Web Access to Environmental Databases: a Database Query and Presentation System for the UK Environmental Change Network.

S.C. Rennie
CEH Merlewood
Grange-over-Sands
Cumbria, LA11 6JU, UK
+44 (0)15395 32264
srennie@ite.ac.uk

A.M.J. Lane
CEH Merlewood
Grange-over-Sands
Cumbria, LA11 6JU, UK
+44 (0)15395 32264
mlane@ite.ac.uk

M. Wilson
CEH Merlewood
Grange-over-Sands
Cumbria, LA11 6JU, UK
+44 (0)15395 32264
mw@nerc.ac.uk

ABSTRACT

This paper describes the development of an interface to an environmental monitoring database using the World Wide Web (WWW) as a front-end. The Environmental Change Network (ECN) is the UK's long-term integrated environmental monitoring program. Many of ECN's users require access to ECN's data for reporting purposes. Since these users are unfamiliar with the use of the Oracle Relational Database Management System used by ECN and the structure of the data in the ECN database, they require guided access to the data. Additionally, these users generally do not require access to high-resolution data. They need access to summary data in order to view trends, display the data for reporting and as an aid to making policy decisions. To support these users, ECN has developed an interface to its summary data over the WWW (http://www.nmw.ac.uk/ecn/Database). The system allows users to query the ECN database by site, measurement and date range and displays the data in cross-tabulations or graphs. Data may also be downloaded to the users local machine via electronic mail.

Keywords

Environmental data, WWW, ADO/ASP, Database, Guided access.

1. INTRODUCTION

Environmental change has emerged as an important issue on the global agenda. Consequently, there is a demand for access to objective, reliable and up-to-date environmental information. The routine collection of data from integrated environmental monitoring programmes can provide a wealth of data for scientific researchers. It is of prime importance that this data is made quickly available for research and policy decisions. The challenge facing managers of environmental databases is to provide network access methods to suit the different styles of requirement. These access methods should give sufficient guidance to users unfamiliar with the structure of the data whilst at the same time providing users with sufficient flexibility in data query and presentation.

The Environmental Change Network (ECN) is the UK's long-term integrated environmental monitoring program. It consists of a series of sites throughout the UK where both biological and

physical aspects of the environment are intensively monitored. This paper describes ECN's approach to allowing access to its summary database through the use of the Internet.

2. THE ENVIRONMENTAL CHANGE NETWORK

The ECN programme is sponsored by a consortium of UK government departments and agencies (see acknowledgements) with an interest in the environment, who contribute to the programme through funding either site monitoring or network co-ordination activities. ECN's primary role is to provide an information resource for scientific research to; (1) identify and quantify natural and man-induced environmental factors; (2) distinguish short-term fluctuations from long-term trends; (3) improve understanding of the causes of change and (4) predict future changes. The programme is also designed to provide more immediate information about trends and early warning of environmental extremes that may directly influence environmental policy.

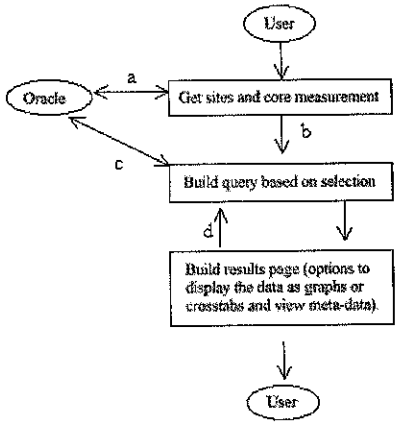
There are currently 12 terrestrial and 42 freshwater sites in the network, selected to cover as far as possible the main range of environmental conditions present in the UK. The monitoring programme includes a wide range of physical, chemical and biological 'driving' and 'response' variables, grouped into 'core measurement' areas, identified as being important for the assessment of environmental change. This suite of variables is measured at the same position within each site at the same time, using standard protocols incorporating standard quality control procedures [4]. The intention is that ECN provides a broad baseline of integrated environmental information relevant to a range of current environmental issues (e.g. climate change, water pollution, biodiversity, atmospheric pollution) and future, less obvious, issues. There are many projects that focus on long term environmental change in particular aspects of the environment, but ECN's principle strength lies in its multidisciplinary approach, enabling the analysis of important relationships between environmental variables and across ecosystem components.

3. THE ECN DATABASE

Databases form the core information resource for long-term monitoring programs like ECN. Long-term environmental research databases must be reliable and stable in terms of data quality, secure over a long time span, accessible but with access controls and allow for spatio-temporal analyses of a range of variables at a range of scales [3]. The use of reliable, well-known and well-supported database software is of paramount importance for long-term security.

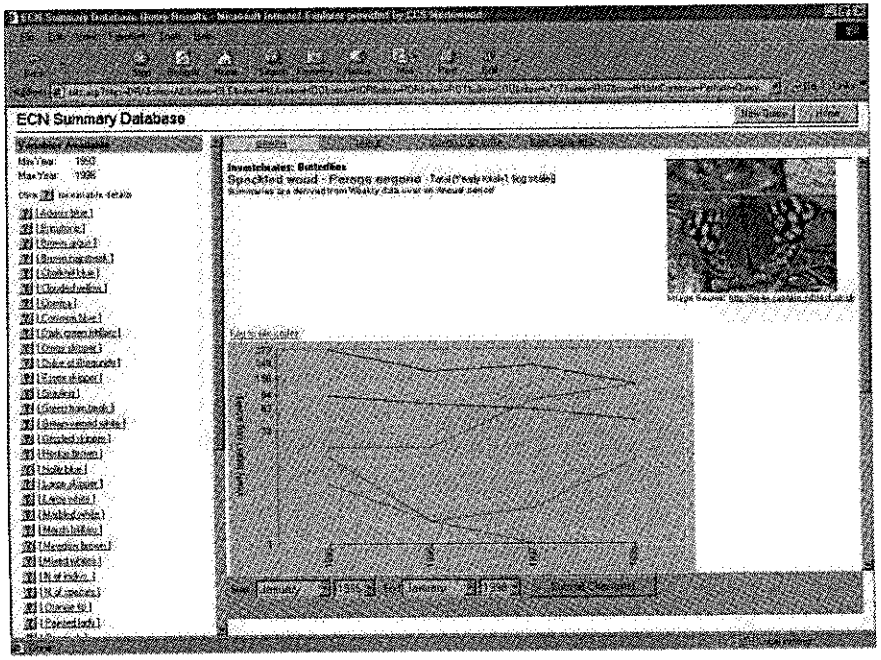
A central ECN database with remote access links was the model agreed by the sponsors as the most appropriate for ensuring a fully integrated system with the required data quality standards [2]. The database is managed by the ECN Central Co-ordination Unit (ECN CCU) at the Centre for Ecology and Hydrology's (CEH) Merlewood Research Station in Cumbria (UK). ECN's strategy is to maintain a centrally managed core database with good network access for users and to establish distributed database links with other related information sources.

The database uses the Oracle relational database management system (RDBMS), with links to the Arc/Info and ArcView, Geographical Information Systems (GIS), for spatial data handling. The database, GIS and software run on Unix-based servers (connected on a local area network) with links to the Internet for remote access and incoming data. The system provides an integrated storage and retrieval facility for all ECN data, incorporating meta-information including dataset descriptions and derivations, units of measurement and quality information. The core database stores raw data at the resolutions specified in the ECN protocols. An associated summary database (also held in Oracle) consists of monthly and/or annual summaries of this data using summary statistics appropriate to each measurement, as advised through ECN's expert committees. Access to the summary data is freely



The system is designed to be as generic as possible, allowing new measurements, types of graph and images are added to the system through meta-database table updates, rather than through the programming code. For example, the appropriate graphing method for each variable is determined from a meta-data code and image references can be stored in the meta-database so appropriate pictures, such as the butterfly species, can be displayed when the data are queried.

Figure 4: Example output from the ECN Summary Database Interface



Since users are accessing an automated system they do not have the opportunity for dialogue with the data originators or database staff therefore it is essential that the system be tailored to provide meta-information on data descriptions, measurement parameters and quality information. Meta-data concerning the derivation of the summary values, for example the number of samplings on which the figure is based, is automatically displayed as part of the cross tabulations and can be toggled on and off. The data (and associated meta-data) may also by downloaded via electronic mail in 'column' format, for import into local software.

Use of the ECN summary interface pages is monitored by IP address of the machine used for access and by the e-mail address entered by the user when downloading data. Monthly usage statistics are generated automatically.

6. SYSTEM EVALUATION AND USER FEEDBACK

The system was conceived and designed through discussions with potential policy users of ECN data, the Department of the Environment, Transport and the Regions (DETR) in particular. Following preliminary meetings, a prototype was developed and demonstrated, and a useful dialogue established for the development of the eventual system.

The ECN CCU hold regular meetings with sections of the user community (e.g. ECN site managers) and make presentations on the system and receive feedback on it. We also monitor the use of the system automatically and there is an opportunity for users to email comments and suggestions. As well as policy users, the system is used by scientific researchers, university and school students and members of the general public.

7. FUTURE DEVELOPMENTS

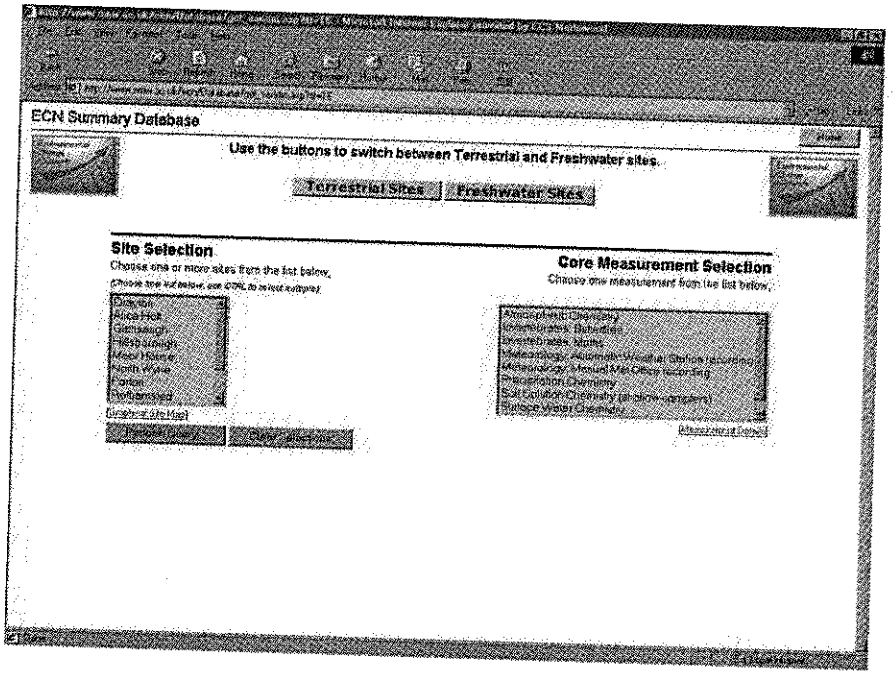
The described WWW interface satisfies many of the information access requirements for ECN by providing a combined information browsing, data presentation and data access system for the summary data. However, even more valuable to policy users would be some interpretation of these patterns, for example whether they are indicative of a long-term trend and whether this trend is likely to continue. Detailed analysis of causal processes is of course a complex task, but some of the key features of time-series data can be determined routinely using relatively simple models. For example, a web-based demonstration system has been developed for water quality data which shows how the relative contributions of the different model components can be extracted and presented to the user (http://www.nmw.ac.uk/eddemo/).

Whilst policy user requirements can be served by access to data summaries and interpretations, research scientists need access to the raw data for detailed analysis of spatial and temporal patterns. At present, any direct access to the database is provided via general-purpose SQL or MS-Access interfaces. This means that users need to understand the sometimes complex structures and relationships within the database, and how to interpret the data and its associated quality parameters. ECN will be evaluating ways of providing a tailored but flexible interface to its core database, which aims to incorporate embedded integrity rules, present data together with important meta-data on quality, and provide access to analytical and modeling tools.

In relation to these planned developments, ECN is beginning to explore the use of data warehouses and on-line analytical processing (OLAP) and data mining analysis tools used to extract data from data warehouses. These tools are designed for information analysis rather than information storage and may have considerable potential for an organization like ECN to support the increasing demand for derived information products and data exploration tools.

available, whereas access to the raw data is more closely controlled.

Figure 1: Example output from the ECN Summary Database Interface



4. INFORMATION ACCESS

It is of prime importance to ECN that data collected over the network are made quickly available for research and policy information purposes. ECN makes use of the Internet to disseminate the most current information and data to its wide user community. The ECN web site (http://www.nmw.ac.uk/ecn) provides higher-level meta-information about the network, as well as links to the ECN database and pages showing interpreted data, for example indicators of climate change and tutorials to teach school children about the weather and climate change. ECN also provides access to 'real-time' data from an automatic weather station (AWS) at the ECN Moor House/Upper Teesdale site in the north Pennines (UK). The AWS generates hourly data that are transmitted to the ECN CCU *via* a modem link. These data are stored in a database and graphs and tabulations of selected data are automatically generated for display on the web. Direct links from AWS's at other ECN sites are planned; as well as from other automatic monitoring instrumentation e.g. water quality loggers.

One particular challenge that ECN is starting to address, through the use of the Internet, is to provide data access methods to suit different styles of use and which can provide sufficient guidance and information about the system to enable little or no initial learning process for the user. The majority of ECN's users are unfamiliar with the use of an RDBMS like Oracle and with the ECN database structures. Additionally, many users, particularly policy-makers within ECN sponsoring organisations, do not generally need access to the high-resolution data. They require guided access to information and summary data for display, extraction and incorporation into reports, information

about indicators of environmental change and the interpretation of trends in the data. With this type of user in mind ECN has developed a 'tailored' web interface to its summary database (http://www.nmw.ac.uk/ecn/Database).

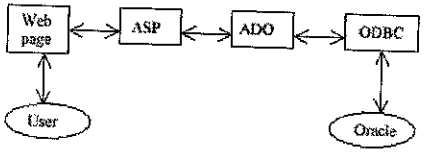
5. THE ECN SUMMARY DATABASE INTERFACE ON THE WEB

The ECN web interface enables users to build their own database query by selecting any combination of ECN sites, core measurement variables and date ranges for instant 'on-the-fly' generation of tables and graphs.

The system was originally written in 1996 using C with Oracle Pro*C embedded SQL to handle the database link [1]. It was one of the few direct Web-to-database interfaces available at that time and proved popular with users. However, significant programming effort was needed to make any changes to the system and it required users to navigate through several web pages before they could get the information they needed. Therefore the system was redesigned in 1998 to be more user-friendly and a lot more generic.

The system works using Active Server Pages (ASP) with an ADO/ODBC connection to the Oracle database. ODBC (Open Database Connectivity) is a C-level application programming interface (API) for SQL data. While developers can code directly to the ODBC API it can be fairly difficult and involve a lot of code. Because of this, Active Data Objects (ADO) are used as the interface to ODBC. ADO is an application-level programming interface. It is a data access object model that allows programmers to access data from any data source from any language including Visual Basic, Java, C++, VBScript and JScript. Figure 2 shows the ADO interface.

Figure 2: The ADO Interface



In the ECN system, ADO and SQL are used to create a record set from the database. This record set is used to create the HTML page using VBScript in ASP. A user chooses the sites and the core measurement that they are interested from lists (fig 1) which have been generated from the Oracle database (fig 3, a). A query is built based on the users choices (fig 3, b) and the data are retrieved from the database (fig 3, c). These data are stored as a record set which is temporarily stored in a 'session variable' so that the database does not need to be re-queried if the user chooses to view the data in a different format. The record set is used to build the main result window (fig 4). The user can choose to view the data in this window as a graph or cross-tabulation. They can also choose to view the meta-data associated with the data displayed or to download the data to their local machine via email. They also have an option to change the date ranges and variables that they have chosen (fig 3, d).

Figure 3: The flow of information through the system.

8. ACKNOWLEDGEMENTS

The UK ECN sponsoring organizations are: Biotechnology & Biological Sciences Research Council; Countryside Council for Wales; Defense Evaluation Research Agency; Department of Agriculture for Northern Ireland; Department of the Environment for Northern Ireland; Department of the Environment, Transport and the Regions; English Nature; Environment Agency; Forestry Commission; Ministry of Agriculture, Fisheries and Food; National Assembly for Wales; Natural Environment Research Council; Scottish Environment Protection Agency; Scottish Executive Rural Affairs Department; and Scottish Natural Heritage

9. REFERENCES

1. Brocklebank, M., Lane, A., Watkins, J. & Adams, S. (1996). *Access to the Environmental Change Network Summary Database via the World-Wide-Web.* (ECN Technical Report 96/1.) Grange-over-Sands: Environmental Change Network. (ECN, Merlewood Research Station, Grange-over-Sands, Cumbria, LA11 6JU, UK).
2. Lane, A.M.J. (1997). The UK Environmental Change Network Database: An Integrated Information Resource for Long-term Monitoring and Research. *Journal of Environmental Management*, 51, 87-205.
3. Michener, W. (1986). Data management and long-term ecological research. *Research Data Management in the Ecological Sciences.* (W K Michener, ed.), 1-8. Belle W. Baruch Institute for Marine Biology and Coastal Research, University of South Carolina.
4. Sykes, J.M. & Lane, A.M.J. (1996). Eds. *The United Kingdom Environmental Change Network: Protocols for standard measurements at terrestrial sites.* London: HMSO.