_LEVEL"
DoCmd.OpenQuery "DELETE_ALL_IN---Dta_RIVER", acViewNormal, acEdit	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_RIVER", "IMPORTED---Dta_RIVER", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_RIVER", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_RIVER"
Dta_SITE data deleted by cascade delete from Dta_RIVER	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_SITE", "IMPORTED---Dta_SITE", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_SITE", acViewNormal, acEdit Forms!Start_Up_Message_Refreshing_Data.Box31.Visible = True	DoCmd.DeleteObject acTable, "IMPORTED---Dta_SITE"
Dta_PROJECT_SITE data deleted by cascade delete from Dta_RIVER	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_PROJECT_SITE", "IMPORTED---Dta_PROJECT_SITE", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_PROJECT_SITE", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_PROJECT_SITE"
Dta_SAMPLE data deleted by cascade delete from Dta_RIVER	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_SAMPLE", "IMPORTED---Dta_SAMPLE", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_SAMPLE", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_SAMPLE"
Dta_TAXA data deleted by cascade delete from Dta_RIVER	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_TAXA", "IMPORTED---Dta_TAXA", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_TAXA", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_TAXA"

Appendix 11

Delete existing data in Front end Database(s)	**Transfer equivalent tables from Back-end database with 'Imported---' at start of table name**	**Append data from Imported tables to empty table structure in Front-end Database**	**Delete 'Imported---' tables**
Dta_CHEMISTRY data deleted by cascade delete from Dta_RIVER	DoCmd.TransferDatabase acImport, "Microsoft Access", Back_End_and_Path, acTable, "Dta_CHEMISTRY", "IMPORTED---Dta_CHEMISTRY", False	DoCmd.OpenQuery "APPEND_ALL_IN---Dta_CHEMISTRY", acViewNormal, acEdit	DoCmd.DeleteObject acTable, "IMPORTED---Dta_CHEMISTRY"

This page is intentionally left blank