

**'Scoping project to investigate options for linking the standing waters database and the Lakes Inventory'**

Contractor: Michael Hughes, ENSIS Ltd. University College London.

m.hughes@ucl.ac.uk

Project Officer: Susan Watt, JNCC.

susan.watt@jncc.gov.uk

## Table of Contents

Summary	3
1. Introduction	4
2. Database description	4
3. Linking with other databases	5
4. Future development	8
5. Development options	10
6. Conclusions	13
7. Contact list	14
8. Links	15
Figures	16
Appendix 1. Example questionnaire	20

## Summary

The JNCC Standing Waters Database is a Microsoft Access database containing macrophyte, invertebrate and other biological survey data as well as water chemistry, geology and physical site data for standing water sites in the UK (England, Scotland and Wales). The Database uses a site/survey concept whereby the fundamental units of data storage are sites and surveys.

The purpose of this scoping study is to provide a scope for the development of the Database including advice on improving accessibility, linking with other databases (e.g. GBLakes database) and an investigation of data provision via the National Biodiversity Network (NBN) Gateway.

An analysis of the Database shows a large bias towards Scotland for sites and macrophyte data. However, other data types only occur for sites in England and Wales.

Problems of linking sites to other databases were investigated. Inconsistencies in site names and grid references cause problems in linking. The full JNCC site list was linked with the GBLakes inventory. 94% of sites were matched to a resolution of 150m; the remaining 6% could not be matched.

Potential links with the NBN Gateway were investigated. Some macrophyte data from the Database have already been uploaded. The NBN Gateway provides useful search and mapping tools but data cannot be identified to a particular water body as yet.

There is currently no active management of the Database. Organisations are continuing to collect data which is being stored in custom databases or proprietary software (Recorder). It is thought that several versions of the Database are in use across these organisations.

Future development of the Database is discussed with reference to 6 key questions. Four options are presented which range from doing nothing to developing the database to focus on macrophyte data with links to GBLakes and a web site. However, there is a need for a more detailed needs assessment to help inform the choice of development option.

## **1. Introduction**

### **1.1 Background**

The JNCC GB Standing Waters Database (referred to in this document as the Database) is a Microsoft Access database containing macrophyte, invertebrate and other biological survey data as well as water chemistry, geology and physical site data for standing water sites in the UK (England, Scotland and Wales). The Database was initially put together to house survey results from Scottish Natural Heritage (SNH) Loch Survey but now also contains some data from England and Wales surveys.

The Database is held by several organisations across the UK including Scottish Natural Heritage (SNH) and the Countryside Council for Wales (CCW).

### **1.2 Aims**

The purpose of this scoping study is to provide a scope for the development of the Database including advice on improving accessibility, linking with other databases (e.g. GBLakes database) and an investigation of data provision via the National Biodiversity Network (NBN) Gateway.

## **2. Database description**

### **2.1 Concept**

The Database uses a site/survey concept whereby the fundamental units of data storage are sites and surveys. A survey is associated with a site, a site can have many associated surveys. Figure 1 shows the database schema for the Site section of the database. Figure 2 shows the database schema for the Survey section of the database. These schema show a structure with a series of data tables linked to either the site or survey table. Data tables may also be linked to dictionary tables.

### **2.2 Analysis**

An analysis of the main tables is given in Table 1. Dictionary tables are excluded. Figure 3 shows the spatial extent of sites.

It is apparent that there are more data for some sites than others. Whilst the bias is on Scotland in terms of sites recorded and macrophyte data, there are some data types for which data only exists for Welsh sites (e.g. zooplankton and diatoms).

Table 1. Main data tables and numbers of records (number in brackets shows number of sites with data).

Table name	Number of records
Site Data	27652 (-)
Geology Data	31711 (27285)
Designation Data	4533 (4388)
SNH Area Data	27318 (27318)
District Data	0 (0)
NHZ Data	27302 (27276)
Reference Data	0 (0)
10km Square Data	27285 (27285)
Survey Data	3776 (3594)
Lake Type Data	3312 (3246)
Substrate Data	259 (63)
Macrophyte Data	80889 (3594)
Invertebrate Data	746 (29)
Open Water Zooplankton	277 (31)
Surface Sediment Diatom Data	1598 (31)
Epilithic Diatom Data	2218 (31)
NVC Data	6183 (2060)
Herbarium Data	3327 (1336)
Secchi Data	4 (4)
Primary Water Chemistry Data	2940 (2836)
Secondary Water Chemistry Data	62 (62)

### 3. Linking with other databases

#### 3.1 Problems

In order to link with another database the site data must be stored in a way that allows matching by grid reference, site name or site code.

Most of the site codes used in the Database were inherited from the original surveys and subsequent sites were coded in a similar way. However, it is not clear that other organizations use the same codes and perhaps only Scottish Natural Heritage (SNH) still recognise the codes used for Scottish sites. English sites are all prefixed EN, Welsh sites CW and Scottish sites use a coding system loosely based on their 6 figure grid reference.

Alternatively, site names may be used to link with external datasets but this has inherent problems too. A site name is not unique. For example, there are more than 30 occurrences each of the names Black Loch, Loch Dubh, Loch Fada, Loch Mor and Loch nan Eun. This problem is most acute in Scotland. Across the UK the problem is exacerbated by the use of local names, alternative spellings and frequent misspellings of names which makes matching on site names impossible unless additional data such as grid reference are also available.

Using grid references is perhaps the most reliable way of identifying water bodies or indeed any spatial object in the UK but is still fraught with difficulties. Many grid references are either not written down correctly at the time of data collection or are incomplete. The most common mistakes encountered are the switching of the X and Y parts of a six-figure grid reference, missing leading zeroes, missing letter identifiers, incorrect number reference due to poor map-reading, incomplete number reference. For example, the grid reference NH452099 Loch nan Eun (Database site code NH4001) might be incorrectly

recorded thus:

NH099452	X and Y parts switched
NH45299	Leading zero of X part omitted
452099	2-letter 100km square identifier missing
NH352099	X part misread

Additionally, the recording of grid references is a subjective process and one recorder may interpret the map differently leading to more than one grid reference for the same site. This is particularly the case for large lakes where the grid reference given is often the site where sampling took place. A six-figure grid-reference is only accurate to 100m.

Clearly, linking the Database with other databases is going to be a lengthy process with much manual checking of site locations and there will inevitably be some sites which will remain unmatched.

### **3.2 Links with the GBLakes database**

The GBLakes database (referred to in this document as GBLakes) uses a unique identifier system whereby each water body (as defined by the project) is given a water body ID (WBID). A water body is defined by its inclusion in the OS PANORAMA dataset and is represented spatially as a non-overlapping, contiguous polygon. GBLakes is the most comprehensive inventory of fresh waters in Great Britain to date and contains data for England, Scotland and Wales.

To explore the potential for linking the Database with sites in GBLakes an attempt was made to create a lookup table to translate site codes used in the Database with GBLakes WBID's. The aim here was to ascertain how easily this could be done and to suggest a methodology for future work in this area.

Matching was performed using a combined manual/automatic method based on grid references and site names in a geographic information system (GIS). Sites from the Database were mapped using their grid references and overlaid onto the GBLakes lake dataset.

Only 9024 of the 27652 JNCC Database sites (33%) were coincident with a GBLakes lake polygon, most of these were in Scotland. Originally, the Database site list was constructed using different methods for Scotland and England and Wales. Sites for Scotland were systematically generated from digital map data whereas for England and Wales they were added only if the site had been surveyed and then the recorded grid reference was used to define the site location. GBLakes was created systematically using digital map data (for the UK, not just Scotland) and so a good correspondence between JNCC Database sites and GBLakes sites for Scotland would be expected. Figure 3 shows a sample 20km square from Scotland. As the map extract shows, there is a consistent spatial shift between the two datasets which makes matching by overlay fail for small sites. Whilst the spatial mismatch is consistent it is not constant and it is not possible to convert one set of grid references to the other using simple formula.

Using sites names and grid references it was possible to assign a further 1942 JNCC Database sites (7%) to GB Lakes water bodies even though the sites grid reference did not exactly overlay a GB Lakes site. Finally, proximity functions in GIS were used to assign a further 14956 of the JNCC Database sites (54%) to the closest GB Lakes water body within a 150m radius (accounting for the fact that grid references were only six-figure and thus accurate to 100m). These matches are clearly not 100% certain but manual checking of each site is not possible.

The remaining 1730 sites (6%) in the Database could not be matched to a GB Lakes site i.e. they were more than 150m from any GB Lakes site.

Common reasons for not being able to match a JNCC site with a GB Lakes sites are:

- The site does not exist in GB Lakes, it may be too small or was never present in the OS PANORAMA dataset
- An incorrect grid reference has been provided for the JNCC site
- An inaccurate grid reference has been provided for the JNCC site

### 3.3 Links to NBN Gateway

The National Biodiversity Network (NBN) Gateway (<http://www.searchnbn.net/>) site allows you to view distribution maps and download UK wildlife data by using a variety of interactive tools. The Gateway can hold geographic and species datasets and includes things like boundaries for designated areas (e.g. RAMSAR, Local Nature Reserves) and species group data (e.g. crayfish, otter, dragonfly). Queries can be carried out using a map interface (see Figure 5).

Some data have already been made available to the NBN Gateway from the Database. Scottish Natural Heritage have provided what they call their 'Standing Waters Database' as well as several other datasets which contain macrophyte data (Rare Plants Database, Scottish Lowland Raised Bog Inventory)

The 'Standing Waters Database' dataset is described thus:

"This dataset comprises the macrophyte records from the JNCC standing waters database. It includes 79,246 records from 3,547 locations from across England, Wales and Scotland and date from 1970 to 2000. The parent database held by SNH also holds information on invertebrates as well as the physical and chemical attributes of the individual standing waters and their respective catchments. The Scottish data, which makes up the majority of the records were collected during the SNH Loch Survey project."

It seems, therefore that the data available through the NBN Gateway at present represent some of the macrophyte data from the Database, but this is a snapshot from some years ago. An email from Geoff Johnson (SNH) who was responsible for uploading the data to the Gateway confirms this.

The other main source of macrophyte data is the 'New Atlas of the British & Irish Flora 2002', described thus:

"The database of 8.9 million records contains the distributions of 4269 taxa of

flowering plants and ferns. The collated databases include those of Botanical Society of the British Isles (BSBI) Vice County recorders and some species specific databases, each of which is itself a collation of many sources (and of each other). Data collected for the 1962 Atlas of the British Flora are also included. Subspecies are omitted but many hybrids are included. The majority of the data have come from BSBI volunteer effort. The collation has been validated at the 10km square level by BSBI Vice-County recorders and the editors of the New Atlas, and is now believed to provide a good representation of vascular plant distributions in Britain and Ireland at this level, based on available data. Availability of data at resolutions higher than 10km square varies between sources.”

Note that data displayed in the NBN Gateway are located geographically by grid-reference and are not attributed to a particular lake or site. Grid-references are degraded to 2km resolution for most users of the site but are available at the original 100m resolution (i.e. 6-figure grid reference) to registered users.

#### **4. Future development**

There are some basic questions that need to be reviewed before a plan for future development of the Database can be put forward:

- What is the purpose of the Database?
- Who provides data for the Database?
- Who will use the Database?
- Who will be the custodian of the Database?
- How will the Database be accessed?
- What other databases will it link to?

##### **4.1. What is the purpose of the Database?**

The database currently holds site data and survey data. Site and survey data are heavily biased towards Scottish sites. The site data are largely duplicated in the GBLakes database, which is now being actively developed. The advantage of GBLakes is that it covers England, Scotland and Wales and was created using a unified methodology and single data source for all countries. Future development of the Database should focus on the collation of survey data with appropriate links to the GBLakes database for site data.

Survey data are mainly macrophyte survey data but other data types exist such as invertebrate, zooplankton and diatom. To what extent are these data types still being collected and/or stored in other databases? Should the Database be solely focused on macrophyte survey data? Should the Database be adapted to hold river survey data or are these data available elsewhere? (E.g. LOCAR River Conservation Database <http://www.nwl.ac.uk/locar/metadata/panglambourn/ecology/3rdparty/ConsRiver.htm>)

The original database was designed to be linked to ArcView. Now that



ArcView has been superseded by ArcGIS (ArcMap) are these links still appropriate? What GIS packages do organisations use now or intend to use in the future? Is a web-based mapping interface such as that used in NBN Gateway preferable?

#### **4.2. Who provides data for the Database?**

This is a crucial question. It seems that since the database was originally constructed there has been no attempt to regularly update it with data from across the relevant organisations (SNH, SEPA, CCW, EN etc...). Anecdotal evidence suggests that some data may have been added since but where these are, who entered them, how they have been validated etc is not known.

The majority of the data in the Database come from the SNH Lochs survey with additional data from the Countryside Council for Wales (CCW) and English Nature (EN) for a smaller number of sites.

To what extent do external organisations such as Environment Agency (EA) and those just mentioned plan to continue data collection and what protocols do they already have in place to do this? In other words, is there actually a need for continuing development of the JNCC Database?

CCW uses Recorder 2000 across the organization to store all biological data. Data are then meant to be uploaded from Recorder to NBN Gateway. However, some macrophyte data gathered as part of their technical support program is entered in their copy of the JNCC Database (T. Hatton-Ellis, *pers. comm.*).

SNH have been collecting new macrophyte data in various one-off surveys and more systematic survey campaigns (such as 2004 Site Condition Monitoring project carried out by CEH). However none of these data have been added to the Database and some data are being stored in a separate database (but using JNCC Database species codes) (M. Hennessy, *pers. comm.*).

The Environment Agency is currently collecting data as part of the Pilot Monitoring Project and collating other data for the LEAFPACS project. They are using the same survey methodology as the SNH 2004 survey. The data could be accommodated in the Database with minimal preparation.

#### **4.3 Who will be the custodian of the Database?**

When the Database was first constructed it appears there was no plan put in place for database management and there was no data custodian formally appointed. Consequently, the Database has undergone sporadic and *ad hoc* revision and there are undoubtedly several versions in use across the country. Several organisations are currently storing macrophyte data in other databases or spreadsheets and there is no attempt at harmonising and collating data across organisations.

Future development of the Database would require a data custodian to act as a focus for data input, database version control, collaboration with NBN and ongoing development.

#### **4.4 Who will use the Database?**

To date, the Database has been used mainly by staff within the originating organisations (SNH, CCW, EN). However, we do not exactly know who uses the database, how many copies have been circulated, who is updating data etc. Are there other organisations (such as university departments) who routinely collect data which could be stored in the Database? More research is needed to answer this question, including the potential cost benefits by data sharing and synergies implementing monitoring for the Water Framework Directive and the Habitats Directive.

#### **4.5 How will the data be accessed?**

The current method of distributing the database is by MS Access database. This is problematic for several reasons. It can result in several organisations holding different versions which they have edited or added data to. Not everyone has the software needed to run the database. The database front-end (interface) is not sufficiently advanced to allow easy addition of large amounts of data.

The NBN Gateway allows access to data to a wide range of end-users from the general public to the data collectors themselves. Is the type of data presentation available via the NBN Gateway (i.e. dot maps at varying resolutions with overlays of protected area boundaries, vice counties etc, see Figure 5) sufficient for macrophyte data or is there demand for a dedicated website or data portal, perhaps linked to GBLakes website?

#### **4.6 What other databases will it link to?**

Links with the GBLakes database have already been discussed. But what other databases are there that may be usefully linked to the JNCC Database? The Environment Agency have a biological data database for internal use.

### **5. Development options**

The previous section has presented a number of questions which require further investigation. The final choice of option for the development of the Database would need to be informed by more detailed database needs assessment. A questionnaire to relevant stakeholders (EA, SNH, SEPA, CCW, CEH, EN etc) would probably be the best way to achieve this. An example is given at Appendix 1. Following this a data pathway can be prepared to clarify the processes for data collection and deposition in the database.

Since the inception of the JNCC Standing Waters Database a number of changes have been made across various organisations which mean that the Database and the way it is managed is no longer viable.

It is especially important to look at how data are collected in the field. SNH and EA are now using the same survey methodology (developed by CEH for

the 2004 Scottish Site Condition Monitoring project). So far these data have been stored in custom databases or spreadsheets with no protocol for getting them into a format suitable for the Database.

There is a clear need for a detailed database needs assessment which would answer the questions posed in section 4 above and set out a framework for future data collection, management and storage. This framework should cover the entire lifetime of the data starting at the first stages of data collection during field survey through to presentation of data to the end-user. Above all, whatever strategy is decided it needs to be coordinated by an appropriate body using existing networks where possible.

One thing that is not clear is whether it is acceptable for each countries statutory organisations to 'do their own thing' or whether there should in fact be a co-ordinated UK approach to data management whether via NBN or some other method.

Based on the assessment of the Database carried out in this report the following options for development are proposed. The choice of option will depend largely on the outcome of a more detailed needs assessment but may also be driven by national and international policy (such as the Water Framework Directive and the lake habitat action plans).

### **5.1 Option 1 – Do nothing**

Individual organisations continue to use the database as they wish.

*Advantages:* No cost, no work. Existing links with organisations GIS will be unaffected.

*Disadvantages:* The database remains with incomplete and inconsistent site data. The interface is not user friendly making data entry difficult. There remain no links with NBN. Data will not be standardised across organisations. Access to data restricted to those who hold a copy of the database.

Analysis and use of datasets at a UK level not possible.

### **5.2 Option 2 – Re-design existing stand-alone database**

The database will be completely re-designed in Microsoft Access. The 'site data' section will be removed and survey sites will relate directly to GBLakes sites. The database will be primarily for macrophyte data although the other data types will remain available. The data custodian will manage a central database and issue versions at a periodic interval. New data will be provided via a network of organisations using standardised spreadsheets.

*Advantages:* A streamlined database will be easier to manage. Versioning system will aid database management. Data could be sent to NBN Gateway periodically.

*Disadvantages:* The problems of a stand-alone distributed database remain. Problems of longevity.

### **5.3 Option 3 – Re-design existing stand-alone database with web access**

As Option 2 but with dedicated website linked to database to act as data portal. Data can be searched, mapped, downloaded through a web browser. A MSAccess version would still be available. Periodic updates would be controlled using versioning system. The data custodian would manage updates, making changes to the master version of the Database and 'releasing' new versions as appropriate. New data could be uploaded via the web interface for the data custodian to quality control before adding to the Database.

*Advantages:* As above with added advantage of easy access to data by public. Data could be linked to GBLakes database / website and NBN Gateway.

*Disadvantages:* Will require most work and skilled web/database programmer. Will require a permanent home on a web server somewhere and a long-term management strategy.

### **5.4 Option 4 – Abandon Database and use Recorder software**

The database data will be archived (macrophyte data updated to NBN gateway) and a new methodology developed. Recorder 2002 which is designed specifically for site survey data collection and data can be uploaded to NBN directly.

*Advantages:* Some organisations already using Recorder (CCW). Data can be uploaded to NBN Gateway easily. Up to individual organisations to develop and manage their own data policy.

*Disadvantages:* Requires each organisation to have software (at a cost) and an experienced operator to input data. Data cannot be linked to GBLakes resulting in repetition of site data. Survey methodology might result in data which are difficult to store in Recorder.

## **6. Conclusions**

The JNCC Standing Waters Database holds a large amount of macrophyte data, mainly from Scottish surveys, as well as some other data types. Today, similar data are being collected across the country by various organisations but there is no process for getting these data in the Database and some organisations have devised their own databases for internal use. Therefore the Database as it stands is nothing more than an archive for a heavily biased set of surveys.

The Database can be linked to GBLakes although there is a small percentage of sites which cannot be linked.

Data from The Database can be uploaded to the NBN Gateway and this has been done for the majority of macrophyte data.

A set of key questions is posed to assist with the choice of option for future development of the Database. Four options for future development are proposed.

Any future development of the Database would require a data custodian to act as a focus for data input, database version control, collaboration with NBN and ongoing development.

## 7. Contacts

*Joint Nature Conservation Committee (JNCC)*

Susan Watt (Coastal and Freshwater Habitats Advisor)

[susan.watt@jncc.gov.uk](mailto:susan.watt@jncc.gov.uk)

Andrew Brewer (NBN technical Liaison Officer)

[Andy.Brewer@jncc.gov.uk](mailto:Andy.Brewer@jncc.gov.uk)

Andrew Cottam (Species Data Custodian)

[Andrew.Cottam@jncc.gov.uk](mailto:Andrew.Cottam@jncc.gov.uk)

*Countryside Council for Wales (CCW)*

Helen Wilkinson (Biodiversity Information Team)

[H.Wilkinson@ccw.gov.uk](mailto:H.Wilkinson@ccw.gov.uk)

James Dargie (Technical Support Program)

[J.Dargie@ccw.gov.uk](mailto:J.Dargie@ccw.gov.uk)

*Scottish Natural Heritage (SNH)*

Sue Alvin (Natural Heritage Data Unit (NHDU) Data Custodian)

[Sue.Alvin@snh.gov.uk](mailto:Sue.Alvin@snh.gov.uk)

Nina Turner (NHDU Data Officer)

[Nina.Turner@snh.gov.uk](mailto:Nina.Turner@snh.gov.uk)

Geoff Johnson

[Geoff.Johnson@snh.gov.uk](mailto:Geoff.Johnson@snh.gov.uk)

Alison Lee (SNH/JNCC)

[Alison.Lee@snh.gov.uk](mailto:Alison.Lee@snh.gov.uk)

Jennifer Davidson (NBN/LRC Co-ordinator)

[Jennifer.Davidson@snh.gov.uk](mailto:Jennifer.Davidson@snh.gov.uk)

*Centre for Ecology and Hydrology*

Laurence Carvalho (2004 Site Condition Monitoring)

[laca@ceh.ac.uk](mailto:laca@ceh.ac.uk)

*Environment Agency (EA)*

Jo-Anne Pitt (Ecology Advisor (Water Framework Directive))

[jo-anne.pitt@environment-agency.gov.uk](mailto:jo-anne.pitt@environment-agency.gov.uk)

*English Nature*

Stewart Clarke (Senior Freshwater Ecologist)

[Stewart.Clarke@english-nature.org.uk](mailto:Stewart.Clarke@english-nature.org.uk)

Nick Stewart (Independent contractor, Stonewort expert)

[nstewart@eurobell.co.uk](mailto:nstewart@eurobell.co.uk)

**8. Links**

GBLakes

<http://ecrc.geog.ucl.ac.uk/gblakes/>

NBN Gateway

<http://www.searchnbn.net/>

Recorder Software

<http://www.nbn.org.uk/information/info.asp?Level1ID=1&Level2ID=1>

<http://www.dorsetsoftware.com/recorder/index.htm>

# JNCC Standing Waters Database Scoping Study

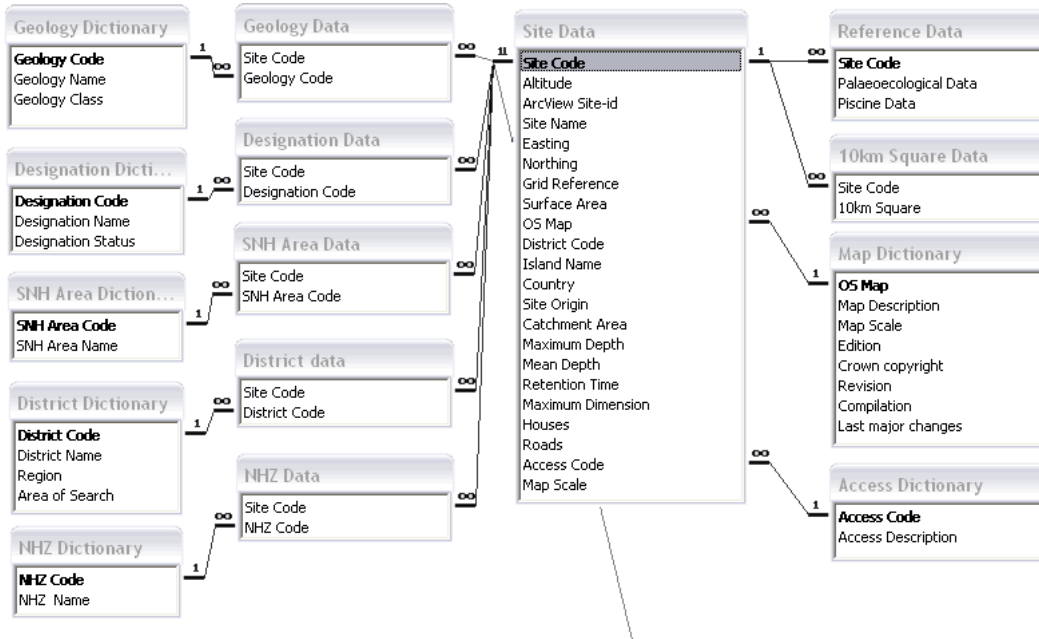


Figure 1. JNCC Database schema – Site Data



# JNCC Standing Waters Database Scoping Study

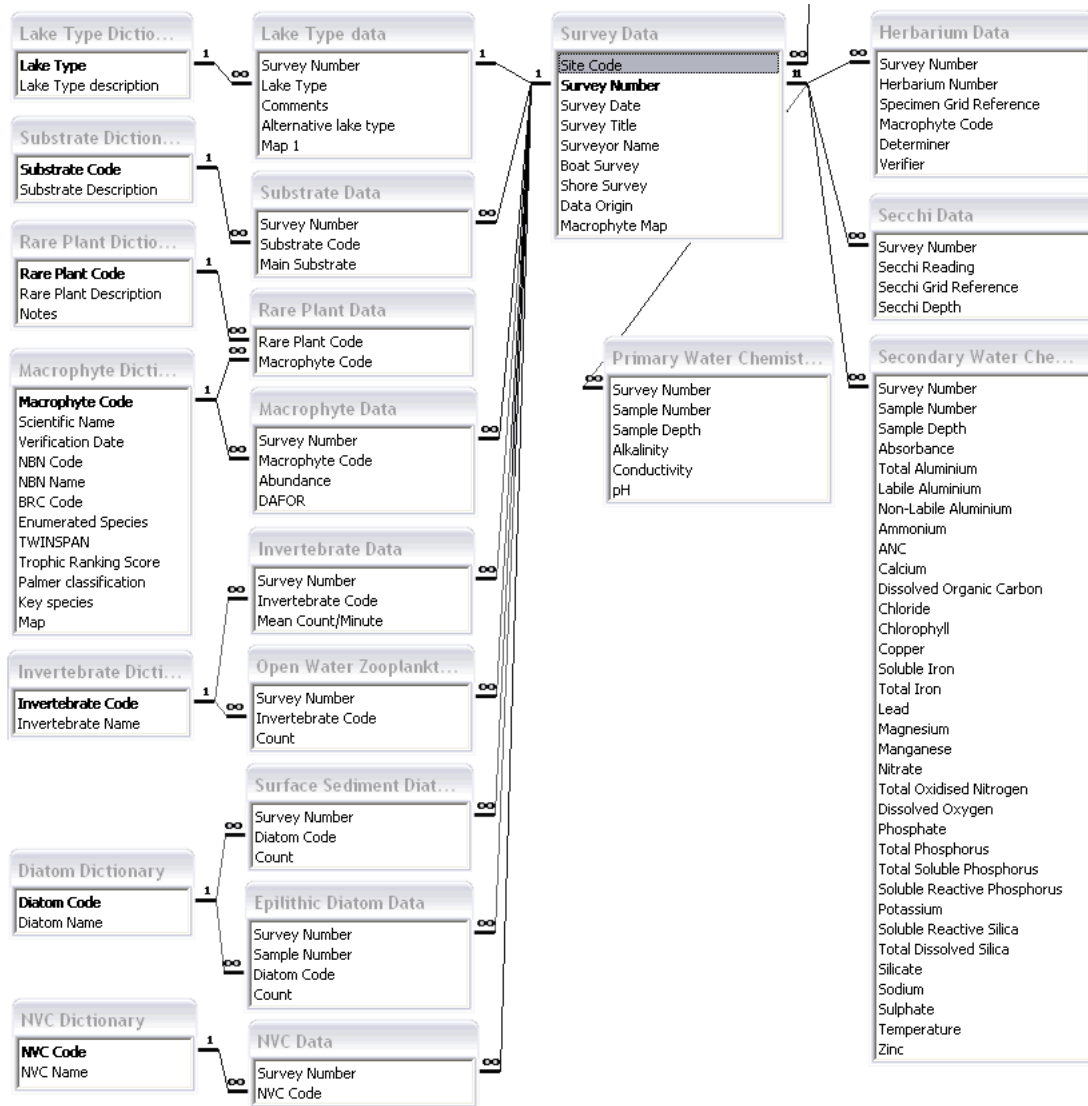


Figure 2. JNCC Database schema – Survey Data

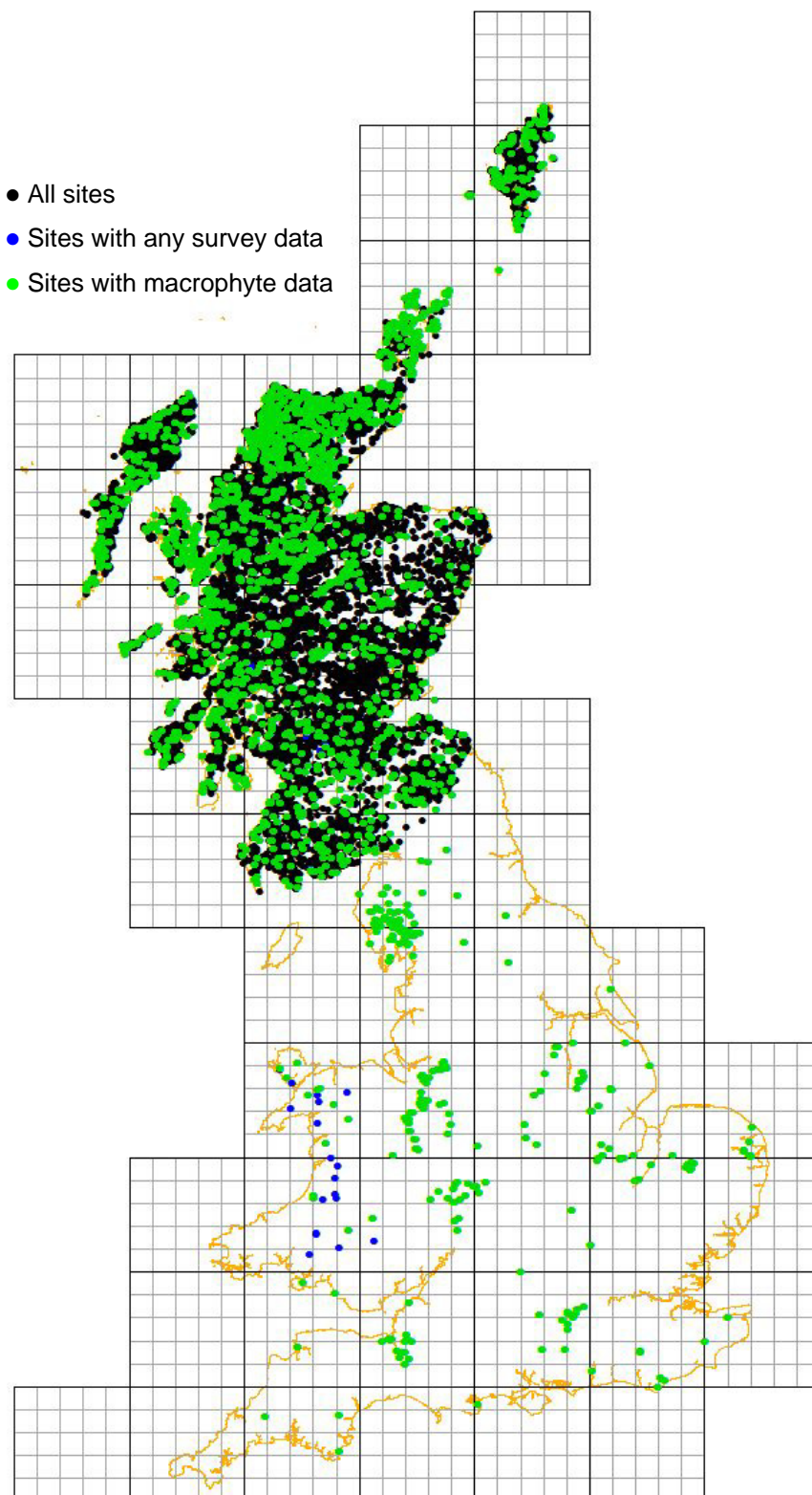


Figure 3. Map showing distribution of JNCC Database sites in UK

## JNCC Standing Waters Database Scoping Study

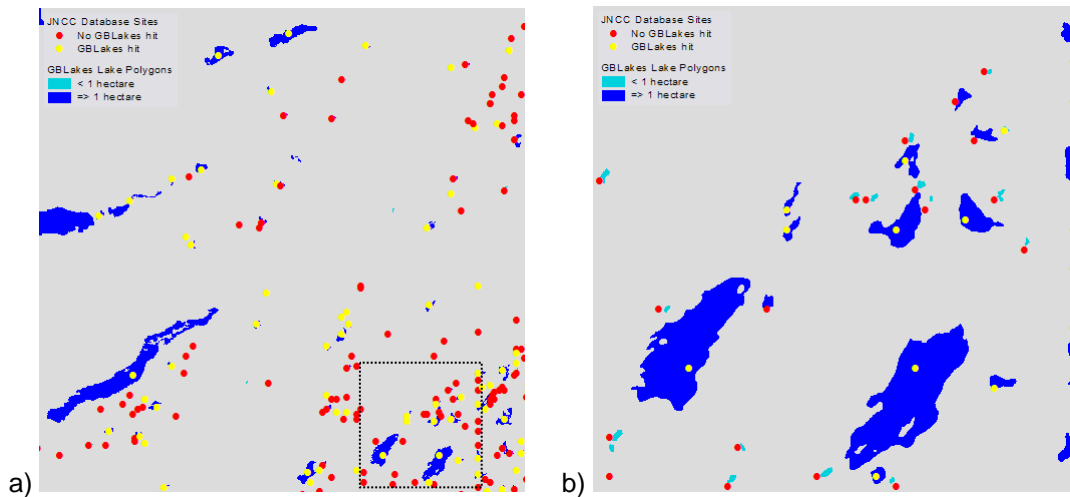


Figure 4. Excerpt from GIS of 20km square NH22 showing coincidence of JNCC Database sites with GBLakes sites for a) entire 20km square and b) subset of 20km square. Red dots are JNCC Database sites which do not overlay a GBLakes Lake Polygon, yellow dots are JNCC Database sites which do overlay a GBLakes Lake Polygon.

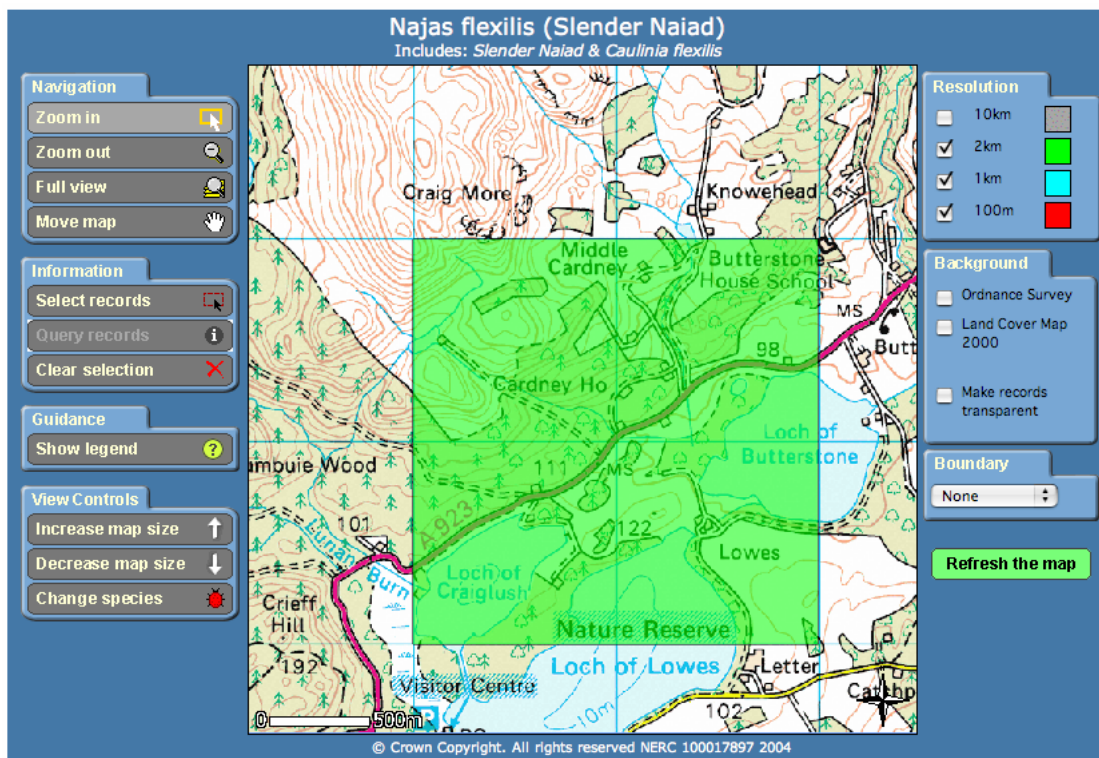


Figure 5. Excerpt from [www.searchnbn.net](http://www.searchnbn.net) showing example data extraction for Slender Naiad from SNH 'Standing Waters Database'. The 2km resolution makes it impossible to see which lake this record refers to. 100m data are available to certain registered users.

## Appendix 1. Example needs assessment questionnaire

The following set of questions could be used to assess the database needs of an organisation or project for either general purposes or more specific data type (e.g. substitute 'data' with 'macrophyte data').

- What kind of *data* do you collect?
- What is the primary purpose of *data* collection?
- What is the frequency of *data* collection?
- How many sites do you collect *data* for?
- What survey methods do you use in the field to collect data?
- How do you identify survey sites? What coding system is used?
- How are *data* validated?
- How are *data* stored?
- Are there any written procedures or policies for *data* collection and storage?
- Assuming *data* are stored electronically what software does your organisation use to store and query *data*?
- What kinds of searches or analyses do you need to perform on your *data*?
- Do you access *data* through a Geographic Information System (GIS)?
- What kind of computing system do you use?
- What level of technical support is available to you?
- What kind of problems do you have managing *data* at your organisation?
- How do you make *data* available to others within your organisation?
- How many other people at your organisation collect similar *data* to you?
- Do you have a statutory obligation to make *data* available to the public?
- Assuming *data* are public domain how does your organisation go about making *data* available to the public?
- How are *data* presented to the public?
- Do you have a *data* custodian who is responsible for managing *data*?
- Do you need to provide an audit trail for the *data*?
- Briefly describe the organisation you work at.