June 2012

TECHNICAL BRIEF

Water & Environment

Upgrading South Africa's premier river ecology database

Funds from the WRC helped to upgrade the functionality of the Biological and Physico-chemical Database (Biobase) and make it accessible via the Rivers Web application, and to incorporate current spatial aspects, thereby making it compatible with more recent spatial classifications and applications.

Ecological aspects of rivers in South Africa

Research on the ecological aspects of rivers in South Africa began in the early 1950s and was followed by a number of studies in the 60s, 70s, 80s and 90s. Many recent studies utilise historical data from earlier studies which enable comparisons of biological and/or physico-chemical data between current and historical conditions to be made and the degree of change, in for example, water quality, to be ascertained. Availability and accessibility of documented studies are often problematic since much of the early work was published in reports that are presently not readily available.

Biobase

In 1998 a database (known as the Biobase, WRC K5/626) was developed comprising biological (macroinvertebrate) and physico-chemical data derived from documented studies of riverine ecosystems within South Africa. The database, which collates a vast amount of information pertaining to riverine macroinvertebrates and water chemistry, has been used extensively in several projects aimed at the protection and management of our lotic ecosystems.

Revised Biobase

The revised version, which incorporates web-based functionality, was released in 2010 (WRC K8/906). In the latest development a technical manual has been created for Biobase, which is now accessible on the River Health Programme's Rivers Database website. Detailed information on viewing and querying data in the Biobase has been provided, together with screen dumps of the actual database. These include the four key viewing components, namely Sites, Biological & Physico-Chemical Data, Taxonomic Data, and Study Reference Information. The Query Master is described and steps for extracting data listed. The utility and limitations of the database have been discussed. These include, among others, deducing ranges of different physico-chemical parameters, for different taxa; assessing changes in community structure, using historical records; examining biotope preferences of specific taxa; assessment of water quality guidelines and input into the ecological Reserve; determination of macroinvertebrate reference conditions; development of conservation targets; and determination of the geographic distribution of taxa.

Collation and application

The database, which collates a vast amount of information pertaining to riverine macroinvertebrates and water chemistry, has several useful applications if used with the awareness of the problems outlined in this report. In establishing this database, one of the problems encountered, that was both difficult to resolve and avoidable, was that caused by the lack of consistency in the way in which different authors present their data. Frequently, useful data are lost, merely because they cannot be compared with others. The data captured in the Biobase include most of the biological data on macroinvertebrates prior to 1998. Since then numerous studies have been undertaken and a wealth of data exists.

Data collection

The database has been constructed using data pertaining to South African rivers and extracted from much of the available literature and unpublished reports, in which biological and physico-chemical data have been collected concurrently.

Biological data

Most of the biological data that are available relate to the benthic invertebrate fauna, although some work has been done on fish taxa. This bias is probably a result of the early recognition of the fact that benthic fauna provide an easy and fairly reliable way of assessing pollution. Records of the invertebrate riverine fauna thus form the biological component of this database. These data include those from intensive studies of individual systems, extensive one-off surveys of regions, ad hoc surveys and impact assessment reports. Thus far 43 studies, of which 40 had associated physico-chemical data, have contributed to the biological records of the database.

Physico-chemical data

The physico-chemical data used have been extracted from the same literature sources as the biological data, but vary between studies in terms of the number of variables analysed. The main criteria for the inclusion of physico-chemical data have been the exact or approximate coincidence of these measurements with those of the relevant invertebrate biological details. A complete list of the study references used is provided. Each of these studies are also described in detail, listing which records were extracted and what, if any adjustments, were made to allow compatibility with other records.

A total of 140 000 biological records have been entered into the database thus far, and most of these are accompanied by records of physico-chemical conditions (between 1 and 48 physico-chemical variables covered in each case) in the river at the time of sampling. The time spanned by the records dates from 1951 to the 1998, and, in the case of some rivers, records are available from several consecutive studies, which provide a historical trace of both biological and physico-chemical conditions.

Access to the Database

The Biobase has been incorporated into the website of the River Health Programme: ***http://www.riv.co.za/Rivers***. For users that are already registered to use the Rivers Database, they simply need log in and select "Biobase" on the menu bar. New users not yet registered need to register as a new user by completing the user registration form and submitting the electronic form to the Rivers Administrator. Once they receive their user information the user can login to the Rivers Database and select "Biobase" on the menu bar. This manual is available for download from the Biobase menu bar option.

Structure of this report

This report takes the form of a manual. It describes the structure of the database, both in terms of the spatial frameworks incorporated and the type and quality of data. It outlines the steps for users to select and view the data online and to extract data by using the Query Centre, which is designed to facilitate extraction of data by users. Pre-defined query frameworks enable querying of biological data, physico-chemical data, and physico-chemical parameters linked to biological ones.

This manual also describes the potential uses of the database and, more importantly, expands on the limits of its applicability. There is always a danger that a large store of data such as this may exude an air of reliability by virtue of its size alone. It is important to understand and be aware of the problems involved in amalgamating records from different sources, relating to data gathered by different authors, at different levels of intensity, and for different purposes.

Structure of the database

The intrinsic variability of biotic and physico-chemical components of riverine ecosystems within South Africa has necessitated the adoption of a hierarchical framework within which biological and physico-chemical data are accessed and queried. The primary level is the regional or geographic framework, the secondary level is at the longitudinal differentiation and the tertiary level is the site. Biological data are further differentiated into the level of biotope. Physico-chemical data are at the level of site. The hierarchical relationship between each level is schematically indicated in a diagramme.

Viewing data in the database

All users are able to view data. In future it may be possible to enable users to add their own data to the database, although aspects such as quality control will need to be resolved prior to this function being made available. All viewing and querying is accessed via the "Biobase" on the menu bar in the Rivers Database website. Data for viewing include Sites, Biological and Physico-chemical Data, Taxonomy and Study references, the Query Master and User Manual.

Querying the data (Query Master)

The Query Master is accessed via the Biobase on the Rivers menu bar by selecting "Query Master – Biobase". The Query Master is an HTML application, which allows the user to construct user-defined queries by combining the available attributes (fields) of registered data components (pre-defined queries). These can be sorted and filtered to control the data that is returned by the query. The Query Master is designed to filter the dataset using set criteria so that the raw data required can be copied to a spreadsheet (such as MS Excel) for further analyses. **Note** for biological data only taxa PRESENT at a site are returned.

Future collections

One of the more important recommendations to emerge from this area of work is that future biological and chemical collections should conform to the standard units of physico-chemical measurements as described in this User Manual and that for biological data details of proportional abundance, as well as factors such as biotope type, should be provided.

In addition, the actual dates on which both biological and physico-chemical data were collected should be available for reference, where they are not actually presented in published reports. Preferably an electronic copy of all data, in its original form should be lodged with a responsible organisation, but funding would be needed to capture data from studies undertaken that are not yet in the Biobase.

Further reading:
To obtain the report, *The Biological and Physico-chemical Database (version 2): User Manual (***WRC Report No. KV262/10***)*, contact Publications at Tel: (012) 330-0340; Fax: (012) 331-2565; Email: orders@wrc.org.za or Visit: www.wrc.org.za to download a free copy.