Standardisation of river classifications:

Framework method for calibrating different biological survey results against ecological quality classifications to be developed for the Water Framework Directive

Contract No: EVK1-CT-2001-00089
4th deliverable, due 31/05/03, entitled

Database selection and overview
Acquisition of existing data

Compiled by Cynthia Davies, John Davy-Bowker & John Murphy
Partner no. 1 (Centre for Ecology and Hydrology)

A project under the 5th Framework Programme
Energy, Environment and Sustainable Development
Key Action 1: Sustainable Management and Quality of Water

Table of contents

List of tables

Database selection and overview: Acquisition of existing data
A Report on STAR Work Package 4 - Data Acquisition: May 2003

1 Objectives (taken from the Description of Work)

To obtain readily available existing data, of sufficient quality, reliability and methodological compatibility, in order to enhance the achievement of the project objectives. To maintain good lines of communication and co-operation with all other relevant European and national sampling and assessment programmes and standardisation approaches.

2 Methodology and scientific achievements related to Workpackage 4, including contribution from partners

2.1 Background

At the project kick-off meeting in December 2001, it was agreed that the acquisition of existing data should be achieved by means of a questionnaire including a request for data. This questionnaire was forwarded to representatives of the STAR consortium in each country for circulation to individuals and agencies with potentially useful data holdings.

For each data type (Biological Quality Element - BQE), the questionnaire included requests for the following information: the availability of existing data sets to the STAR project; sampling methodologies; field and laboratory procedures; quality control; the availability of additional site description parameters (especially chemical data and variables used in assigning Water Framework Directive (WFD) site typology).

Suitable data sets relating to macroinvertebrates, phytobenthos, macrophytes and hydromorphology were requested. Preferential interest was expressed in high quality electronic data sets of substantial size that contained data collected using consistent methodologies. The ideal data sets would be those that dealt with two or more organism groups from the same sites with supporting site description parameters.

2.2 Data acquisition

During 2002 and early 2003, the responses from the questionnaires were collated and the data received were recorded by the Centre for Ecology and Hydrology (CEH). Renewed requests for data were issued where necessary. The results of the questionnaire responses are summarised in Table 1. The availability of the data types other than invertebrate (i.e. phytobenthos, macrophyte and hydromorphology) is poor.

Organisations holding data had expressed concerns about how their data might be used prior to its release. To circumvent this potential resistence, it was considered that a data licence agreement, including a statement of ‘end-use’ would encourage supply of existing data. Accordingly, licence agreements were drawn up and sent to participating countries. Response has been variable; some data sets have been released on signature of a licence agreement, while other partners are still awaiting agreement from the data suppliers.

Data sets that have been received are in the process of being evaluated and documented prior to loading into a central databases. Invertebrate data sets are being documented in the first instance. Field names and contents are being documented so that the availability of different variables can be used to determine a practicable combined database structure. Provisional database structures have been designed to take the four different types of data (macroinvertebrates, phytobenthos, macrophytes and hydromorphology) but these structures cannot be finalised until documentation of data sources is complete.

Table 2 compares the advised availability of data with what has actually been received by CEH Dorset. Table 3 provides a detailed report on data supply by each Partner. Table 4 summarises the data received by BQE.

Table 1 Questionnaire responses and data indicated as being available

	Questionnaire returned?	Invertebrates	Phytobenthos	Macrophytes	Hydromorphology
Austria	YES	YES	NO	NO	NO
Czech Rep.	NO	YES	NO	NO	NO
Denmark	YES	YES	NO	YES	NO
France	YES	YES	NO	NO	NO
Germany	YES	YES	YES	YES	YES
Greece	NO	YES	YES	NO	NO
Italy	NO	YES	NO	NO	NO
Latvia	NO				
Netherland	YES	YES	NO	YES	YES
Poland	YES	YES	YES?	NO	NO
Portugal	NO	YES	YES	YES	NO
Slovakia	NO				
Sweden	YES	YES	YES	NO	YES
UK	YES	YES	YES	YES	YES

Table 2 Data indicated as available / received by CEH Dorset ('Part' indicates that more data are expected)

Data available / received	Invertebrates	Phytobenthos	Macrophytes	Hydromorphology
Austria	YES / YES	NO / -	NO / -	NO / -
Czech Rep.	YES / NO	NO / -	NO / -	NO / -
Denmark	YES / YES	NO / -	YES / YES	NO / -
France	YES /YES	NO / -	NO / -	NO / -
Germany	YES / YES	YES / NO	YES / NO	YES / YES (Part)
Greece	YES / NO	YES / YES	NO / -	NO / -
Italy	YES / NO	NO / -	NO / -	NO / -
Latvia	NO / -	NO / -	NO / -	NO / -
Netherlands	YES / NO	NO / -	YES / NO	YES / NO
Poland	YES / NO	YES(?) / NO	NO / -	NO / -
Portugal	YES / NO	YES / NO	YES / NO	NO / -
Slovakia	No information	No information	No information	No information
Sweden	YES / YES	YES / YES	NO / -	YES / YES
UK	YES / YES (Part)	YES / YES	YES / YES	YES / YES

Note: The AQEM data set received from Germany includes some invertebrate data from Italy, Portugal, Greece, Germany, Netherlands, Czech Republic, Austria and Sweden. This is in addition to the information given in the table above.

Table 3 Detailed reports by partners

Detailed report by Partner
Partner 1 CEH Dorset UK (CEH, NERC) *The following data sets have been received:* File name Clean.mdb (Bowburn Consultancy) Code number - BQE Phytobenthos Format Access database Data set name Environmental Change Network File name ecnfin.mdb; (ecnfwc.mdb–environmental variables) Code number - BQE Invertebrates Format Access database Data set name Environmental Change Network File name ecnfma.mdb (Environmental Change Network) Code number - BQE Macrophytes Format Access database The ECN databases hold data for each element from the same site. The invertebrate database includes site details. Data set name River Conservation Macrophyte Database (JNCC) File name Cons_rivers.mdb Code number - BQE Macrophytes Format Access database Data set name Mean Trophic Rank database File name MTR Code number - BQE Macrophytes Format Access database A questionnaire response has been received for the Acid Water Monitoring Network (invertebrates, phytobenthos and macrophytes). RIVPACS data (invertebrates) and other CEH data will also be made available.

Partner 2. University of Duisburg-Essen (Germany – Christian Feld)

The following data sets have been received:

Data set name Strukturgütedatenbank NRW
File names 0212_GSG_NRW_GF.mdb;
0212_GSG_NRW_KF.mdb;0212_GSG_NRW_KF_flach.mdb
Code number DE/Hyd/002
BQE Hydromorphology
Format Access database
Data Licence Agreement signed

Data set name Gewässerstrukturgütedaten (RP)
File names Struka V (große Gewässer); Struka V Stammdatenverwaltung;
Struka V (kleine Gewässer)
Code number DE/Hyd/003
BQE Hydromorphology
Format GIS application
Data Licence Agreement signed

Data set name UBA Saprobic database
File name Saprobie_Datenbank_IDART.mdb
Code number DE/Inv/004
BQE Invertebrates
Format Access database
1315 samples containing 44352 invertebrate records

Data set name AQEM (Data exported from AQEM for all contributing countries)
File name AQEM_Europe_database-export.mdb
Code number DE/Inv/003
BQE Invertebrates
Format Access database

Requests for the following data sets and licence agreements were sent at the end of February but data have not yet been received.

phytobenthos & macrophytes	PHYLIB database
hydromorphology	Hesse GESIS database

Partner 3. BOKU (Austria)

The following data set has been received:

Data set name WGEV - Austria
File name Star_Aut
Code number AU/Inv/007
BQE Invertebrates
Format Access database
270 sites containing 49284 invertebrate records

Data on phytobenthos, macrophytes and hydromorphology are not available in a readily accessible form and no further requests for data are anticipated.

Partner 4. ALTERRA Green World Research (Netherlands)

The following data sets have been promised:

Invertebrates	Water District Managers Bekentypologie database
Invertebrates	ALTERRA data
Macrophytes	ALTERRA Turboveg database
Hydromorphology	ALTERRA data set

Requests for these data sets and licence agreements were sent at the end of February but no data have been received as yet. No phytobenthos data are available.

Partner 5. Swedish University of Agricultural Sciences (Sweden)

The following data sets have been received:

Data set name	Swedish National Stream Survey 1995
File name	Swedish_benthic_fauna.xls
Code number	SE/Inv/002
BQE	Invertebrates
Format	Excel spreadsheet

Data set name	Swedish National Stream Survey database
File name	Swedish_chemistry_env_data.xls
Code number	SE/Hyd/004
BQE	Hydromorphology
Format	Excel spreadsheet

Data set name	Phytobenthos – AQEM sites (Swedish EPA data)
File name	PhytobenthosSweden
Code number	SE/Phy/002
BQE	Phytobenthos
Format	Spreadsheet

The diatom data are from AQEM sites.

No macrophyte data are available. No further requests for data are anticipated.

Partner 6. Masaryk, BRNO (Czech Republic)

The following data set has been promised:

macroinvertebrates	PERLA data

No questionnaire response has been received from the Czech Republic. Preparation of a macroinvertebrate data set from the PERLA database took place in February 2002, but this has not been yet been received. Licence agreements were sent with a further request for these data in March 2003. A reminder was sent in mid-May and data are expected imminently.

Partner 7. National Centre for Marine Research, Institute of Inland Waters (NCMR, IIW) (Greece)

The following data sets have been received:

Data set name	DIATOMA
File name	DIATOMA.xls (Summary data set. Raw data have been requested)
Code number	GR/Phy/003
BQE	Phytobenthos
Format	Excel spreadsheet

File name	MASTERTHESISVOURDOUMPAMACROINVERTEBRATES.xls
Code number	-
BQE	Invertebrates
Format	Spreadsheet

No questionnaire response was received from Greece. Invertebrate data from AQEM project have been received as part of AQEM_Europe_database-export.mdb. The masters thesis has been supplied as hard copy with only species lists for 19 sites supplied electronically. (Supporting River Habitat Survey and other data are available in the thesis. This has been excluded from the summaries in this report.) There are no other data available on hydromorphology. No further requests for data are anticipated.

Partner 8. IRSA-CNR (Italy)

No questionnaire response was received from Italy. It is understood that invertebrate data are available but there are no data on phytobenthos, macrophytes or hydromorphology. Detailed negotiations regarding the use of WP4 data were underway. Copies of the Licence Agreement and a further request for data were sent in February 2003, but as yet there has been no response. Some invertebrate data from AQEM project have been received as part of AQEM_Europe_database-export.mdb.

Partner 9. University of Évora (Portugal)

No questionnaire response was received from Portugal. Some invertebrate data from AQEM project have been received as part of AQEM_Europe_database-export.mdb. It has been confirmed that further data on macroinvertebrates, phytobenthos and macrophytes are available subject to licence agreement. A further request for data and copies of the licence agreement were sent in February 2003.

Partner 10. NERI (Denmark)

The following data sets have been received:

Data set name	Danish Stream Fauna Index AQUA database; Intensive Biological Programme database
File name	fauna98-00.mdb
Code number	DK/Inv/001
BQE	Invertebrates
Format	Access database

Data set name	Intensive Biological Programme database
File name	bioveg98-00.mdb
Code number	
BQE	Macrophytes
Format	Access database

(No site data included)

There are no phytobenthos or hydromorphology data available and no further requests for data are anticipated.

Partner 11. The Environment Agency (UK)

The following data sets have been received:

Data set name	NRA General Quality Assessment (1990, 1995)
File names	(Held as part of larger database)
Code number	-
BQE	Invertebrates
Format	Access database

Data set name	EA General Quality Assessment (Biology. 2000)
File names	EA_GQA_2000.mdb; GQA_2000_JMB.mdb; All_2000_GQA.mdb (GQA data)
Code number	-
BQE	Invertebrates
Format	Access database

Data set name	River Habitat Survey Database
File name	RHS database
Code number	-
BQE	Hydromorphology
Format	Access database

No further requests for data are anticipated

Partner 12. (Masaryk Water Research Institute) N/A

Partner 13. (Province of Bolzano) N/A

Partner 14. University of Metz (France)

The following data set has been received:

Data set name	French R.N.B. Data set (Réseau National de Bassin)
File name	Frenchdata.xls (Coord-FrenchSta546GB.xls (site details incl. co-ordinates for Invertebrate data))
Code number	FR/Inv/005
BQE	Invertebrates
Format	Excel spreadsheet

1414 site records (929 quantitative records, 485 qualitative records)for 17 of the 23 regional DIREN agencies.

No further requests for data are anticipated.

Partner 15. (Research Institute Senckenberg) N/A

Partner 16. (CEN TC230/WG2) N/A

Partner 17 University of Lódz , Institute of Biology and Environmental Protection (Poland)

Partner 18 Agricultural University of August Cieszkowski (Poland)

Partner 19 Institute of Environmental Protection (Poland)

The following data set has been promised:

Invertebrates	BEZKREGOWCE (Invertebrates)

A questionnaire response has been received indicating that macroinvertebrate data are available. Phytobenthos data may also be available subject to licence agreement. A request for data and copies of the licence agreement were sent in February 2003.

Partner 20 Institute of Biology, University of Latvia (Latvia)

A questionnaire was sent in February 2003. No questionnaire has been completed, but first indications are that data may not be readily available.

Partner 21 Slovak Academy of Sciences. Institute of Zoology (Slovakia)

Partner 22 Faculty of Natural Sciences, Department of Ecology Comenius University (Slovakia)

A questionnaire was sent in February 2003. No response has been received.

Table 4 Summary of data sets received by Biological Quality Element

BQE	Country	Data set	Format	Licence agreement?
Hydromorphology				
	Germany	Gewässerstrukturgütedaten (RP)	GIS application	Yes
	Germany	Strukturgütedatenbank NRW	Database	Yes
	Sweden	Swedish National Stream Survey 1995	Spreadsheet	
	UK	River Habitat Survey Database	Database	
Invertebrates				
	All Europe (via Germany)	AQEM_Europe_database-export	Database	
	Austria	WGEV - Austria	Database	
	Denmark	Danish Stream Fauna Index AQUA database / Intensive Biological Programme database	Database	
	France	French R.N.B. Data set	Database	
	Germany	UBA Saprobic database	Database	
	Sweden	Swedish National Stream Survey 1995	Spreadsheet	
	UK	ECNFIN	Database	
	UK	EA General Quality Assessment (Biology. 2000)	Database	
Macrophytes				
	UK	River Conservation Macrophyte Database	Database	
	Denmark	Intensive Biological Programme database	Database	
	UK	ECNFMA	Database	
	UK	Mean Trophic Rank database	Database	
Phytobenthos				
	Greece	DIATOMA	Spreadsheet	
	Sweden	Phytobenthos – AQEM sites (EPA)	Spreadsheet	
	UK	CLEAN	Database	
NONE				
	Czech Rep.	NONE		
	Italy	NONE		
	Latvia	NONE		
	Netherlands	NONE		
	Poland	NONE		
	Portugal	NONE		
	Slovakia	NONE		

Note: Files comprising only environmental variables linked to a BQE are omitted from the summary table.

3 Discussion and conclusion

This workpackage (WP) has been successful in acquiring (or obtaining promises to supply) a considerable number of data sets from most of the partner countries. Continued efforts to obtain data will ensure that WP4 can contribute useful data sets to other parts of the STAR project. However, it has become increasingly obvious that the time required to obtain and collate the available data will extend the time-scale in the project plan. The addition of the Newly Associated States (NAS) to the STAR project resulted in an extension to the deadline for Work Package 4 to June 2003 but an ongoing programme of work is planned to meet the WP objectives in full. This deliverable will be updated as further progress is made.

Data have been provided in a variety of formats and data sets are largely undocumented. Language differences apart, many field names require interpretation as they use abbreviations or codes. It is not possible to decide on a structure for a combined database until the availability of different variables is known. Work on documentation and analysis of the structure of contributing data sources has begun initially with invertebrate data sets. Additionally, not all species names used in the source data link directly to the European lists to be used in Work Package 10. Species and code lists are being extracted from each invertebrate data set and a combined look-up table is being constructed to cross-refer between names and codes. These lists will need to be referred back to the data suppliers for confirmation and clarification. It is clear from an initial inspection of several data sets that geo-referencing of survey sites is variable and often missing from the data supplied. Further requests for this type of supporting information are likely to be needed if BQE's are to be cross-referenced.

There is a particular need to obtain data sets that deal with 2 or more organism groups from the same sites with supporting site description parameters. Data sets which should meet these criteria have been identified and a further set is promised: e.g. the Danish Intensive Biological Programme database (invertebrates and macrophytes), UK ECN data (phytobenthos, invertebrates and related environmental variables), Swedish National Stream Survey 1995 (invertebrates, hydromorphology and related environmental variables) and Swedish Phytobenthos – AQEM sites (phytobenthos from AQEM sites (invertebrates)). The German PHYLIB database (phytobenthos & macrophytes) has been promised but not yet received. There are linked data for a limited number of sites in Greece available at present only in hard copy.

Licence agreements have been sent to all partners but so far only two signed agreements have been returned. It remains to be seen whether licence agreements will facilitate the release of additional data sets. A copy of the Licence Agreement is appended (Annex I)

While some countries are already engaged in substantial monitoring programmes, utilising all of the quality elements required by the Directive, others have small monitoring programmes utilising fewer quality elements. It is therefore likely that some countries will have more work to do than others in contributing to the full and consistent implementation of the Framework Directive across Europe.

4 Plans and future objectives

Documentation of data sets received by CEH Dorset and cross-referencing of species names and code lists used in source data will continue. Renewed requests for data to the existing STAR partners and the NAS states (including the Data Licence Agreement) will be made where appropriate. Questionnaire collation and evaluation will continue. Review of the variables included in source data will continue to enable construction of combined databases. Where possible, data structures will be compatible with those developed for Work Package 10.

ANNEX I DATA LICENCE AGREEMENT

DATA LICENCE AGREEMENT

LICENSOR(s):

LICENSEE(s): The EU Project STAR consortium (22 partner institutes from 14 countries)

Led by –

Centre for Ecology and Hydrology, CEH Dorset, Winfrith Technology Centre, Winfrith Newburgh, Dorchester, Dorset DT2 8ZD, United Kingdom (a component of the Natural Environment Research Council, Polaris House, North Star Avenue, Swindon, SN2 1EU, United Kingdom).

DATA:

Date of Issue:

Period of Issue: Open ended

I, THE LICENSOR, HAVE READ THE PARTICULARS OF THIS DATA LICENCE AGREEMENT AND I AGREE TO GRANT A NON-EXCLUSIVE, NON-TRANSFERABLE LICENCE FOR THE DATA MENTIONED ABOVE.

Licensor's name:……..…….
(Print clearly)
Licensor's authorised signature:……..

Position:……..……..

I, THE LICENSEE, HAVE READ THE PARTICULARS OF THIS AGREEMENT AND I AGREE TO ABIDE BY THE TERMS AND CONDITIONS SET OUT IN THIS DATA LICENSE AGREEMENT.

Licensee's name:…….MICHAEL TRISTRAM FURSE……..……..
(Print clearly)
Licensee's authorised signature:……..……....
(On behalf of the STAR consortium)
Position:……..STAR Project Coordinator……...

ANNEX I DATA LICENCE AGREEMENT

TERMS AND CONDITIONS

1.0 Background to STAR (EU Project EVK1-CT-2001-00034)

The EU funded project STAR (Standardisation of River Classifications) is a research project supported by the European Commission under the Fifth Framework Programme and contributes to the implementation of the Key Action "Sustainable Management and Quality of Water" within the Energy, Environment and Sustainable Development Programme. Specifically, STAR seeks to provide a framework for calibrating different biological survey results against ecological quality classifications in each of the partner states. The STAR project will collect standardised samples from each country and calibrate these with national methods for determining water quality so that the results of biological monitoring programmes in each partner country can be calibrated and compared on a European scale. In addition to the new inter-calibration samples collected by the STAR research programme, access to existing data-sets is needed for the reasons set out in section 2.0.

2.0 Use of data sets by STAR

The primary need for data sets in STAR is to test the concordance of different organism groups and environmental stress (both natural and human-generated).

Although few studies have evaluated the spatial concordance among different organism groups, present-day concepts contend that different organism groups may respond differently to natural as well as human-generated stress. For example, fish due to their high mobility and relatively long life-span, might be expected to reflect environmental conditions across broad spatial and temporal scales. By contrast, benthic organisms such as macroinvertebrate or phytobenthos might reflect changes occurring at much smaller spatial (microhabitat) and temporal (response time of weeks to months) scales. Accordingly, we might hypothesize that these groups would show a high degree of concordance if relevant time-lag responses are accounted for. However, difference organism groups supposedly react differently to stress. For example, whereas fish or macroinvertebrates might be an appropriate indicator of changes in temperature or oxygen, they might not be considered as “first” choice indicators of nutrient enrichment. In contrast, phytobenthos or phytoplankton and even subsequently macrophytes might be considered as a better “first” choice for detect of ecological change due to nutrient enrichment. Hence, a robust biomonitoring programme would incorporate knowledge of how different organism groups react to different human-generated stressors and scale (both spatial and temporal).

Work package 12 (linking organism groups) is designed to test the conceptual model of different organism groups and scale. However, the few number of sites that are part of the STAR project prohibits a thorough analysis of these concepts (i.e. due to economic constraints only a few ecotypes are included in the STAR programme). To improve the statistical power of these tests, other “complementary” data are needed. Data accrued from work package 4 (acquisition of existing data) will hopefully complement the STAR data set resulting in more robust and reliable test of the conceptual model. In particular, we are interested in obtaining data sets that consist of data from two or more organism groups as well as background physical-chemical descriptors. These data will hopefully improve our ability to tease out the effects of scale and human-generated stressors on aquatic organism response.

ANNEX I DATA LICENCE AGREEMENT

3.0 Terms and Conditions

The LICENSOR hereby agrees to permit use of the DATA as set out below.

3.1 The LICENSOR permits the LICENSEE to use the data set(s) (hereafter referred to as the DATA) only for the purposes expressly set out in section 2.0 of this Data Licence Agreement.

3.2 The LICENSOR permits the LICENSEE to analyse and manipulate the DATA to produce derivative data, such as indices, maps and graphs in order to assist in the achievement of the objectives set out in section 2.0 of this Data Licence Agreement.

3.3 To publish papers, reports and other materials that use the DATA so long as these have been produced in accordance with this Data Licence Agreement (see clauses 3.2, 3.7, 3.9 and 3.10 below).

The LICENSEE hereby agrees to abide by the following terms and conditions.

3.4 The LICENSEE agrees that the DATA shall only be used for the purposes set out in this Data Licence Agreement.

3.5 The LICENSEE agrees only to make copies of the DATA when this is needed for data back-up or archive purposes and that effective control will be kept on the number of copies of the DATA.

3.6 The LICENSEE agrees to accept responsibility for installing, and examining the DATA on their own computer systems and for installing and maintaining any software necessary to access the DATA.

3.7 The LICENSEE agrees to acknowledge the source of the DATA in all publications that involve its use.

The LICENSEE agrees not to undertake any of the following actions.

3.8 Transfer, reassign, rent, lease, sell, give or otherwise distribute the DATA to other parties beyond the LICENSEES unless permitted to do in writing by the LICENSOR.

3.9 Publish any raw (unprocessed) DATA e.g. species lists (see clause 3.2 above).

3.10 Publish any derivative DATA (for example, tables of summary statistics or box and whisker plots displaying the DATA from several sampling sites etc.) without the prior written permission of the LICENSOR, sought In Addition to this LICENCE.

3.11 Release any product or tradable commodity based in whole or in part upon the DATA or on data derived from the DATA unless the prior written permission of the LICENSOR has been obtained, In Addition to this LICENCE.