Final Report **Project WFD46**

RIVPACS Database Documentation

February/2007

SNIFFER
SCOTLAND & NORTHERN IRELAND FORUM FOR ENVIRONMENTAL RESEARCH

Dissemination status

Unrestricted

Research contractor

This document was produced by:

John Davy-Bowker
Ralph Clarke
Mike Furse
Cynthia Davies
Tracy Corbin
John Murphy
Nick Kneebone

Centre for Ecology & Hydrology
Winfrith Technology Centre
Dorchester DT2 8ZD
United Kingdom

SNIFFER's project manager

SNIFFER's project manager for this contract is:

Robin Guthrie, Scottish Environment Protection Agency

SNIFFER's project steering group members are:

Robin Guthrie, Scottish Environment Protection Agency
John Murray-Bligh, Environment Agency
Peter Hale, Environment & Heritage Service

SNIFFER
First Floor, Greenside House
25 Greenside Place
EDINBURGH EH1 3AA

www.sniffer.org.uk

EXECUTIVE SUMMARY

WFD46: RIVPACS Database & WFD Screening (February, 2007)

Project funders/partners: SNIFFER

Background to research

With the advent of the EU Water Framework Directive the concept of the 'reference condition' has become explicit within the legislative framework of the European Union. Reference condition has been established as a quality standard against which assessments of biological degradation must be compared. It is therefore essential that Member States can demonstrate that the biological datasets used to define their reference conditions meet the criteria of the WFD. The RIVPACS reference site dataset is therefore central to the definition of reference conditions for macroinvertebrates in streams and rivers in the United Kingdom.

Objectives of research

- To establish the ownership of the RIVPACS reference site dataset
- To liaise with all stakeholders of the dataset to establish unhindered access to the RIVPACS reference site dataset for the UK agencies (in perpetuity)
- To deliver the RIVPACS reference site dataset to the UK agencies and to the public domain in a readily accessible database together will its accompanying physicochemical variables (both existing and newly collated as part of this project), historical and current anthropogenic stress data, and a range of calculated biotic indices

Key findings and recommendations

Ownership of the RIVPACS dataset resides with no single organization and several different organizations consider that they own different portions of the dataset. Formal permissions to release the dataset into the public domain have been obtained from all twelve extant organizations that have been identified as having funded various phases of RIVPACS research. In addition, CEH/NERC has also agreed to release the RIVPACS dataset to the public domain. Terms and conditions relating to the end use of the RIVPACS dataset have now been established. The RIVPACS database has been assembled in Microsoft® Access and can now be downloaded from the CEH web site. This report details the terms and conditions that apply to all end users of the database and it documents the tables given in the database, their structure and the origin of their data. A separate Pressure Data Analysis report describes the screening of the RIVPACS sites in terms of the current and emerging definitions of reference condition.

Key words: RIVPACS database, database documentation, Water Framework Directive

TABLE OF CONTENTS

List of Tables

List of Figures

APPENDICES

Appendix I A sample of the letter sent by CEH seeking formal written consent for release the RIVPACS dataset into the public domain

Appendix II End use terms and conditions set by the stakeholders of the RIVPACS dataset

Appendix III Taxonomy and rules for taxon recording

1. INTRODUCTION

With the advent of the EU Water Framework Directive (Council of the European Communities, 2000) the concept of the 'reference condition' has become explicit within the legislative framework of the European Union. The reference condition has been established as a quality standard against which assessments of biological degradation must be compared. It is therefore essential that Member States can demonstrate that the biological datasets used to define their reference conditions meet the criteria of the WFD. The WFD describes reference conditions as follows:

There are no, or only very minor, anthropogenic alterations to the values of the physicochemical and hydromorphological quality elements for the surface water body type from those normally associated with that type under undisturbed conditions.

The values of the biological quality elements for the surface water body reflect those normally associated with that type under undisturbed conditions, and show no, or only very minor, evidence of distortion.

RIVPACS (River InVertebrate Prediction And Classification System), developed by CEH, the Environment Agency, and their predecessors is the principle methodology used by UK government environment agencies to assess the biological quality of rivers. RIVPACS assesses the biological condition at a test site by comparing the observed fauna with a fauna predicted to occur at the site in the absence of pollution or other significant environmental stress. Biological quality is assessed by in terms of two biotic indices (Number of Taxa and Average Score Per Taxon) expressed as their Observed/ RIVPACS Expected (O/E) ratios, which are equivalent to WFD Ecological Quality Ratios (EQRs).

The RIVPACS reference site dataset is central to the definition of reference conditions for macroinvertebrates in streams and rivers in the UK. The agencies with responsibility for implementation of the WFD, the Environment Agency (in England and Wales), the Scottish Environment Protection Agency (in Scotland) and the Environment and Heritage Service (in Northern Ireland), need to secure unrestricted access to the RIVPACS reference site dataset. This is essential if the agencies are to be able to carry out WFD implementation activities such as intercalibration with other national definitions of reference conditions across Europe and the setting of class boundaries. A further important requirement is that the datasets underpinning the definitions of reference states in the United Kingdom need to be transparent to end users and those that might be charged with remedial works to improve river quality. This project has therefore sought to:

- Establish the ownership of the RIVPACS reference site dataset.
- Liaise with all stakeholders to establish unhindered access to the RIVPACS reference site dataset for the UK agencies (in perpetuity).
- Deliver the RIVPACS reference site dataset to the UK agencies and to the public domain in a readily accessible database together will its accompanying physicochemical variables, historical and current anthropogenic stress data, and a range of calculated biotic indices.

The ownership of the RIVPACS dataset has always been complicated and difficult to establish. This project has therefore set out to seek written permission to place the dataset in the public domain from all those extant organisations that have either provided

funding support to the various phases of RIVPACS data collection or might have legitimate claim to the intellectual property rights of the dataset.

This project has also provided funding to support transfer the RIVPACS dataset from the CEH National Invertebrate Database to a separate and devoted RIVPACS database that can be made readily available from an Internet download page. The project has also funded the collation of additional data that might assist in the screening of the RIVPACS sites in terms of the new definitions of reference condition that arising from the implementation of the Water Framework Directive (e.g. the definitions of reference condition set out in the Framework Directive itself, the more recent REFCOND guidance notes (Wallin et al., 2005) and the work going on in the Geographical Intercalibration Groups – GIGs). While the exact definition of reference condition in terms of physicochemical, hydrological and hydromorphological conditions has still to be established, this project set out to collate all existing data and new data that might assist in the screening of the RIVPACS sites in terms of the anticipated new definitions of reference condition.

The overall twin aims of this project have therefore been:

- To establish permission to release the data to the public domain and to transfer the existing data to a new RIVPACS database (reported here)

- To collate new physicochemical, hydrological and hydromorphological data and add these to the database and to perform analyses to assess the quality of the reference sites in terms of the WFD definitions of reference state (this aspect of the project is reported in a separate pressure data analysis report)

2. PERMISSIONS

The 835 current RIVPACS sites were sampled between March 1978 and May 2002 in a series of phases, many of which were supported by different combinations of funding bodies. The collection and analysis of the various phases of RIVPACS samples has also been supported by varying proportions of external and internal (CEH/NERC) funding. A list of extant organisations that we have collectively regarded at stakeholders in terms of seeking permission to release the dataset to the public domain is given below. Organisation names are shown in italics for clarity. Organisation names are given firstly as the names they had at the time they provided support and their current names and intermediate names are also given chronologically in brackets:

- *Central Water Planning Unit* (became part of the *Department of the Environment*, then the *Department of the Environment, Trade and the Regions*, and is now part of the *Department for Environment, Food and Rural Affairs*)

- *Department of the Environment* (became part of the *Department of the Environment, Trade and the Regions*, and is now part of the *Department for Environment, Food and Rural Affairs*)

- *Department of the Environment (Northern Ireland)*

- *Environment and Heritage Service* - an Agency within the *Department of the Environment (Northern Ireland)*

- *Freshwater Biological Association*

- *Industrial Research and Technology Unit* - an Agency within the *Department of Economic Development (Northern Ireland)*
- *Manders, Raikes, Marshall* (no longer extant)
- *National Rivers Authority* (became the *Environment Agency*)
- *Natural Environment Research Council*
- *Nature Conservancy Council* (now *English Nature*)
- *Nature Conservancy Council for Scotland* (now *Scottish Natural Heritage*)
- *Scottish Development Department* (became part of the *Scottish Office*, which is now part of the *Scottish Executive*)
- *Scottish Environment Protection Agency*
- *Scottish Office* (now part of the *Scottish Executive*)
- *South West Water plc*
- *Welsh Office* (now part of the *Welsh Assembly Government*)

These 16 organisations can be reduced to a list of 12 current extant organisations (Table 1). These are the organisations from which CEH sought formal written consent to release the RIVPACS dataset into the public domain. CEH wrote to each of the organisations in Table 1 after first making contact with individuals within each organisation who were willing to deal with our request. A sample of the letter we wrote to each organisation is given in Appendix 1.

Table 1 - Organisations contacted by CEH to release the RIVPACS dataset

Organisation
Freshwater Biological Association (FBA)
Environment Agency (EA)
Scottish Environment Protection Agency (SEPA)
Environment and Heritage Service (EHS)
Department for Environment Food and Rural Affairs (DEFRA)
Welsh Assembly Government (WA)
Scottish Executive (SE)
English Nature (EN)
Scottish Natural Heritage (SNH)
Countryside Council for Wales (CCW)
South West Water (SSW)

In our contact with the various organisations (Table 1) it was clear that ownership of the dataset resides with no single organisation and that several different organisations consider that they 'own' different portions of the dataset where they have provided some funding support. We have therefore sought to obtain permissions from all of the organisations in Table 1 to release the dataset into the public domain so that each may set out their own conditions on its subsequent end use.

All of the organisations in Table 1 provided formal written consent to release the RIVPACS dataset into the public domain, although several set out particular terms and conditions. These terms and conditions are given in full in Appendix II and also on the web page for disseminating the database (see section 4). The original copies of these formal letters of consent are stored in the River Communities research group at CEH Dorset and copies have also been copied to SNIFFER.

In addition, CEH/NERC engaged in its own internal enquiries to secure it's own formal permission to release the RIVPACS dataset and has similarly agreed to allow the dataset to be released into the public domain (terms and conditions are given in Appendix II and on the web page for disseminating the database.

The permissions from all relevant organisations to release the RIVPACS dataset into the public domain are now therefore regarded as having been obtained and organisations wishing to utilise the RIVPACS dataset in a particular way should consult the terms and conditions set out in Appendix II. These terms and conditions are also reproduced on the RIVPACS download page (section 4) and on the first form of the RIVPACS database).

3. DATABASE TABLES

The RIVPACS database has been assembled in Microsoft® Access 2000 and consists of 25 tables storing a variety of site, sample and macroinvertebrate data together with a wide range of associated environmental data and biotic index values, some of which have been newly collated for this project. Where appropriate the tables have been linked together using MS Access relationships and referential integrity (Figure 1). The structure of the database can be most easily understood by first considering the sites table. This table contains 835 records (one for each RIVPACS reference site) together with columns storing data that is relevant to each site. The samples table is linked to the sites table by the RIVPACS site identification number (the field Site ID) and contains 2,505 records (3x835) representing the samples from spring, summer and autumn that have been collected at each site. Linked to the samples table is the Taxa (all) table. This table stores 178,597 records of macroinvertebrate taxa recorded in each sample. The site, sample and taxa tables therefore form the backbone of the RIVPACS database and other tables are linked to these where appropriate. The structure of each table and the origin of the data in the tables is described in the sections that follow together with the precise meaning and units for each field (column).

3.1 Sites

The sites table (Table 2) gives details of the 835 sites that belong to the various RIVPACS models. The first few columns contain the country, RIVPACS site number, river name and site name. Next there are various spatial referencing fields (National Grid Reference, latitude and longitude and easting and northing). There then follows a set of fields containing RIVPACS variables (Mean Air Temp to Mean Substratum). The next four fields identify which of the current RIVPACS models use each site (some sites are used in both the RIVPACS III+ GB and Highlands models). The next 4 fields identify the Water Framework Directive (WFD) System-A stream types for each site. The last six fields store the agency (EA, SEPA and EHS) stretch or site codes that have been linked to the RIVPACS sites by CEH in order to extract data from the respective datasets. Table 2 gives the names of the fields in the sites table, their data type (e.g. text, number, date), and a description of each field. These descriptions are also entered in the properties of each field in the database and can be viewed in the bottom left corner of the MS Access window when the cursor is moved into a field in a table.

Figure 1 - RIVPACS database relationships diagram

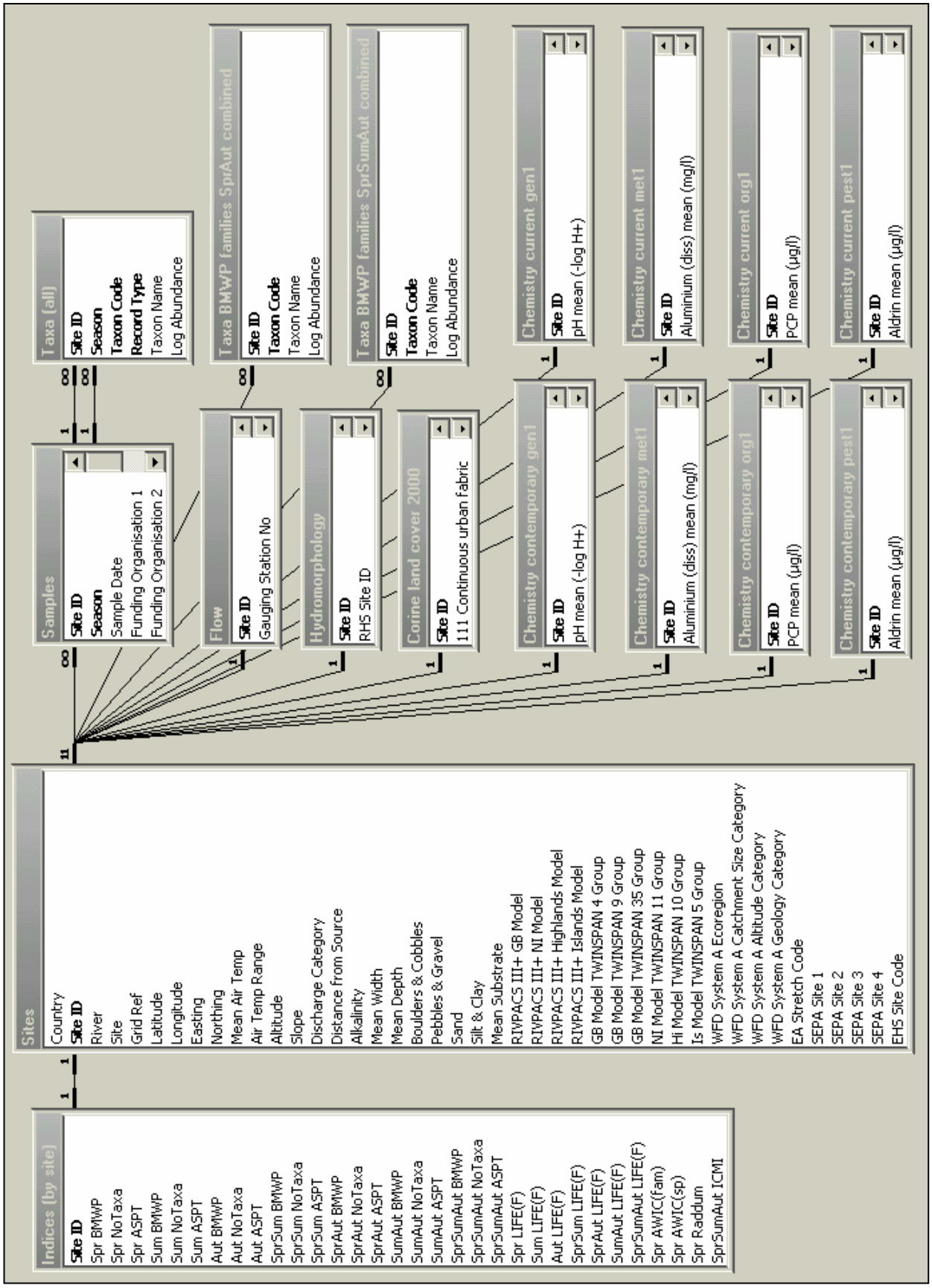


Table 2 - The site table

Field	Type	Description
Country	Text	Country Name
Site ID	Text	Site Number
River	Text	River Name
Site	Text	Site Name
Grid Ref	Text	National Grid Reference
Latitude	Text	Latitude (Decimal Degrees)
Longitude	Text	Longitude (Decimal Degrees)
Easting	Text	Easting
Northing	Text	Northing
Mean Air Temp	Text	Mean Air Temperature (Degrees Celsius)
Air Temp Range	Text	Air Temperature Range (Degrees Celsius)
Altitude	Number	Altitude of site (m)
Slope	Number	Slope of site (m/km)
Discharge Category	Number	Discharge Category in Cubic Metres/Sec (cumecs): 1=<0.31, 2=0.31-0.62, 3=0.62-1.25, 4=1.25-2.5, 5=2.5-5, 6=5-10, 7=10-20, 8=20-40, 9=40-80, 10=80-160
Distance from Source	Number	Distance from Source (km)
Catchment Area	Number	Catchment Area (km2) from ArcMap GIS
Alkalinity	Number	Alkalinity (mg/l CaCO3)
Mean Width	Number	Mean Width of site (m)
Mean Depth	Number	Mean Depth of site (cm)
Boulders & Cobbles	Number	% Boulders & Cobbles (>64mm)
Pebbles & Gravel	Number	% Pebbles & Gravel (2mm - 64mm)
Sand	Number	% Sand (0.06mm - 2mm)
Silt & Clay	Number	% Silt & Clay (<0.06mm)
Mean Substratum	Number	Mean Substratum composition (Phi scale)
RIVPACS III+ GB Model	Text	Site part of RIVPACS III+ GB Model
RIVPACS III+ NI Model	Text	Site part of RIVPACS III+ NI Model
RIVPACS III+ Highlands Model	Text	Site part of RIVPACS III+ Highlands Model
RIVPACS III+ Islands Model	Text	Site part of RIVPACS III+ Islands Model
GB Model TWINSPAN 4 Group	Text	RIVPACS III+ GB Model, TWINSPAN End Group (at the 4 Group Level)
GB Model TWINSPAN 9 Group	Text	RIVPACS III+ GB Model, TWINSPAN End Group (at the 9 Group Level)
GB Model TWINSPAN 35 Group	Text	RIVPACS III+ GB Model, TWINSPAN End Group (at the 35 Group Level)
NI Model TWINSPAN 11 Group	Text	RIVPACS III+ NI Model, TWINSPAN End Group (at the 11 Group Level)
Hi Model TWINSPAN 10 Group	Text	RIVPACS III+ Scot. Highlands Model, TWINSPAN End Group (at the 10 Group Level)
Is Model TWINSPAN 5 Group	Text	RIVPACS III+ Scot. Islands Model, TWINSPAN End Group (at the 5 Group Level)
WFD System A Ecoregion	Text	Water Framework Directive, WFD System A, Ecoregion

WFD System A Catchment Size Category	Text	Water Framework Directive, WFD System A, Catchment Size Category (derived by the UK Agencies): Small 10-100km2; Medium >100-1,000km2; Large >1,000-10,000km2. NB Very small = <10km2 (outside WFD)
WFD System A Altitude Category	Text	Water Framework Directive, WFD System A, Altitude Category (derived by the UK Agencies): Low <200m; Medium 200-800m; Large >800m
WFD System A Geology Category	Text	Water Framework Directive, WFD System A, Geology Category (derived by the British Geological Survey for the UK Agencies)
EA Stretch Code	Text	Linked Environment Agency Stretch Code
SEPA Site 1	Text	Linked Scottish Environment Protection Agency Site Code
SEPA Site 2	Text	Linked Scottish Environment Protection Agency Site Code
SEPA Site 3	Text	Linked Scottish Environment Protection Agency Site Code
SEPA Site 4	Text	Linked Scottish Environment Protection Agency Site Code
EHS Site Code	Text	Linked Environment & Heritage Service Site Code

3.2 Samples

The samples table contains 2,505 records, one for each of the samples collected in spring, summer and autumn at each RIVPACS reference site together with the date the samples were collected (Table 3). This table also shows the organisations that have provided funding support for particular batches of RIVPACS samples (three fields) and those organisations that also have a potential claim to intellectual property rights on batches of samples that are part of the current RIVPACS dataset (2 columns). All extant organisations represented in this table were consulted to obtain permissions to release the dataset into the public domain (see section 2).

Table 3 - The sample table

Field	Type	Description
Site ID	Text	Site Number
Season	Text	Season in which the macroinvertebrate sample was taken
Sample Date	Date	Date macroinvertebrate sample taken
Funding Organisation 1	Text	Funding Organisation (1)
Funding Organisation 2	Text	Funding Organisation (2)
Funding Organisation 3	Text	Funding Organisation (3)
Additional IPR 1	Text	Organisation with additional (IPR 1)
Additional IPR 2	Text	Organisation with additional (IPR 2)

3.3 Taxa (all)

The taxa (all) table contains the macroinvertebrate raw data from each of the 2,505 spring, summer and autumn RIVPACS samples (Table 4). This table has the largest number of records in the RIVPACS database (178,597).

Table 4 - The Taxa (all) table

Field	Type	Description
Site ID	Text	Site Number
Season	Text	Season in which the macroinvertebrate sample was taken
Taxon Code	Text	Code for the macroinvertebrate taxon
Record Type	Text	Type of Record (PA = Presence/Absence; IFELG = Log10 Abundance categories)
Taxon Name	Text	Name of the macroinvertebrate taxon
Log Abundance	Number	Log10 abundance of the taxon (only for Record Type IFELG)

Taxa in RIVPACS samples were recorded at two levels, 'family' and 'species'. The families are denoted with the record type 'IFELG' while species records are denoted 'PA'. Most families are recorded simply at family level, however families that are part of the BMWP scoring system are recorded at BMWP 'family' level. For example, taxa such as Rhyacophilidae and Glossosomatidae are aggregated into a single record for the taxon "Rhyacophilidae (incl. Glossosomatidae)". Families recorded in this way are known as artificial taxon groups (see italicised taxa in Table 5). Family level records are also recorded with $\log_{10}$ abundances where 1=1-9 individuals, 2=10-99, 3=100-999, 4=1000-9999 and 5=≥10000. Species records are recorded at presence/absence level so that the presence of a record indicates that that species was present in the sample. All taxa (both family and species level) are given with both a taxon name and code. The coding system is based on the Maitland coding system (for further details see Appendix III).

Table 5 - Families recorded in the RIVPACS database (artificial groups italicised)

Spongillidae	Baetidae	Philopotamidae
Hydridae	Heptageniidae	Polycentropodidae
Dendrocoelidae	Leptophlebiidae	Hydropsychidae
Planariidae (incl. Dugesiidae)	Potamanthidae	*Psychomyiidae (incl. Ecnomidae)*
Nematomorpha	Ephemeridae	Phryganeidae
Chordodidae	Ephemerellidae	Brachycentridae
Nematoda	Caenidae	Lepidostomatidae
Ectoprocta	Taeniopterygidae	Limnephilidae
Gastropoda	Nemouridae	Goeridae
Neritidae	Leuctridae	Beraeidae
Viviparidae	Capniidae	Sericostomatidae
Valvatidae	Perlodidae	Odontoceridae
Hydrobiidae (incl. Bithyniidae)	Perlidae	Molannidae
Physidae	Chloroperlidae	Leptoceridae
Lymnaeidae	Platycnemididae	Lepidoptera
Planorbidae	Coenagriidae	Pyralidae
Ancylidae (incl. Acroloxidae)	Calopterygidae	Tipulidae
Succineidae	Gomphidae	Psychodidae
Zonitidae	Cordulegasteridae	Ptychopteridae
Margaritiferidae	Aeshnidae	Dixidae
Unionidae	Libellulidae	Chaoboridae
Sphaeriidae	Mesovelidae	Culicidae
Dreissenidae	Hydrometridae	Thaumaleidae
Aeolosomatidae	Veliidae	Ceratopogonidae

Lumbriculidae	Gerridae	Simuliidae
Haplotaxidae	Nepidae	Tanypodinae
Enchytraeidae	Naucoridae	Diamesinae
Naididae	Aphelocheiridae	Prodiamesinae
Tubificidae	Notonectidae	Orthocladiinae
Lumbricidae	Corixidae	Chironomini
Piscicolidae	Haliplidae	Tanytarsini
Glossiphoniidae	Gyrinidae	Bibionidae
Hirudinidae	*Dytiscidae (incl. Noteridae)*	Stratiomyidae
Erpobdellidae	*Hydrophilidae (incl. Hydraenidae)*	Rhagionidae
Hydracarina	Scirtidae	Tabanidae
Cladocera	Dryopidae	Empididae
Ostracoda	Elmidae	Dolichopodidae
Copepoda	Chrysomelidae	Syrphidae
Argulidae	Curculionidae	Sciomyzidae
Astacidae	Sialidae	Ephydridae
Asellidae	Osmylidae	Muscidae
Corophiidae	Sisyridae	Lonchopteridae
Gammaridae (incl. Crangonyctidae & Niphargidae)	Hydroptilidae	
Siphlonuridae	*Rhyacophilidae (incl. Glossosomatidae)*	

3.4 Taxa BMWP families SprAut combined

The table Taxa BMWP families SprAut combined contains only BMWP family level data that has been aggregated into spring and autumn sample data combined. These data have been generated by converting the spring and autumn sample family level data from the table Taxa (all) into BMWP families only and then by combining the separate season samples and taking the maximum (highest) $\log_{10}$ abundance in either season as the combined abundance. For example, the families Heptageniidae and Caenidae at site 0007:

Site ID	Season	Taxon Code	Record Type	Taxon Name	Log Abundance
0007	Autumn	40130000	IFELG	Heptageniidae	2
0007	Autumn	40510000	IFELG	Caenidae	1
0007	Spring	40130000	IFELG	Heptageniidae	2
0007	Spring	40510000	IFELG	Caenidae	2

Spring and autumn combined $\log_{10}$ abundance of Heptageniidae = 2
Spring and autumn combined $\log_{10}$ abundance of Caenidae = 2

Readily available RIVPACS reference site spring and autumn combined BMWP family level data are particularly useful for the EA, SEPA and EHS because these agencies currently base their assessments of the biological condition of rivers on spring and autumn sampling. The fields in the table Taxa BMWP families SprAut combined are described in detail in Table 6.

Table 6 - The Taxa BMWP families SprAut combined table

Field	Type	Description
Site ID	Text	Site Number
Taxon Code	Text	Code for the macroinvertebrate taxon
Taxon Name	Text	Name of the macroinvertebrate taxon
Log Abundance	Number	Log10 abundance of the taxon

3.5 Taxa BMWP families SprSumAut combined

The table Taxa BMWP families SprSumAut combined contains only BMWP family level data that has been aggregated into spring, summer and autumn sample data combined. These data have been generated by converting the spring, summer and autumn sample family level data from the table Taxa (all) into BMWP families only and then by combining the separate season samples. However, in the case of three seasons combined data a slightly different rule is used for calculating abundances. This is known as the max+1 rule and it is used for consistency with the biological classifications that underpin the RIVPACS models (which also use this rule). Max+1 abundance is calculated by taking the maximum $\log_{10}$ abundance in either season as the combined abundance, except where a family occurs in all three seasons and all three abundances are the same, in which case the abundance is increased by 1 $\log_{10}$ unit. For example using the same example site as in section 3.4, the families Baetidae and Heptageniidae at site 0007:

Site ID	Season	Taxon Code	Record Type	Taxon Name	Log Abundance
0007	Spring	40120000	IFELG	Baetidae	3
0007	Spring	40130000	IFELG	Heptageniidae	2
0007	Summer	40120000	IFELG	Baetidae	2
0007	Summer	40130000	IFELG	Heptageniidae	2
0007	Autumn	40120000	IFELG	Baetidae	3
0007	Autumn	40130000	IFELG	Heptageniidae	2

Spring and autumn combined $\log_{10}$ abundance of Baetidae = 3
Spring and autumn combined $\log_{10}$ abundance of Heptageniidae = 3

The fields in the table Taxa BMWP families SprSumAut combined are the same as those given in Table 6.

3.6 Indices (by site)

The table Indices (by site) contains a selection of observed values of biotic index values that have been calculated for each site, and where appropriate, in each separate season and season combination (Table 7). The three BMWP indices (BMWP, Number of Taxa and ASPT) and the family level LIFE index are given in all separate seasons and season combinations. The acidity indices AWIC(fam), AWIC(sp) and Raddum are calculated only for spring samples. Both AWIC indices are specifically designed for use with spring samples and Raddum is also only given for spring because this is the season in which the impact of acidity on the macroinvertebrate community is generally considered to be most apparent. Also included in the table are the Intercalibration Common Metric index (ICMi) values calculated by John Murray-Bligh (EA) using ASTERICS software.

Table 7 - The Indices table

Field	Type	Description
Site ID	Text	Site Number
Spr BMWP	Number	Spring – BMWP[1]
Spr NTaxa	Number	Spring - Number of Taxa (Number of BMWP Scoring Families)[1]
Spr ASPT	Number	Spring - Average Score Per Taxon (BMWP/NTaxa)[1]
Sum BMWP	Number	Summer - BMWP[1]
Sum NTaxa	Number	Summer - Number of Taxa (Number of BMWP Scoring Families)[1]
Sum ASPT	Number	Summer - Average Score Per Taxon (BMWP/NTaxa)[1]
Aut BMWP	Number	Autumn - BMWP[1]
Aut NTaxa	Number	Autumn - Number of Taxa (Number of BMWP Scoring Families)[1]
Aut ASPT	Number	Autumn - Average Score Per Taxon (BMWP/NTaxa)[1]
SprSum BMWP	Number	Spring & Summer combined - BMWP[1]
SprSum NTaxa	Number	Spring & Summer combined - Number of Taxa (Number of BMWP Scoring Families) [1]
SprSum ASPT	Number	Spring & Summer combined - Average Score Per Taxon (BMWP/NTaxa) [1]
SprAut BMWP	Number	Spring & Autumn combined - BMWP[1]
SprAut NTaxa	Number	Spring & Autumn combined - Number of Taxa (Number of BMWP Scoring Families)[1]
SprAut ASPT	Number	Spring & Autumn combined - Average Score Per Taxon (BMWP/NTaxa)[1]
SumAut BMWP	Number	Summer & Autumn combined - BMWP[1]
SumAut NTaxa	Number	Summer & Autumn combined - Number of Taxa (Number of BMWP Scoring Families)[1]
SumAut ASPT	Number	Summer & Autumn combined - Average Score Per Taxon (BMWP/NTaxa)[1]
SprSumAut BMWP	Number	Spring, Summer & Autumn combined - BMWP[1]
SprSumAut NTaxa	Number	Spring, Summer & Autumn combined - Number of Taxa (Number of BMWP Scoring Families)[1]
SprSumAut ASPT	Number	Spring, Summer & Autumn combined- Average Score Per Taxon (BMWP/NTaxa)[1]
Spr LIFE(F)	Number	Spring - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
Sum LIFE(F)	Number	Summer - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
Aut LIFE(F)	Number	Autumn - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
SprSum LIFE(F)	Number	Spring & Summer - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
SprAut LIFE(F)	Number	Spring & Autumn - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
SumAut LIFE(F)	Number	Summer & Autumn - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
SprSumAut LIFE(F)	Number	Spring, Summer & Autumn - Lotic-invertebrate Index for Flow Evaluation (Family)[2]
Spr AWIC(fam)	Number	Spring - Acid Water Indicator Community (Family Level) Index[3]
Spr AWIC(sp)	Number	Spring - Acid Water Indicator Community (Species Level) Index[4]
Spr Raddum	Number	Spring - Raddum Index[5]
SprAut ICMI indices	Number	Spring & Autumn combined - Intercalibration Common Metrics - calculated by John Murray-Bligh using ASTERICS software. Based on $\log_{10}$ family level data assigned abundances 3, 30, 300, 3000 & 30000 then summed across spring and autumn[6]

1 - Armitage et al. (1983) Water Research 17: 333-347
2 - Extence et al (1999) Regulated Rivers: Research & Management 15: 543-574
3 - Davy-Bowker et al. (2005) Archiv für Hydrobiologie 163: 383-403
4 - Murphy et al. (unpublished)
5 - Fjellheim & Raddum (1990) Science of the Total Environment 96: 57-66
6 - Intercalibration Common Metrics - calculated by John Murray-Bligh using ASTERICS software

3.7 Indices (means by TWINSPAN group)

The table Indices (means by TWINSPAN group, Great Britain 614), and the three tables that follow, give the mean values of the observed biotic index values for each of the TWINSPAN end groups that form part of the four current RIVPACS predictive models (Great Britain, Northern Ireland, Scottish Highlands and Scottish Islands). The TWINSPAN classification underpinning the 614 site Great Britain model has 35 TWINSPAN end groups. The same TWINSPAN classification is also sometimes used at the coarser 9-end group and 4-end group levels. For the Great Britain table, mean biotic index values have therefore been calculated for the 4, 9 and 35 group TWINSPAN levels, and for all seasons and season combinations (where appropriate – see section 3.6 above). The Northern Ireland, Scottish Highlands and Scottish Islands RIVPACS models all use single TWINSPAN classification levels (11-group, 10-group and 5 group respectively). A table showing the fields in the Indices (means by TWINSPAN group, Great Britain 614) table is given below (Table 8). The Northern Ireland, Scottish Highland and Scottish Islands versions of this table all have similar structures.

Table 8 - The Indices (means by TWINSPAN group, Great Britain 614) table

Field	Type	Description
Season	Text	Season in which the macroinvertebrate sample was taken
TWINSPAN Classification	Text	RIVPACS III+ GB Model, TWINSPAN Classification Group Level (4 Group, 9 Group or 35 Group)
End Group	Number	RIVPACS III+ GB Model, TWINSPAN End Group
BMWP	Number	BMWP[1]
No Taxa	Number	Number of Taxa (Number of BMWP Scoring Families)[1]
ASPT	Number	Average Score Per Taxon (Average BMWP Score per BMWP Scoring Family) [1]
LIFE(F)	Number	Lotic-invertebrate Index for Flow Evaluation (Family)[2]
AWIC(fam)	Number	Acid Water Indicator Community (Family Level) Index[3]
AWIC(sp)	Number	Acid Water Indicator Community (Species Level) Index[4]
Raddum	Number	Raddum Index[5]

1 - Armitage et al. (1983) Water Research 17: 333-347
2 - Extence et al (1999) Regulated Rivers: Research & Management 15: 543-574
3 - Davy-Bowker et al. (2005) Archiv für Hydrobiologie 163: 383-403
4 - Murphy et al. (unpublished)
5 - Fjellheim & Raddum (1990) Science of the Total Environment 96: 57-66

3.8 Indices (means by WFD System-A types)

The table Indices (means by WFD System-A Types, Great Britain 614), and the three tables that follow, give the mean values of the observed biotic index values for each of the Water Framework Directive system-A stream types that are given in the sites table. Within the respective WFD Ecoregions (Great Britain or Ireland), WFD system-A streams types are derived by a combination of geology, altitude and catchment size which are categorised as follows:

Geology (calcareous, siliceous or organic)
Altitude (Low <200m; Medium 200-800m; High >800m)
Catchment size (Small 10-100km^2; Medium >100-1,000km^2; Large >1,000-10,000km^2)

Geology categories were derived by the British Geological Survey for the UK Agencies. Altitude and catchment size categories were derived by the EA, SEPA and EHS (supplemented by CEH map work for a small number of sites where these categories could not be calculated). Seven of the 835 RIVPACS reference sites were judged to have catchment sizes smaller than the WFD 10-100km^2 category – denoted 'Very small' in the sites table (these sites were not included in the calculation of mean index values

for WFD system-A stream types). As in the tables presenting mean biotic index values within TWINSPAN end groups, mean biotic index values were calculated for all appropriate separate seasons and season combinations (see section 3.6 above). Details of the fields used in the mean biotic index tables for WFD system-A stream types are given in Table 9 and are the same for all four tables (Great Britain, Northern Ireland, Scottish Highlands and Scottish Islands).

Table 9 - The Indices (means by WFD System-A types, Great Britain 614) table

Field	Type	Description
Season	Text	Season in which the macroinvertebrate sample was taken
Geology	Text	Water Framework Directive, WFD System A, Geology Category (derived by the British Geological Survey for the UK Agencies)
Altitude	Text	Water Framework Directive, WFD System A, Altitude Category (derived by the UK Agencies). Low <200m; Medium 200-800m; High >800m
Catchment Size	Text	Water Framework Directive, WFD System A, Catchment Size Category (derived by the UK Agencies). Small 10-100km2; Medium >100-1,000km2; Large >1,000-10,000km2
BMWP	Number	BMWP[1]
No Taxa	Number	Number of Taxa (Number of BMWP Scoring Families)[1]
ASPT	Number	Average Score Per Taxon (Average BMWP Score per BMWP Scoring Family) [1]
LIFE(F)	Number	Lotic-invertebrate Index for Flow Evaluation (Family)[2]
AWIC(fam)	Number	Acid Water Indicator Community (Family Level) Index[3]
AWIC(sp)	Number	Acid Water Indicator Community (Species Level) Index[4]
Raddum	Number	Raddum Index[5]

1 - Armitage et al. (1983) Water Research 17: 333-347
2 - Extence et al (1999) Regulated Rivers: Research & Management 15: 543-574
3 - Davy-Bowker et al. (2005) Archiv für Hydrobiologie 163: 383-403
4 - Murphy et al. (unpublished)
5 - Fjellheim & Raddum (1990) Science of the Total Environment 96: 57-66

3.9 Chemistry contemporary/current gen1

The tables Chemistry contemporary/current gen1, Chemistry contemporary/current met1, Chemistry contemporary/current org1 and Chemistry contemporary/current pest1 present chemical data (where available) for each of the 835 RIVPACS reference sites. These data were collated from a variety of sources, some of which were held by CEH and some of which have been newly collated from EA, SEPA and EHS datasets supplied specifically for this project. The file names of the data sources used in these tables are given in Table 10 below.

Table 10 - Sources of data used in the chemical tables

CEH Vax files from RIVPACS phases I, II and III
CEH database - H&I Data Entered from MTF Paper Files.mdb
SEPA_NCW_database_v2.2
CEH spreadsheet GQA_CHEMISTRY_DATA_88-90_93-95.mdb
CEH spreadsheet IFE_Data.xls
CEH spreadsheet GQA_1995_data.mdb
CEH spreadsheet GQA_2002_data.mdb
Environment and Heritage Service data from 1986-1995
Environment and Heritage Service data from 1986-1995 (Garvary Road)
Environment and Heritage Service data from 1987-1997
Environment and Heritage Service data from 1996-2005

There are eight chemical data tables in all. The tables store four types of chemical data - the general determinands such as pH, Oxygen, BOD, nutrients etc (gen1), the metals (met1), the organic chemicals (org1) and pesticides (pest1). The tables are further differentiated in to contemporary and current tables. Contemporary indicates that where time series data were available from one or more sources, the data were selected that most closely matched the sampling date of each RIVPACS sampling site (although much of this may actually be current data where this was the oldest available). Current indicates that data were selected that most closely matched the current (2005) date. For each chemical determinand a mean value is given for each site. The number of samples used to generate the mean (n) is also given, together with the sampling dates between which the chemical samples contributing to the mean were collected (these dates were estimated for some of the older data). The origin of the data (see Table 10 above) is also shown in the origin fields for each determinand. The fields in the tables Chemistry contemporary gen1 and Chemistry current gen1 are given in Table 11 below (the fields n, from, to and origin apply to all determinands but for brevity are shown for here for pH only).

Table 11 - The Chemistry contemporary/current gen1 tables

Field	Type	Description
Site ID	Text	Site Number
pH mean (-log H+)	Number	Mean pH (-log H+)
pH n	Number	Number of pH measurements used to calculate mean
pH from	Date	Date of first pH measurement
pH to	Date	Date of last pH measurement
pH origin	Text	Origin of pH measurements
Total Alkalinity mean (mg/l CaCO3)	Number	Mean Total Alkalinity (mg/l CaCO3)
Total Hardness mean (mg/l CaCO3)	Number	Mean Total Hardness (mg/l CaCO3)
Calcium mean (mg/l Ca)	Number	Mean Calcium (mg/l Ca)
Conductivity 20˚C mean (µS/cm)	Number	Mean Conductivity 20˚C (µS/cm)
Conductivity 25˚C mean (µS/cm)	Number	Mean Conductivity 25˚C (µS/cm)
Dissolved Oxygen mean (mg/l)	Number	Mean Dissolved Oxygen (mg/l)
Oxygen mean (%sat)	Number	Mean Oxygen (%sat)
BOD mean (mg/l O2)	Number	Mean BOD (mg/l O2)
BOD ATU mean (mg/l O2)	Number	Mean BOD ATU (mg/l O2)
Soluble Reactive Phosphorus mean (mg/l P)	Number	Mean Soluble Reactive Phosphorus (mg/l P)
Total Phosphorus mean (mg/l P)	Number	Mean Total Phosphorus (mg/l P)
Nitrate mean (mg/l N)	Number	Mean Nitrate (mg/l N)
Nitrite mean (mg/l N)	Number	Mean Nitrite (mg/l N)
Free & Saline Ammonia mean (mg/l N)	Number	Mean Free & Saline Ammonia (mg/l N)
Ammonia non ionic mean (mg/l NH3)	Number	Mean Ammonia non ionic (mg/l NH3)
Total Oxidised Nitrogen mean (mg/l N)	Number	Mean Total Oxidised Nitrogen (mg/l N)
Kjeldahl Nitrogen mean (mg/l)	Number	Mean Kjeldahl Nitrogen (mg/l)
Dissolved Silicate mean (mg/l SiO2)	Number	Mean Dissolved Silicate (mg/l SiO2)
Dissolved Sulphate mean (mg/l SO4)	Number	Mean Dissolved Sulphate (mg/l SO4)
Dissolved Chloride mean (mg/l Cl)	Number	Mean Dissolved Chloride (mg/l Cl)
Dissolved Fluoride mean (mg/l F)	Number	Mean Dissolved Fluoride (mg/l F)
Suspended Solids mean (mg/l 105˚C)	Number	Mean Suspended Solids (mg/l 105˚C)
Suspended Solids non volatile mean (mg/l 500˚C)	Number	Mean Suspended Solids non volatile (mg/l 500˚C)
Turbidity FTU mean	Number	Mean Turbidity FTU

For brevity the fields n, from, to and origin are shown for pH only

In the building of the chemical tables considerable care and attention had to be devoted to conversion and standardisation of the units of each determinand from different data sources. The fields storing chemical determinand means also include details of the units (bracketed part of the field name). Each determinand in all eight chemical data tables has also examined by scatter plots to check for extreme values that high indicate orders of magnitude errors in units. This process has resulted in several doubtful wild values being removed from the data used to calculate the mean chemical values.

3.10 Chemistry contemporary/current met1

The tables Chemistry contemporary/current met1 store metals data (differentiated into dissolved or total values) that have been linked to each RIVPACS reference site. The fields in these tables are given in Table 12 below (the fields n, from, to and origin are shown for Aluminium (dissolved) only).

Table 12 - The Chemistry contemporary/current met1 tables

Field	Type	Description
Site ID	Text	Site Number
Aluminium (diss) mean (mg/l)	Number	Mean Aluminium (dissolved) (mg/l)
Aluminium (diss) n	Number	Number of Aluminium (dissolved) measurements used to calculate mean
Aluminium (diss) from	Date	Date of first Aluminium (dissolved) measurement
Aluminium (diss) to	Date	Date of last Aluminium (dissolved) measurement
Aluminium (diss) origin	Text	Origin of Aluminium (dissolved) measurements
Aluminium (total) mean (mg/l)	Number	Mean Aluminium (total) (mg/l)
Arsenic (diss) mean (mg/l)	Number	Mean Arsenic (dissolved) (mg/l)
Arsenic (total) mean (mg/l)	Number	Mean Arsenic (total) (mg/l)
Cadmium (diss) mean (mg/l)	Number	Mean Cadmium (dissolved) (mg/l)
Cadmium (total) mean (mg/l)	Number	Mean Cadmium (total) (mg/l)
Copper (diss) mean (mg/l)	Number	Mean Copper (dissolved) (mg/l)
Copper (total) mean (mg/l)	Number	Mean Copper (total) (mg/l)
Chromium (diss) mean (mg/l)	Number	Mean Chromium (dissolved) (mg/l)
Chromium (total) mean (mg/l)	Number	Mean Chromium (total) (mg/l)
Iron (diss) mean (mg/l)	Number	Mean Iron (dissolved) (mg/l)
Iron (total) mean (mg/l)	Number	Mean Iron (total) (mg/l)
Lead (diss) mean (mg/l)	Number	Mean Lead (dissolved) (mg/l)
Lead (total) mean (mg/l)	Number	Mean Lead (total) (mg/l)
Magnesium (diss) mean (mg/l)	Number	Mean Magnesium (dissolved) (mg/l)
Magnesium (total) mean (mg/l)	Number	Mean Magnesium (total) (mg/l)
Manganese (total) mean (mg/l)	Number	Mean Manganese (total) (mg/l)
Mercury (diss) mean (mg/l)	Number	Mean Mercury (dissolved) (mg/l)
Mercury (total) mean (mg/l)	Number	Mean Mercury (total) (mg/l)
Nickel (diss) mean (mg/l)	Number	Mean Nickel (dissolved) (mg/l)
Nickel (total) mean (mg/l)	Number	Mean Nickel (total) (mg/l)
Potassium (diss) mean (mg/l)	Number	Mean Potassium (dissolved) (mg/l)
Potassium (total) mean (mg/l)	Number	Mean Potassium (total) (mg/l)
Sodium (diss) mean (mg/l)	Number	Mean Sodium (dissolved) (mg/l)
Sodium (total) mean (mg/l)	Number	Mean Sodium (total) (mg/l)
Vanadium (diss) mean (mg/l)	Number	Mean Vanadium (dissolved) (mg/l)
Vanadium (total) mean (mg/l)	Number	Mean Vanadium (total) (mg/l)
Zinc (diss) mean (mg/l)	Number	Mean Zinc (dissolved) (mg/l)
Zinc (total) mean (mg/l)	Number	Mean Zinc (total) (mg/l)

For brevity the fields n, from, to and origin are shown for Aluminium (dissolved) only

3.11 Chemistry contemporary/current org1

The tables Chemistry contemporary/current org1 store organic compound data that have been linked to each RIVPACS reference site. The fields in these tables are given in Table 13 below (summarised by showing the fields n, from, to and origin for PCP only). All organic chemistry units have been standardised in these tables to µg/l.

Table 13 - The Chemistry contemporary/current org1 tables

Field	Type	Description
Site ID	Text	Site Number
PCP mean (µg/l)	Number	Mean PCP (µg/l)
PCP n	Number	Number of PCP measurements used to calculate mean
PCP from	Date	Date of first PCP measurement
PCP to	Date	Date of last PCP measurement
PCP origin	Text	Origin of PCP measurements
PCB con 028 mean (µg/l)	Number	Mean PCB congener 028 (µg/l)
PCB con 052 mean (µg/l)	Number	Mean PCB congener 052 (µg/l)
PCB con 101 mean (µg/l)	Number	Mean PCB congener 101 (µg/l)
PCB con 118 mean (µg/l)	Number	Mean PCB congener 118 (µg/l)
PCB con 138 mean (µg/l)	Number	Mean PCB congener 138 (µg/l)
PCB con 153 mean (µg/l)	Number	Mean PCB congener 153 (µg/l)
PCB con 180 mean (µg/l)	Number	Mean PCB congener 180 (µg/l)
1,2,3-Trichlorobenzene mean (µg/l)	Number	Mean 1,2,3-Trichlorobenzene (µg/l)
1,2,4-Trichlorobenzene mean (µg/l)	Number	Mean 1,2,4-Trichlorobenzene (µg/l)
1,3,5-Trichlorobenzene mean (µg/l)	Number	Mean 1,3,5-Trichlorobenzene (µg/l)
1,1,1-Trichloroethane mean (µg/l)	Number	Mean 1,1,1-Trichloroethane (µg/l)
1,2-Dichloroethylene mean (µg/l)	Number	Mean 1,2-Dichloroethylene mean (µg/l)
2,4-Ethenoic mean (µg/l)	Number	Mean 2,4-Ethenoic (µg/l)
2-Chlorophenol mean (µg/l)	Number	Mean 2-Chlorophenol (µg/l)
Tetrachloromethane mean (µg/l)	Number	Mean Tetrachloromethane (µg/l)

For brevity the fields n, from, to and origin are shown for PCP only

3.12 Chemistry contemporary/current pest1

The tables Chemistry contemporary/current pest1 store pesticide data that have been linked to each RIVPACS reference site. The fields in these tables are given in Table 14 below (the fields n, from, to and origin are shown for Aldrin only). All organic chemistry units have been standardised in these tables to µg/l.

Table 14 - The Chemistry contemporary/current pest1 tables

Field	Type	Description
Site ID	Text	Site Number
Aldrin mean (µg/l)	Number	Mean Aldrin (µg/l)
Aldrin n	Number	Number of Aldrin measurements used to calculate mean
Aldrin from	Date	Date of first Aldrin measurement
Aldrin to	Date	Date of last Aldrin measurement
Aldrin origin	Text	Origin of Aldrin measurements
Atrazine mean (µg/l)	Number	Mean Atrazine (µg/l)
Azinphos mean (µg/l)	Number	Mean Azinphos (µg/l)

Azinphos-ethyl mean (µg/l)	Number	Mean Azinphos-ethyl (µg/l)
Chlorfenvinphos mean (µg/l)	Number	Mean Chlorfenvinphos (µg/l)
DDD pp mean (µg/l)	Number	Mean DDD pp (µg/l)
DDE pp mean (µg/l)	Number	Mean DDE pp (µg/l)
DDE pp' mean (µg/l)	Number	Mean DDE pp' (µg/l)
DDT op mean (µg/l)	Number	Mean DDT op (µg/l)
DDT op' mean (µg/l)	Number	Mean DDT op' (µg/l)
DDT pp mean (µg/l)	Number	Mean DDT pp (µg/l)
DDT pp' mean (µg/l)	Number	Mean DDT pp' (µg/l)
Diazinon mean (µg/l)	Number	Mean Diazinon (µg/l)
Dichlorvos mean (µg/l)	Number	Dichlorvos mean (µg/l)
Dieldrin mean (µg/l)	Number	Mean Dieldrin (µg/l)
Dimethoate mean (µg/l)	Number	Mean Dimethoate (µg/l)
Endosulfan A mean (µg/l)	Number	Mean Endosulfan A (µg/l)
Endosulfan Total mean (µg/l)	Number	Mean Endosulfan Total (µg/l)
Endrin mean (µg/l)	Number	Mean Endrin (µg/l)
Fenitrothion mean (µg/l)	Number	Mean Fenitrothion (µg/l)
Fenthion mean (µg/l)	Number	Mean Fenthion (µg/l)
HCB mean (µg/l)	Number	Mean HCB (µg/l)
HCBD mean (µg/l)	Number	Mean HCBD (µg/l)
HCH α mean (µg/l)	Number	Mean HCH α (µg/l)
HCH β mean (µg/l)	Number	Mean HCH β (µg/l)
HCH δ mean (µg/l)	Number	Mean HCH δ (µg/l)
HCH γ mean (µg/l)	Number	Mean HCH γ (µg/l)
Isodrin mean (µg/l)	Number	Mean Isodrin (µg/l)
Linuron mean (µg/l)	Number	Mean Linuron (µg/l)
Malathion mean (µg/l)	Number	Mean Malathion (µg/l)
MBAS mean (µg/l)	Number	Mean MBAS (µg/l)
Mecoprop mean (µg/l)	Number	Mean Mecoprop (µg/l)
Mevinphos mean (µg/l)	Number	Mean Mevinphos (µg/l)
Parathion mean (µg/l)	Number	Mean Parathion (µg/l)
Parathion-methyl mean (µg/l)	Number	Mean Parathion-methyl (µg/l)
Propetamphos mean (µg/l)	Number	Mean Propetamphos (µg/l)
Simazine mean (µg/l)	Number	Mean Simazine (µg/l)
TDE pp' mean (µg/l)	Number	Mean TDE pp' (µg/l)
Triazophos mean (µg/l)	Number	Mean Triazophos (µg/l)
Trifluralin mean (µg/l)	Number	Mean Trifluralin (µg/l)

For brevity the fields n, from, to and origin are shown for Aldrin only

3.13 Corine land cover 2000

The table Corine land cover 2000 provides Corine 2000 land cover percentages (at the Corine level 3 label level) summarised for the watersheds draining into each RIVPACS reference site. These data were collated in ARC Map version 9 by using the CEH digital terrain model (DTM) to generate watersheds (catchments) that drain into each RIVPACS reference site and then by extracting the Corine land cover 2000 data coincident with these watersheds. The data are summarised as percentage cover for each of the 44 Corine level 3 land cover types and can be further aggregated to generate Corine level 2 and level 1 land cover types. The fields used in this table are described in Table 15 below (the field descriptions essentially resemble the field names and are therefore omitted).

Table 15 - The Corine land cover 2000 table

Field	Type
Site ID	Text
111 Continuous urban fabric	Number
112 Discontinuous urban fabric	Number
121 Industrial or commercial units	Number
122 Road and rail networks and associated land	Number
123 Port areas	Number
124 Airports	Number
131 Mineral extraction sites	Number
132 Dump sites	Number
133 Construction sites	Number
141 Green urban areas	Number
142 Sport and leisure facilities	Number
211 Non-irrigated arable land	Number
212 Permanently irrigated land	Number
213 Rice fields	Number
221 Vineyards	Number
222 Fruit trees and berry plantations	Number
223 Olive groves	Number
231 Pastures	Number
241 Annual crops associated with permanent crops	Number
242 Complex cultivation patterns	Number
243 Agriculture, with significant areas of natural vegetation	Number
244 Agro-forestry areas	Number
311 Broad-leaved forest	Number
312 Coniferous forest	Number
313 Mixed forest	Number
321 Natural grasslands	Number
322 Moors and heathland	Number
323 Sclerophyllous vegetation	Number
324 Transitional woodland-shrub	Number
331 Beaches, dunes, sands	Number
332 Bare rocks	Number
333 Sparsely vegetated areas	Number
334 Burnt areas	Number
335 Glaciers and perpetual snow	Number
411 Inland marshes	Number
412 Peat bogs	Number
421 Salt marshes	Number
422 Salines	Number
423 Intertidal flats	Number
511 Water courses	Number
512 Water bodies	Number
521 Coastal lagoons	Number
522 Estuaries	Number
523 Sea and ocean	Number

3.14 Flow

The flow table contains data on flow from National Water Archive gauging stations that have been linked to 443 RIVPACS reference sites. The flow table provides the NWA

gauging station number, the distance between the gauging station and the RIVPACS reference site, whether the gauging station is upstream or downstream of the RIVPACS site and the dates between which flow data were available. Also included is the mean flow (m^3/s) over the whole period when data were available and the flow in the year in which the RIVPACS samples were taken. Percent flow is derived by dividing the flow in the RIVPACS sampling year by the flow over the whole period over which flow data were available (Table 16).

The flow table also contains observed/RIVPACS-expected LIFE ratios in each of spring, summer and autumn and a field to indicate whether flow stress was thought to have been significant in the year in which the RIVPACS samples were collected. This was derived by assessing both the observed/expected values and the Percent flow. Further details of this assessment can be found in the report by Clarke et al. (2002).

Table 16 - The Flow table

Field	Type	Description
Site ID	Text	Site Number
Gauging Station No	Text	NWA (National Water Archive) gauging station number
GS Distance	Number	Distance of gauging station to RIVPACS site (km)
GS US DS	Text	Is the gauging station upstream or downstream of the RIVPACS site (U=upstream, D=downstream)
Flow from (year)	Text	First year of flow records
Flow to (year)	Text	Last year of flow records
Flow mean (m^3/s)	Number	Mean of the summer flows in the years recorded (m^3/s)
RIVPACS sampling year	Text	Year the RIVPACS site was sampled
Flow in RIVPACS sampling year (m^3/s)	Number	Mean summer flow in the year of RIVPACS sampling (m^3/s)
Percent flow	Number	(Mean summer flow in the year of RIVPACS sampling/Mean of the summer flows in the years recorded)x100 (m^3/s)
LIFE O/E spr	Number	Observed/expected LIFE index (spring samples)
LIFE O/E sum	Number	Observed/expected LIFE index (summer samples)
LIFE O/E aut	Number	Observed/expected LIFE index (autumn samples)
Evidence of Flow Stress in RIVPACS sampling year	Text	Evidence of Flow Stress in the RIVPACS sampling year in: Clarke RT, Armitage PD, Hornby D, Scarlett P & Davy-Bowker J. 2002. Investigation of the relationship between the LIFE index and RIVPACS - putting LIFE into RIVPACS. Environment Agency, Bristol.

3.15 Hydromorphology

The Hydromorphology table contains data on 269 RIVPACS reference sites that have been linked to River Habitat Surveys (carried out at various dates). These sites were linked by their respective grid references as being within 500m of each other and by checking the agreement of the river names. The 269 linked sites include 41 of the 110 RIVPACS sites in Northern Ireland. Each RHS also has an accompanying Habitat Modification Index and HMI class (which range from 1 for no or minimal modification

sites to 5 for highly modified sites). Three season combined observed/expected Number of Taxa and RIVPACS model end group details are also provided for all of the sites with linked HMI values (Table 17).

The second part of the Hydromorphology table provides a further set of 447 RIVPACS reference sites for which the original RIVPACS field survey sheets have been reassessed for hydromorphological influences. This gives a total of 537 RIVPACS sites at which either one of both of the HMI or RIVPACS field sheet based data have been collated. The RIVPACS field sheet derived fields (prefixed RP) give information on bridges, weirs, bank modifications, channel straightening and dredging that are closely linked to the precise locations at which the RIVPACS samples were collected (i.e. drawn on the original sampling site maps or noted on the sheets as being in close proximity). Accompanying the field sheet derived data are columns giving the three season combined Number of Taxa and the three season combined observed/RIVPACS-expected Number of Taxa together with RIVPACS model end group details for all of the sites with field sheet based hydromorphological reassessments.

Table 17 - The Hydromorphology table

Field	Type	Description
Site ID	Text	Site Number
RHS Site ID	Text	River Habitat Survey Site Identification Number
RHS NGR	Text	River Habitat Survey National Grid Reference
RHS River	Text	River Habitat Survey River
RHS Date	Date	River Habitat Survey Date
RHS HMI	Number	River Habitat Survey Habitat Modification Index
RHS HMI class	Text	River Habitat Survey Habitat Modification Index Class (1=no or minimal modification, 5=highly modified)
SprSumAut O/E NoTaxa	Number	Spring, Summer & Autumn combined - Observed/Expected Number of Taxa (Number of BMWP Scoring Families)[1]
RIVPACS model & EndGroup	Text	RIVPACS model and TWINSPAN End Group (at the 9 Group Level for GB)
RP Field Sheet	Text	RIVPACS field survey sheet reassessed for hydromorphological modifications
RP Bridge	Text	Proximity of bridges (excluding footbridges) to the sampling site (from RIVPACS field sheet)
RP Weir	Text	Proximity of weirs to the sampling site (from RIVPACS field sheet)
RP Bank Modification	Text	Bank modification (hard engineering) at the sampling site (from RIVPACS field sheet)
RP Channel Straightened	Text	Channel straightened (from RIVPACS field sheet)
RP Dredged (within 1 year)	Text	Dredged within the last year (from RIVPACS field sheet)
RP Hydromorphology	Text	Hydromorphological modification. Based on the RIVPACS field sheet (weir downstream or within site; 1 or both banks modified; channel straightened or site dredged in last year)
SprSumAut NoTaxa	Number	Spring, Summer & Autumn combined - Number of Taxa (Number of BMWP Scoring Families) [1]
SprSumAut O/E No Taxa	Number	Spring, Summer & Autumn combined - Observed/Expected Number of Taxa (Number of BMWP Scoring Families) [1]
RIVPACS model & End Group	Text	RIVPACS model and TWINSPAN End Group (at the 9 Group Level for GB)

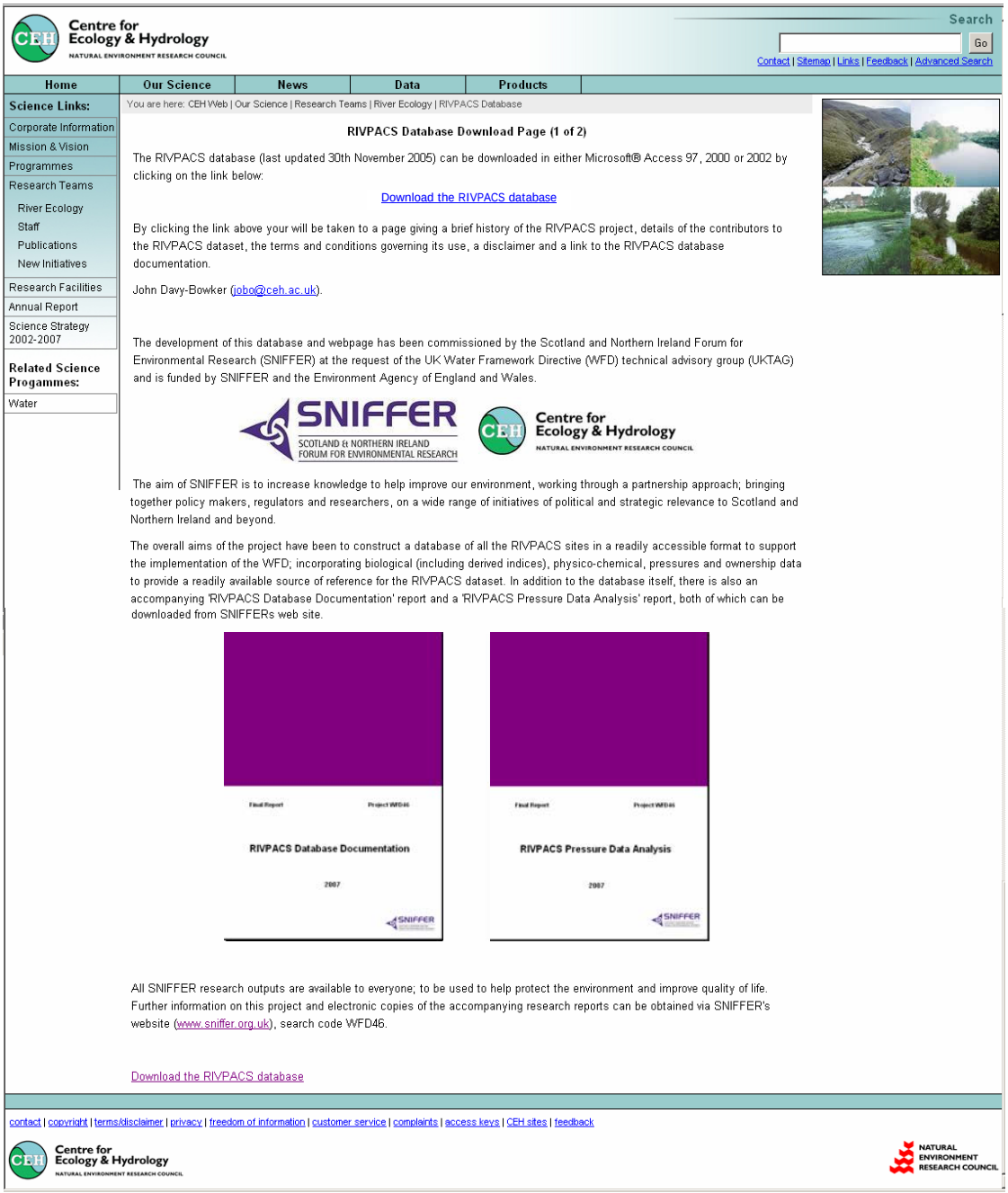
1 - Armitage et al. (1983) Water Research 17: 333-347

4. DISSEMINATION

The RIVPACS database can be downloaded from the CEH web page below:
http://www.ceh.ac.uk/sections/re/rivpacs_database_1.html

The database is provided in Microsoft® 97, 2000 and 2002. By clicking the link to download, you will be taken to a page providing a brief history of the RIVPACS project, details of the contributors to the RIVPACS dataset, the terms and conditions governing its use and a disclaimer. Users will be required to agree to the terms and conditions and to supply some basic name, address and email address details. From time to time there may be a need to update the database. The last updated date shown on the web page indicates when the database was last updated and the opening screen of the database itself gives details the changes.

Figure 2 - RIVPACS database download page



5. ACKNOWLEDGEMENTS

We would like to thank the Environment Agency, the Scottish Environment Protection Agency and the Environment and Heritage Service for their assistance in sending data on stresses. We would also like to acknowledge the assistance of the Centre for Ecology & Hydrology at Monk's Wood for supplying the 44-class CORINE 2000 national land cover map and the Centre for Ecology & Hydrology at Wallingford for the flow data extracted from the National Water Archive (NWA). We would also like to thank the various stakeholders in the RIVPACS dataset for their assistance and we also acknowledge the Ordnance Survey.

6. REFERENCES

Armitage, P.D., Moss, D., Wright, J.F., Furse, M.T. (1983) The performance of a new water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. *Water Research* 17: 333-347.

Clarke R.T., Armitage, P.D., Hornby, D., Scarlett, P. & Davy-Bowker, J. (2002) Investigation of the relationship between the LIFE index and RIVPACS - putting LIFE into RIVPACS. Environment Agency, Bristol.

Council of the European Communities, 2000. Directive 2000/60/EC, Establishing a framework for community action in the field of water policy. European Commission PE-CONS 3639/1/100 Rev 1, Luxembourg.

Davy-Bowker, J., Murphy, J.F., Rutt, G.P., Steel, J.E.C. & Furse, M.T. (2005) The development and testing of a macroinvertebrate biotic index for detecting the impact of acidity on streams. *Archiv für Hydrobiologie* 163: 383-403.

Extence, C.A., Balbi, D.M. & Chadd, R.P. (1999) River flow indexing using British benthic macroinvertebrates: a framework for setting hydroecological objectives. *Regulated Rivers: Research & Management* 15: 543-574

Fjellheim, A. & Raddum, G.G. (1990) Acid precipitation: biological monitoring of streams and lakes. *Science of the Total Environment* 96: 57-66.

Wallin, M., Wiederholm, T. & R.K., Johnson (2005) Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. A report by the EU Water Framework Directive Common Implementation Strategy working group 3.2 (REFCOND), version 7 (final), 2003-03-05.

APPENDICES

Appendix I A sample of the letter sent by CEH seeking formal written consent for release the RIVPACS dataset into the public domain

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

CEH Dorset
Winfrith Technology Centre
Winfrith Newburgh
Dorchester
Dorset DT2 8ZD

Telephone: 01305 XXXX
Fax: 01305 213600

DATE

Mr/Ms x
ORGANISATION

Dear XXXXX,

Release of freshwater biological data into the public domain

Since 1977 the River Communities section Centre for Ecology and Hydrology has worked on a series of contracts sponsored by a variety of agencies and organisations that include Defra, the Environment Agency (EA), the Scottish Environment Protection Agency (SEPA), the Environment and Heritage Service (EHS) (Northern Ireland) the Scottish Executive, the Welsh Office and Scottish Natural Heritage and several of their preceding organisations. During this period the name of my own organisation changed firstly from the Freshwater Biological Association (FBA) to the Institute of Freshwater Ecology (IFE) and then to CEH. Each of the contracts in question have been jointly funded by the Natural Environment Research Council (NERC), CEH's parent organisation, or by the FBA

These contracts, including those sponsored by [NAME OF ORGANISATION] have involved the collection and interpretation of freshwater macro-invertebrate communities of streams and rivers. Although the specific objectives of individual contracts varied in detail, a common theme of the data collection was to develop a new procedure for assessing the ecological status of running waters. We achieved this through the production of a software system called RIVPACS (River InVertebrate Prediction and Classification System). RIVPACS is based on data from 838 sites throughout the UK. It is now extensively used by the EA, SEPA and EHS for monitoring and assessment purposes. RIVPACS also provided the basis for many of the principles of the Water Framework Directive (WFD). In particular, it introduced the concept of the reference condition for assessing the Ecological Status of water bodies.

The data that we have collected are now seen as of particular importance to the implementation of the WFD and for the definition of reference conditions, setting of class boundaries and for the statutory inter-calibration exercise. Working groups have been established to tackle these issues at UK level. CEH are pleased to be able to assist in this process. In order to do so, one practical measure that we can adopt is to make our data widely available to the working groups and to other interested parties. Making these data freely available in the public domain will also be of general benefit to government departments and agencies and also the wider

academic community and is coherent with current freedom of information legislation.

We have therefore been contracted by SNIFFER (Scotland and Northern Ireland Forum For Environmental Research) to facilitate the release of the data into the public domain. However, in order to achieve this objective, we also need the formal consent to make the data publicly available of all parties with IPR in the data.

The purpose of this letter is therefore to provide you with the background to my request and to formally seek your written consent to make the data wholly or partially collected under contract to you, freely and openly available in the public domain via the CEH website

I am grateful for any assistance that you can give and will be pleased to provide further information if this will be helpful. This can include references to the contract numbers and reports relating to your sponsored contracts and a full list of the relevant sites and sampling dates for which we hope to release data. Please contact me if you would like this or any other background information.

Yours sincerely,

Mike T Furse

Appendix II End use terms and conditions set by the stakeholders of the RIVPACS dataset

Terms and conditions set on the end use of the RIVPACS dataset by:
Natural Environment Research Council/Centre for Ecology and Hydrology

CLICK-WRAP LICENCE

WARNING. BEFORE PROCEEDING PLEASE READ THE TERMS OF THIS LICENCE AGREEMENT CAREFULLY. YOU MAY OBTAIN A COPY OF THIS DATABASE BY DOWNLOADING IT FROM THE CEH WEB PAGES. ALL INTELLECTUAL PROPERTY RIGHTS IN THE RIVPACS DATABASE AND THE DATA WHICH CONSTITUTE THE DATABASE (''THE DATABASE'') ARE AND REMAIN THE PROPERTY OF THE LICENSOR AND OTHER STAKEHOLDERS. YOU ARE LICENSED TO USE THE DATABASE ON THE CONDITION THAT YOU ACCEPT ALL THE TERMS CONTAINED IN THIS LICENCE AGREEMENT.

BY CLICKING ON THE ACCEPTANCE BUTTON WHICH FOLLOWS THIS LICENCE AGREEMENT YOU ARE CONFIRMING YOUR ACCEPTANCE OF THE TERMS OF THIS LICENCE AGREEMENT. SUCH ACCEPTANCE IS EITHER ON YOUR OWN BEHALF OR ON BEHALF OF ANY CORPORATE ENTITY WHICH EMPLOYS YOU OR WHICH YOU REPRESENT ("CORPORATE LICENSEE"). IF YOU DO NOT ACCEPT THESE TERMS, YOU SHOULD CLICK ON THE CANCEL BUTTON.

1 DEFINITIONS

"Contractor" means a third party with whom you or the corporate entity which employs you has entered into a contract.

"Internal Business Use" means use of The Database within the corporate entity that employs you.

"Personal Use" means use by you acting on your own behalf.

"Stakeholders" means organisations who have contributed to the compilation of the Database: Centre for Ecology and Hydrology; Countryside Council for Wales; Department of the Environment, Food and Rural Affairs; English Nature; Environment Agency; Environment and Heritage Service; Freshwater Biological Association; Scotland and Northern Ireland Forum for Environmental Research; Scottish Environment Protection Agency; Scottish Executive; Scottish Natural Heritage; South West Water; Welsh Assembly Government.

"The Licensor" means Centre for Ecology and Hydrology (CEH), a component part of the Natural Environment Research Council (NERC).

"You" means the reader and any Corporate Licensee (the Licensee).

2 GRANT OF LICENCE

The Licensor grants to you a non-exclusive, non-transferable, free licence to use The Database providing the use falls within the permitted use set out below and for no other purpose. Your licence does not permit you to sub-license.

3 PERMITTED USE

3.1 You may use The Database for your own Personal Use.

3.2 You may use The Database for Internal Business Use.

3.3 You may make The Database available on your intranet.

3.4 You may disseminate publications and reports based upon The Database to third parties.

3.5 You may supply The Database to your Contractors with a copy of The Database (or derivations thereof). Your Contractors may only use The Database in connection with work for you.

4 RESTRICTIONS ON USE

4.1 You must retain the original automatic database start-up form contained within The Database and must not disrupt its operation.

4.2 You may not make The Database available on any internet site.

4.3 You may not use The Database for any commercial purpose, financial profit or gain.

4.4 You may not build products for commercial gain based either wholly, or in part, on The Database.

4.5 You may not rent, lease, sell, sublicense or otherwise distribute The Database to a third party.

4.6 You may not rent, lease, sell, sublicense or otherwise distribute products based either wholly, or in part, on The Database to a third party.

4.7 You may not assign or transfer The Database.

4.8 Your Contractors may not make The Database available on any internet site; may not use The Database for any commercial purpose, financial profit or gain; may not rent, lease, sell, sublicense or otherwise distribute The Database or products based either wholly or in part on The Database to a third party; and may not assign or transfer The Database.

5 UNDERTAKINGS

5.1 You undertake to ensure that any third party is made aware of the terms of this Licence Agreement before using The Database.

5.2 You undertake to hold all data and all other information relating to The Database confidential and not at any time disclose the same, during this licence or after its expiry or whether directly or indirectly, to any third party without our consent.

5.3 In any publications or reports arising from use of The Database, you undertake to acknowledge the following organisations for their contributions towards the compilation of The Database: Centre for Ecology and Hydrology and other Stakeholders/Centre for Ecology and Hydrology, Countryside Council for Wales, Department of the Environment, Food and Rural Affairs, English Nature; Environment Agency, Environment and Heritage Service, Freshwater Biological Association, Scotland and Northern Ireland Forum for Environmental Research, Scottish Environment Protection Agency, Scottish Executive, Scottish Natural Heritage, South West Water, Welsh Assembly Government.

The Licensor reserves the right to joint authorship of any publications in which material based on or derived from The Database constitutes a significant component.

5.4 You undertake to include the following citation in the reference list of any publications or reports arising from the use of The Database:

Wright, J. F. (2000). An introduction to RIVPACS. In: Assessing the biological quality of fresh waters: RIVPACS and other techniques (Eds. Wright, J. F., Sutcliffe, D. W. and Furse, M. T.), pp 1-24. Freshwater Biological Association, Ambleside, UK.

6 OWNERSHIP

The Licensor and other Stakeholders (listed in clause 1) retain ownership of The Database and related documentation and all copies of The Database at all times, which are copyright works and are also protected under applicable database laws.

The following copyright notice shall be used on all publications and reports:

You are not granted any rights in connection with any logos or trademarks of The Licensor, SNIFFER or The Database.

7 DISCLAIMER

All guarantees, representations and warranties of any kind, whether express or implied, including, without limitation, the implied warranties of satisfactory quality, merchantability and fitness for a particular purpose or ability to achieve a particular result are hereby excluded, so far as such exclusion or disclaimer is permitted under the applicable law. You assume the entire risk as to the quality and performance of the Database. Should the Database prove defective, you (and not the Licensor) assume the entire cost of all necessary servicing or correction. The Licensor does not warrant that The Database will meet your requirements or that its operation will be uninterrupted or error free. This Licence Agreement does not affect your statutory rights.

8 EXCLUSION OF LIABILITY

8.1 In no event shall the Licensor be liable to you for any damages, including loss of business, loss of opportunity, loss of data, loss of profits or for any other indirect or consequential loss or damage whatsoever arising out of the use of or inability to use The Database, even if the Licensor has been made aware of the possibility of such damages.

8.2 Nothing in this Licence Agreement shall exclude or limit the liability of the Licensor for fraudulent misrepresentation or for death or personal injury resulting from the negligence of the Licensor.

9 TERMINATION

9.1 This Licence Agreement will terminate automatically if you breach any of its terms or if you destroy The Database and any copies or return The Database to the Licensor.

9.2 Upon termination all rights you have to use The Database will cease and you must destroy or delete The Database and all copies from all storage media in your control.

10 SEVERABILITY

If any provision of this Licence Agreement is held invalid, illegal or unenforceable for any reason by any court of competent jurisdiction such provision shall be severed and the remainder of the provisions hereof shall continue in full force and effect as if this Licence Agreement had been agreed with the invalid illegal or unenforceable provision eliminated.

11 ENTIRE AGREEMENT

This Licence Agreement contains the entire Agreement between us relating to the subject matter and supersedes all proposals, representations, understandings and prior agreements, whether oral or written, and all other communications between us relating to that subject matter.

12 GOVERNING LAW AND JURISDICTION

This agreement shall be governed by and construed in accordance with English law and the parties submit to the exclusive jurisdiction of the English courts.

Terms and conditions set on the end use of the RIVPACS dataset by: **Environment Agency** *The Environment Agency have approved the licence conditions above*
Terms and conditions set on the end use of the RIVPACS dataset by: **Countryside Council for Wales** *Countryside Council for Wales require that all users of the data acknowledge their funding support*
Terms and conditions set on the end use of the RIVPACS dataset by: **Scottish Natural Heritage** *Scottish Natural Heritage require that all users of the data acknowledge their funding support*
Terms and conditions set on the end use of the RIVPACS dataset by: **Department for Environment Food and Rural Affairs** *Department for Environment Food and Rural Affairs require that all users of the data acknowledge their funding support*
Terms and conditions set on the end use of the RIVPACS dataset by: **English Nature** *English Nature require that all users of the data acknowledge their funding support*
Terms and conditions set on the end use of the RIVPACS dataset by: **South West Water** *South West Water require to be acknowledged as funder or co-funder (as the case may be)*
Terms and conditions set on the end use of the RIVPACS dataset by: **Freshwater Biological Association** *No restrictions*
Terms and conditions set on the end use of the RIVPACS dataset by: **Scottish Environment Protection Agency** *No restrictions*
Terms and conditions set on the end use of the RIVPACS dataset by: **Environment and Heritage Service**

No restrictions
Terms and conditions set on the end use of the RIVPACS dataset by: **Welsh Assembly Government** *No restrictions*
Terms and conditions set on the end use of the RIVPACS dataset by: **Scottish Executive** *No restrictions*

Appendix III Taxonomy and rules for taxon recording

This appendix describes the current taxonomy of the RIVPACS database and expands upon the description of taxa recording given in section 3.3

Taxonomy

The taxonomy used in the RIVPACS database generally conforms to:

> Furse, M, I. McDonald, & R. Abel, 1989. A revised coded checklist of freshwater animals occurring in the British Isles.

This nomenclature is very dated. Mike Furse (CEH) maintains a comprehensive checklist of freshwater macroinvertebrates and is very familiar with the latest taxonomy. For some time CEH have been planning to apply an update to the taxonomic data stored within our National Invertebrate Database (from where the taxa data in the RIVPACS Database have been extracted). We aim to perform a taxonomic upgrade to both the National Invertebrate Database and the RIVPACS Database. CEH aim to do this by the end of 2006. We then need to apply a similar upgrade to the taxonomy used in RIVPACS itself.

Rules for Recording Taxa

The RIVPACS database contains data recorded at two levels (family level - with a $\log_{10}$ abundance category) and species level (those species present are simply listed). The data is therefore double entered (e.g, three species of Baetis in a sample would be recorded as one record for the family Baetidae (with $\log_{10}$ abundance), and three separate species records without abundances (one for each of the 3 species).

The RIVPACS data have been collected and processed over a very long period of time and there are cases where the approach to data recording may have changed (e.g. through the availability of new keys).

At family level, all families have been recorded (with $\log_{10}$ abundances) with the following exception. Data recording has always used the BMWP artificial families. CEH have never attempted to identify the individual constituents of these artificial taxa (and the RIVPACS software does not therefore have the capacity to predict the abundances of the constituent families of the artificial BMWP groups). For example, while the RIVPACS software can predict the $\log_{10}$ abundance of Hydrobiidae including Bithyniidae, it cannot predict the abundance of either Hydrobiidae or Bithyniidae as separate families. The same is true for Planariidae including Dugesiidae or the other 6 artificial groups. Therefore even if in the raw data a poorly preserved specimen is identified to Planariidae as a separate family, distinct from Dugesiidae, but cannot be identified more precisely than that, at family level this specimen is treated as Planariidae including Dugesiidae

At species level, specimens have always been identified and recorded to the most precise level possible. In many cases this is not at species level. Species data therefore comprise true species determinations, artificial taxon groups/complexes, genera and families. This approach has been consistent across all RIVPACS samples, the only exception being Hydracarina that were originally identified to species by Terry Gledhill at the River lab, but in more recent times have only been identified as Hydracarina. There are however, sources of variation in this approach. For example, as new keys have emerged, CEH may have become able to identify certain species groups to species level. There may also have been cases where originally distinct species have in more recent keys been aggregated as complexes or the reliable identification of certain group to species has been called in the question so that these can now only reliably be recorded at higher taxonomic levels.

In the RIVPACS software, CEH have attempted to minimise the impact of these problems by standardising to a common and consistent level of identification. However, the inability to identify every specimen to species has consequences for the species predictions produced by RIVPACS. For example, if the only specimen of Rhyacophila in a RIVPACS sample could not be identified beyond genus level but

was in reality a R. dorsalis, the failure to identify it would mean that R. dorsalis was predicted with a lower probability at some test sites than if it had been possible to identify the specimen to species. This creates a tendency to under-predict difficult to identify species. However, it is also important to realise that the same effect exists in test samples, so that the two tend to balance out. This problem is non existent or at worst very trivial at family level because almost all RIVPACS specimens are identified to at least this level.

A related issue is the lack of abundance data for species level records and the fact that family level records are only enumerated using the $\log_{10}$ scale rather than as absolute counts. CEH have recently investigated the possibility of obtaining more detailed abundance data from the original laboratory species sheets filled out at the time the samples were identified. Unfortunately very few samples have abundances recorded for all families and species present in a sample so it is extremely unlikely that abundance data at this level ever existed.

For any potential new RIVPACS samples we should consider whether any new approaches to recording taxa are required. We should also consider recording the taxonomic key used for all determinations so that the identity of each record is completely clear over time.