

Using the BSBI Distribution Database

[Version date 1st March 2015]

Compiled by Andy Amphlett

Contents (Ctrl + click to go to that section in document)

Introduction	Page 3
Getting started	Page 4
A note on web browsers	
Things you can use the DDb for without a login	
How to login to the DDb	
Message board	Page 7
Maps	Page 8
Searching the DDb	Page 11
Introduction	
Explanation of terms in query window	
Taxon	
Grid reference	
County / region	
Date recorded	
Checklists / attributes	
Bounded area	
More options	
Group by, distinct and subdivide by	
Locality (precision)	
Data source and dataset	
Grid reference and vice-county validation	
Join with a sub-query	
A note on searches for large numbers of records	
FAQs – how to find what you want, and other useful stuff, using the DDb	Page 24
Lists	Page 25
How do I create a list of all species / taxa recorded in a vice-county?	
How do I create a species list for a vice-county subdivided by BSBI date class?	
How do I create a species list for a vice-county comparing post 2000 to pre 2000?	
How do I create a list of species with earliest and most recent year recorded?	
How do I create a species list for a hectad listing most recent year each taxon has been recorded?	
How do I create a list of all taxa in a vice-county with the number of tetrads / monads with records since a specified year?	
How do I create a list of all infraspecific taxa in a vice-county?	
How do I create a list of all subspecies in a vice-county?	
How do I create a list of all records for Nationally Rare and Nationally Scarce taxa in a vice-county in a year?	
How do I list all records for a vice-county added to the DDb since a date?	

<u>Maps</u>	<u>Page 31</u>
<u>How do I create a map showing tetrad survey coverage in a vice county for the period since 2000?</u>	
<u>How do I create a map of survey effort across a vice-county?</u>	
<u>Altitudes</u>	<u>Page 33</u>
<u>How do I find the altitude of plant locations?</u>	
<u>How do I find the maximum and minimum altitude for all taxa in a vice-county or group of vice-counties?</u>	
<u>How do I summarise species frequency in altitude class intervals?</u>	
<u>New vice-county / hectad / tetrad / monad records</u>	<u>Page 37</u>
<u>How do I create a list of all new vice-county records in a year?</u>	
<u>How do I create a list of all new hectad records in a vice-county in a year?</u>	
<u>How do I create a list of all new tetrad records in a vice-county in a year?</u>	
<u>How do I create a list of all new monad records in a vice-county in a year?</u>	
<u>How do I create a list of new hectad records of Nationally Rare or Nationally Scarce taxa in a vice-county in a year?</u>	
<u>Data sources and datasets</u>	<u>Page 40</u>
<u>How do I view all DDb records for a year, originally entered into an individual copy of MapMate?</u>	
<u>How do I create a list of all data sources and data sets with records in the DDb for a vice-county?</u>	
<u>How do I view records of a taxon or group of taxa, which are restricted to or exclude individual data sources and / or data sets?</u>	
<u>Rare plant registers</u>	<u>Page 41</u>
<u>How do I use a list of Rare Plant Register taxa in a query?</u>	
<u>Grid references and vice-county errors</u>	<u>Page 43</u>
<u>How do I view all records for a vc, where the site grid reference is outside the vc boundary?</u>	
<u>How do I view records for another vc, that have erroneous grid references that fall within 'my' vc?</u>	
<u>How do I find out the correct (or inferred) vice-county for records misattributed to a vice-county?</u>	
<u>How do I take data analysis and mapping further?</u>	
<u>Feedback and future versions of this guidance</u>	<u>Page 45</u>

Introduction

The distribution database (DDb) is maintained by the Botanical Society of Britain and Ireland, to store and make accessible the ever-growing pool of biological records collected or collated by the society. The DDb holds c.37 million records including 14 million from the Vascular Plant Database (VPDB) at BRC, 14 million records from MapMate users, along with many county databases and data from specialist surveys. New and historic datasets continue to be added. Since 2000, between 436,000 and 1.1 million new records have been added to the database each year (median 846,000). 23 vice-counties have more than 500,000 records on the DDb, and three vice-counties have more than 1 million records each.

The DDb provides vice-county recorders, research workers, Agencies and NGOs with on-line access to large, detailed data sets and, where appropriate, allows data to be edited to correct mistakes or to add annotations. The system provides sophisticated search and reporting tools. Additional capabilities are being added frequently.

Distribution maps of taxa, and read-only access to the 'message board' are available to anyone, you do not have to have a login password. However all access to view, download, or if you have the necessary access permissions, to edit records, requires a login. Requests to register for access to the DDb should be made via this link <http://bsbidb.org.uk/createuser.php>.

These guidance notes are designed to help new users of the DDb get started. They explain how to view distribution maps, set up and run queries, and download data. The notes are illustrated by examples, chosen to be typical of the sorts of questions that a VCR or other user might ask of the database. Guidance is not comprehensive, but sufficient for someone to be able to make good and effective use of the DDb. Familiarity with these notes will equip someone to tackle more complex tasks themselves, or to ask for help via the DDb 'message board' (see below). Functionality of the DDb is also constantly being refined; bugs fixed, new options added etc., so the guidance notes may, in time, become slightly out of date.

The DDb has begun to change the way that BSBI works. In the past, VCRs worked in comparative isolation, collecting and collating records for their vice-county, and only periodically passing these, or a summary, to BSBI in support of a specific project, eg an Atlas. With the widespread adoption of MapMate as a recording database by many VCRs, electronic submission of records to BSBI is now universal, and MapMate records are incorporated into the DDb every two to three weeks. Other datasets, in a wide range of formats can also be incorporated, though this has to be undertaken by Quentin Groome or Tom Humphrey.

Rather than taking a somewhat insular (vice-county) perspective, the DDb facilitates botanists taking a more holistic view, across any geographic area, or time period, that they wish. So 'my' records add value and provide perspective to my neighbouring VCR's records, and vice-versa. A user of the DDb can quickly gain an overview of where a species is being recorded, and who is recording it; hence the database becomes a shared resource, where each of our inputs, of new records, or by validating or correcting existing records, benefits all users.

Getting started

A note on web browsers

The DDb, most specifically the maps, require that you use a modern web browser, eg. Google Chrome, Firefox or Internet Explorer (version 9 or later). It is recommended that you use the latest version of whichever browser you use. To check if you have the latest version go to http://www.whatbrowser.org/intl/en_uk/

Things you can use the DDb for without a login

The DDb provides access for all, to distribution maps and a range of tools for checking and formatting records. Tools and maps are accessed via the tabs at the top left of the home page, highlighted in yellow in the screenshot below.



Maps

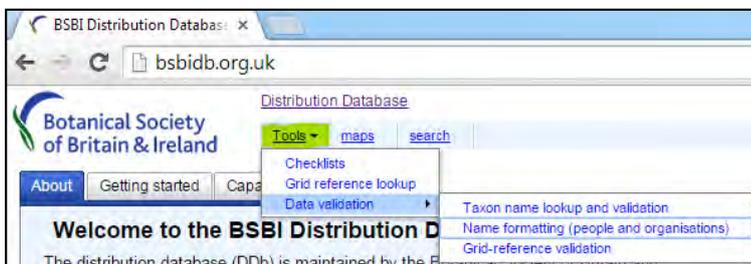
Hectad and tetrad distribution maps for all taxa are viewable, updated daily from the live database. All functionality is available except that:

- zoom level is restricted to tetrad. With a login, monad and 100m grid square resolution records are viewable.
- you cannot search the database by clicking on the map (login required).

To find out more about DDb mapping, go to [Maps](#)

Tools

Checklists, Grid reference lookup and several Data validation tools are accessed via the Tools tab.



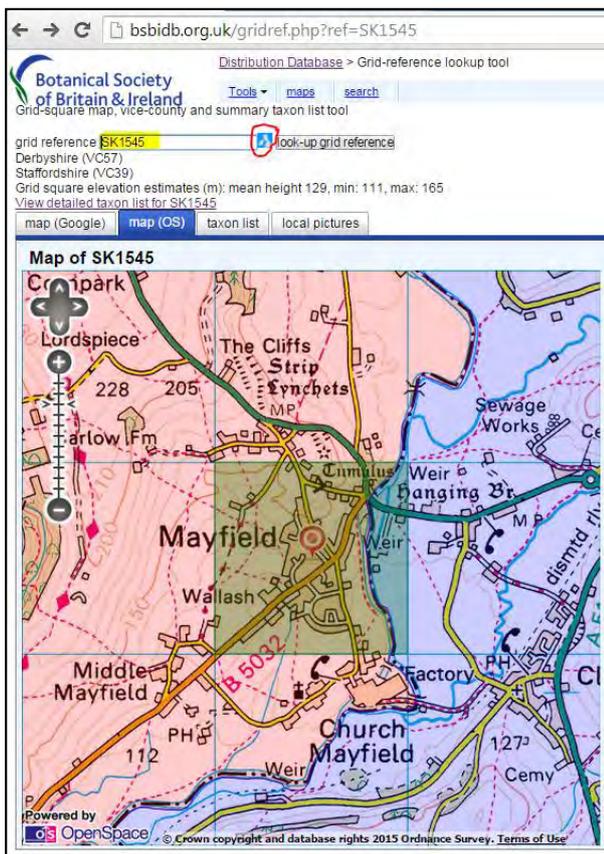
Checklists

More than 60 Checklists, available for use in DDb queries can be accessed via this tab. They can be downloaded as csv files for use in spreadsheets. As used by the DDB, 'Checklists' provide a general-purpose way of grouping or applying attributes to records when running queries. Checklists consist of a set of 'key' values which link to a field (e.g. a taxon name). Each item in the checklist set may optionally have a data value associated with each key.

Grid-reference lookup

The grid reference lookup tool can be accessed directly at <http://bsbidb.org.uk/gridref.php>. This is the best available tool to check the location of a single grid reference, and to find out which vice-county the grid reference is within or overlaps. In addition, the mean, minimum and maximum altitude for the grid square are given.

Enter a grid reference in the box highlighted in yellow, below, and click on 'look-up grid reference'. Alternatively the grid reference picker tool can be used, circled in red below; (see [Grid reference](#) for more details). The displayed map highlights the grid square defined by the grid reference, and the vice-county within which the grid reference falls, is shaded. If the grid reference overlaps the boundary of two or more vice-counties, all intersected vice-counties are shaded in different colours. In the screenshot below, monad SK1545 is shown to overlap the vc39 / 57 boundary; vc37 to the west of the River Dove, vc57 to the east of the river. The mean altitude of the monad is 129m (range 111 - 165m).



The default is the Ordnance Survey base map, but aerial photography can be used as a backdrop via the 'map (Google)' tab. The 'local pictures' tab is a very nice touch, linking to photographs on <http://www.geograph.org.uk/>. Currently (17/02/2015) this website has photographs taken in more than 273,000 monads (82% of all monads in GB and Ireland).

Data validation



Three data validation tools are available; taxon names, recorder names and grid references. These work in a similar way. Records are submitted as tab delimited text files. In each case you have to enter the column number for the values to be validated. NB the first column is column 0!

Please note! Results are returned as a .csv file. To open these particular downloads in Excel, you must open Excel first, then open the file, which initiates the text import dialogue box. The file opens correctly when the delimiter is set to 'tab', rather than 'comma' as you would expect. However, if using LibreOffice, the .csv file opens directly from the download. This appears to be a bug in Excel.

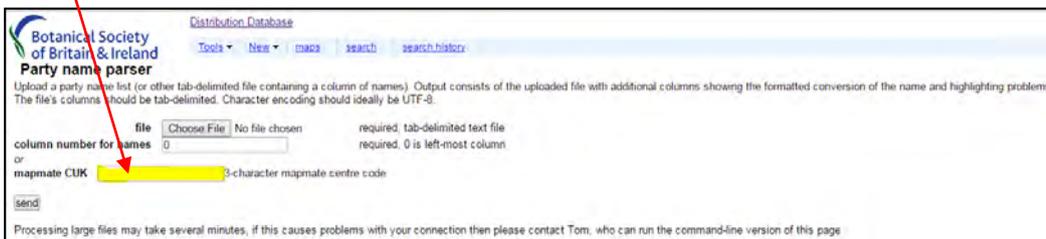
Taxon names

Matches taxon names from other sources against accepted names in the DDb taxon lists (largely a variant of the 3rd edition of Stace's Flora). You can specify a range of additional output columns, including the matching MapMate taxon and vernacular name. Names which cannot be parsed are indicated and should be reformatted manually.

Recorder name formats

This tool attempts to apply standard formatting (e.g. **Surname**, Forename or **Surname, A.;** **Other**, A.N.) to a list of names of people or organisations. Names which cannot be parsed are indicated and should be reformatted manually. The DDb attempts to split strings of multiple recorders into individual entries but the process fails if punctuation is used inconsistently. The DDb tool is also available direct at <http://bsbidb.org.uk/partynameparser.php>

By entering a MapMate CUK rather than uploading an external file, the tool can test existing names, created in that copy of MapMate, and for which records are held on the DDb.



This will highlight any names that the DDb can't split up correctly and is a simple way to check for and amend names in MapMate, so that when MapMate is next re-synced with the DDb the recorder names can be re-evaluated. ('Failed' names still appear in the DDb, but the problem is that the system can't tell if they are people or organizations (or a combination of both) so sorting, searching and formatting of the entries works less well.)

Grid-references

Checks for malformed references and matches grid-references to vice-county boundaries. Grid references that cannot be parsed are indicated and should be reformatted manually.

Outputs from the validation tools will require some subsequent manipulation in a spreadsheet. Note carefully the advice at the top of the previous page on how to open output files in Excel.

How to login to the DDb

Either via the BSBI home page <http://bsbi.org.uk/>, where there is a Quick Link to 'Distribution Database' or at <http://bsbidb.org.uk/>.

Click on login, enter username and password. By default your password is not saved; each time you go to the DDb website, you will have to re-enter your username and password.



The screenshot shows the login form on the DDb website. The username field contains 'AndyAmphlett' and the password field is masked with dots. The 'save details option' dropdown menu is set to 'shared computer (do not save password)'. There are 'login' and 'register' buttons at the top left, and 'Tools', 'maps', and 'search' links at the top right. A link to 'click here' is provided for forgotten passwords.

Assuming you are not using a shared computer, changing the save details option to 'stay logged in for 30 days' speeds up use of the database considerably. In this case, do not log-out when exiting the DDb.



This screenshot is identical to the previous one, but the 'save details option' dropdown menu is now set to 'stay logged in for 30 days'. A red arrow points to this dropdown menu.

The DDb can be used with a variety of web browsers, but Google Chrome is recommended, and is most consistently reliable.

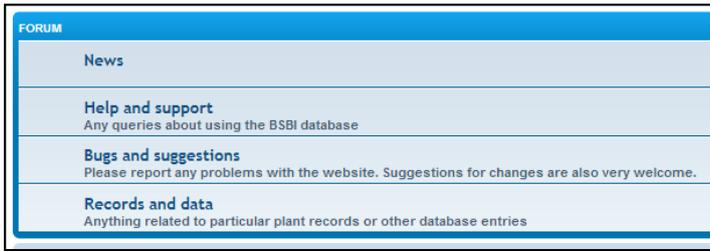
Message board

At the top right of the DDb home page, is a tab for the 'message board'.



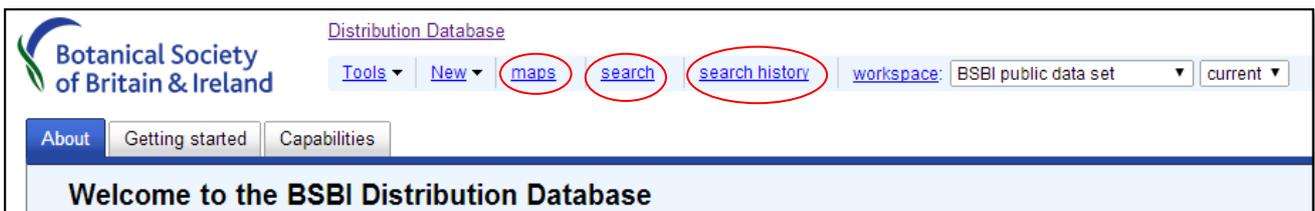
The screenshot shows the navigation bar at the top right of the DDb home page. It contains links for 'message board', 'my details', and 'logout Andy(AndyAmphlett)'. The 'message board' link is circled in red.

Postings are listed under four headings.



The message board is becoming a very useful repository of tips on how to do things, any problems users face (and solutions), and suggestions for improvements. Please have a browse through the posts and replies, and if you have a log in, please ask questions via this route, as the replies may be of help to others.

On the DDb home page several menu options are listed in blue. These guidance notes mainly refer to 'maps', 'search' and to 'search history'; one of the options under 'tools' is also discussed.



Maps

Distribution maps are accessed via the maps tab, or directly via <http://bsbidb.org.uk/maps/>.

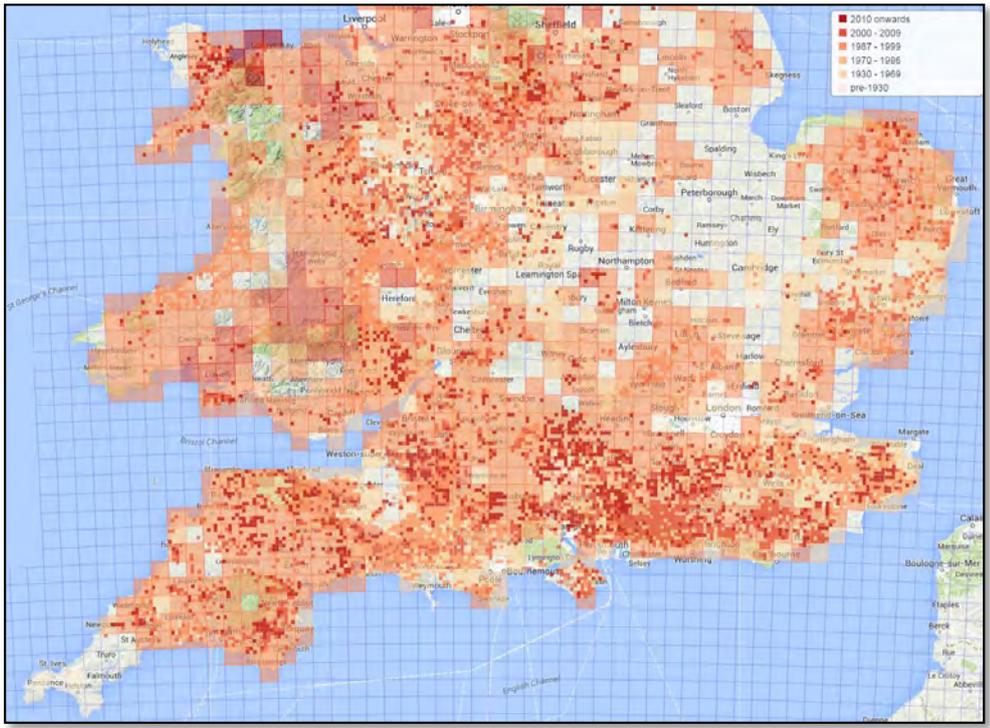
Anyone can access hectad and tetrad distribution maps for all taxa, no login is required. Maps are updated daily from the live database. All functionality is available except that:

- zoom level is restricted to tetrad, and
- you cannot search the database by clicking on the map.

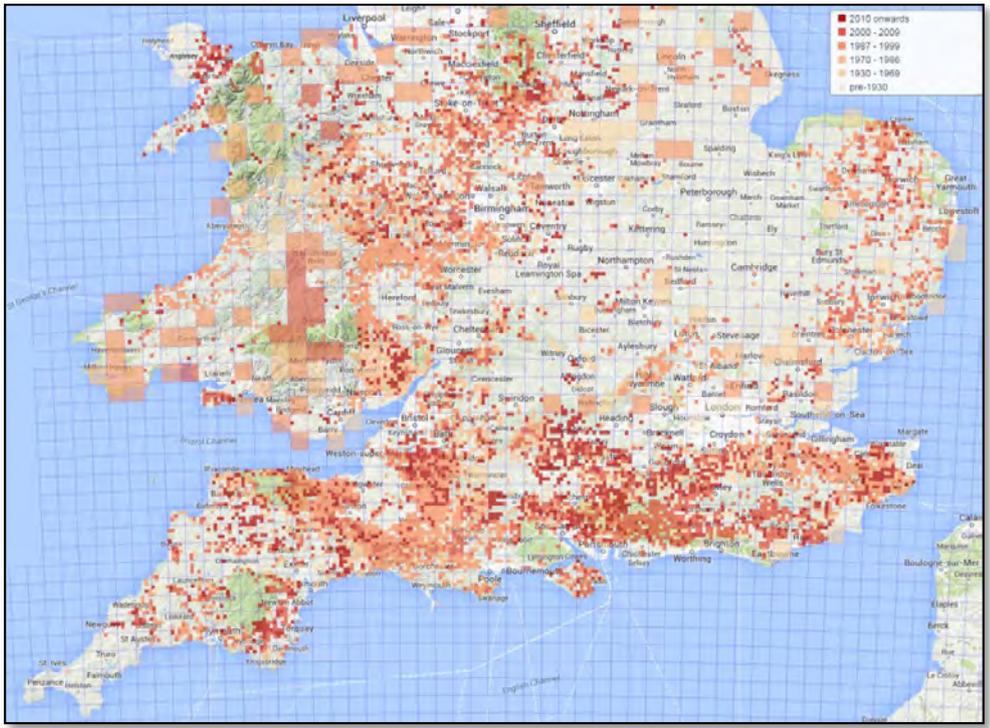
With a login, monad and 100m grid square resolution records are viewable on the maps, and you can search for records by clicking on the map.

The maps tab provides the optimum way to access distribution maps for the whole of GB and Ireland. Similar mapping can be accessed via the 'search' tab, but creating maps via searches is much slower, especially for common taxa. For the commonest species (with >50,000 records in the DDb), the resultant map is incorrect, as the web browser display limit is set at 50,000, and any map is therefore based on incomplete data. Using the 'maps' tab circumvents this issue.

Maps appear very quickly, and are first drawn at hectad resolution. Zooming in using the mouse wheel reveals more detailed resolution records. By default maps are drawn with the zoom option set to 'zoom to best available' and partition by set to 'date class'. The screenshot below shows the default options for *Adoxa moschatellina*, zoomed part way in to show tetrads.

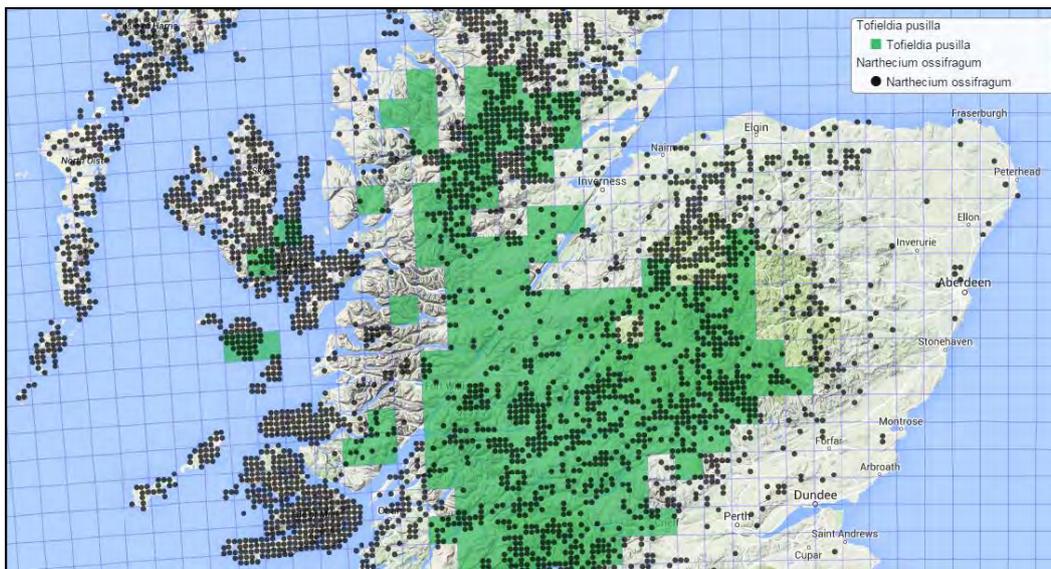


Changing the zoom option to 'zoom to best available (without residue)', hides less detailed records as you zoom in, and provides the clearest mapping. See screenshot below, again for *Adoxa*.



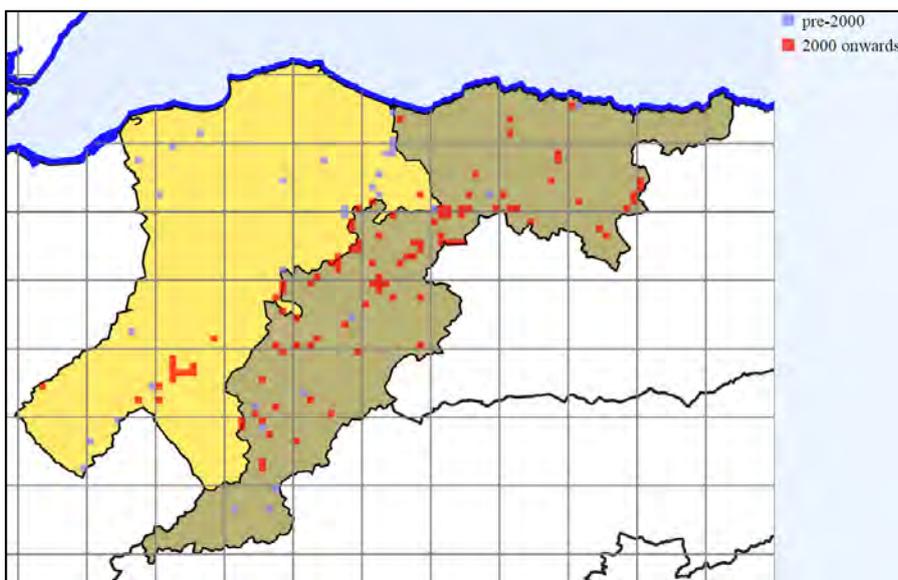
The maps are highly customizable, and the options selected are retained within a single browsing session.

More than one taxon can be mapped, click on 'add another data series'. The example below overlays the tetrad distribution of *Narthecium ossifragum* over the hectad distribution of *Tofieldia pusilla*.



Distribution data can also be overlain over air photos.

Printable map options are fairly sophisticated, eg the map below shows the monad distribution of *Alchemilla xanthochlora*, pre and post 2000, in vc95 and vc94; the two vcs are coloured differently and their boundaries shown, along with the boundaries of the adjacent vcs.



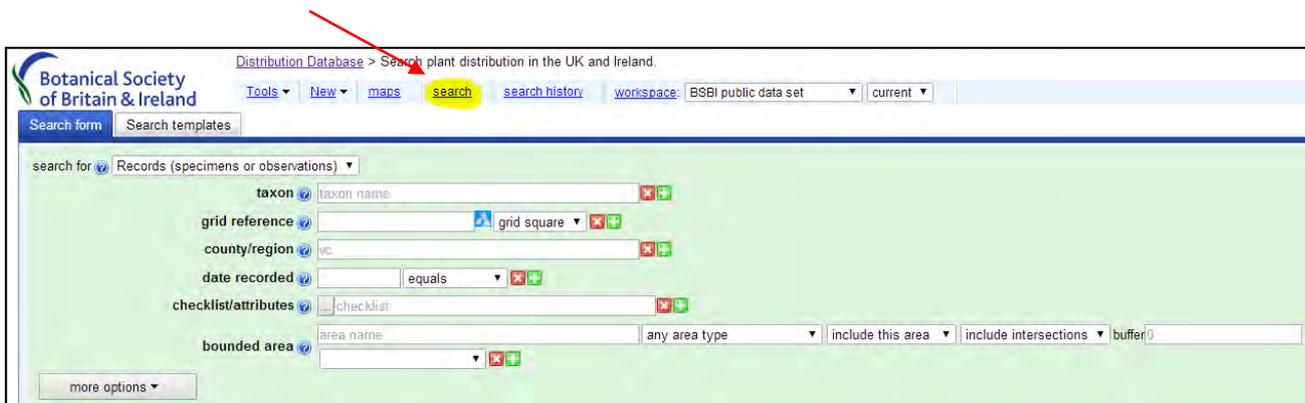
Once printing options are set, they are retained within a single browsing session. Printable maps can be saved as pdf and png files, and in several other formats. Such maps are suitable for use in presentations, eg as slides. However, creation of printable maps is rather slow, and best used for creation of a small number of maps. For many less critical purposes, using the zoomable maps and copying or saving part of the map using the Windows 'Snipping tool' is a satisfactory way to create maps.

Searching the DDb

Introduction

The DDb allows for great flexibility in searching for, viewing, mapping, summarising and downloading records. **To carry out searches you must be logged in.**

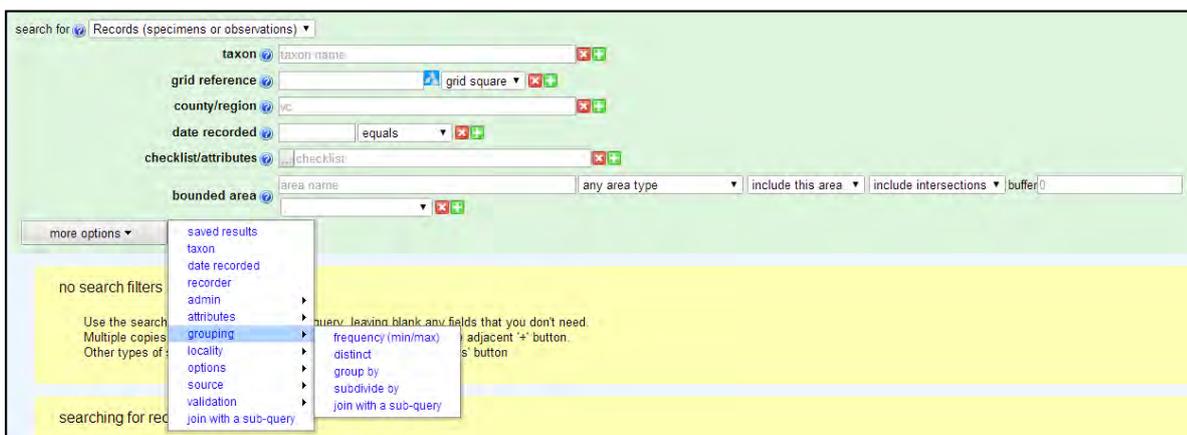
Click on the 'search' tab.



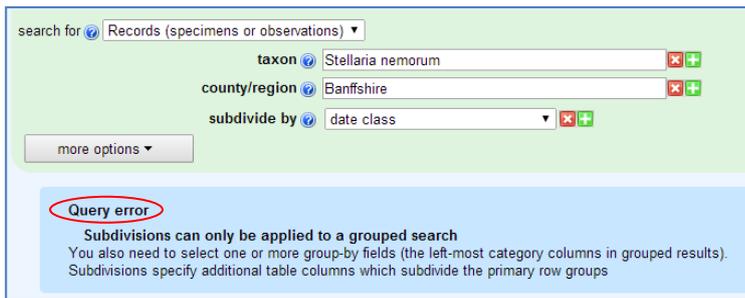
This opens the query window:



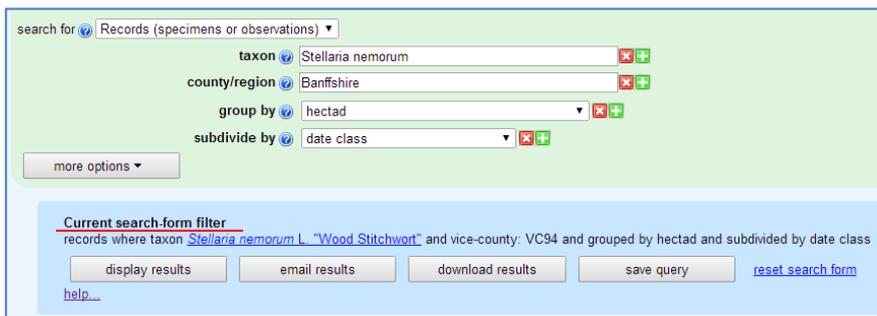
Clicking on the 'more options' drop down arrow reveals a number of additional options, some with further drop down arrows:



Experimentation is recommended! Don't worry, if you create a query that the DDb is not able to run, it gives a 'Query error' message, eg.



If the query is a valid one, a text description of the query is given as the 'Current search-form filter'.



Explanation of terms in query window

In this section, many of the terms used in the query window are explained.

Taxon

The DDb has more than 8000 taxa in its species dictionary, and while its default taxonomy is the 3rd edition of Stace's Flora, names in the 2nd edition are included as synonyms. Starting to type in a name generates a list of possible taxa; select the correct one as it appears; it is not a scrollable list. If you enter a synonym, eg starting to type *Potentilla pal*, the window prompts you to select *Potentilla palustris* = *Comarum palustre*. Selecting this option, the taxon tab shows the name *Potentilla palustris*, but when the query is run (click on the 'display results' tab), the displayed records will have *Comarum palustre* as the taxon name.

You can search for more than one taxon at a time by either; clicking on the green plus (+) sign at the far right end of the taxon line; or by entering additional taxon names separated by commas (,) or semi-colons (;).

Genus and Family can also be used in searches.

Grid reference

In queries, grid references are accepted in OS / OSI formats. Tetrads searches are allowed, in DINTY format, but searches for records within Quadrants (5 x 5km squares) are not possible. Grid references can be entered manually or by using the pop-up grid reference selection tool, which also includes a gazetteer search capability.

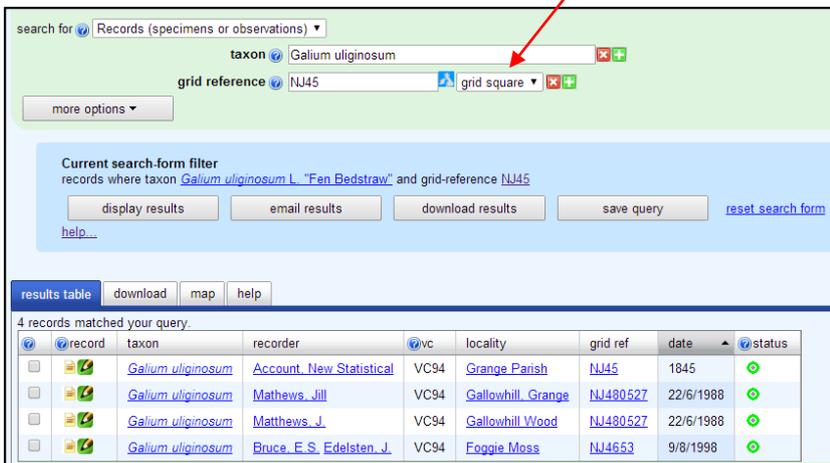


The pop-up window can be resized and zoomed in or out and the basic mapping can be changed to an air photo. Clicking on the map or image creates a list of grid references at varying precision; from hectad, tetrad, monad, 100m down to 10m (8 figure), from which you can select the grid reference to use.



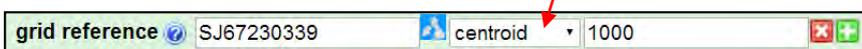
The selected point on the air photo above is the car park at Pitlochry train station.

The default is to use a grid reference as a grid square:



In this case only the four records of *Galium uliginosum* whose grid references fall within the square defined by NJ45 are returned when the query is run. Additional grid references can be added to a query by clicking on the green plus (+) sign at the far right end of the grid reference row in the query. The query then finds records of a taxon in any of the squares defined by the list of grid references. Alternatively, a list of grid references, separated by commas (,) or semi-colons (;) can be entered.

Grid reference can also be used as the centre (centroid) of a search area.



SJ67230339 is the centre of the world famous 'Iron Bridge' in Ironbridge, Shropshire (vc40).

The radius of the circle is in metres. The default is 5000m; change this to your preference.

Running a query to find all records within 1km of this grid reference returns (as of 4th May 2014), 8796 records from 114 unique grid references – table below.

GridRef precision	Number of unique gridrefs
10m	37
100m	64
Monad	8
Tetrad	4
Hectad	1

Using grid reference as a centroid finds all records whose site grid reference (which defines a square), falls within or overlaps the circle defined by the centroid and radius. This is why in the above example (a 1km radius search area) there are records for 4 tetrads and a hectad. These large squares overlap the smaller search area. If you wish, records with coarse precision site grid references can be excluded from a query, (see section on [Locality \(precision\)](#) below).

County / region

As well as vice-county, there are options for grouped vice-counties, eg. Perthshire (vc 87, 88 and 89) or Yorkshire (vc 61 – 65), and for the individual countries, Scotland, Wales, England and Ireland. The latter refers to the whole island of Ireland (vc H1 – H40). Running a query for records from one or more vice-counties returns records with that vice-county listed in the vice-county field in the record. Alternatively, and really only for more specialist purposes such as record validation, vice-county can be entered as a [Bounded area](#), see below.

Date recorded

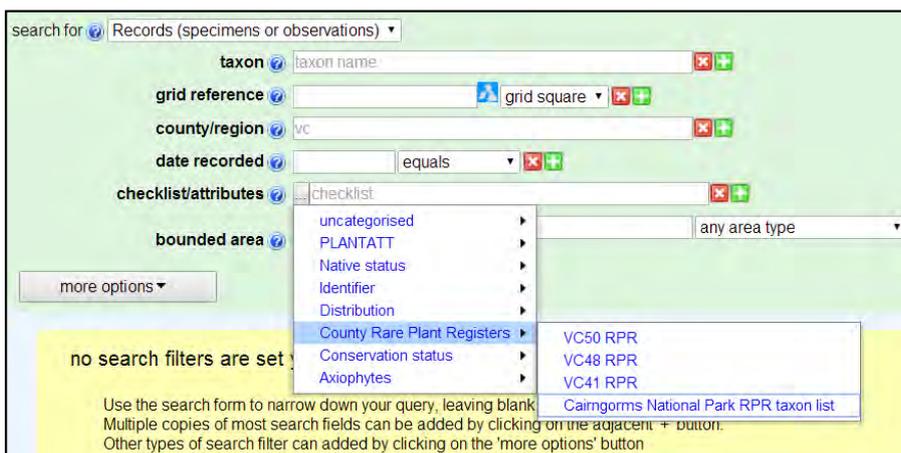
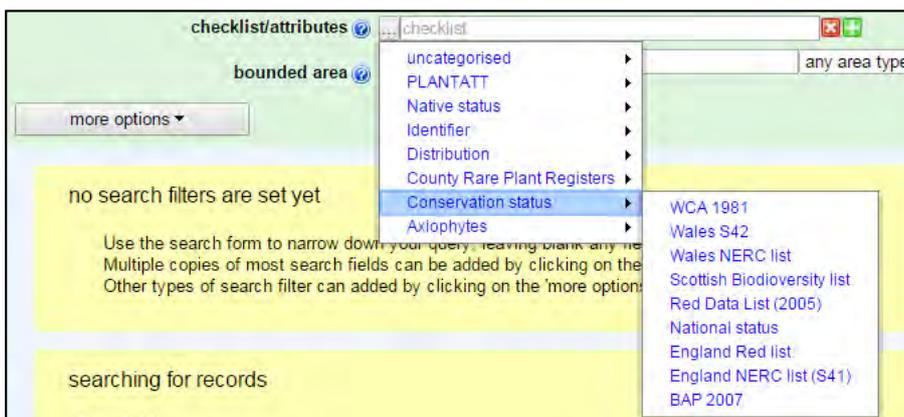
Date in dd/mm/yyyy, mm/yyyy or yyyy formats. The database copes with formats such as 1/2/13 meaning 01/02/2013. Date can be an exact date, as well as the start or end date, and a date range can be used. The drop down list also includes standard BSBI date classes.

Checklists / attributes

A wide range of Checklists are available, some (eg PLANTATT) with a long list of Attributes that can be selected. The list of available Checklists is viewed by clicking on the grey rectangle with three dots.



Two sets of available Checklists of particular interest to VCRs will be 'Conservation status' and the (at present under-used) 'County Rare Plant Registers'. Available Checklists under these two headings are shown below.

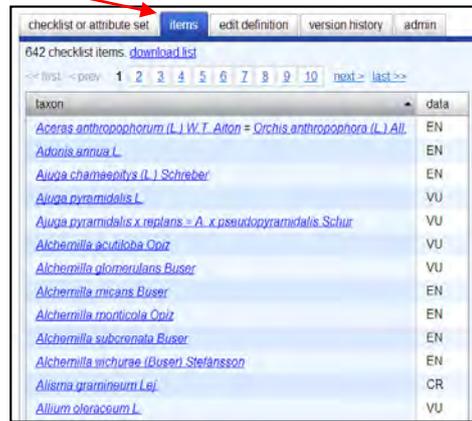


The National status Checklist includes GB Nationally Rare and GB Nationally Scarce.

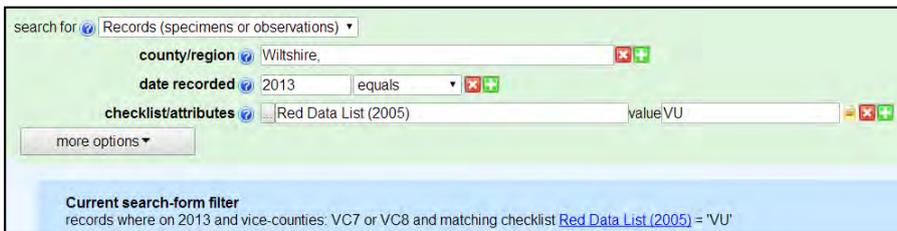
Once a Checklist has been selected, a 'show checklist details' tab appears.



Clicking on this opens up a new window. Clicking items displays a list of included taxa or attributes.



In the data column are listed the 'values' that can be included in a query. So to find records of all taxa that are listed as 'Vulnerable' in the Red Data List (2005) you would enter 'VU' in the 'value' field of the query. For example



If you wanted all taxa listed in the Red Data List (2005) you would leave the 'value' field blank.

To find all records of Nationally Rare or Nationally Scarce taxa, recorded in 2013 in Moray (vc95), the query would be:



As of 15/2/2015 this query returns 249 records. This query will return records of all listed taxa, irrespective of whether they are native or not in that vice-county. Some approaches to tackling this issue are described in the FAQ at [How do I create a list of all records for Nationally Rare and Nationally Scarce taxa in a vice-county in a year?](#)

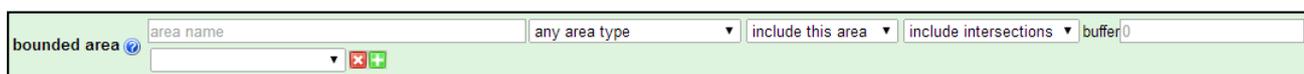
The PLANTATT Checklist refers to 'Attributes of British and Irish Plants: Status, Size, Life History, Geography and Habitats' by Mark Hill et al. Please refer to the online version for explanation of coding and meaning of terms - <http://nora.nerc.ac.uk/9535/1/PLANTATT.pdf>

At the time of writing Axiophyte Checklists are only available for two vice-counties, and Rare Plant Register taxon lists for three vice-counties (as well as for the Cairngorms National Park). These should always be combined with the relevant vice-county (or bounded area, below) within queries. It would not usually be appropriate to search for records of taxa listed in a CRPR that are in a different vice-county.

Bounded area

Bounded areas are polygons defining the boundary of an area. Currently the DDb has boundaries for all SSSIs, ASSIs, NNRs and for two National Parks (Cairngorms and Loch Lomond and the Trossachs). In addition, vice-county boundaries can be used as bounded areas, which is useful in some types of searches. It is intended that a more extensive range of bounded areas will be provided in the future. At the moment it is not possible for the DDb user to upload their own bounded areas, but this may be an option in the future.

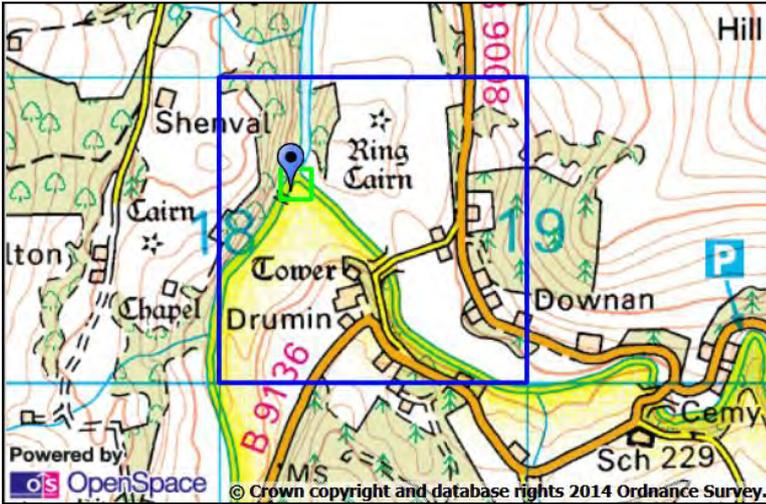
The DDb can carry out spatial queries using bounded areas, in a similar way to GIS. Options available are:



- Area type. Allows searches of all bounded areas within a category. Fully implemented at the time of writing are SSSIs, NNRs and Natural Improvement Areas.
- Include this area (default) or exclude this area.
- Include intersections (default) or contained within. Intersections refers to overlap of the grid square defined by the record's site grid reference and the boundary of the bounded area.
- Buffer. Add a buffer (in metres) around a bounded area (default is no buffer).
- Set grid square resolution at which to test for overlap. Default is maximum; options are 1km or 100m square resolution. As currently implemented, setting this option, eg. to 1km, does two separate things. It restricts the query to returning records at 1km precision or better, and returns records that are outside the boundary, if they fall within a 1km grid square that overlaps the boundary.

A couple of examples will help to clarify how these bounded area options work in queries. (Results were correct when the queries were originally run in May 2014).

Monad NJ1830 (in vc94) is bisected by the Cairngorms National Park boundary (map below).



Monad NJ1830 (blue square); 100m grid cell NJ182306 (green square); Cairngorms National Park Boundary (yellow). (Source <http://www.bnhs.co.uk/focuson/grabagridref/html/index.htm>)

A simple query for all records within the monad returns 295 records. Adding 'Cairngorms National Park' as a bounded area, and accepting the default of 'include intersections' returns 283 records.

 A screenshot of a search form. The search criteria are: 'Records (specimens or observations)', grid reference 'nj1830', and bounded area 'Cairngorms National Park'. The 'include intersections' dropdown is selected. Below the search criteria, the current filter is shown as 'records where grid-reference NJ1830 and area intersecting Cairngorms National Park'. Buttons for 'display results', 'email results', 'download results', 'save query', and 'reset search form' are visible.

Hence there are 12 records within this monad that fall entirely outside the National Park boundary. To view these records, change 'include this area' to 'exclude this area'.

 A screenshot of the same search form, but the 'bounded area' dropdown is now set to 'exclude this area'. The current filter is updated to 'records where grid-reference NJ1830 and excluding areas intersecting Cairngorms National Park'. The other search criteria and buttons remain the same.

To view those records whose site grid reference falls entirely within the National Park, change 'include intersections' to 'contained within'. This returns just 10 records.

search for

bounded area

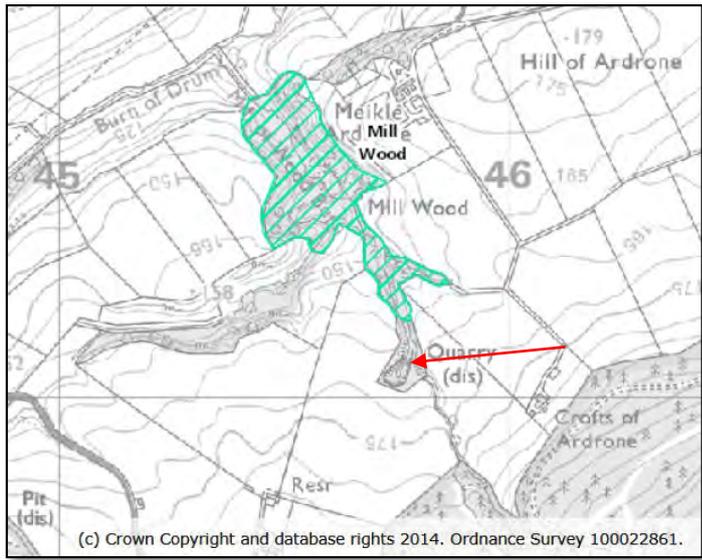
more options ▾

Current search-form filter
 records where grid-reference [NJ1830](#) and area within [Cairngorms National Park](#)

[reset search form](#)

[help...](#)

A set of queries for records from Mill Wood, a SSSI in vc94, help illustrate how buffers work. The wood lies entirely within monad NJ4550. A search for records from that monad returns 714 records, and of these all but 3 records have 'Mill Wood' in the site name. However the wood is more extensive than the SSSI itself. Searching on Mill Wood as a bounded area and choosing 'contained within' returns just 51 records.



Mill Wood SSSI (green hatch); *Equisetum hyemale* (Rough Horsetail) site (red arrow). Source: <http://www.magic.gov.uk/MagicMap.aspx>

A recent notable re-find here was that of *Equisetum hyemale* in 2012, the first record for at least 100 years. Looking at the results of the query showed that this record was missing. Knowing that the person who made the record would have taken a detailed GPS grid reference, suggested that the record was actually outside the SSSI. Retaining the 'contained within' option in the query, but adding a buffer of 50 metres returned 494 records, including the 2012 record of *E. hyemale*.

search for

bounded area

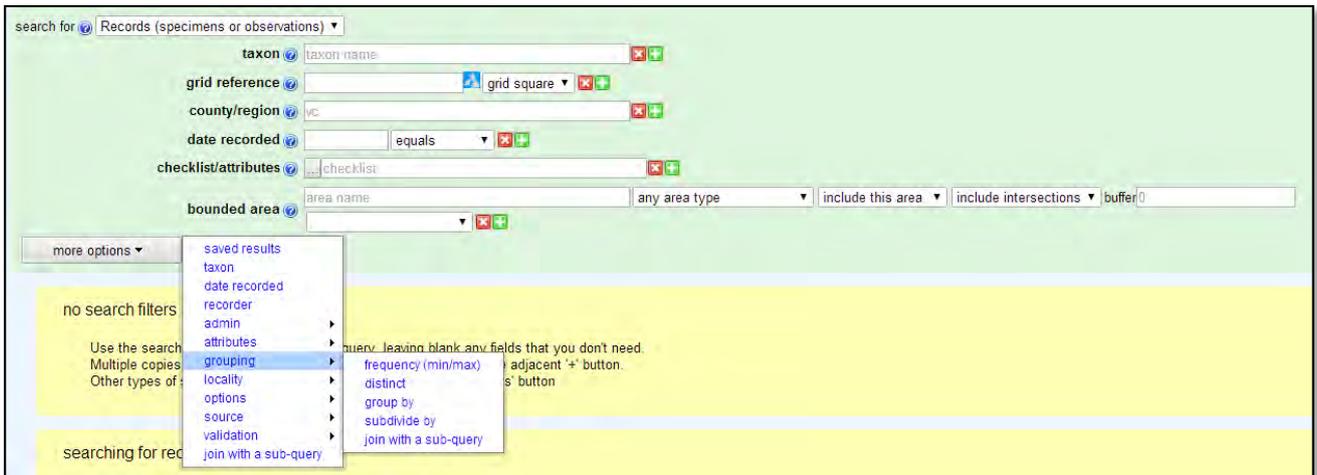
more options ▾

Current search-form filter
 records where area within [Mill Wood](#) + 50m buffer

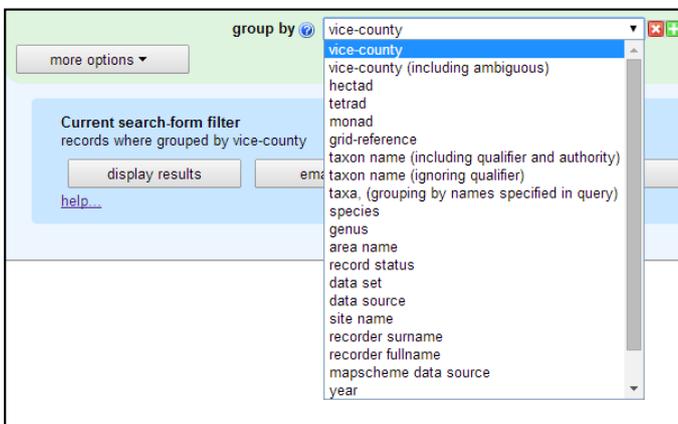
The *Equisetum* has a 10 figure site grid reference, and varying the buffer in small steps shows that the plant is actually 42 metres outside the SSSI boundary!

More options

Clicking on the 'more options' drop down arrow reveals a number of additional options, some with further drop down arrows.



When one of these is selected, a new row is added to the query screen, often with a drop down arrow, which opens up more options, eg. for 'group by'.



It is not the purpose of these guidelines to explain and illustrate in detail, all functions of the DDb. If nothing else, the guidelines would probably become too lengthy and cumbersome to use. However it is helpful to draw attention to some of the available options, and the FAQ section illustrates many examples.

Group by, distinct and subdivide by

Queries typically return actual records, with one row in the web-browser, or in a downloaded file being a single record. While that may often be what is required, just as often some sort of summary of the records is required. Rather than having to download records and manipulate them in a spreadsheet, the summarising can much more easily, and speedily be carried out within the DDb. This is done by using options found via 'more options' > 'grouping'.

'Group by' when used alone, counts the number of records within a category, eg per vice-county, hectad, or tetrad.

When an additional 'distinct' clause is added to a 'group by' query, the frequency count returned refers to the number of distinct variants there are of the field specified by the 'distinct' clause.

'Subdivide by' adds columns to the output table, subdividing the records, eg. into date class or year. An example will help clarify this.

County Waterford (vcH12) had 242,390 records on the DDb (as of 8th May 2014). This is a huge dataset, and in its raw form, rather unwieldy to work with. But it can be summarised easily, in a variety of ways to meet the needs of the DDb user.

In the screenshot below, adding 'group by' hectad, lists hectads with records in the left hand column. Adding 'distinct' species, counts the number of species and hybrids (sub-species are lumped into their parent species) in each hectad, and that figure is in the right hand column. Adding 'subdivide by' date class, adds 6 columns, one for each BSBI date class, and the values are the number of species & hybrids in each hectad in each date class. The first row, with no hectad listed, refers to records for the whole vice-county or other locations, with no grid reference given in the original record.

The screenshot shows a search interface with the following filters:

- search for: Records (specimens or observations)
- county/region: Co. Wexford
- group by: hectad
- distinct: species
- subdivide by: date class

Current search-form filter: records where grouped by hectad and counting distinct species and subdivided by date class and vice-county: VCH12

Buttons: display results, email results, download results, save query, reset search form

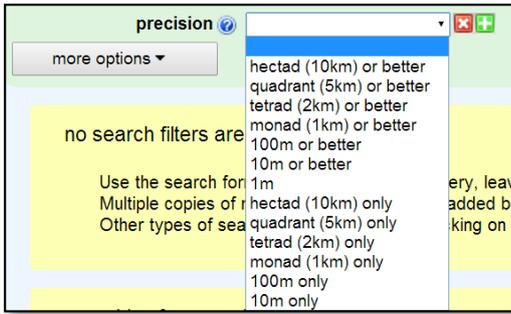
Results table: 44 groups matched your query.

hectad	pre 1930	1930-1969	1970-1986	1987-1999	2000-2009	2010-	total freq
	127	35	8	924	5	0	961
L65	1	0	0	0	0	0	1
S61	53	252	6	363	334	409	567
S62	8	85	0	317	198	324	454
S70	61	347	170	399	343	510	666
S71	51	247	212	433	446	597	737
S72	42	333	101	443	528	538	703
S73	26	283	1	437	419	413	611
S74	3	5	0	156	0	22	166
S80	46	214	137	431	337	395	562
S81	273	295	169	446	374	520	655

'Distinct' and 'subdivide by' can only be used in combination with 'group by' in a query.

Locality (precision)

Under 'more options > locality > precision', the drop down menu lists many options for minimum or exact grid reference precision. This allows you to narrow down the precision of records you want to view or summarise.

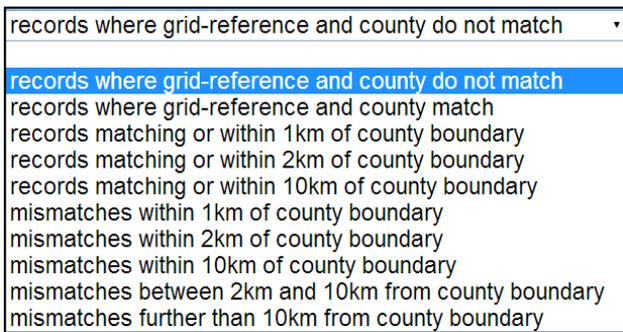


Data source and dataset

Under 'more options > source', options include MapMate centre (ie CUK), dataset and data source. A dataset or data source can be selected, or you can choose to exclude it from a query. Dataset and data source can also be used as groupings in 'grouped queries', to answer questions such as what data sets are there for a vice-county, and how many records are there in each?

Grid reference and vice-county validation

Under 'more options > validation' there is the 'grid-ref and vc validation' option. This checks that the vice-county listed for each record matches the actual vice-county boundary. There are various options to choose from:



In many queries, it is good practice to include a clause that restricts records to those where grid-reference and vice-county match, or perhaps 'matching or within 1km of county boundary', to exclude records that may have a wildly incorrect grid reference, and which have not yet been corrected.

Join with a sub-query

Sub-queries add considerable sophistication to searches. Some examples are given in the FAQ section, below. Consistently being able to create sub-queries which do what the user wants requires practice. Please ask for assistance if you get stuck.

A note on searches for large numbers of records

While the DDb allows access to all plant records that the BSBI holds, there are some limits on the number of results that can be viewed or downloaded. Up to 50,000 can be viewed in the on-line results table. Users without extra admin permissions can download or have e-mailed to them, up to 150,000 records in one go. If you need to exceed these limits then please ask for assistance by email to tom.humphrey@bsbi.org. The only likely scenario for this would be a VCR wanting to work on a whole county data set, eg for a Flora, Checklist or Rare Plant Register. The resultant very large files can be supplied via a file sharing website, eg. Dropbox or similar.

Tom Humphrey has provided the following guidelines to help ensure that large searches work reliably (rather than failing due to web-browser time-outs or lack of memory) and that they don't negatively affect the running of the system.

- Use the 'Email search results' option, rather than downloading the results directly. This avoids problems with web browser time-outs and lets the server run your search task more efficiently. You can email results straight away without having to run the query first in the web browser (where the results will have been capped at 50,000 anyway).
- If possible download large result sets as a csv (text) file rather than in Excel or Ods format. Delimited-text can still be easily loaded into a spreadsheet, but is a far simpler and quicker format to generate on the server.
- Please avoid rerunning the same query multiple times. If you are unsure if your query has worked then, before restarting it, check in your search history, to see if your search is still running or if there is a previous result set waiting that you can retrieve quickly. If a slow query fails more than once then please seek assistance rather than retrying it - having failed once it will probably fail every time.
- Consider whether your query can sensibly be broken up into smaller parts - but only within reason: splitting a slow query into a few parts (e.g. partitioning by county or date-class) is probably quick and worthwhile, whereas having to run it as many separate searches is a waste of time - in that case please seek advice about how best to run the search in one go).
- If you don't need individual records, only a summary then it's usually quicker to search for the grouped summary directly rather than downloading a large record set and analysing it off-line in a spreadsheet.
- Consider if you can weed out data that you don't need before searching, e.g. you may need only recent records (rather than those dating back to 1700) or you may be able to exclude a large chunk of data that you already know about (e.g. records already in MapMate, or from a recent flora data set that you already have access to).

The last bullet point, above, is from a practical viewpoint perhaps the most important. In many circumstances, it is much more useful to just view or download a subset of records, or a summary of them. It is also quicker. Searching for records of *Adoxa moschatellina* returns (as of 16/2/2015) 29,961 records, the query taking 40 to 55 seconds. That is a bit of a wait, and do you really want to wade through that many records to find the ones of interest to you? Probably not. Restricting the search to a single vice-county (vc94) returns 190 records, and the query runs in about 2 seconds. Restricting the query to vc94 and to records since 2000, returns 104 records, and the query runs in less than one second; so both much faster and easier to find the records of particular interest to you. Working with a broadband connection that (on a good day) manages a download speed of 7Mbps, allows many such restricted DDb queries to return results as fast or faster than those run directly from MapMate on my computer's hard drive. However, some queries are slower to run, eg those using bounded areas, and particularly complex queries can take a few minutes to run.

NB You do not have to wait for queries to run; results can be retrieved via 'search history'. Therefore a sequence of individual queries can be created and run, retrieving results later.

FAQs – how to find what you want, and other useful stuff, using the DDb

This section of the guidance document illustrates how to use the DDb using actual examples. Links to all the queries are provided below the screenshots.

An especially helpful feature is being able to email a link to a DDb query, or include it in a document. This is done by clicking on 'download', right clicking on 'this link' (highlighted in yellow below), and selecting 'copy link address'. If the link doesn't work, copy and paste into your browser window, and press return.



If downloading large datasets, opt for the e-mail option (which creates a csv file) or download as a csv file (comma separated values). If the query returns more than 50,000 records, you must opt for the e-mail option. This format is compatible with any spreadsheet, and is created more rapidly by the DDb.

While the DDB provides many powerful tools to select, summarise and analyse plant records, for many purposes the subset of records returned by a query will require further manipulation in a spreadsheet. Some examples are given where additional analysis using (Excel) has or can be undertaken.

LISTS

How do I create a list of all species / taxa recorded in a vice-county?

Creating a list of taxa recorded in a vice-county is a simple matter of entering the vice-county of interest, and adding a 'group by' clause via 'more options > grouping'. You can choose from several taxon options.



Selecting 'taxon name (including qualifier and authority)' provides the greatest detail; 'taxon name (ignoring qualifier)' lumps some taxa together eg. *Agrostis canina* s.s and s.l., but retains subspecies; 'species' lumps subspecies into the parent species total, but retains hybrids as separate entries.

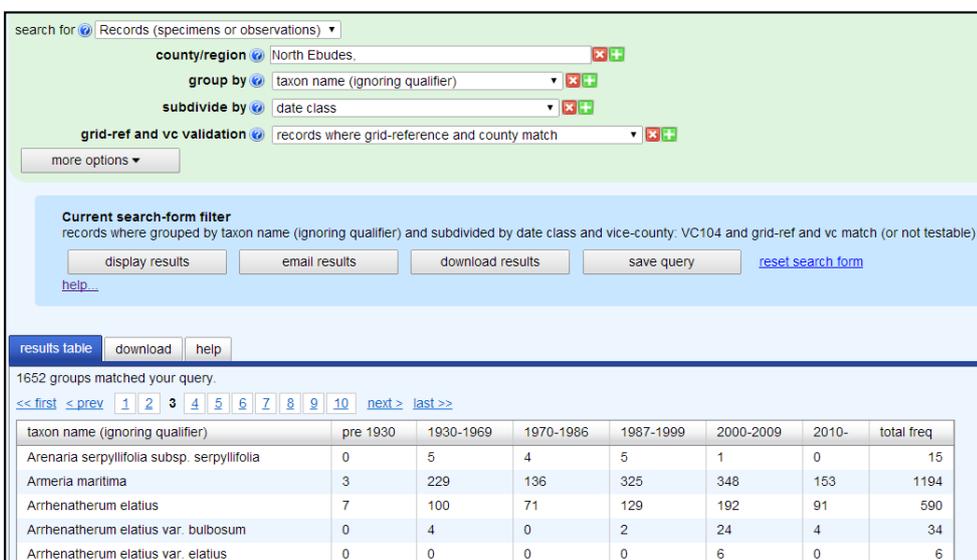


<http://bsbidb.org.uk/search.php#query=2a73807a1bfb380939226c4f8d14c341>

So for the West Mayo example, the three versions of this query return 1127, 1100 and 1005 taxa. Each may be the appropriate one to use in a given circumstance.

How do I create a species list for a vice-county subdivided by BSBI date class?

The query above can be modified by adding a 'subdivide' clause via 'more options > grouping'.



taxon name (ignoring qualifier)	pre 1930	1930-1969	1970-1986	1987-1999	2000-2009	2010-	total freq
<i>Arenaria serpyllifolia</i> subsp. <i>serpyllifolia</i>	0	5	4	5	1	0	15
<i>Armeria maritima</i>	3	229	136	325	348	153	1194
<i>Arrhenatherum elatius</i>	7	100	71	129	192	91	590
<i>Arrhenatherum elatius</i> var. <i>bulbosum</i>	0	4	0	2	24	4	34
<i>Arrhenatherum elatius</i> var. <i>elatius</i>	0	0	0	0	6	0	6

<http://bsbidb.org.uk/search.php#query=738e5ecbb1f35f45ae89109d3423ddb0>

'Subdivide' adds additional columns to the output of the query, in this case for vc104 (Mid Ebudes), giving the number of records for each taxon in each date class period.

This query can be modified to give a count of hectads for each species in each date class, by a 'distinct' clause to the query, and selecting hectad.

The screenshot shows a search interface with the following parameters:

- search for: Records (specimens or observations)
- county/region: North Ebudes
- group by: taxon name (ignoring qualifier)
- distinct: **hectad** (indicated by a red arrow)
- subdivide by: date class
- grid-ref and vc validation: records where grid-reference and county match

Current search-form filter: records where grouped by taxon name (ignoring qualifier) and counting distinct hectad and subdivided by date class and vice-county: VC104 and grid-ref and vc match (or not testable)

Buttons: display results, email results, download results, save query, reset search form, help...

Results table navigation: results table, download, help

1650 groups matched your query.

Navigation: << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >>

taxon name (ignoring qualifier)	pre 1930	1930-1969	1970-1986	1987-1999	2000-2009	2010-	total freq
Carex paniculata	1	4	4	5	4	2	11
Carex pauciflora	9	18	12	11	9	4	30
Carex pendula	0	0	1	1	0	2	2
Carex pilulifera	5	34	18	42	34	24	47
Carex pulicaris	4	39	18	44	45	30	48
Carex remota	1	9	10	11	12	9	22
Carex rostrata	7	35	27	37	37	26	44
Carex rupestris	0	0	1	1	1	1	1
Carex sylvatica	4	13	16	14	7	4	23

<http://bsbidb.org.uk/search.php#query=efa464763c091439df655b9bd19e7dac>

Because old historic records are often only available at hectad resolution (at best) running this query at the hectad precision level is sensible as it allows comparison over many decades.

How do I create a species list for a vice-county comparing post 2000 to pre 2000?

Since 2000, BSBI date classes are decadal, while Atlas 2020 will use the 20 year period 2000-2019 as its current date class. The DDb provides an option within the 'subdivide by' drop down list to split records before 2000 and since 2000. This is found via 'more options > grouping'.

Using North Ebeudes again, and counting number of hectads, using a 'distinct' clause, the modified query is:

search for

county/region

group by

distinct

subdivide by

grid-ref and vc validation

more options ▾

Current search-form filter
records where grouped by taxon name (ignoring qualifier) and counting distinct hectad and subdivided by date class (pre/post 2000) and vice-county: VC104 and grid-ref and vc match (or not testable)

display results email results download results save query [reset search form](#)
[help...](#)

results table download help

1650 groups matched your query.

<< first < prev 1 2 3 4 5 6 7 8 9 10 next > last >>

taxon name (ignoring qualifier)	pre 2000	2000-	total freq
Carex paniculata	9	6	11
Carex pauciflora	25	11	30
Carex pendula	1	2	2
Carex pilulifera	44	38	47
Carex pulicaris	46	48	48
Carex remota	15	17	22
Carex rostrata	43	39	44
Carex rupestris	1	1	1
Carex sylvatica	23	9	23

<http://bsbidb.org.uk/search.php#query=d07af12be70d662e70e072665bdb1db5>

So for example, *Carex sylvatica* has been recorded in 23 hectads in vc104; all hectads have pre 2000 records, but since 2000, this species has only been recorded in 9 hectads (so far).

How do I create a list of species with earliest and most recent year recorded?

search for

county/region

group by

subdivide by

grid-ref and vc validation

more options ▾

<http://bsbidb.org.uk/search.php#query=01994f8ce41d072879373e3239575bc5>

In this query 'subdivide by date range' adds columns for the earliest and latest year with records. NB this query uses the most recent year of any records with date range as the most recent year recorded. This may be incorrect. For the most recent year only, use 'subdivide by latest date' (see query below. This uses the start year of any date ranges as the most recent year, and is the more conservative measure.

How do I create a species list for a hectad listing most recent year each taxon has been recorded?
 There is a 'latest date' option available as a 'subdivide' clause, found by clicking on 'more options > grouping'.

The screenshot shows a search interface with the following settings:

- search for: Records (specimens or observations)
- grid reference: SW84 (grid square)
- group by: taxon name (ignoring qualifier)
- subdivide by: latest date
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=33e086de300dbc7ebed4b7d50c8bd21c>

Hectad SW84 in Cornwall has a huge number of records on the DDb (more than some vice-counties!). The DDb's 'group by' and 'subdivide by' functions, make light work of extracting the data summary. This query returns latest year for 1479 taxa. Changing 'latest date' to 'date range (min/max date)' only returns 1466 taxa (query below).

The screenshot shows a search interface with the following settings:

- search for: Records (specimens or observations)
- grid reference: SW84 (grid square)
- group by: taxon name (ignoring qualifier)
- subdivide by: date range (min/max date)
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=9bf75b17b6b9d0a76c8ba3f57ec6aabdb>

This difference is because some dates have vague date ranges, eg. 'before 1987', which has an end year but not a start (minimum) year, hence is excluded from the query.

How do I create a list of all taxa in a vice-county with the number of tetrads / monads with records since a specified year?

To count the number of tetrads for each taxon, you add a 'distinct' clause to the grouped query. In the example below, the number of tetrads for each taxon within Kent (vice-counties 15 and 16) since 1987 is listed. Monad can be substituted instead of tetrad within the 'distinct' clause.

The screenshot shows a search interface with the following settings:

- search for: Records (specimens or observations)
- county/region: Kent
- date recorded: from 1987 (for vague record date ranges use start date)
- group by: taxon name (ignoring qualifier)
- distinct: tetrad
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=c51845694beadd8f15c22b6684daf72a>

How do I create a list of all infraspecific taxa in a vice-county?

This query has two parts, signified by the 'join with' part of the screenshot below. The first part of the query creates a list of taxa recorded in the vice-county, Cardiganshire in this instance. The records are sub-divided by BSBI date class. The second part of the query restricts the output of the

query to taxa at subspecies rank or below. It therefore returns a list of all infraspecific taxa; subspecies, varieties, forms and cultivars.

The screenshot shows the search interface for the BSBI database. The search criteria are as follows:

- search for: Records (specimens or observations)
- county/region: Cardiganshire
- group by: taxon name (ignoring qualifier)
- subdivide by: date class
- grid-ref and vc validation: records where grid-reference and county match
- join with: Taxa
- join using: exact taxon match
- taxon rank: subspecies
- include ranks below: (checked)
- include subdivisions: (checked)

<http://bsbidb.org.uk/search.php#query=f473772e751d865229cf43c0d1cae87c>

How do I create a list of all subspecies in a vice-county?

This query is identical to that used for infraspecific taxa, but the second part of the query restricts output to subspecies rank only, and does not include ranks below.

The screenshot shows the search interface for the BSBI database. The search criteria are as follows:

- search for: Records (specimens or observations)
- county/region: Cardiganshire
- group by: taxon name (ignoring qualifier)
- subdivide by: date class
- grid-ref and vc validation: records where grid-reference and county match
- join with: Taxa
- join using: exact taxon match
- taxon rank: subspecies
- include ranks below: (unchecked)
- include subdivisions: (checked)

A red arrow points to the 'include ranks below' dropdown menu, which is currently set to 'this rank'.

<http://bsbidb.org.uk/search.php#query=02c532ed5561bf77efd3fa20829815ac>

How do I create a list of all records for Nationally Rare and Nationally Scarce taxa in a vice-county in a year?

The three queries below, using Moray (vc95) in 2013 as an example, illustrate different approaches that may be taken.

The first query lists all records for taxa listed as Nationally Rare or Nationally Scarce in vc95 in 2013. It does so irrespective of whether the taxa are native in vc95, or in that hectad. As of 25/5/2014 this query returns 222 records.



<http://bsbidb.org.uk/search.php#query=9b214497e5f354a1bb24e837340e2d64>

The query can be restricted to only returning records where the taxa are listed as native in each hectad in Atlas 2000. The modification to the query is included via 'more options' > 'attributes' > 'status (external scheme)'.



<http://bsbidb.org.uk/search.php#query=0a95ca8699d0c83cc6b107456e7ce546>

This query returns (as of 25/5/2014) 125 records. But, this query excludes any records for hectads where the taxon had not been recorded in Atlas 2000. To list these records you need to modify this query, changing status field to 'no status information'.

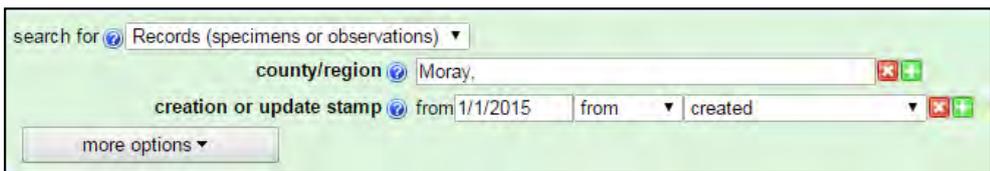


<http://bsbidb.org.uk/search.php#query=971af102ba187171a676333bebb4a201>

This last query returns 19 records. Some of these will be new hectad records of native species, while others will be new hectad records of alien species (in this vice-county).

How do I list all records for a vice-county added to the DDb since a date?

It is useful to be aware of what records are being added to the DDb, not all of which may have come via the VCR's copy of MapMate. This query finds all records 'created' (ie newly added to the DDb) since 1st January 2015 in Moray (vc95).



<http://bsbidb.org.uk/search.php#query=cab0377f841f86fa64a94b1c6a19ac98>

Run on 1/3/2015 it returned 3095 records (of which only 123 were records for 2015 itself), the remainder being for previous years.

MAPS

How do I create a map showing tetrad survey coverage in a vice county for the period since 2000?

BSBI guidance for the Atlas 2020 project is that, wherever possible, a minimum of 5 tetrads are surveyed (reasonably thoroughly) in each hectad in the period 2000 – 2019. A map showing species recorded per tetrad in a vice-county can be created using the DDb.

The class intervals displayed on the map are automatically set within the DDb mapping software, but some flexibility is provided by adding a 'frequency' clause to the query; available via 'more options > grouping'.

This query returns a list of tetrads with at least 50 species recorded since 2000 in vc96.

search for Records (specimens or observations)

county/region

date recorded from for vague record date ranges

group by

distinct

frequency (min/max) at least

<http://bsbidb.org.uk/search.php#query=32e941350abe167b40b603a23e1bea63>

It may be useful to view recording coverage in adjacent vice-counties, and this can be particularly helpful where hectads overlap the vice-county boundary. Add additional vice-counties to the query by clicking on the green plus symbol to the right of 'county/region'.

search for Records (specimens or observations)

county/region

county/region

date recorded from for vague record date ranges

group by

distinct

frequency (min/max) at least

<http://bsbidb.org.uk/search.php#query=b4386f30c8a49302ba837385b7fecfa5>

Results from this query can be mapped by clicking on the zoomable or printable map tabs. using the printable map, the mp area is restricted to the vcs of interest by clicking on the 'areas and boundaries' tab, and entering the individual vcs, in this case vc96 and vc106.

dataset partitioning | marker styles | line styles | **areas and boundaries** | page layout | save or print

region to show

Channel Islands

Northern Isles

marked or constrained boundaries

type clip

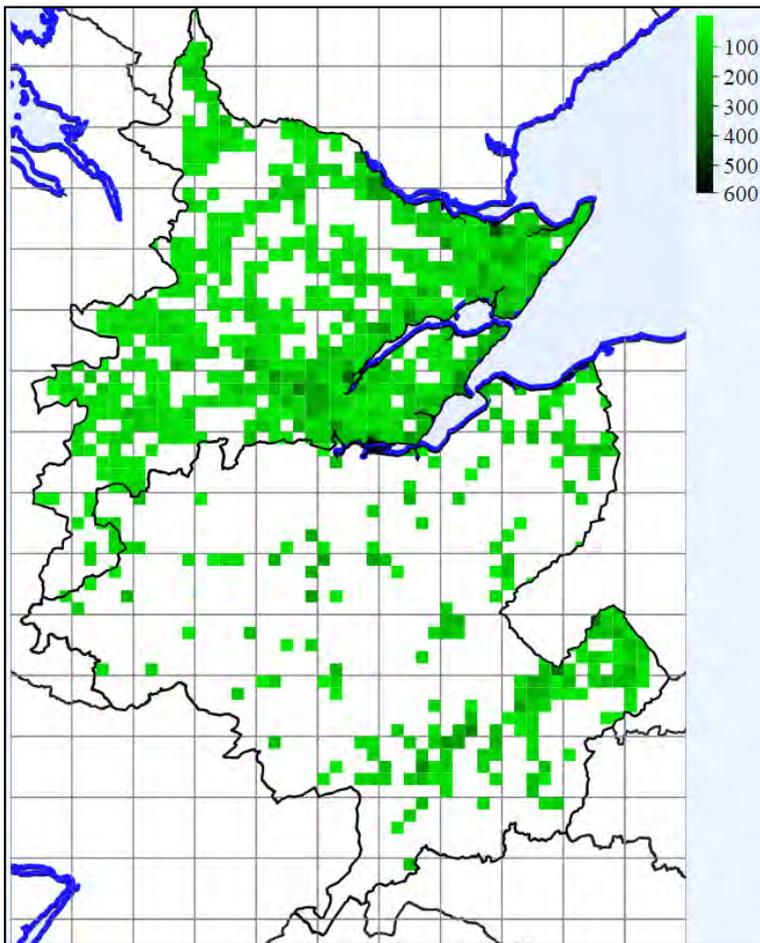
background colour outline style

type clip

background colour outline style

show county boundaries as background

The OS grid is added automatically; which grid depends on the area of the map.



The map above shows tetrads with >50 species in vcc 96 & 106, since 2000

The DDb mapping doesn't label grid squares, but labels can be added in Word or in image editing software. For the purposes of survey planning, the maps from the DDb are quite adequate.

If you want to see all tetrads with records, change the 'frequency' vale to 0 (zero).

How do I create a map of survey effort across a vice-county?

The query below counts the number of days (minimum 5 days), with records per tetrad in Cheshire (vc58) since 2000.

search for

county/region

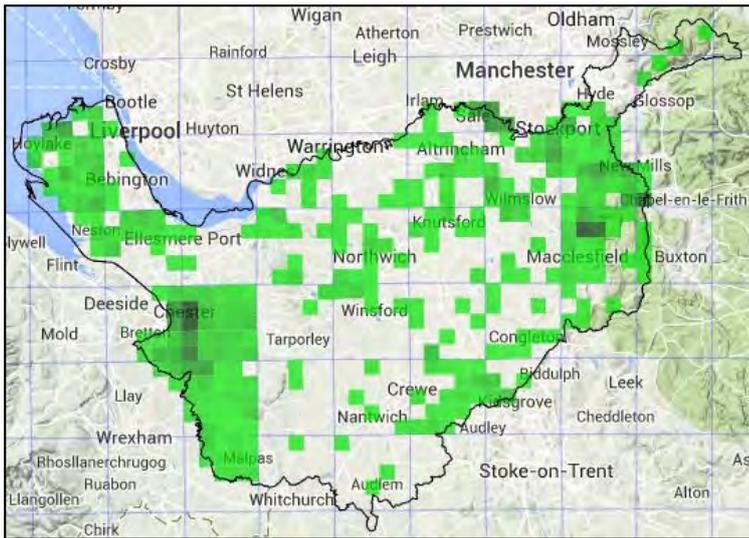
date recorded from for vague record date ranges use start date

group by

distinct

frequency (min/max) at least

<http://bsbidb.org.uk/search.php#query=1179ec5d4e33c8ca374c0900c028c2a3>



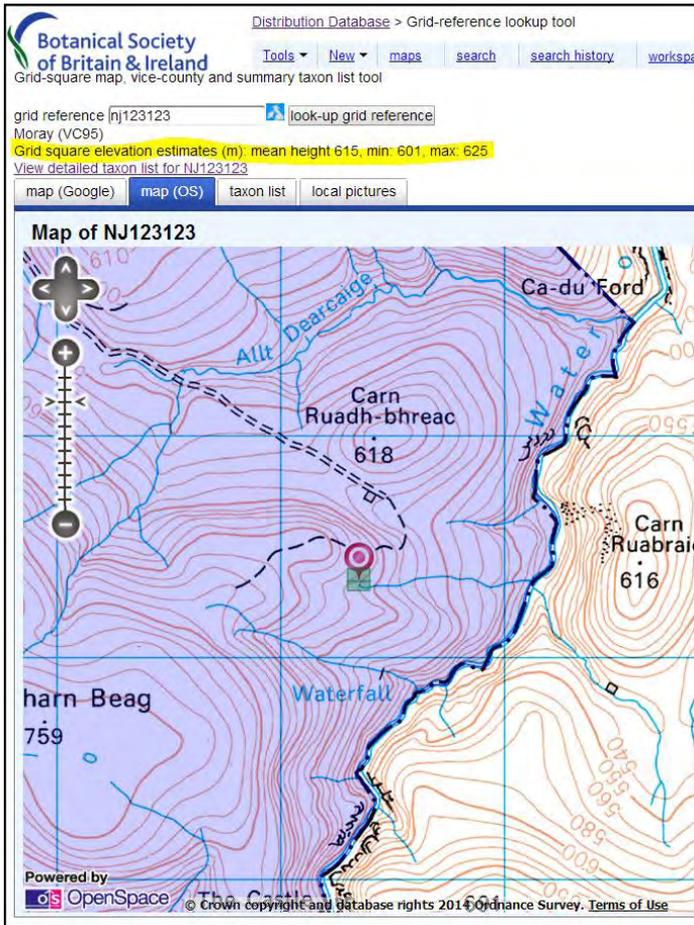
Removing the 'frequency' part of the query would return all tetrads recorded since 2000, with number of days with records.

Altitudes

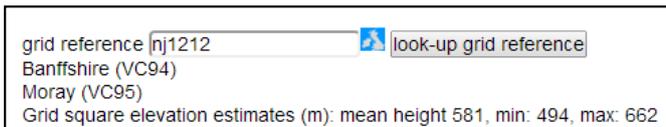
How do I find the altitude of plant locations?

You do not need to be logged in to the DDb to use the 'Grid-reference lookup tool'. This is found via the 'Tools' tab on the DDb home page, or at <http://bsbidb.org.uk/gridref.php>.

Entering the 100m grid reference NJ123123 and clicking on 'look-up grid reference' opens a new frame in the browser window (below) showing the grid reference plotted as a green square, over a map backdrop. In GB this is OS mapping, but in Ireland is Google maps. Google maps (with aerial photography), is also available as an option for GB. Vice-county boundaries are plotted, and the OS version shows them in different (transparent) colours. Mean, minimum and maximum altitude within the grid square defined by the grid reference are shown (highlighted in yellow below).



Changing the grid reference to that of the monad, NJ1212, calculates the mean, minimum and maximum altitude for the whole monad.

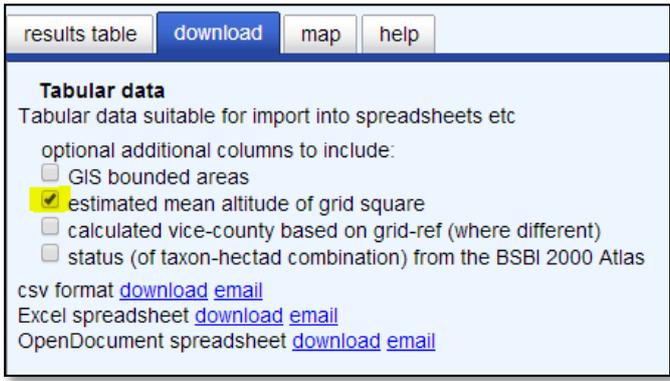


Note that in this instance the grid reference lookup tool has detected that a very small corner of the monad is within vc94, whereas the 100m grid reference was entirely within vc95.

Altitudes are estimated by the DDb using a very high resolution point altitude GIS dataset.

When logged in to the DDb, when viewing any records in the web browser window, clicking on a site grid reference, opens a new tab with a similar window to that described above.

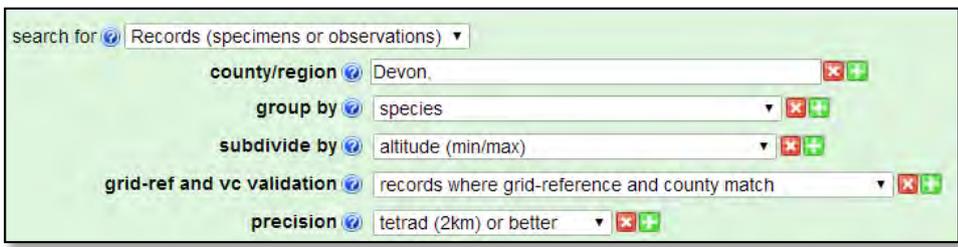
Estimated mean altitude can be included as a column in a download. Tick the box in the download options screen (highlighted in yellow below).



Remember, the mean altitude is calculated for the whole square as defined by each record's site grid reference. In areas with marked topography, the mean altitude for coarse precision grid references, eg a hectad will not be meaningful. In such situations it will be better to restrict the query, or your spreadsheet analysis, to higher precision records.

How do I find the maximum and minimum altitude for all taxa in a vice-county or group of vice-counties?

The DDb has an option to 'subdivide by altitude (min/max)', available via the 'More options > grouping' tabs. Because it is 'subdivide by', there has to be a 'group by' criterion within the query. In this case it is 'species', although taxon is also available as an option.



<http://bsbidb.org.uk/search.php#query=af311f5afe4257fab3802b6c947450a7>

The query (for Devon, vc3 and vc4) is restricted to records with precision at tetrad or better, to exclude coarse precision records, for which altitude ranges may be spurious.

Important note. Currently (May 2014) the way this query works, and that for 'subdivide by altitude (grouping)' discussed below, is to (at best) only use monad resolution altitudes. This approach was implemented for reasons of database efficiency. The precise altitude values of every site grid reference are not currently stored in the database - but are calculated by interpolating from values in a look-up table every time a precise altitude is needed, whereas altitudes for all monads are pre-cached separately so are available to queries quickly. This approach may be modified in the future.

Therefore minimum and maximum attitudes returned by this query are likely to return values that are outside the actual altitudinal range for some species. The more variable the topography, the more this needs to be kept in mind.

If it is critical to investigate the actual altitudinal range of a species, download the records for that species into a spreadsheet, ticking the appropriate box in the download options screen (see above).

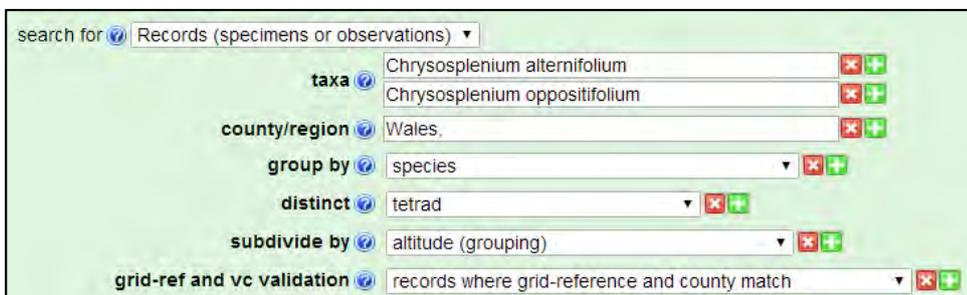
This does calculate the actual mean altitude for that grid square, including those with precision greater than monad.

How do I summarise species frequency in altitude class intervals?

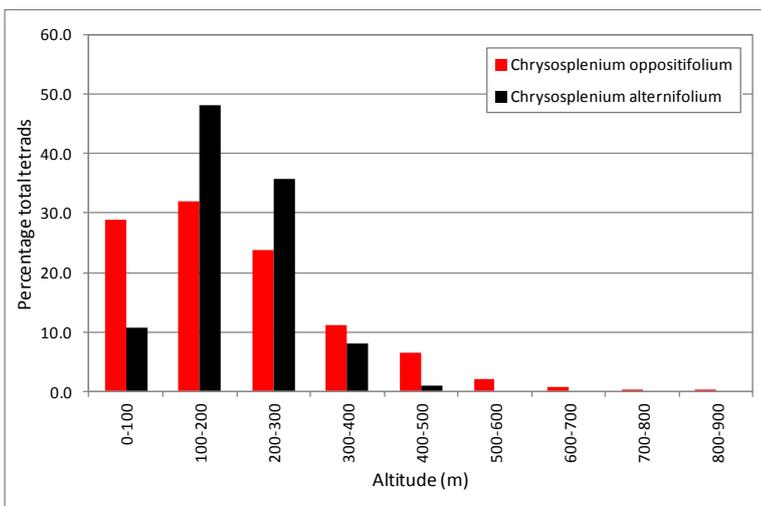
The DDb has an option to ‘subdivide by altitude (grouping)’ available via the ‘More options > grouping’ tabs. Because it is ‘subdivide by’, there has to be a ‘group by’ criterion within the query. In this case it is ‘species’, although taxon is also available as an option. Altitudes are grouped in 100m altitude class intervals from 0-100m to 1300-1400m.

Two examples are given, each with pairs of species in the same genus. In both cases a ‘distinct’ clause is included in the query, specifying tetrad. This means that the query counts the number of tetrads in each altitudinal band. If you remove this part of the query, (click red X at end of the line), and re-run the query, you will get the number of individual records in each altitudinal band. Specifying tetrad in the ‘distinct’ clause also has the effect of excluding any records with site grid reference at a coarser precision than tetrad.

In each example, the data has been downloaded to a spreadsheet, and the percentage of the total tetrads for each species, in each altitudinal band shown as a graph.



<http://bsbidb.org.uk/search.php#query=5d6b5ae07cc637437c0f1a143deec4e5>



Chrysosplenium species in Wales: percentage of total tetrads occupied by each species x 100m altitude class intervals

In Wales, as elsewhere in GB, *Chrysosplenium oppositifolium* has a wider altitudinal range than *C. alternifolium*, and is relatively more frequent at both low and higher altitudes. *C. alternifolium* is particularly associated with intermediate altitudes (100 – 300m AOD).

The *Alchemilla* query below, uses a ‘bounded area’ rather than a named vice-county or country. The default option of ‘include intersections’ has been kept, meaning that any records whose grid references are within or overlap the National Park Boundary are included in the analysis.

search for Records (specimens or observations) ▾

taxa Alchemilla glabra
 Alchemilla alpina

bounded area Cairngorms National Park include this area ▾ include intersections ▾

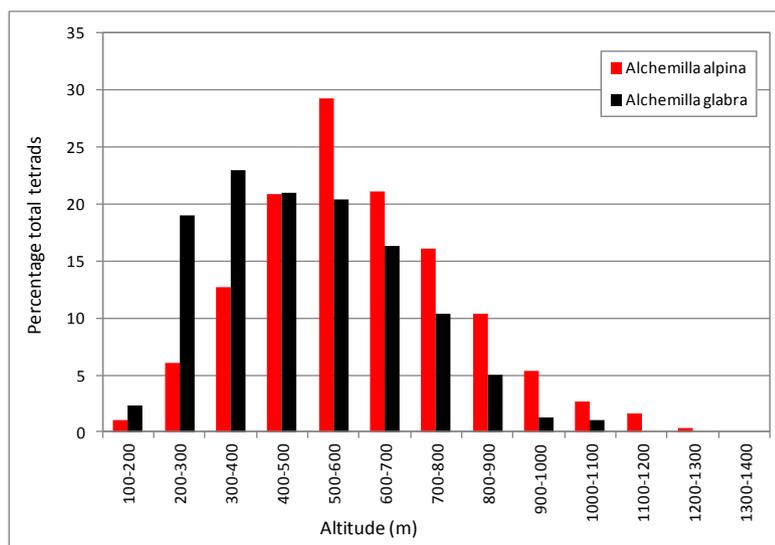
group by species ▾

distinct tetrad

subdivide by altitude (grouping) ▾

grid-ref and vc validation records where grid-reference and county match ▾

<http://bsbidb.org.uk/search.php#query=f729f2c3ed9422b504edf2c3276f4066>



Alchemilla alpina and *A. glabra* in the Cairngorms National Park: percentage of total tetrads occupied by each species x 100m altitude class intervals

Both *Alchemilla* species occupy wide altitudinal ranges; *A. glabra* at slightly lower altitudes (modal altitude 300 – 400m) cf *A. alpina* (modal altitude 500 – 600m).

New vice-county / hectad / tetrad / monad records

How do I create a list of all new vice-county records in a year?

To create this list requires a second, sub-query, to be used. This is added by clicking on ‘more options > join with a sub-query’. The example below, finds all records of any taxa new to County Wexford in 2013. The first part of the query lists all taxa recorded in 2013 (‘group by’). The second part of the query excludes any taxa that had been recorded up to, and including 2012. What you are left with is just the new vc records. The query can be adapted for different years by changing the year values in both parts of the query.

search for

county/region

date recorded equals

group by

join with

join using exclude

county/region

date recorded to for vague record date ranges

<http://bsbidb.org.uk/search.php#query=5c0da1da413d281fb1d6e86329a153d4>

How do I create a list of all new hectad records in a vice-county in a year?

To create this list also requires a second, sub-query, to be used. This is added by clicking on 'more options > join with a sub-query'. The example given below, finds all new hectad records in County Wexford in 2013. The first part of the query lists all taxa recorded in 2013 ('group by'), with a list of all the hectads with records ('subdivide by'), and counts up the number of hectads (distinct). The second part of the query excludes any combination of taxon and hectad that had been recorded up to, and including 2012. What you are left with is just the new hectad records. The query can be adapted for different years by changing the year values in both parts of the query.

search for

county/region

date recorded equals

group by

subdivide by

distinct

join with

join using exclude

county/region

date recorded to for vague record date ranges

<http://bsbidb.org.uk/search.php#query=3e31de102dfb46b523d3f26a3cbd3c31>

In the example, there were 317 new hectad records, of 238 taxa in County Wexford in 2013.

How do I create a list of all new tetrad records in a vice-county in a year?

The query, above, for new hectad records can easily be adapted to list all new tetrad records. The required changes are highlighted in yellow.

The screenshot shows a search query interface with the following settings:

- search for: Records (specimens or observations)
- county/region: Co. Wexford
- date recorded: 2013 equals
- group by: taxon name (including qualifier and authority)
- distinct: tetrad
- subdivide by: tetrads (aggregate)
- more options: [button]
- join with: Records (specimens or observations)
- join using: taxon-tetrad
- exclude: [checked]
- county/region: Co. Wexford
- date recorded: to2012 to for vague record date ranges use end date

<http://bsbidb.org.uk/search.php#query=699bd2d0c6a5e0884ad0e1ef9879159b>

Again, in County Wexford in 2013, there were 6810 new tetrad records of 721 taxa.

How do I create a list of all new monad records in a vice-county in a year?

The hectad / tetrad queries (above) can be further modified to list all new monad records in a year. This shows that in County Wexford in 2013, there were 12,506 new monad records of 745 taxa (!). the query is below:

The screenshot shows a search query interface with the following settings:

- search for: Records (specimens or observations)
- county/region: Co. Wexford
- date recorded: 2013 equals
- group by: taxon name (including qualifier and authority)
- distinct: monad
- subdivide by: monads (aggregate)
- more options: [button]
- join with: Records (specimens or observations)
- join using: taxon-monad
- exclude: [checked]
- county/region: Co. Wexford
- date recorded: to2012 to for vague record date ranges use end date

<http://bsbidb.org.uk/search.php#query=6cd268cf56a8a86ecbb8065de10b2f24>

How do I create a list of new hectad records of Nationally Rare or Nationally Scarce taxa in a vice-county in a year?

The query for new hectad records in a year is simply adapted by adding in 'National Status' in the 'checklist/attributes' field. Leaving the 'value' field blank, means that the query returns both NR and NSc taxa.

The screenshot shows a search query configuration in the BSBI database. The search criteria are: county/region: South Hampshire; date recorded: 2013 equals; checklist/attributes: National status (with an empty value field); group by: taxon name (including qualifier and authority); distinct: hectad; and subdivide by: hectads (aggregate). A second query is joined with the first, using 'taxon-hectad' as the join key, with criteria: county/region: South Hampshire; date recorded: to 2012 to for vague record date ranges use end date.

<http://bsbidb.org.uk/search.php#query=f46eef1902add5cb040ee346e337abea>

This query is likely to return taxa which are not native in your vice-county, so will require scrutiny. (There is an option to restrict queries to taxon hectad combinations listed as native in Atlas 2000, but this should not be used within this query; it will prevent new hectad records from being listed).

Data sources and datasets

How do I view all DDb records for a year, originally entered into an individual copy of MapMate?

An individual MapMate CUK can be entered into the 'mapmate centre' field. This is found by clicking on 'more options > source > mapmate centre'.

The screenshot shows a search query configuration with the following criteria: date recorded: 2014 equals; and mapmate centre: 2dd.

<http://bsbidb.org.uk/search.php#query=5840a4894ed5d2a70389a50a264c59f7>

How do I create a list of all data sources and data sets with records in the DDb for a vice-county?

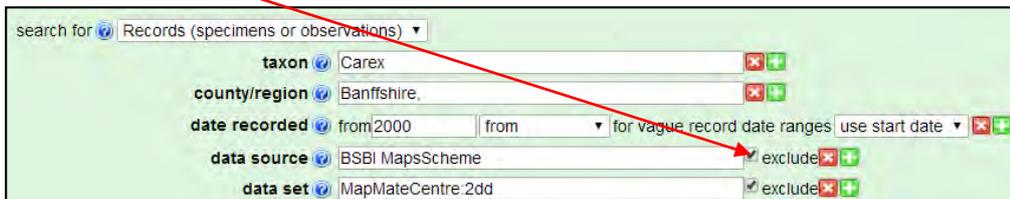
Data source and data set are available as options via 'more options > grouping > group by'. Using both options in a single query, lists data sources broken down into individual data sets. In the case of South Northumberland, the DDb has records from 104 data sets.

The screenshot shows a search query configuration with the following criteria: county/region: South Northumberland; and group by: data source and data set.

<http://bsbidb.org.uk/search.php#query=b2df1c17efa40e3f5e93052e72a73dcb>

How do I view records of a taxon or group of taxa, which are restricted to or exclude individual data sources and / or data sets?

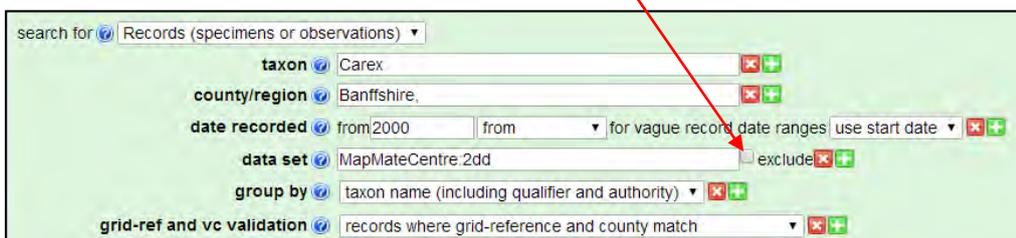
'Data source' and 'data set' can be added to queries via 'more options > source'. Ticking the 'exclude box' excludes any records from these sources, when the query is run.



<http://bsbidb.org.uk/search.php#query=35ab99c498bbf51671304fe164c59999>

The query above therefore, returns all records of *Carex* species or hybrids, from vc94, since 2000, excluding records entered into the VCRs copy of MapMate (CUK 2dd) and from the BSBI MapsScheme.

If you want to restrict a search to just those records from a particular data source or data set, add that option, but leave the 'exclude box' un-ticked.



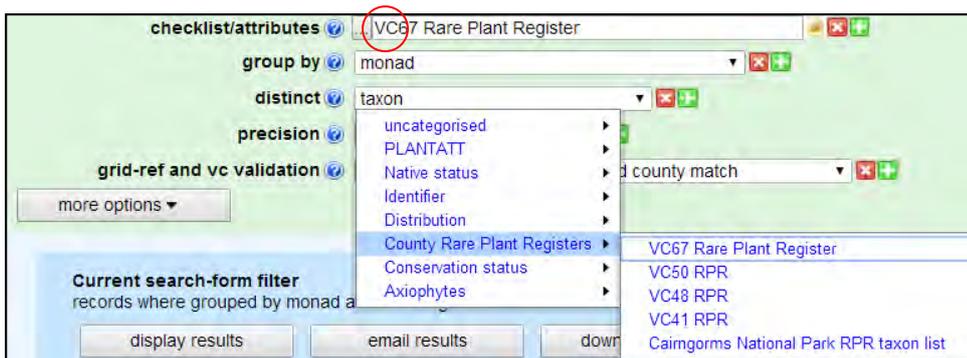
<http://bsbidb.org.uk/search.php#query=8797a193240e064253d31325e546bd9f>

This example lists all *Carex* taxa (with number of records), recorded in vc94 since 2000, which had been entered into an individual copy of MapMate (CUK 2dd).

Rare plant registers

How do I use a list of Rare Plant Register taxa in a query?

The list of taxa included in a CRPR can be selected by clicking on the grey square with '...' to the right of 'checklist/attributes' and selecting 'County Rare Plant Registers'.



At the time of writing there are only four CRPR lists available, along with that for the Cairngorms National Park (a bounded area). RPR taxa lists can be sent to Tom Humphrey or Quentin Groome to be added to the list. This is strongly recommended, as the RPR list can be used in any DDb query. A few examples are given below to illustrate what a powerful tool this provides.

Using vc67 (South Northumberland) as an example, note that you must enter a vice-county into the 'county/region' field as well as selecting the vc67 RPR in the 'checklist/attributes' field; normally the vc selected should match the CRPR. The query (below) lists the number of RPR qualifying taxa recorded per monad since 2000.

The screenshot shows the search interface with the following parameters:

- search for: Records (specimens or observations)
- county/region: South Northumberland
- date recorded: from 2000
- checklist/attributes: VC67 Rare Plant Register
- group by: monad
- distinct: taxon
- precision: monad (1km) or better
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=0c2a0abbeba1e782e1b4ea2694612af0>

The query, below, lists the number of monads in which each RPR taxon has been recorded in vc67 since 2000.

The screenshot shows the search interface with the following parameters:

- search for: Records (specimens or observations)
- county/region: South Northumberland
- date recorded: from 2000
- checklist/attributes: VC67 Rare Plant Register
- group by: taxon name (including qualifier and authority)
- distinct: monad
- precision: monad (1km) or better
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=04347b3c06b736a43d32b584b24f86bd>

The CRPR taxa list can be used in conjunction with a bounded area; in this example all SSSIs in vc48 (Merionethshire). The query below returns all records of taxa in the vc48 RPR that have site grid references at 100m precision or better, and that fall within or overlap the boundary of any SSSI.

The screenshot shows the search interface with the following parameters:

- search for: Records (specimens or observations)
- county/region: Merionethshire
- date recorded: from 2000
- checklist/attributes: VC48 RPR
- bounded area: area name, SSSI, include this area, include intersections
- precision: 100m or better
- grid-ref and vc validation: records where grid-reference and county match

<http://bsbidb.org.uk/search.php#query=19ffe216a89bf0739fcf7d1aa5ad6f10>

This query can be changed to return all records that do not overlap any SSSI by changing 'include this area' in the 'bounded area' clause to 'exclude this area' (highlighted in yellow below).

<http://bsbidb.org.uk/search.php#query=0aed1c9bb2cbe496e8270d56094db0c5>

The RPR taxon list can be used within a bounded area defined as a radius around a central grid reference. The query below returns all records of taxa listed in the Cairngorms National Park RPR, that have site grid reference precision at 1km or better, and that are within or overlap a 1.5km radius of the National Park office in Granttown on Spey.

<http://bsbidb.org.uk/search.php#query=c9cbaf6bf3864d5c94c4bbac4ddabb12>

Grid references and vice-county errors

How do I view all records for a vc, where the site grid reference is outside the vc boundary?

There are two types of site grid reference errors that the DDb can check for. The first is where a record, belonging to a vice-county has a site grid reference that falls outside the vice-county boundary. The query below returns all records which have vc94 (Banffshire) listed in the vice-county field of the record, but with site grid references that fall outside, and do not overlap, the vice-county boundary.

<http://bsbidb.org.uk/search.php#query=10a980983c6ca49df3f19a6c8f75f5b3>

Specifying precision at hectad or better, excludes records without a grid reference, which the query may display as errors. These will include records for the whole vice-county or a parish that cannot be allocated to a site grid reference.

The next query summarises these erroneous records by the data set they belong to.

<http://bsbidb.org.uk/search.php#query=7c368e0525a272e140ef068f823fe48f>

The second type of site error, is where a record for a different vice-county, has a site grid reference within 'your' vice-county. This is described in the next FAQ.

How do I view records for another vc, that have erroneous grid references that fall within 'my' vc?

Records, belonging to another vice-county, may have site grid reference errors that mean that the record falls within 'your' vice-county. To search for such errors, you need to use your vice-county as a 'bounded area'. This makes the DDb undertake a spatial (GIS type) query.

<http://bsbidb.org.uk/search.php#query=5da24297fa77f27ab03dfb1bcad47154>

How do I find out the correct (or inferred) vice-county for records misattributed to a vice-county?

When you download these records, tick the option to add an additional column showing 'calculated vice-county based on grid-ref (where different)'.

How do I take data analysis and mapping further?

While the DDb provides powerful tools to explore, summarise and (to a degree) analyse records, it would be a mistake to expect it to be able to do all that you might want. The ease with which records or record summaries can be downloaded to a spreadsheet, allows much more to be done with the data, by additional exploration and analysis in Excel or a similar spreadsheet programme.

Downloads of records can also be imported into GIS. For those dealing with GB records (but not Orkney or Shetland) site grid references can be easily converted to centroid or south-west corner points using the Excel conversion utility that can be downloaded from

<http://naturedata.blogspot.co.uk/2012/08/blurring-and-plotting-os-grid-references.html>. This is a brilliant tool, creating points suitable for importing into GIS. For other vice-counties there are ways to carry out the conversion using a spreadsheet.

Feedback and future versions of this guidance

Please feed back any comments via the DDb message board. Revised versions of these guidelines will, hopefully, be made available at intervals.

Back to [Contents](#).