3 BugsCEP Database System Description

BugsCEP, the welcome screen of which is shown in Figure 3.1, is a complicated piece of software, consisting of over 30 000 lines of program code in addition to the c. 200 000 individual data records. The *data* and the *program* together constitute an advanced research and teaching tool far more useful than the data alone could possibly be. The data and program are described in this chapter, with minimal attention paid to the applications of the system, some of which are described through case studies in Chapter 6. Note that that the descriptions are orientated more towards users than developers, and no attempt is made to follow any particular system analytical schemes[i]. Some basic entity-relationship models – conceptual diagrams that explain aspects of data and program flow – are included where their use is more practical than written descriptions alone. The second half of this chapter, from section 3.4 onwards, functions as a descriptive manual, and may be used in combination with the inline help when using the program. Section 3.5 provides examples of the reports described in the preceding sections, along with basic instructions on how to create them.

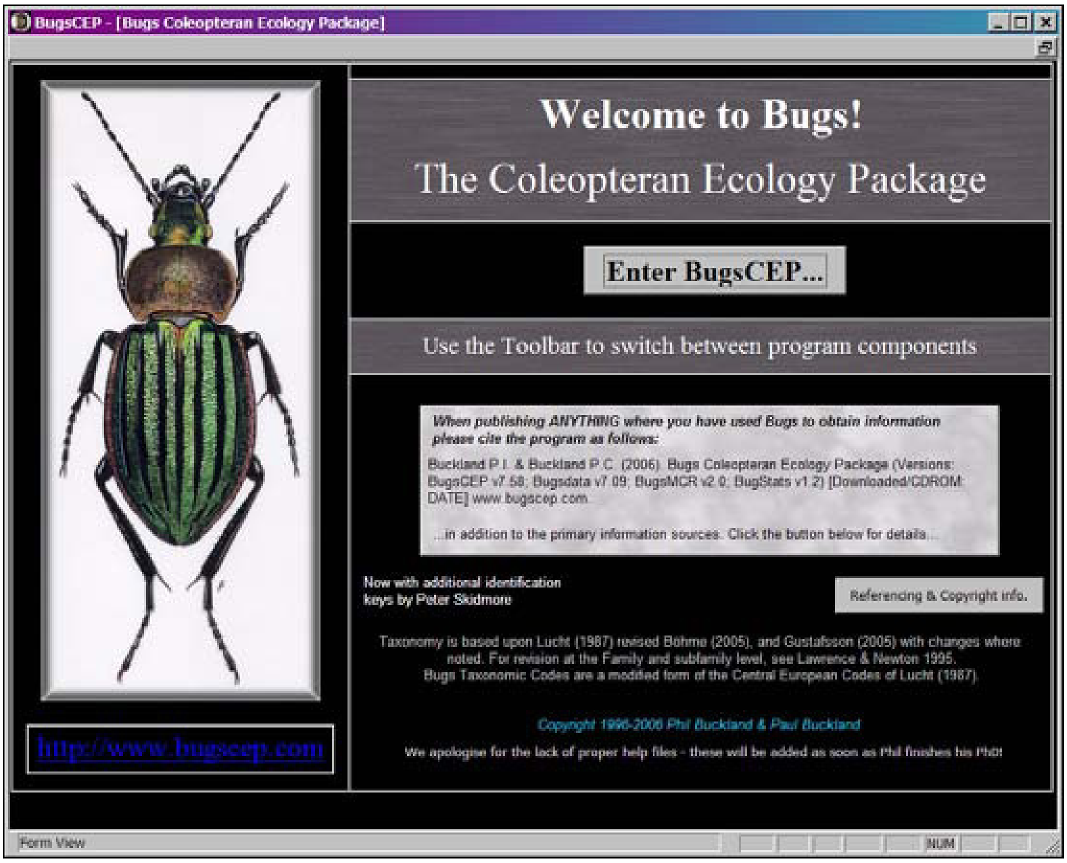


Figure 3.1. Screenshot: BugsCEP start-up screen.

3.1 Data Area Descriptions

The data in BugsCEP can, for the sake of practicality, be divided into three areas: common data, modern reference/calibration data, and fossil/site related data. This is illustrated in the database structure diagram (Figure 2.3), which also shows how the tables mentioned below are related. Data are distributed among the 43 tables in a logic that minimises the amount of duplication of data and empty cells, and groups of tables are used together to store combinations of data items that can be extracted as information. These groups are described below, using the major data area divisions outlined above,

[i] There are innumerable software development schemes to choose from, search the Internet for tips.

and with an aim towards providing an insight into the way data are organised with respect to the information accessible via the user interfaces, as described in section 3.4 and onwards.

3.1.1 Common data area

This area contains data that are of relevance to both modern and fossil specimens, or of a structural nature. The master taxonomic index, identification keys, synonyms, taxonomic notes and bibliography form the bulk of this data.

3.1.1.1 Master species list and taxonomic code

The *INDEX* table is a nodal element in the BugsCEP database structure, and stores the master taxonomic list to which all other taxon data are in some way linked. The CODE field – the primary key for this table – is used to identify taxa uniquely, and changes in this cascade to all related tables. This means that only the taxonomic CODEs, and not full species names, need to be stored in any of the related tables. This enables the developer to easily see which taxon is attached to which data throughout the database when debugging or constructing queries. The taxonomy is based on Lucht (1987), as revised in Böhme (2005) and Gustafsson (2005), although some major revisions, such as the repositioning of Dytiscidae before Carabidae as followed by Gustafsson 2005, have not yet been implemented.

The form of the taxonomic code remains the same as that used in the original Bugs database (Sadler *et al.,* 1992), and is derived from the Central European Code of Lucht (1987) as follows:

CODE: 4.0310040 =

4.	031	004	0
Family	Genus	Species	‘Flag’
This example:			
Dytiscidae	*Dytiscus*	*marginalis*	-

The ‘flag’ is an extra digit that allows for the insertion of species not in the Central European list, and can be used to cope with a limited amount of taxonomic readjustment. The use of fractional numbers allows for the relatively easy addition of further ‘flag’ digits should it be needed, although some reprogramming and interface adjustments would be needed to handle this in the software. Note that the authority, L. in the above example, is not a component of the taxonomic code.

The gymnastics necessary around the taxonomic code in order to encompass other regional faunas are evident in the Egyptian database (Buckland *et al., in press*). This situation will have to be resolved in subsequent versions, and it may be necessary to include support for parallel taxonomic systems to provide greater international support.

3.1.1.2 Synonyms

As discussed earlier, insect taxonomy is complex, and undergoes constant revision. This leads to the fact that the same species may have been described under different names by different authors and subsequently synonomised. BugsCEP caters for this by allowing the storage of any number of synonyms for a taxon, along with supporting references and notes. These synonyms are searchable in

[ii] See Chapter 3, on the development of BugsCEP’s database structure.

various areas of the program, and can be used as an alternative navigation method on the main screen (3.4.1.4).

3.1.1.3 Taxonomic notes

Notes on taxonomy, such as those relating to identification or the splitting of species, are stored as paired reference-data couplets, where the reference always gives the source of the data.

3.1.1.4 Measurable attributes

Measurable attributes are stored in a table designed for flexibility which can be adapted to any quantifiable trait. At the time of writing, this table is only used to store body length data[iii] in the format shown in Table 3.1.

Table 3.1. Example of size data stored in the measurable attributes table.

CODE	Attribute Type	Attribute Measure	Value	Units
01.00100200	Length	Min	15	mm
01.00100200	Length	Max	19	mm

The example shows the minimum and maximum recorded body lengths of the tiger beetle *Cicindela sylvatica* L., measured in millimetres. At the moment there are no constraints to the entry of attributes – as long as the Value field contains a number – and users are free to enter any useful data. As any number of quantifiable attributes can be stored for each taxon, this feature has great potential for morphometric studies, and even environmental studies if, for example, data such as pollutant tolerance limits were to be included.

3.1.1.5 Identification keys (provided by Peter Skidmore[iv])

A number of dichotomous identification keys to elements of the British fauna are stored in the *TKeys* table as text (memo) data. These and the taxonomic notes also contain comments on closely related species so far not recorded from the UK, a useful adjunct where fossil material is being considered.

3.1.1.6 Bibliography

The *TBiblio* table contains full references for all information sources cited in BugsCEP. It forms a bibliography of over 3 300 scientific articles and books on modern biology and distribution, fossil record, and to a small extent methodology. About 20 % of the cited works concern palaeo work. Notes may be attached to references, but this feature has not been used much. References are stored as three fields, as shown in Table 3.2

[iii] These data were extracted by from a number of written and online sources, including Fauna Entomologica Scandinavica and Harde (1998). Additional data were extracted from Skidmore (*unpubl.*).

[iv] Skidmore (*unpubl.*)

Table 3.2. Bibliography table fields, with example.

Field	Description	Example
Reference (or Ref)	Abbreviated citation used by BugsCEP to link tables, similar to the Harvard System's recommendation for citations in the body of a text. By convention, references to sources on fossil work include the date in brackets, and references to work on modern specimens are without brackets.	Lindroth 1985
Author	Names and initials of all authors and date of publication.	Lindroth, C.H. (1985)
Title	Article/book title, journal, volume, page numbers, publisher etc.	The Carabidae (Coleoptera) of Fennoscandia and Denmark. Fauna Entomologica Scandinavica, 15,1. E.J.Brill, Leiden.

With hindsight, this data should have been stored as individual fields per data item, e.g. Authors; date; title; journal etc., to allow for easier export and more advanced searching. This is, to a large part a legacy problem, in that although a semi-automated conversion would have been possible, it would also have entailed the manual checking of at least 2 500 references imported from the previous version of Bugs. It was decided not to do this, and that the system was efficient enough for its purpose in BugsCEP, with full title searching being sufficiently fast with the small number of records involved.

3.1.2 Modern reference/calibration data

Bugs was initially designed to allow palaeoecologists to analyse their fossil beetle faunas with reference to the modern ecology of the species. The data reflect this, in that the significant bulk of them describes the known biology, distribution and ecology of living insects. Ecological summary data are included in the form of two coded habitat and ecology descriptors – Bugs EcoCodes, and Koch ecology codes (Koch, 1989-92). UK RDB classifications are included for about a quarter of the species. A large number of synonyms are included, which is necessary where the referenced literature is international and of varying age. Thermal envelopes are included for over 400 taxa, enabling at least as good a possibility for climate reconstruction as any previous software. The envelopes have been imported from the existing MCR software RECON (see chapters 2 and 5).

3.1.2.1 Biology and distribution text abstracts

Similar in structure to taxonomic notes above, these are stored as paired reference-data couplets, where the reference always gives the source of the data. If a text abstract refers to another work then the latter work is given its own record, with the data "[See X]" – where 'X' is the reference to the secondary referrer. For example, the two records shown in Table 3.3 are amongst the biology data for the dytiscid *Ilybius vittiger* (Gyll.). The later authors have commented on the work of an earlier author. Where the primary data for the species includes an abstract from the primary reference, the "[See X]" data are omitted.

The use of the term 'biology' rather than 'ecology', denotes the more general nature of the information in the modern data area when compared to the more ecology specific information in the 'Ecology Summary Codes' described below. The biology records can vary from broad generalizations to specific local observations. One should *always* take note of the source of the data, and in particular the geographical scope of the work. For example, Nilsson & Holmen (1995) are cited in the example with respect to the "...aquatic Adephaga (Coleoptera) of Fennoscandia and Denmark.", whereas in the biology data for another dytiscid *Nartus grapii* (Gyll.) are abstracted form Duff (1993) on the "Beetles of Somerset...". BugsCEP includes a broad range of specific and general sources, but may be found lacking for certain species and geographical areas, the data for which need expanding.

Table 3.3. Two biology records for the water beetle *Ilybius vittiger* (Gyll.).

CODE	Ref	Data
04.02401010	Nilsson & Holmen 1995	It occurs in small stagnant water-bodies shaded by trees or boulders and often containing Sphagnum mosses. Mainly in wet spruce forests. Often at bog-margins or, especially in the north, in open bogs. The larva was described by Nilsson (1983)
04.02401010	Nilsson 1983a[v]	[See Nilsson & Holmen 1995]

The distribution data are in the form of abstracted records of occurrence. This is a poor substitute for a GIS system with individual find location coordinates, but provides a broad overview of the recent geographical range of species. In addition, there exist a number of national and international systems purposefully built for recording the geographical location of insect and beetle finds. For example: in the UK, MapMate and Recorder[vi] have been used by many entomologists; Sweden is developing the ArtPortalen[vii] for centralizing species observations; and internationally, GBIF[viii] is becoming a strong contender for a global species information centre, including the recording of museum collection data. Currently, none of these systems provide adequate resolution for detailed, pan-European, studies.

Both biology and distribution texts can be searched using the search interface (see 3.4.5).

3.1.2.2 Ecology summary codes (Bugs EcoCodes & Koch ecology codes)

Two sets of tables hold coded ecological summary values, which can be used to either summarize the ecology of a single taxon or a group of taxa. These tables are fully normalized and allow for any number of codes to be assigned to any taxon. The coding systems – Bugs EcoCodes (created by the author of this thesis and Paul Buckland) and Koch (1989-92) codes are described in detail in Chapter 4.

3.1.2.3 Red Data Book (RDB) classifications

The Red Data Books are an internationally recognised set of publications describing the rarity of species in different countries, and are highly valued in conservation research and planning. RDB status is a measure of the security of a species in a particular country, and the different systems provide classes for a number of degrees of rarity, from 'extinct' (EX) to 'least concern' (LC), for example, in the IUCN 2001 3.1 system (IUCN, 2001). A flexible set of normalized tables have been put in place to allow for multiple country and international rarity indicators to be stored (Figure 3.2). Although keys are only currently included for the UK, Swedish and international IUCN RDB classifications, it would be easy to add further systems and countries. RDB systems are linked to the bibliography through the 'Ref' field in the *TRDBSystems* table, and species are linked to the main *INDEX* table through the 'CODE' field in the *TRDB* table.

[v] Nilsson (1983) in the references for this thesis.
[vi] MapMate (2000) © Teknica Ltd - http://www.mapmate.co.uk
Recorder 2002 (2004) © NBN Trust - http://www.nbn.org.uk/recorder
[vii] ArtPortalen (2006) (The Species Gateway). Swedish Species Information Centre and Swedish Environmental Protection Agency - http://www.artportalen.se/
[viii] GBIF (2006) http://www.gbif.org/

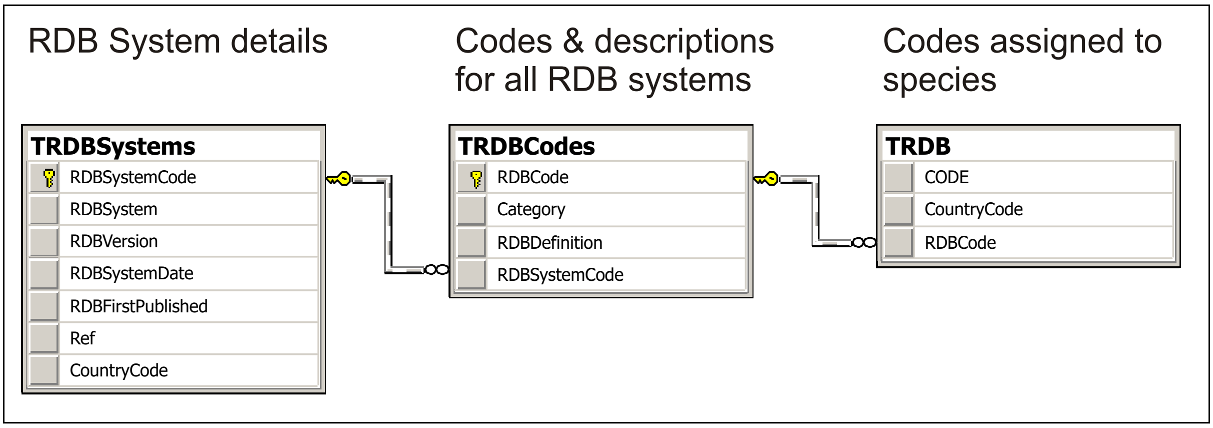


Figure 3.2. Structure of the Red Data Book (RDB) system and code (classification) data area.

3.1.2.4 Species associations

Taxa may be associated with any number of other taxa, and the nature of the relationship specified, along with a primary reference stored. This was a late addition to the database, and as such contains very little data – just over 500 records – which reflect the research activities of the developers during the final two years prior to this publication. This should be considered a work in progress, and the 'Association Types' available are not formalized at the moment. The field can even be left blank, indicating an unspecified general association, most likely by habitat, although this is not recommended. The majority of association data included so far is from Lindroth (1985, 1986).

3.1.2.5 Seasons of adult activity

Data on the seasons during which adults of a species are active are of use to both collectors and palaeoecologists alike, in that it provides a proxy for seasonality, and can be of use when identifying modern specimens. Currently, only UK data for just over 2 200 species are included, but the structure allows for seasonality data for any number of countries to be added[ix].

3.1.2.6 Climate calibration data (thermal envelopes and summaries)

The calibration *data* will be briefly summarized here with respect to how it is stored in the database. Chapter 5 discusses in detail the BugsMCR system for climate calibration from Coleoptera, which is an implementation of the Mutual Climatic Range (MCR) method (Atkinson *et al.,* 1986).

BugsCEP includes the 'MCR Birmingham' data from Russell Coope's group working in Birmingham (Atkinson *et al.,* 1986; Perry, 1986). To simplify the import process, the original binary grids, which represent one-degree Celsius cells, were imported and converted into a structure where each record represents a one-degree TMax (mean temperature of the warmest month) interval for a taxon. Each record includes 60 fields, which represent the total extent of TRange (difference between the averages of the warmest and coldest months) values for all taxa in one-degree cells. In addition, two further fields allow the records to be assigned to a taxon, and sorted in the correct order.

Information is stored documenting the conversion of each of the taxa in the original MCR dataset into BugsCEP taxa. This information is of no use in BugsCEP, but is stored in the name of backward compatibility and would allow BugsMCR results to be easily compared with those generated by RECON.

[ix] A minor alteration would also allow the storage of other seasonality data than adult activity – such as larval activity and breeding season.

The use of binary cells allows easy and fast manipulation of the data using SQL. Unfortunately, the use of queries in MS Access leads to file bloating, and in the case of the record and field rich MCR calculations, so much occurred that this convenient method had to be abandoned in favour of VBA calculation. The latter, although it requires more complex programming, is potentially a more flexible and powerful method.

Data for the taxa thermal envelopes, including the outer limits of TMax, TMin and TRange, along with a centre of gravity measure (COG) for each envelope are stored in a summary table. The outer limits are used in the 'BugsMCR Predict' program component (see 3.4.4 and Chapter 5), in which species presences can be predicted from temperature values. The COG values are not currently used by the program, and are part of ongoing development and enhancement of the MCR functionality (see Chapter 5).

3.1.3 Fossil/site data area

BugsCEP has the capability of holding abundance data for both fossil sites – the results of palaeoentomological analyses, and modern sites – the results of entomological collection or survey. For this reason, this data area will be generally referred to as fossil/site data. Sites descriptions can be stored, along with references, the location of specimens, geographical information and other metadata. Samples, site species lists, and abundance or presence/absence data are created and displayed as countsheets, i.e. cross-tabulations of numbers of individuals per species/sample (see Table 1.1).

BugsCEP contains the greater part of the published palaeoentomological data for Europe, and a few extra-European sites, where records, for example of ectoparasites or pests of stored products, are relevant to overall interpretation. In addition to abundance data, an amount of descriptive, geographical and other metadata is stored for sites. There are a few important omissions, such as Coppergate, York (Hall *et al.*, 1992; Kenward & Hall, 1995), where it has not yet proved possible to extract the data in a coherent form[x], and several sites where there are single identifications, sometimes of doubtful validity (e.g. Mjöberg, 1915). Site records, with essential metadata and references have been included for many of these sites, and it should be possible to add them in the near future. All sites are fully referenced, and a simple yes/no indication of the availability of other proxy data for sites is also included. Dating evidence is stored for samples where available, and is divided into calendar, radiometric or period dates to allow for the variable nature of archaeological and Quaternary site dating.

The logical hierarchy of Quaternary, and modern sampled beetle abundance data was introduced in Chapter 1. To an extent, BugsCEP utilizes this logic in its structure, with samples and species lists nested under sites. To improve database efficiency however, spreadsheet data are normalized into a number of tables. As this data area is more complex than the other areas, and indeed, is probably of more interest to the majority of readers, it will be explained here in greater detail, including a detailed enumeration of records in the following section (3.2.1).

3.1.3.1 Sites

A site is defined in BugsCEP as the geographical focal point of an investigation. The vast majority of sites stored in BugsCEP are either the results of palaeoenvironmental or archaeological investigations. Sites contain both the metadata to enable the original sources to be located, and the abundance data relating to the investigation, as described in the following sections.

[x] Although the basic species list has been extracted.

3.1.3.2 Site descriptive data and metadata

In addition to the site location, in latitude and longitude and national grid systems, BugsCEP can store a limited amount of descriptive information for sites. The aim was not to produce a system specifically for storing site metadata, as there are already systems in existence for this purpose (e.g. ADS – Archaeological Data Service[xi]), but to provide enough information for the users to be able to get an overview of the current interpretation of the site, locate it geographically, and trace the original work. All of the above could be considered metadata to the species lists, samples, abundances and dates. Site interpretations are necessarily summaries made by past and present authors, and are at some risk of oversimplifying or possibly miss-interpreting the original work. As always, original sources should be consulted for full details.

Site data are stored in the *TSite* table (Figure 3.3), and site references are catered for by an additional linking table, *TSiteRef*, between *TSite* and *TBiblio* (the bibliography), which allows for an unlimited number of references per site. (*TSiteRef* and *TBiblio* have been omitted from Figure 3.3 for clarity, see the full structural diagram, Figure 2.3, for relationship details).

A further table, *TSiteOtherProxies*, is a simple, non-normalized table of yes/no fields to record the presence of other proxy data for a site. This provides a useful overview of the scope of non-beetle analyses at the site. It could potentially be used at a later date as a handle for linking BugsCEP to other proxy databases. References to these other data are progressively being added to the bibliography.

3.1.3.3 Countsheets: species lists, samples & abundance data

A countsheet is a collection of species abundance data for a group of samples at a site (the spreadsheet inset of Figure 3.3). A site can have any number of countsheets, and these are displayed and exported by BugsCEP in a spreadsheet like form (see section 3.4.2.3). Data are not actually stored in this form, as a spreadsheet is a highly inefficient structure in that it can potentially contain many blanks, and it is difficult to assign unique identifiers to each abundance value (i.e. spreadsheet cell). Figure 3.3 shows how data and metadata are stored in relation to sites and countsheets. Note that *TFossil* is joined to the central *INDEX* table through the (taxonomic) CODE field, which provides the easy display of species names in the spreadsheet display.

Countsheet records are used to store descriptive metadata on the sets of samples, species lists and abundance data for sites, in this way allowing samples to be grouped into meaningful subsets of all those investigated at a site. The metadata also allow countsheets to be defined in terms of *fossil/modern*, and further qualified as *abundance*, *partial abundance*, *presence/absence* or *other* data for statistical purposes.

Samples may be given X, Y and two Z coordinates (top and bottom) to enable easier spatial analyses. The Z coordinates are used to order stratigraphic sequence samples by depth on the dates screen, and will later be available when choosing the sort order for exported spreadsheets and diagrams. As archaeological sites are not often sampled in columns, these data will potentially be sorted in a different manner, and the X, Y coordinates may be more relevant here. An '*Other Palaeo*' category is provided for fossil sites that do not fit either of the above.

[xi] ADS http://ads.ahds.ac.uk, Arts and Humanities Data Service, York, UK.

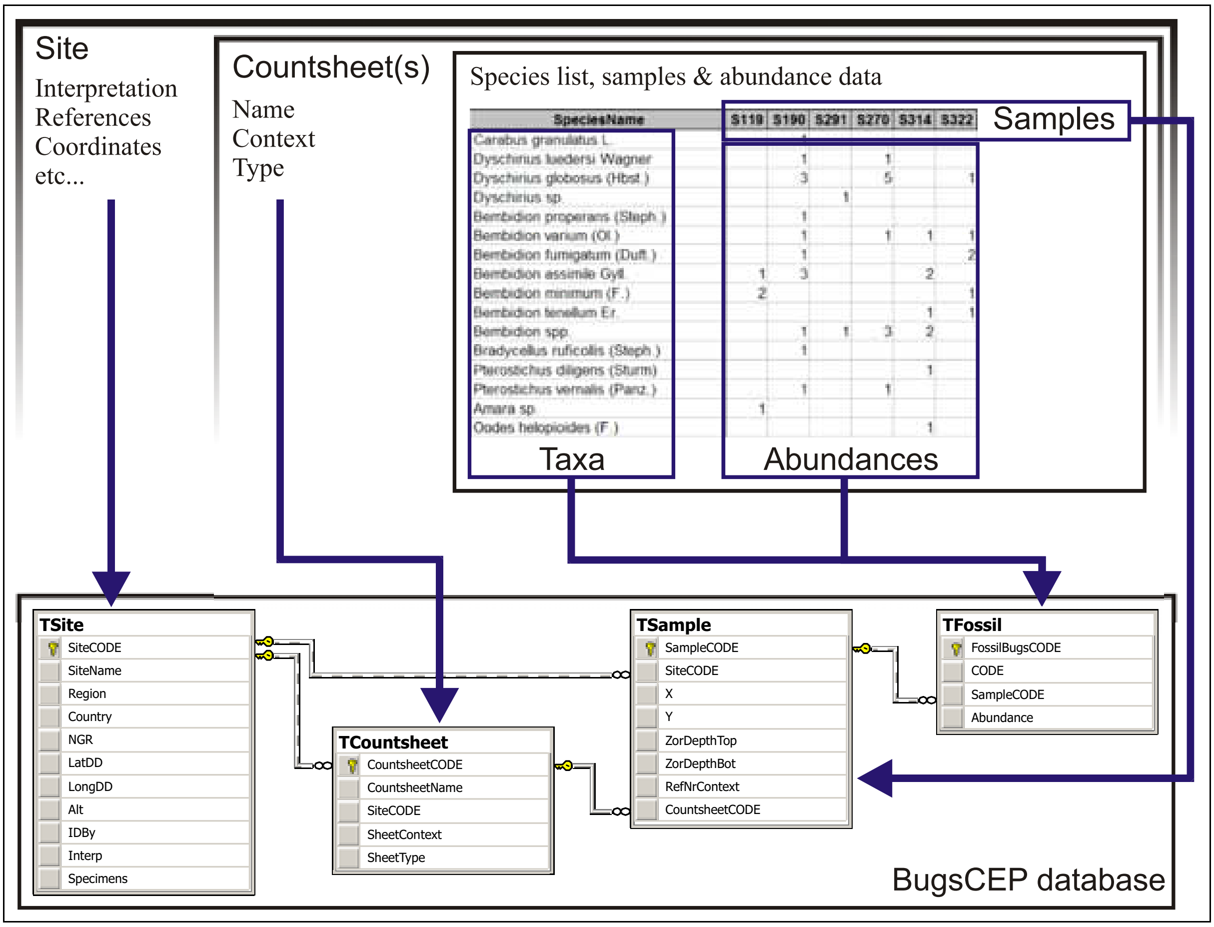


Figure 3.3. Structure of the site and samples area of the database, illustrating how data items are displayed (top) and saved (bottom). Note that the empty cells in the countsheet are not saved in the *TFossil* table.

3.1.3.4 Sample dates

Dates must be attached to samples, and can be defined in terms of three types: calendar, radiometric and period, as described below. Any number of dates can be assigned to a sample, and ranges and uncertainty can be specified using a number of flags as shown in Table 3.4. This allows for a great deal of flexibility in assigning dates, which is especially useful for archaeological sites where dates tend to be limited to contexts containing artefacts or hearths. Previous versions of Bugs only stored dates as lists for sites, and the new version provides much more advanced possibilities for data retrieval and querying.

1. Calendar – dates in specific years or ranges in BP, BC or AD in calendar years. These are most often relevant to archaeological sites, where dates may have been obtained from artefacts or specific events. A somewhat macabre example is that of the single individual of the 'graveyard beetle' *Rhizophagus parallelocollis* Gyll. found in the lead coffin of Anne Mowbray[xii], which was sealed in 1482 (Stafford, 1971; Girling, *pers. comm.* to Paul Buckland).
2. Radiometric – primarily used for radiocarbon dates, but a number of other radiometric (or similar) dating methods are selectable from the 'Dating Method' box. These dates are given in the standard form of *Lab Number* followed by Age BP ± Error, and *always* in radiocarbon (or other method) years.

[xii] Anne de Mowbray – the tragically short lived 8th Countess of Norfolk, Duchess of York, and Duchess of Norfolk, was buried initially in Westminster Abbey, London in 1482.

3. Period – the least specific dating category, which allows samples to be assigned to a number of age categories. These categories vary in their range from shorter archaeological periods, such as, the 150 year Late Roman UK period from AD 250-400, to considerably longer geological periods such as the Holocene, spanning the last 10 000 years.

Table 3.4. Available sample date range and uncertainty flags.

Type	Flag	Date types
Range	From	Calendar, Period
	To	Calendar, Period
	>	Radiometric
	<	Radiometric
Uncertainty	Ca	Calendar, Period
	?	Calendar, Period
Range with uncertainty	From ca	Calendar, Period
	To ca	Calendar, Period

All dates require that the dating method is chosen from a drop down box, and notes can be added should further clarification be necessary.

3.2 Record Enumeration & Database Size

There are many ways of structuring databases, and so enumerating record counts is often of limited real comparative value between systems. They are, however, useful for comparing the relative amounts of data in different sections of a database if a uniform approach to its construction has been followed. Table 3.5 summarises the record counts for the main areas of BugsCEP's data. It should be noted that Bugs is updated continually, and that these counts represent the counts at the time of writing of this thesis. There is also a considerable amount of support data, such as look up tables, definition lists etc., not listed here that enable both the user interface and statistical methods to function.

Table 3.5. Summary of BugsCEP data, showing number of records per data area (support tables excluded).

Species related	
5 845	Taxa
4 698	Synonyms
528	Records of species association
1 719	UK RDB rarity records
5 407	Species attribute records (mainly size data)
Bibliographic	
3 373	References
Fossil record and other sample based	
601	Sites
656	Abundance datasets/countsheets
89 917	Fossil record entries (which make up the 656 countsheets)
4 871	Recorded dates, of which: 646 radiometric; 3 685 period; 540 calendar
Climate and Ecology	
436	MCR thermal envelopes
22 884	Biology text excerpts
22 331	Distribution text excerpts
8 302	BugsEcoCode assignments
11 245	Koch Ecology Code assignments
9 726	UK Monthly activity records
Other	
267	Identification keys (by Peter Skidmore)

Earlier papers have presented the physical size of the database in terms of disk space. At the present day these numbers are somewhat irrelevant, due to variations in data compression/compaction methods, database structure (normalization issues in particular), and other software/hardware issues which can cause the same amount of data to occupy different amounts of disk space. The numbers are, in fact, only interesting in terms of nostalgia (cf. older versions) and practical installation requirements; they are presented here in Table 3.6 for the sake of completeness.

Table 3.6. Physical disk space size of BugsCEP database (see text for disclaimer)

BugsCEP program part	18 MB
Bugsdata database part	34 MB

For comparison, the original Bugs software (Sadler *et al.,* 1992) totalled 8 MB, Bugs2000v.4 was 37 MB, and Bugs2000 v.5 (Buckland, 2000) – the version immediately preceding BugsCEP – was about 150 MB. At around 52 MB, despite a massive increase in data, BugsCEP represents a significant improvement in data storage efficiency over previous versions[xiii].

3.2.1 Sites overview

Aside from the fact that BugsCEP is Eurocentric, the geographical distribution of sites primarily reflects the activities of palaeoentomologists, rather than the availability of suitable sediments. The access to cut or eroded peat bogs, frozen lakes[xiv], and waterlogged archaeological sites are, however, particularly advantageous for palaeoecology. Naturally, the scope of database also reflects the interests of the developers. Figure 3.4 shows the number of sites per country, and England, as the country where palaeoentomology has been most extensively used and most intensively developed, is clearly the most thoroughly palaeoentomologically investigated.

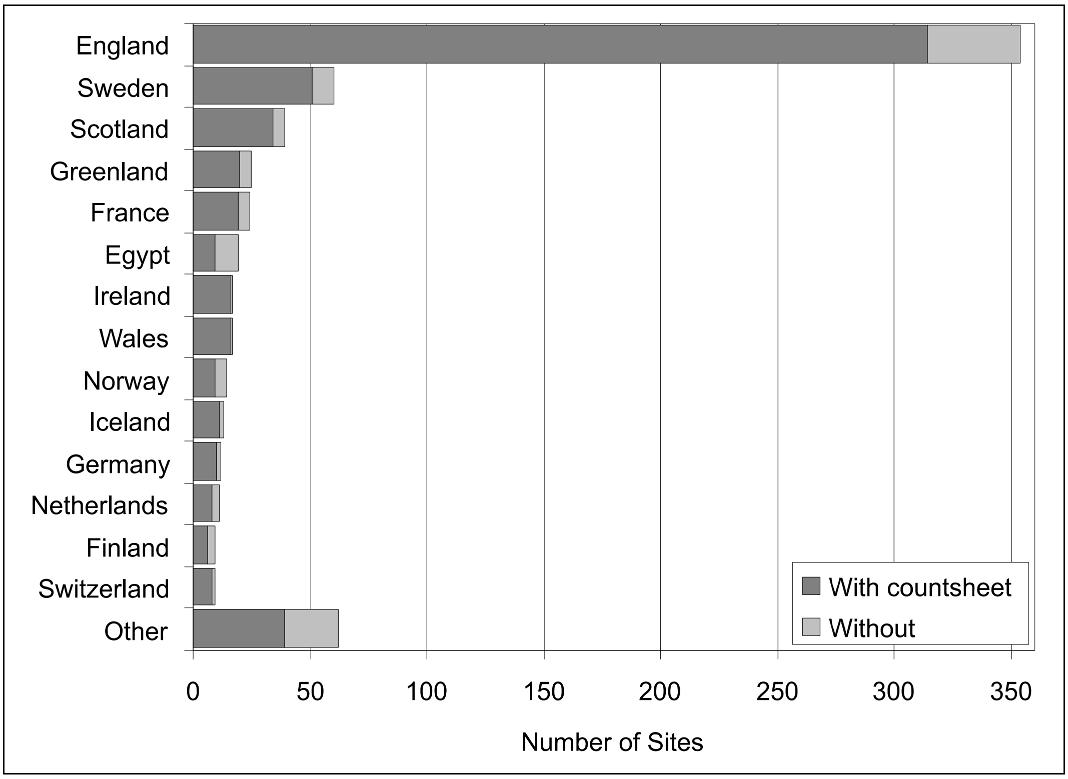


Figure 3.4. Bar graph showing number of sites per country. Lighter parts of the bars represent sites with only summary information and no abundance data. 'Other' shows the total sum for all countries with less than nine sites per country.

The majority of sites have been assigned latitude/longitude, and can thus potentially be useful in GIS analyses. Figure 3.5 shows the location of the sites which currently have coordinates, and lists the number of sites per country for which they are lacking. The latter are mainly due to a surprising number of published papers that do not give coordinates for the sites. The map clearly shows the primary concentration of sites in the British Isles.

[xiii] Note that due to the use of external library files and the MS Access database runtime engine, the actual installation size can be anywhere between 60-150 MB.

[xiv] Although it can be difficult to obtain sufficient material from the latter.

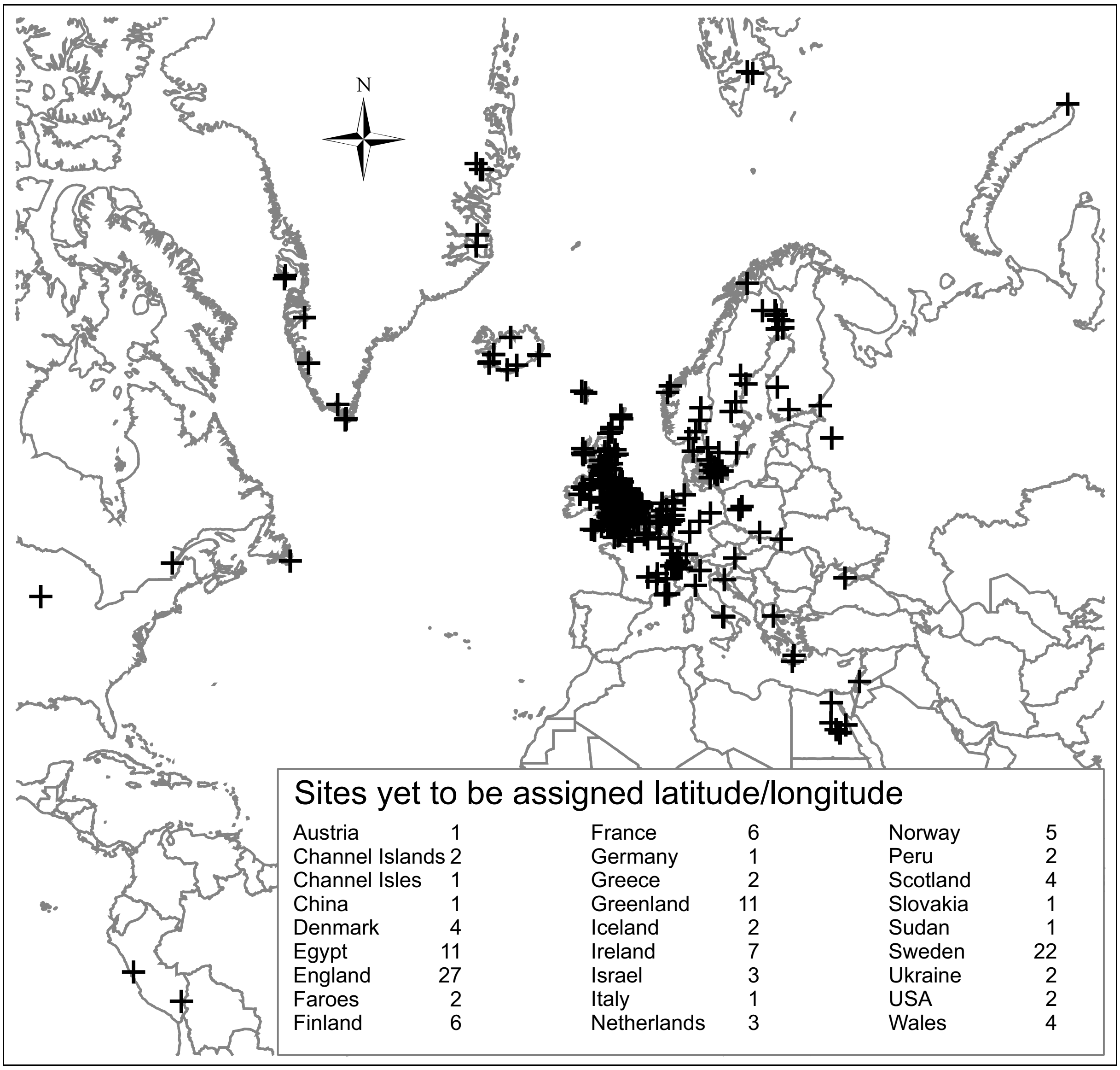


Figure 3.5. Geographical location of sites in the BugsCEP database. Mercator projection.

3.2.2 Countsheets: species lists, samples & abundance data

Table 3.7 shows the number of countsheets stored, at the time of writing, for each of the possible metadata qualifiers which describe the context of the data. Stratigraphic sequences and archaeological datasets make up the bulk of the countsheets, and there is a clear need for the entry of more modern collection data, which would allow for comparative studies between fossil and modern assemblages to be more easily undertaken. The two modern, pitfall trap datasets are discussed in Chapter 6.

Table 3.7. Available countsheet context descriptions and enumeration in BugsCEP.

Timeframe	Context	Count
Fossil	Archaeological contexts	258
	Stratigraphic sequence	345
	Other Palaeo	11
Modern	Pitfall traps	2
	Other Modern	0

The spatial distribution of sites in the core geographical area is shown in Figure 3.6, with stratigraphic sequences and archaeological contexts differentiated (sites without lat/long have been omitted). An interesting point is that Sweden and France have a significantly higher number of sites with stratigraphic sequences (41 and 14 resp.) than archaeological contexts (11 and 4), which most probably reflects the position of palaeoentomology in Quaternary science, rather than archaeology in these countries.

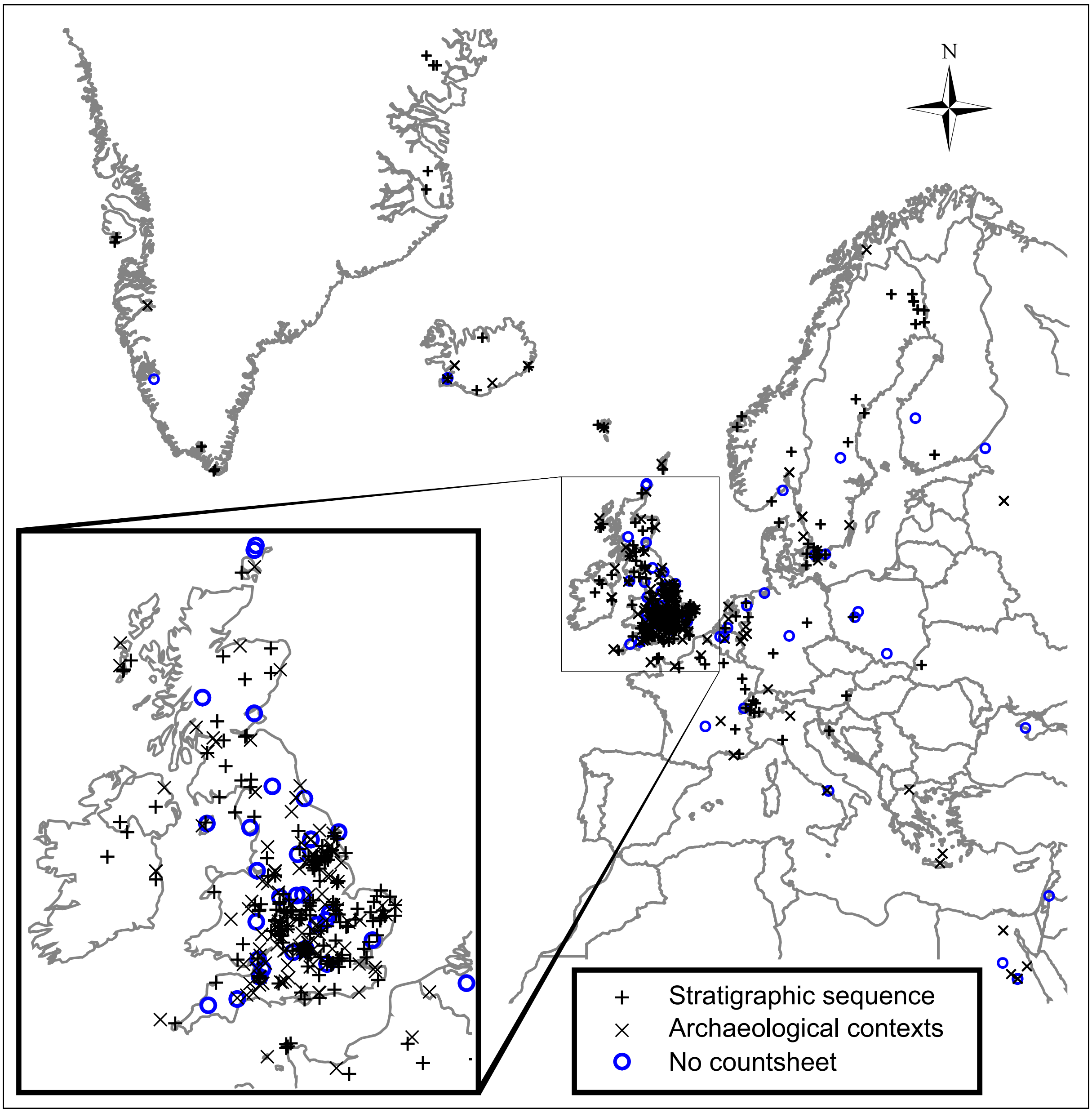


Figure 3.6. Map showing location of sites with stratigraphic sequences, archaeological contexts and no countsheets in the BugsCEP core region.

3.2.3 Sample dates

The enumeration of sample dates is currently a complex matter, as the dating search engine is incomplete, but an approximate summary by dating type is provided here. The final search engine will be able to use a lookup system to search fluidly over all three dating types. Note that a sample may have more than one date of any kind, and that this will affect the histograms below.

3.2.3.1 Radiometric dates

The clear dominance of Holocene and Lateglacial research is indicated in Figure 3.7, although this is of course greatly influenced by the limits of radiocarbon dating. Only 25 dates older than 100 000 BP are currently stored, and have been omitted from the histogram for reasons of clarity. The apparent lack of early-mid Holocene dates is largely a reflection of the poor number of mid-Holocene dated samples (Figure 3.8). Although this may to an extent reflect dating strategies (researchers tend to date the top and bottom of sequences), a cursory examination of the undated samples in BugsCEP suggest that the pattern does indeed reflect a lack of samples as well as dates. This is a significant problem in palaeoentomological research, and makes it hard for the existing data to be used to address contemporary research problems such as the structure and development of mid-Holocene forests, and the impact of Palaeolithic and Mesolithic peoples on the landscape (Buckland [*et al.*], 2005). These time slices are a clear target area for future research, and this illustrates the strategic importance of BugsCEP in identifying areas in need of further research.

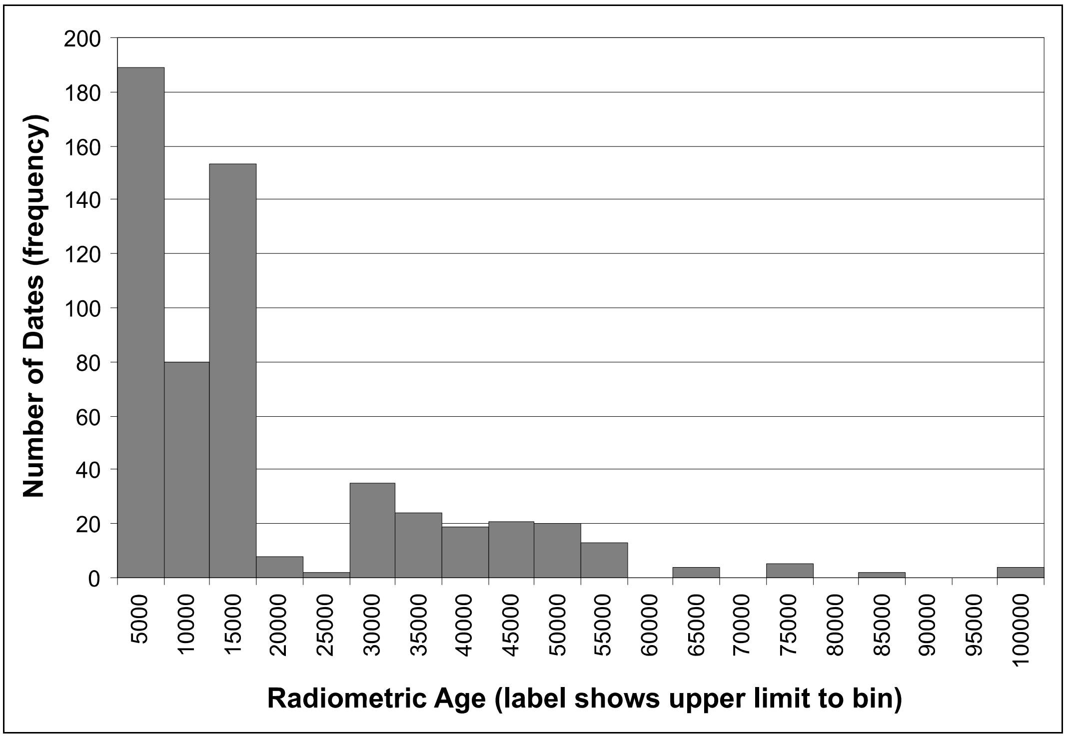


Figure 3.7. Histogram showing number of radiometric dates per 5 000 year period (bin) over the last 100 000 radiometric years.

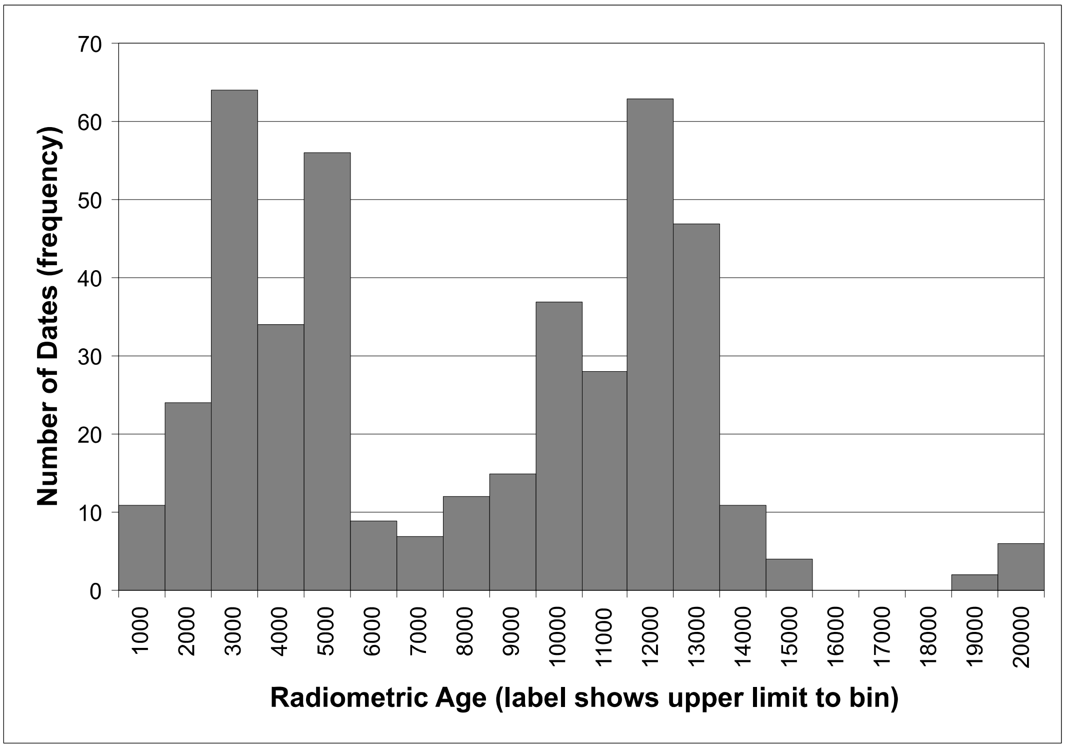


Figure 3.8. Histogram showing number of radiometrically dates per 1 000 year period (bin) over the last 20 000 radiometric years.

3.2.3.2 Calendar dates

Calendar dates are predominantly archaeological in nature, and primarily reflect the focus of archaeological investigations in the British Isles on the medieval and Roman periods (Figure 3.9). 53 dates fell beyond the 5 000 year range of this histogram, their spread lending little to the observed pattern and they were thus omitted from the figure to improve resolution.

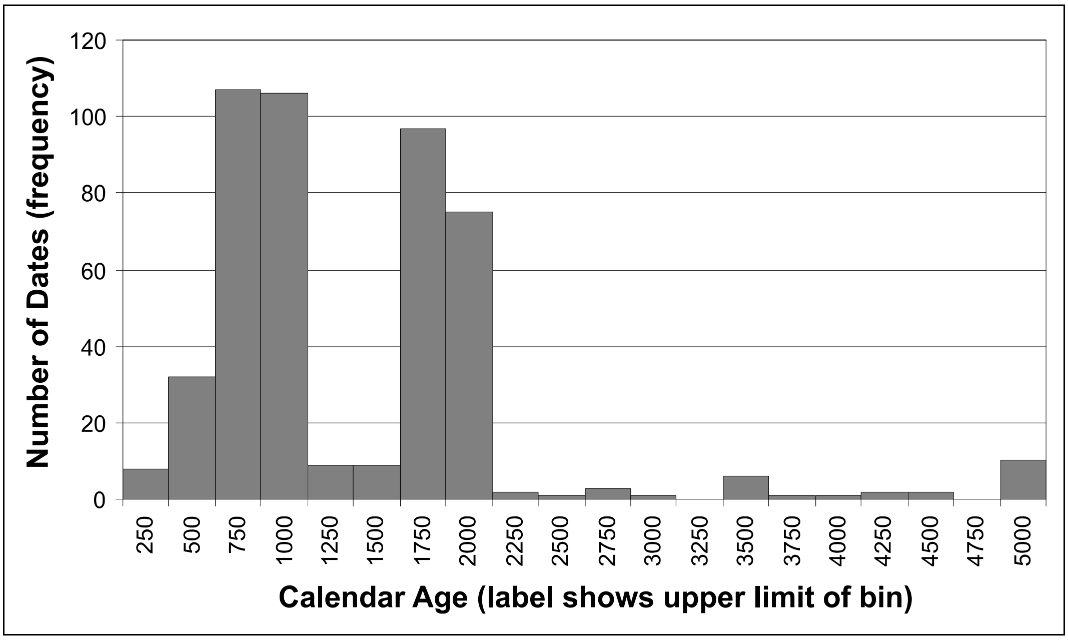


Figure 3.9. Histogram showing number of calendar dates per 250 year period over the last 5 000 years (corrected to calendar BP where present = 1950 for limited comparability with radiometric ages).

3.2.3.3 Period dates

The period date category is still under development, in that the categories still need to be consolidated and assigned calendar or radiometric equivalents to enable cross searching. The assignment of period dates has been somewhat arbitrary, and most often reflects the context of an investigation, for example, a sample from an archaeological investigation is more likely to be classified as post-medieval than Late Holocene. There are also a number of synonymous periods, such as the Eemian, Ipswichian and (Marine) Oxygen Isotope Stage 5e (Bell & Walker, 2005), and it has been decided to use the name given by the original author to avoid potential misinterpretations.

Period dates have been assigned to a large number of samples that have not formally been dated, where their context gives reason to believe a particular period is appropriate. The primary focus of palaeoentomological research on Holocene sediments is again visible in Figure 3.10.

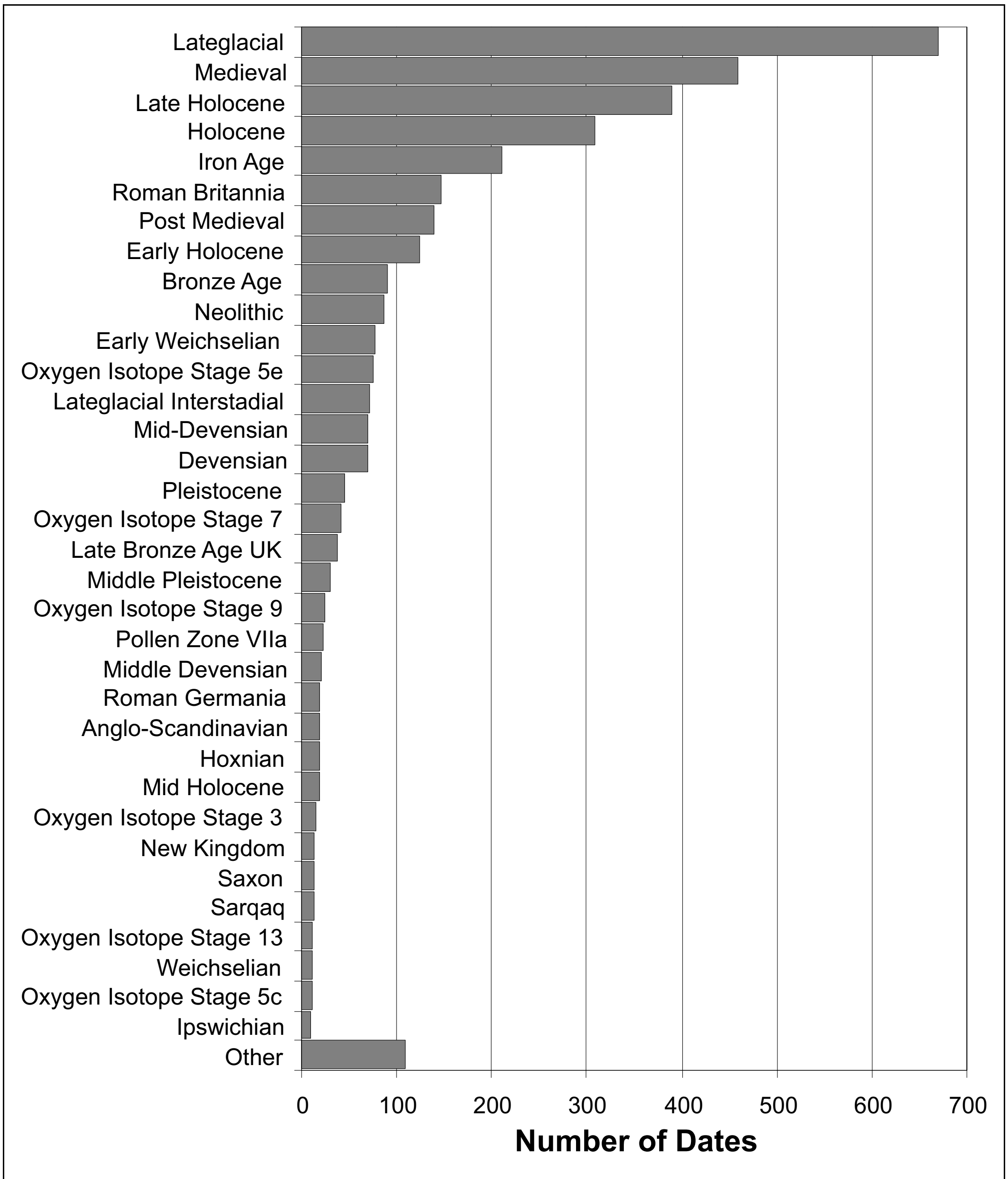


Figure 3.10. Bar chart showing number of dates stored per time period. The 'Other' category is the sum of all periods with less than ten dates.

3.3 User Base

Bugs has never required registration or payment for use, and as such has been freely available for download since 1998, prior to which it was distributed on floppy disks. The tracking of downloads to specific users, a difficult and time consuming task, has been avoided, and as such there is no real record of all users. As mentioned earlier in Chapter 2, there is a general lack of consistency in citing databases and software in scientific publications, despite being covered by the same copyright laws as any other written publication. Quaternary entomology is a small field, with less than 200 researchers world wide, and so the primary user base for the *fossil* side of the database will inevitably be small. However, there has been an increasing amount of interest from modern ecologists, particularly those interested in landscape change or biogeography, and this broadens the potential audience considerably. The Bugs project has taken this interest on board, and enhancements over the previous versions of the software have been implemented into BugsCEP which could increase the usefulness for ecologists considerably.

We have, however, kept track of those users that have specifically requested the database, have requested technical support, or contacted us on other Bugs matters. This information is used here to (anonymously) summarize the proportion of users in different fields and regions, although the numbers and classifications may be approximate due to individuals being active in more than one field. Of the 50 known users, 38 are based at academic institutions, seven are private individuals and five are employed by government institutions including museums. Table 3.8 shows the spread of users by country. Bugs is known to be used in research, teaching and consultancy, and has been implemented as part of teaching modules in the Quaternary in Edinburgh (Scotland), Bournemouth (UK) and Umeå (Sweden). The BugsMCR component has been evaluated for teaching purposes at Royal Holloway (University of London) and East Anglia.

Table 3.8. International distribution of known Bugs users.

Country	Count
UK	33
Sweden	7
Ireland	2
Spain	2
USA	2
Germany	1
Greece	1
Norway	1

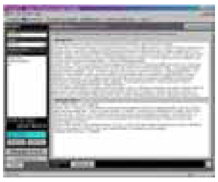
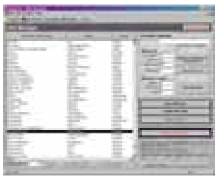
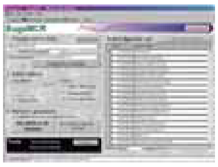
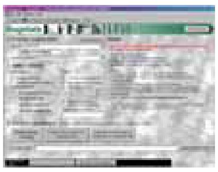
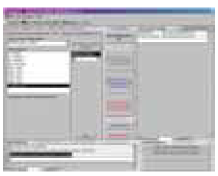
In future versions steps will be made to require user registration so that the use of Bugs can be more accurately approximated. Bugs has previously only been passively marketed, that is to say that there has been a limited amount of promotional material produced, a situation that is not ideal, but is a result of the limited availability of time and funds for marketing. This will hopefully be rectified in the future. This aside, the number of discrete visits to the www.bugscep.com website have steadily increased from around 500 to over 800 per month between November 2006 and March 2007, showing that even passive marketing has some effect.

As an interesting side note, the announcement of BugsCEP on two Coleoptera user groups (british-beetles@yahoogroups.com and coleoptera@yahoogroups.com) was closely followed by 40 downloads. This clearly demonstrates the importance of digital networking and marketing. A similar number of downloads followed the presentation of the software at the Late Neogene Ecosystems Database Workshop (LNED, see Chapter 1), demonstrating the similar importance of making software visible at relevant scientific meetings. None of these downloads are included in the figures above.

3.4 The BugsCEP Program - Detailed Description

BugsCEP can essentially be divided into five interface areas (Table 3.9), which allow the user to undertake a set of specific tasks. An environmental or palaeoecological investigation would inevitably use several, if not all, of these components, but any one can provide independently useful information.

Table 3.9. A summary of the BugsCEP main interface areas. I/O = Input/Output.

1.		**Area:** Primary form: Data: Functions: I/O:	**BugsCEP Main Screen** BugsCEP Main Species specific information: ecology, distribution, taxonomy summary of fossil record associations with other species bibliography identification keys species based data retrieval add/edit species based data add/edit references export data for selected species export selected references
2.		**Area:** Primary form: Data: Functions: I/O:	**Site & Countsheet Management** Site Manager Site specific information: site summaries (interpretation, lat/long, references, etc.) countsheets (species, sample and abundance data) dating evidence create/edit sites and summaries add/edit species lists, samples and abundance data add/edit sample based dating evidence import abundance data export abundance data generate site reports
3.		**Area:** Primary form: Data: Functions: I/O:	**Climate Reconstruction** BugsMCR Species thermal envelopes and summary values thermal reconstructions from site abundance data prediction of species presence from temperatures (under development: jackknife calculations) export reconstruction results export reconstruction graphs export climate space maps and sample species lists
4.		**Area:** Primary form: Data: Functions: I/O:	**Environmental Reconstruction** BugStats Coded species habitat summaries environmental reconstructions from site abundance data correlation coefficients explore habitat codes by species export reconstruction results export reconstruction graphs generate sample level reconstruction breakdown report export correlation matrices
5.		**Area:** Primary form: Functions: I/O:	**Bugs Search Explorer** BugStats search for species by habitat, distribution, RDB, ecology export search results (species lists and information) generate report of sites where species present

In addition to the above, there are a number of pop-up and utility forms that are either reused in various areas, such as the Bibliography Browser, or serve maintenance functions. The Bibliography Browser displays references relevant to the database section that it was accessed from, as well as providing a searchable interface to the complete BugsCEP bibliography of over 3 300 books and articles.

3.4.1 BugsCEP Main Screen – basic data retrieval

Upon starting BugsCEP, users are presented with a welcome screen (Figure 3.1) (and on the first run will have to help the program locate the data file and accept the licensing agreement). Clicking [Enter BugsCEP] opens the BugsCEP main screen, which primarily provides access to information on individual species. Initially, this screen shows the seemingly blank screen for the taxon Carabidae indet.[xv], but by pressing the [Next Sp. ▼] the user can update the display to show the data for *Cicindela sylvatica* L. (Figure 3.11), which is the first species in the BugsCEP taxonomic order. By default, the main screen shows Biology & Distribution data. This always consists of a reference and data component – where the reference is an abbreviated citation to the data, which itself is a text abstract, or quote from the cited work. The first item of biology data for *C. sylvatica*, for example, describes the species as being found "on sandy ground, especially heathland, but generally local and rare", with the information being credited to "Harde 1984"[xvi].

Clicking on the [Bibliography for panel] button will open the pop-up **Bibliography Browser**[xvii] and display the full references for all sources cited for the information displayed. The pop-up lists references alphabetically by first author, with "Harde 1984" being the second record in this case (Figure 3.12). The Bibliography Browser allows users to copy individual references with the click of the [Copy Ref.] button, or export a list of all the references for the current species (see section 3.5). The [Show Entire Bibliography] switches display to the full bibliography, and is replaced by a [Show Specific Bibliography] button when clicked. In this way users can easily switch between viewing only the bibliographic information for the current species and the entire BugsCEP bibliography. The browser is also fully searchable, and clicking the [Search...] button opens the 'Search in references' dialog form. This can be used to find the first and subsequent occurrences of search terms in the abbreviated references or full titles.

[xv] Carabidae indet. is the taxon used to label individuals identified only to the level of Carabid family, and as such is taxonomically hierarchically above all Carabid species. The record has therefore no ecological or distribution information itself, as this would be far to generalized to be of any use. It does, however, have a considerable fossil record due to the frequency of fossil specimens that can only be identified to this level.
[xvi] Harde (1984) in this thesis.
[xvii] Interface forms will be written in **bold** the first time they are mentioned.

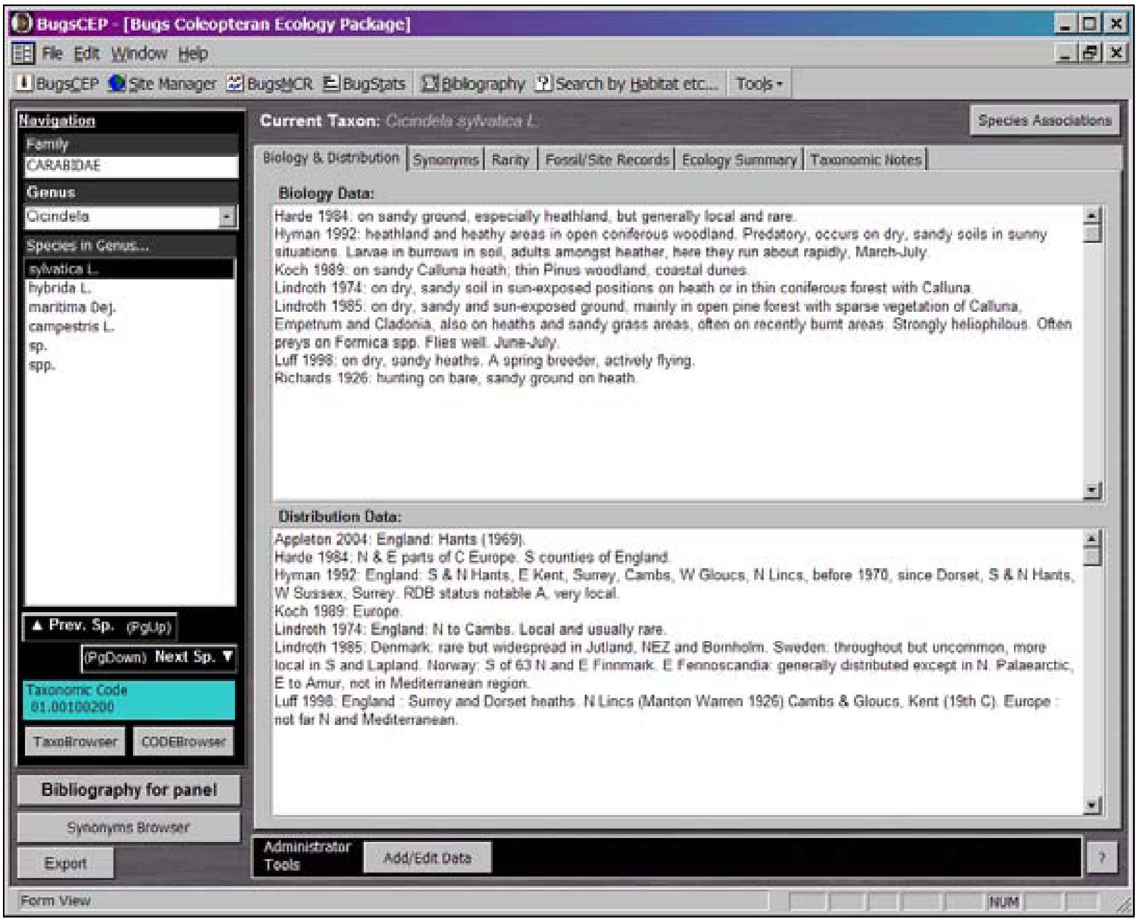


Figure 3.11. Screenshot: BugsCEP main screen showing biology and distribution data for *Cicindela sylvatica* L.

Figure 3.12. Screenshot: Bibliography pop-up showing references for *Cicindela sylvatica* L.

By using paired reference and data items, BugsCEP ensures that there is never any doubt as to the primary source of the information displayed. Also, by using abstracted texts, it removes any possibility of unintentional reinterpretation that would be a problem when summarizing the texts of several authors into a more compact description. This is, of course, always a problem when describing the ecology of species, and is discussed more thoroughly in the context of summary codes later in this thesis (Chapter 4).

The main screen can essentially be divided into five sections, as illustrated in Figure 3.13, which are used to either navigate, display data or activate functions.

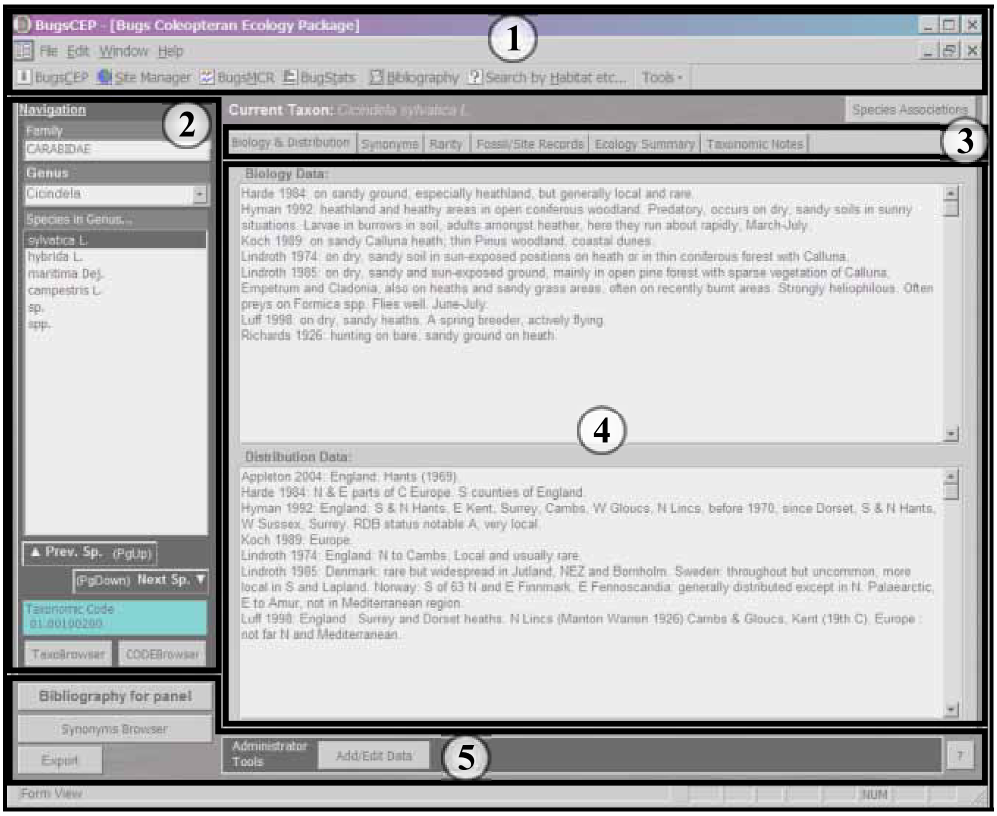


Figure 3.13. BugsCEP Main Screen areas: (1) Title bar, menu bar and component buttons; (2) Navigation panel; (3) Information tabs; (4) Information area; (5) Additional buttons and administrative controls. (The background image is the same as Figure 3.11).

3.4.1.1 Section (1): Title bar, menu bar and component buttons

This area is available from all major interfaces, and allows users to easily switch between program components, and access a number of common functions. The title bar provides the title of the program component, or window, which is currently active. Likewise, the menu items available on the menu bar underneath this will reflect the currently active component, and provide access to functions specific to the latter as well as more general functions available from almost anywhere. Help is available through the menu on most screens, or by pressing the <F1> key on any screen. The Window menu allows users to switch between the program components that are currently open, although this function is unavailable during some critical operations, where its use could lead to loss of data. Standard copy and paste functions are available from the Edit menu[xviii]. Below this, the BugsCEP toolbar provides access to the major components of the program (see Table 3.9) through a number of buttons.

3.4.1.2 Section (2): Navigation panel

This is specific to the main screen, and allows the user to find any species by a number of means. For most users, the most common navigational path will be to type the required genus into the [Genus]

[xviii] Please note that the Find and Find Next functions on the Edit menu only allow you to search within text boxes, such as those for biology and distribution. The Search by Habitat function, as described in section 3.4.5, allows users to searching the descriptions of all species.

'drop down' box on the left at the top of the panel. BugsCEP will search for the genus as the letters are typed, and pressing the <Enter> key will move to the currently displayed genus. When a genus is selected, the [Species in Genus] list below it will be updated to display all the taxa within that genus, including placeholders for specimens identified to generic level. The first species, taxonomically, in the genus will be highlighted, and the information area (4) will be updated to display data for that species. The [Current Taxon] box just under the toolbar will always show which taxon is currently selected, and the [Taxonomic Code] box will display its numerical equivalent. Note that the [Family] box is not selectable, and always displays the family of the currently selected taxon.

Clicking on the [Prev. Sp.] or [Next Sp.] buttons will move the display to the previous or next taxon respectively in taxonomic order, as will pressing the <Page Up> or <Page Down> keys when the cursor is not in an information box[xix].

The Taxonomic Explorer (Figure 3.14) is activated by pressing the [TaxoBrowser] button. This opens a pop-up window which displays three hierarchical panels with all families, genera and species in taxonomic order. Clicking on the [Select Family:] panel will update the [...then Genus:] panel to display the genera within the selected family, which in turn will update the [...then Species] panel to display a list of the taxa within the selected genus. The example shows the navigational path for the Elaterid species *Agriotes pallidulus* (Ill.). Double clicking on the taxon in the final panel, or clicking the [Goto Species...] button will update the main screen to display the available data for the taxon. The pop-up can then either be closed, or moved out of the way so that the main screen is fully visible.

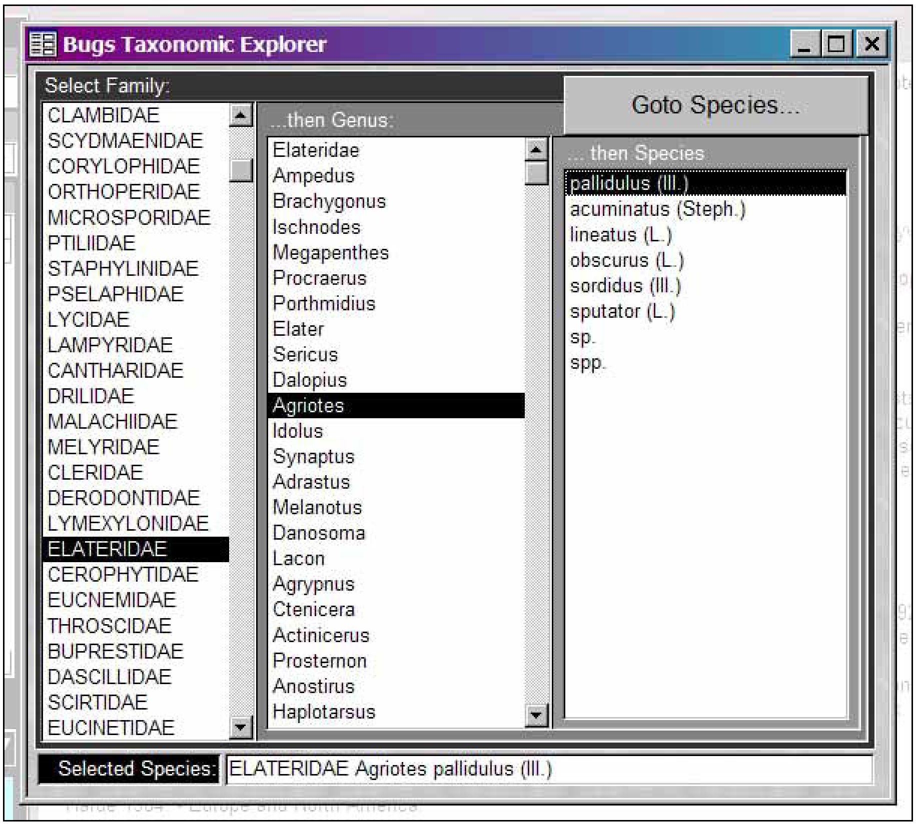


Figure 3.14. Screenshot: Taxonomic Explorer (Main Screen Navigation).

[xix] Navigation with the keyboard requires that the containing interface, and not a subform or data box, has the focus. Using the keys when a text box is selected, for example, will scroll the selected text box, and not move the display to another taxon.

The Taxonomic Code Browser (Figure 3.15) fulfils a similar function, but is aimed more at users with knowledge of, or use for the taxonomic codes used to uniquely identify taxon in BugsCEP. It is particularly useful for quickly locating species where the code is available in an exported report. The [Find a code...] button on this pop-up opens the standard MS Access find dialog box (lower half of Figure 3.15), which can be used to search for a specific code rather than browsing using the scroll bars in the main pop-up. The example shows *A. pallidulus* (Ill.) selected. As in the Taxonomic Explorer, double clicking on the taxon, or clicking the [Goto Species...] button will update the main screen to display the available data for the taxon. The pop-up and dialog can then be either closed or moved.

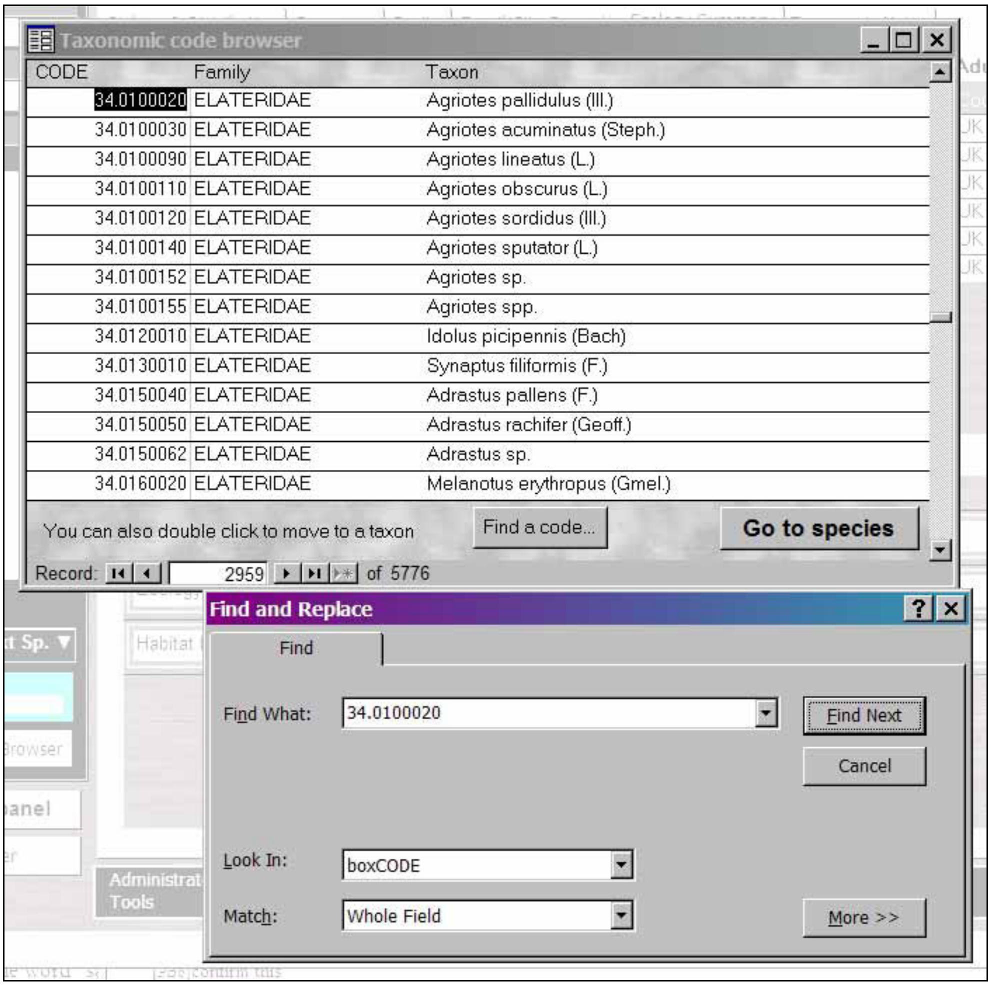


Figure 3.15. Screenshot: Taxonomic code browser (Main Screen Navigation).

3.4.1.3 Section (3): Information tabs, and section (4) Information area

Sections (3) and (4) of the main screen are intimately related in that clicking the information tabs in section (3) dictates the information displayed in section (4). Section (3) is a series of tabs which activate information panels within the information area.

The [Synonyms] panel is quite self explanatory (Figure 3.16), and lists the known synonyms for the selected species, with references to the source for each synonym. Note that whilst BugsCEP does refer to the synonym authority[xx] of either genus or species, it does not always provide the primary source

[xx] Although BugsCEP is primarily a database of ecology and fossil data, and is not designed to hold complex taxonomic information.

for the subjection of the taxon and users should use *Fauna Europaea*[xxi], or other information sources for detailed synonymy. Clicking on the [Bibliography for panel] button, as always, will bring up full details of the references cited (see Figure 3.12). The [Notes] field is used for any important information specific to a particular synonym use, which does not really belong under the [Taxonomic Notes] data for the species in general. Note also that there remain many problems over synonymic pathways in different countries, requiring frequent updates to BugsCEP. The authors would be grateful for notice of any errors or relevant species' reviews.

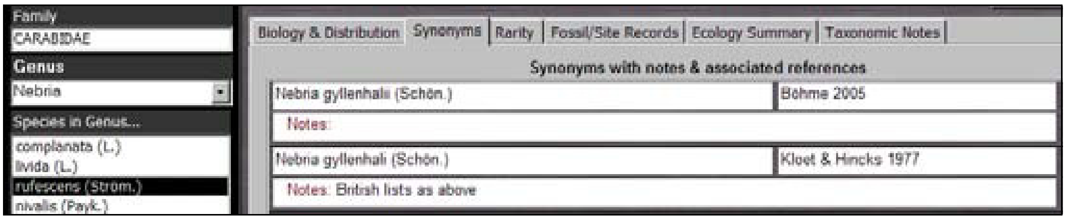


Figure 3.16. Screenshot: Main screen, synonyms panel showing data for *Nebria rufescens* (Ström.).

Synonyms are continually being revised, and the taxonomic designation used in BugsCEP has been standardized as far as possible on Böhme (2005), with a few corrections (Marshall to Marsham throughout, for example), as the most recent published European checklist; other checklists are covered in the synonyms. The use of the taxonomic code as a unique identifier means that the designations can periodically be updated without taxonomic order or links between data being lost[xxii]. The complete list of synonyms stored for all species can also be browser using the [Synonyms Browser] button at the bottom left of section (5).

The [Rarity] panel shows the RDB status for the selected species, along with references and keys for three RDB systems, as shown in Table 3.10. The authors of BugsCEP recognize that these are not the most up to date sources or classifications, but there has not been time to update the data.

Table 3.10. Red Data Book classification systems available in BugsCEP.

System name in BugsCEP	Geographical Range	Reference
UKRDB	United Kingdom	Hyman (1992)
IUCN 2000	Sweden	Gärdenfors (2000)
IUCN 2001 3.1	International	IUCN (2001)

BugsCEP currently includes UKRDB entries for about 1 700 species. The other systems are provided for future use, and allow multiple classifications for any species so that its rarity in any number of countries can be ascertained. This could provide extremely useful in studies analysing the changes in biodiversity caused by environmental perturbations, including climate change.

The Fossil/Site Records panel is perhaps one of the most significant improvements of BugsCEP over previous versions, and perhaps one of its most useful features for palaeoecologists. The panel provides a rapid overview of the sites from which the currently selected species is known (provided that the site is stored in the database, of course). Currently, these sites are almost entirely palaeo in nature, but could equally be modern habitat studies. The current primary function of this panel is to give palaeoecologists a summary of the fossil record of a taxon, the example shown in Figure 3.17 being part of that for the cold tolerant ground beetle *Nebria rufescens* (Ström.). As mentioned previously, samples dates are stored in three categories in BugsCEP: calendar; radiometric; period. The categories

[xxi] http://www.faunaeur.org/
[xxii] The complications and structural issues surrounding the use of a taxonomic code are discussed in Chapter 2.

are displayed as a second row of tabs towards the top of the panel, and can be clicked on to reveal the different kinds of dating evidence. By default, BugsCEP will only display dated fossil records, that is to say, only information from where the current taxon has been found in a dated sample. The records are summarized as: Site name, region, country, sample name and date; and listed alphabetically by site name. Selecting the [All Samples] option button updates the display to show all occurrences of the taxon in samples, irrespective of whether they are dated or not. This is useful for getting a broader view of the occurrence of the taxon in poorly dated or modern sites. Selecting the [Dated Samples Only] switches the display back to the default view.

There were initially plans to include age depth models in BugsCEP, to allow the interpolation and extrapolation of dates between samples in stratigraphic sequences. This is under development at the time of writing, but the structural foundations have been implemented, and users can define a set of samples as a 'Stratigraphic sequence' when entering data (see section 3.2.2).

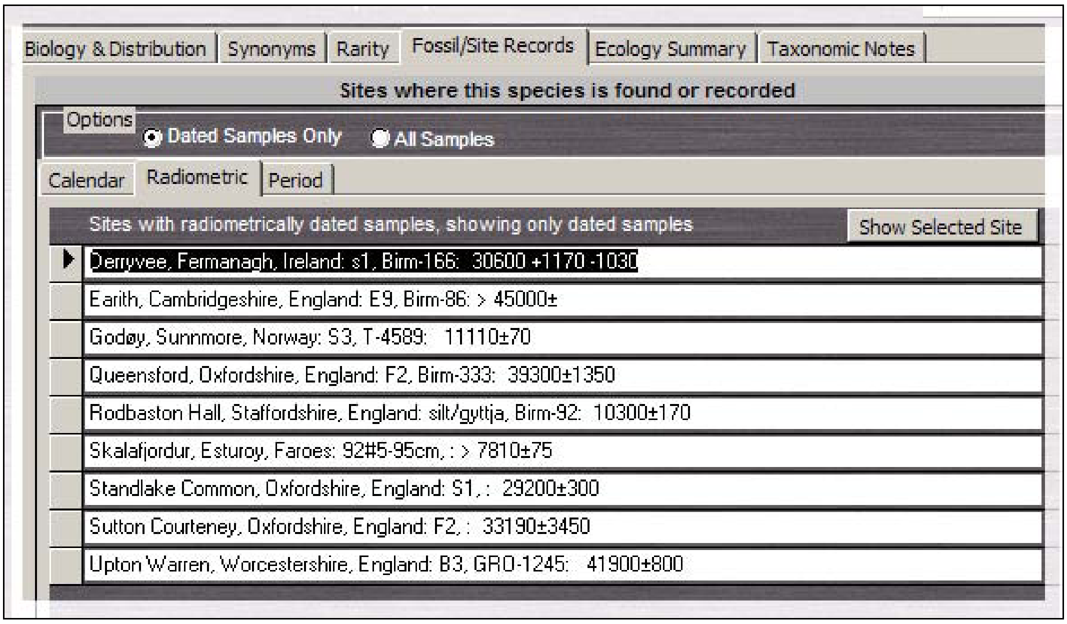


Figure 3.17. Screenshot: Fossil/Site Records panel, showing sites with radiometrically dated samples containing *Nebria rufescens* (Ström.).

The **Ecological Summary** panel (Figure 3.18) shows coded habitat and ecology summaries of the current species, along with information on the seasonal activity of the imagines. The two habitat/ecology classification systems, Bugs EcoCodes and Koch's ecology codes, are summarized in Table 3.11, and described in detail in Chapter 4. The majority of palaeoentomologists have classified their species lists in terms of the modern habitats that they represent, but until now no software has been designed to apply a classification system consistently to any number of sites.

Bugs EcoCodes can be used to create a diagram summarizing the environments represented by any sample(s). This is extremely useful for summarizing environmental change over time, or comparing the signals provided by fossil insects from either different parts of an archaeological site or different horizons within a sequence. This feature is available from the BugStats program component, and is described below (3.4.3), and discussed in detail, with examples, in Chapters 4 and 6.

Seasonality data, which is also shown on this panel.

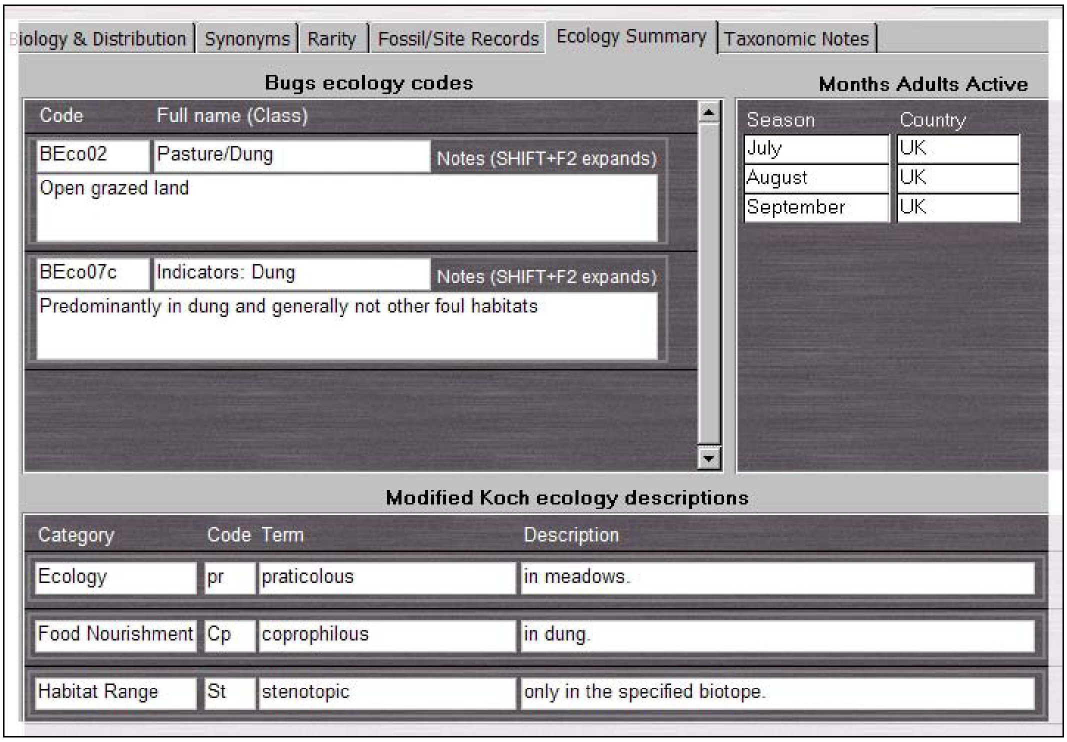


Figure 3.18. Screenshot: Ecological Summary panel for *Aphodius foetens* (F.).

Table 3.11. Basic outline of the ecology classification systems used in BugsCEP.

Ecology code system	Description
Bugs EcoCodes	Habitat summary codes, 22 classes. Devised primarily to be of use to palaeoecologists and quite generalized. Devised by Paul Buckland and Philip Buckland.
Modified Koch's ecology codes	Descriptive ecology codes, 127 classes divided into 6 categories. Give a detailed description of the ecology of the species. Devised by Koch (1989-92), and translated by Paul Buckland and Eva Panagiotakopulu. Computerized by Philip Buckland.

The **Taxonomic Notes** panel displays referenced, abstracted texts which contain information specific to the taxonomy or identification of taxa, rather than their ecology or distribution. The majority of these notes are information on the availability of identification keys, identification tips, and notes on taxonomic inconsistencies or disagreements. For example, in the taxonomic notes for the mould beetle *Latridius anthracinus* (Mann.) Hodge & Jones (1995) note that it was "...previously confused with *Lathridius*[sic] *minutus*, separated on subtle structural differences and genitalia". Clicking the [Bibliography for panel] button shows the full reference details.

The panel also shows **Measurable Attributes**, which currently includes body length data for about 2 700 taxa.

3.4.1.4 Section (5): Additional buttons and administrative controls

Section (5), in addition to holding the [Bibliography for panel] button, provides a number of buttons enabling features for both users and data managers. In the bottom right hand corner of the main screen is a small [?] button, which opens the help file for this screen (the same effect as pressing the <F1> key).

The [**Synonyms Browser**] button opens a pop-up window (Figure 3.19) which lists all the synonyms currently stored in BugsCEP. It allows a user to browse or search for a species by its synonym, and then move to the species in the main screen by clicking the [Goto Species in Main] button. Clicking the [Find synonym…] button opens a custom search dialog, with which the user can search for names, or name parts, within the synonyms list. The example in Figure 3.19 shows how '*gyllenhalii*' is searched for. It should be noted that the Find Synonym dialog box may seem to disappear after clicking on one of the [Find …] buttons, but it is simply hidden behind the Synonyms Browser as the latter is given the focus. Users can position the two pop-ups for more convenient viewing[xxiii].

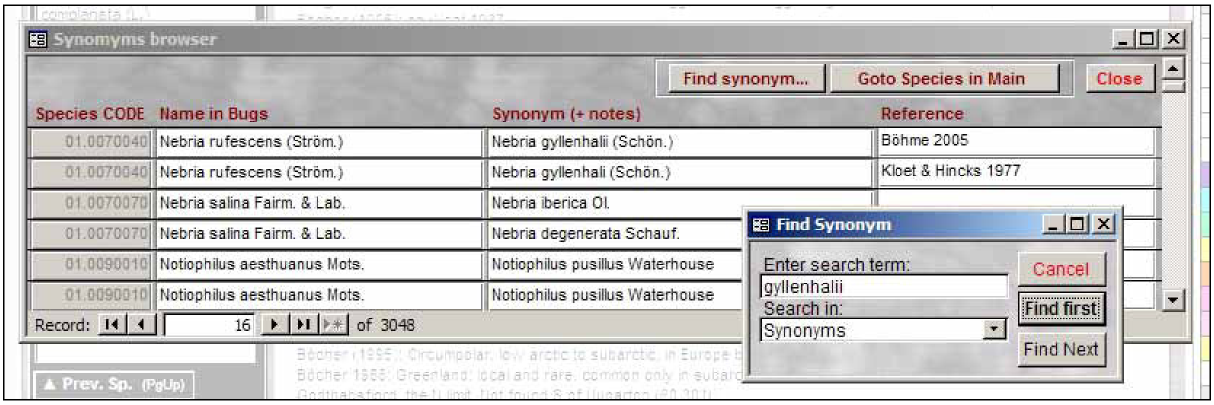


Figure 3.19. Screenshot: Synonym Browser, showing search for '*gyllenhalii*'. The Find dialog has been positioned to overlap the browser for ease of use.

The [Export] button on the main screen provides a variety of options for outputting the data for the current species to an external file or program. A dialog box displays the available reporting options, and the third reporting option, which exports the species name, synonyms and taxonomic notes may be of most use to taxonomists. The fourth option, which outputs the species name, biology and distribution texts along with associated references will undoubtedly be the most useful for ecologists and palaeoecologists as it provides the most descriptive habitat information, with directions to the original sources. The possibilities available on previewing the reports, along with the other report types available are described later in this chapter (3.5).

The final button visible in section (5) is the [Add/Edit Data] button, which is password protected to prevent accidental data alteration. Entering **edit mode** provides an additional [Go to CODE] button, along with buttons for adding, deleting or altering taxon designations. An [Edit Bibliography] button also appears, which is a function duplicated on the [Tools] drop down on the toolbar in section (1). Due to the large scope of BugsCEP's data there is some variety in how different data items are added. Some items must be typed (or pasted from the clipboard), and others chosen from drop down lists. Others, such as entering references are slightly more complex and involve a number of validation checks. Currently, all data are entered, on request, by Paul Buckland on the master copy of the database, and thus the majority of users will not need to enter species related data. The issues surrounding species data entry are briefly explained in section 3.6.

Outside of the numbered screen sections (Figure 3.13), two further buttons, only one of which is visible for all taxa, are provided for accessing information on **species associations and identification keys**. The latter have been provided by Peter Skidmore (Skidmore, unpubl.), and are standard dichotomous keys. The [Keys] button is only displayed when there is a key to the current genus

[xxiii] This is by design. Both pop-ups are set to appear in the middle of the BugsCEP window, and thus the larger one will obscure the smaller one if it is on top of it. This setting eliminates problems that could otherwise arise from different screen sizes, resolutions and numbers on different computers influencing the physical position of pop-up windows.

available, and is only active when the sp. taxon placeholders are selected[xxiv]. The key to click beetle genus *Agriotes* (Figure 3.20), for example, is only available when *Agriotes sp.* is selected.

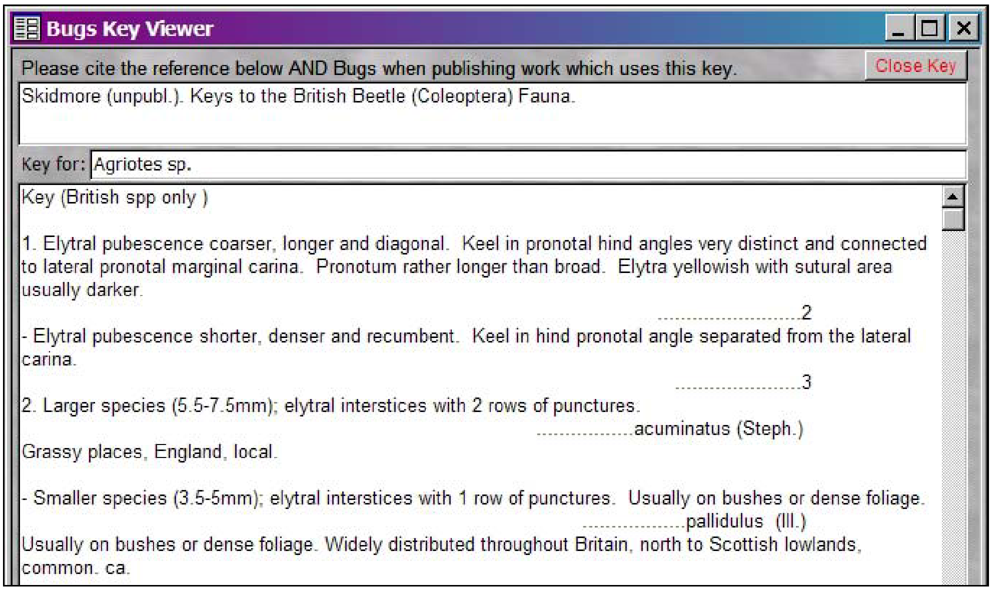


Figure 3.20. Screenshot: First two couplets of the key to *Agriotes* species.

A list of **Species Associations** is provided on clicking the [Species Associations] button, which opens a pop-up listing known relationships between the current species and others. Figure 3.21 shows the four associations stored for the halotolerant wetland/coastal species *Dyschirius nitidus* (Dej.), and clearly indicates its predatory relationship with rove beetles of the genus *Bledius*. Each and every association is referenced, and clicking the [Bibliography] button provides full details of the listed references. In this case, only Lindroth (1985) has been used to provide the raw data.

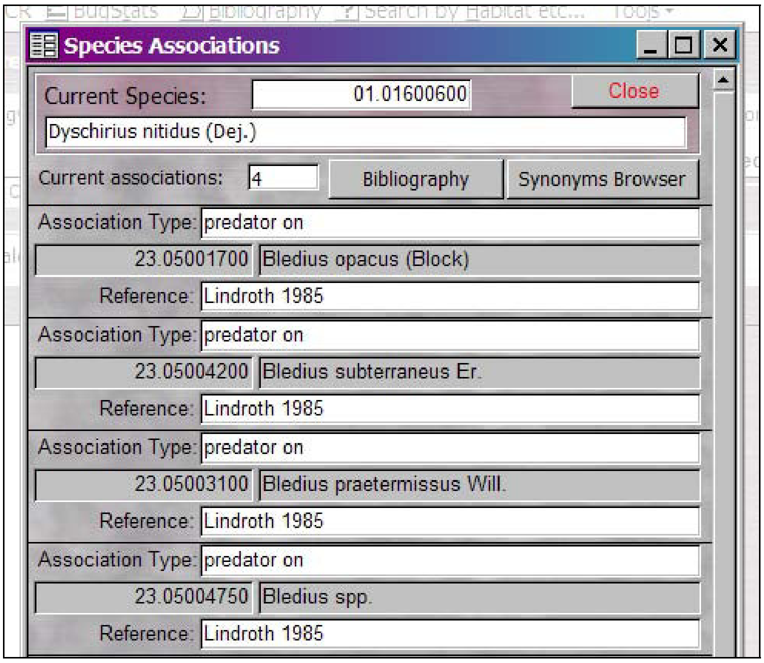


Figure 3.21. Screenshot: Species associations for *Dyschirius nitidus* (Dej.).

[xxiv] This is not an ideal solution, and there are plans to revise the entire identification key system to include hyper-linking and images.

Clicking on the [Species Associations] button when viewing the main screen information for the first of the predated species, *B. opacus* (Block), reveals a reciprocal link to *D. nitidus* (Dej.) along with the information that it is also preyed upon by the closely related, but more eurytopic *D. politus* (Dej.) (Koch, 1989-92).

A large number of the species associations currently included are undefined, indicating non-specific associations between the species. In addition, the large number of non-coleopteran species associations is not catered for by the structure. The majority of associations included are by way of habitat, and it has not yet been decided how to standardize and store the large number of possible relationships into a manageable set. This is clearly an area for future development, and could be useful in studying the differential impacts of environmental change and introductions on species with differing positions in the same food web. For example, the Harlequin ladybird[xxv], *Harmonia axyridis*, introduced into the US from eastern Asia to control aphids "...in 1988, and now the most widespread ladybird on the continent"[xxvi], is now spreading through western Europe, and reached Britain in 2004 (Majerus *et al.*, 2006). It is now well established, out competing native ladybirds and being known to predate on a wide range of invertebrates, including other ladybirds, and is considered a serious threat to the native fauna of the UK. Monitoring of this species is underway, and BugsCEP could help by providing a system for documenting the ecology of the invasive species in its new territories.

3.4.2 Site and Countsheet Management – adding/retrieving sites and abundance data

All data manipulation relating to site data are performed through the **Site Manager** (Figure 3.22), which is available from the BugsCEP toolbar towards the top of the screen.

3.4.2.1 The Site Manager

This screen lists the name, region and country of every site stored in BugsCEP, along with a number of summary counts and buttons for accessing more specific site data. The English Bronze Age site at Brigg, Lincolnshire (Buckland, 1981), can be seen selected in the screenshot, as indicated by the highlighting in the left hand list and the site name in the 'Currently Selected' panel on the right. The first three summary boxes, under 'Number of...' enumerate the total number of countsheets, samples, abundance values (occurrences) for the site. Clicking the [Manage Countsheets & Create Reports] button opens the Countsheet Manager, which is described below in section 3.4.2.3.

[xxv] Also called the Multicoloured Asian Lady Beetle, among others.
[xxvi] The Harlequin Ladybird Survey: http://www.harlequin-survey.org/

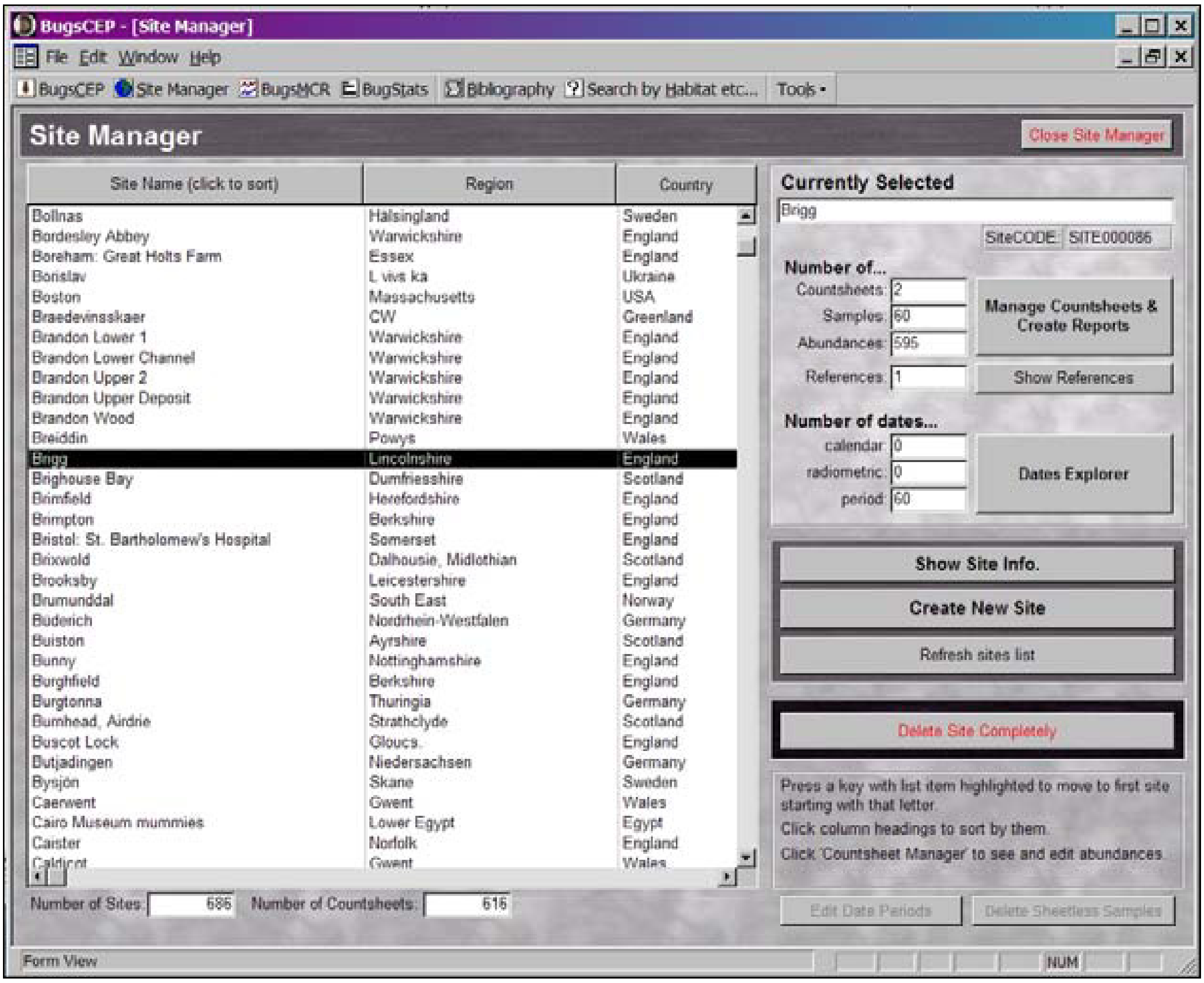


Figure 3.22. Screenshot: BugsCEP Site Manager, with the site 'Brigg' selected.

Below these summary boxes, the number of bibliographic references for the site is shown, and clicking the [Show References] button opens the standard BugsCEP Bibliography Browser with full citations. The number of dates currently assigned to the samples at the site is given in the boxes underneath the 'Number of dates' header. These are enumerated by date category, and clicking the [Dates Explorer] button will enable them to be examined in more detail (see section 3.4.2.5).

3.4.2.2 Site Information screen

Clicking on the [Show Site Info.] button opens the **Site Information** screen (Figure 3.23) with the information for the currently selected site displayed. This screen provides a basic summary of the site – its current interpretation, the names of those who did the identification work, where the specimens are stored (if known), geographical location and a simple check list of other proxy data available for the site. The display of latitude/longitude can be switched between decimal and degrees, minutes and seconds with the DD/DMS toggle button. For countries with an independent national grid, for example Ireland and the UK, a box is provided to enter this. In addition, all of the site's references are listed, as well as the countsheet records for the site, with summary data (bottom left). As before, the [Show Full References] button will open the Bibliography Browser, and the [Manage Countsheets & Create Reports] button will open the Countsheet Manager. The [Show Dates] button opens the Dates Explorer.

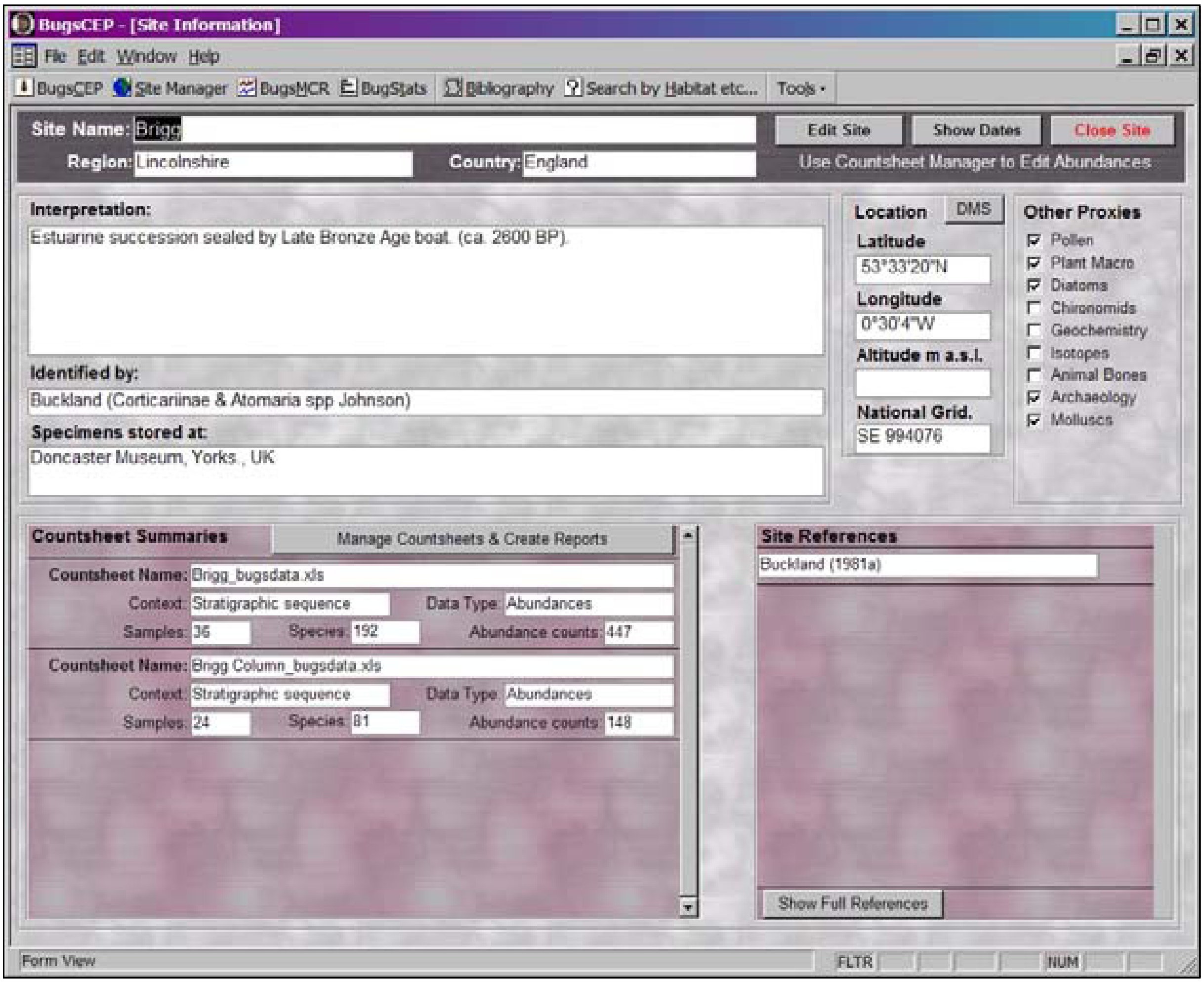


Figure 3.23. Screenshot: Site Information Screen showing details for Brigg (Buckland, 1981).

3.4.2.3 Managing countsheets and entering abundance data

The **Countsheet Manager** (Figure 3.24), accessible from the Site Manager or Site Information screens, allows sets of related samples to be compiled, stored and viewed within a site. A site may have any number of countsheets, which can be thought of as spreadsheet-like collections of species names, samples and abundance values (i.e. the abundance of species in a group of samples). If no countsheet exists for a site then the user is given the option of either creating or importing one (see next section). New countsheets must be given a name, assigned 'context' and 'data type' metadata, and then saved. Only then can samples, taxa and abundance data be added.

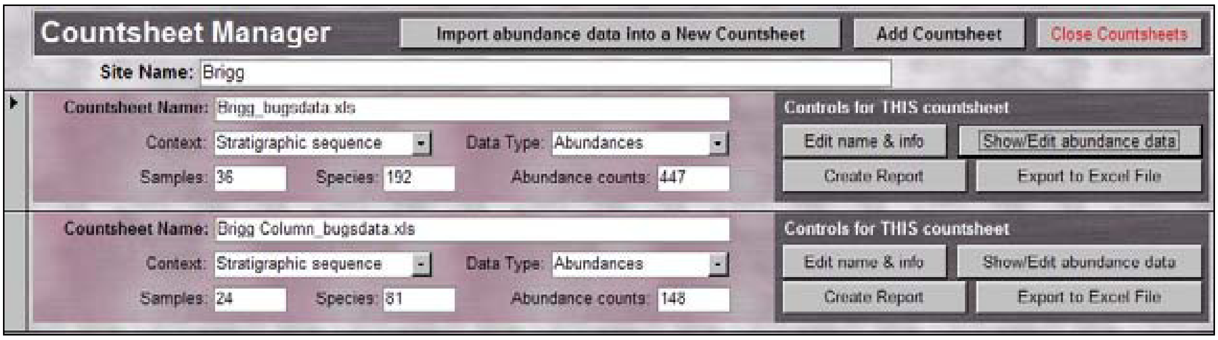


Figure 3.24. Screenshot: The Countsheet Manager showing the two countsheets stored for Brigg.

Countsheets can be examined, created, imported, exported or reported from the Countsheet Manager, and metadata describing the context and type of data stored can be edited. The Countsheet Name identifies a countsheet, although there is no demand on uniqueness for these labels. They can be named as seen fit by the user, and should be chosen so as to help others identify the data stored within them. The majority of countsheets included in BugsCEP on installation are named after the MS Excel files in which they were stored in earlier versions of Bugs. For each countsheet, three summary counts are provided: number of samples, species and abundance counts. Countsheet data can be exported as an MS Excel file by pressing the [Export to Excel File] button, but perhaps the most powerful function is accessed by pressing the [Create Report] button. Site reporting allows the species biology and distribution data, along with site summary and all relevant references, for all the species represented in a countsheet to be printed or exported. The available reports are described in section 3.5.3.

The **Countsheet Explorer** (Figure 3.25) is activated by clicking the [Show/Edit abundance data] button on the countsheet record. This is a spreadsheet-like editing area for viewing or changing species abundance data, where sample names form the column headers, apart from the first column which contains taxa names[xxvii]. Numerical abundance data (or a '1' indicating presence only) can be entered directly into the data cells.

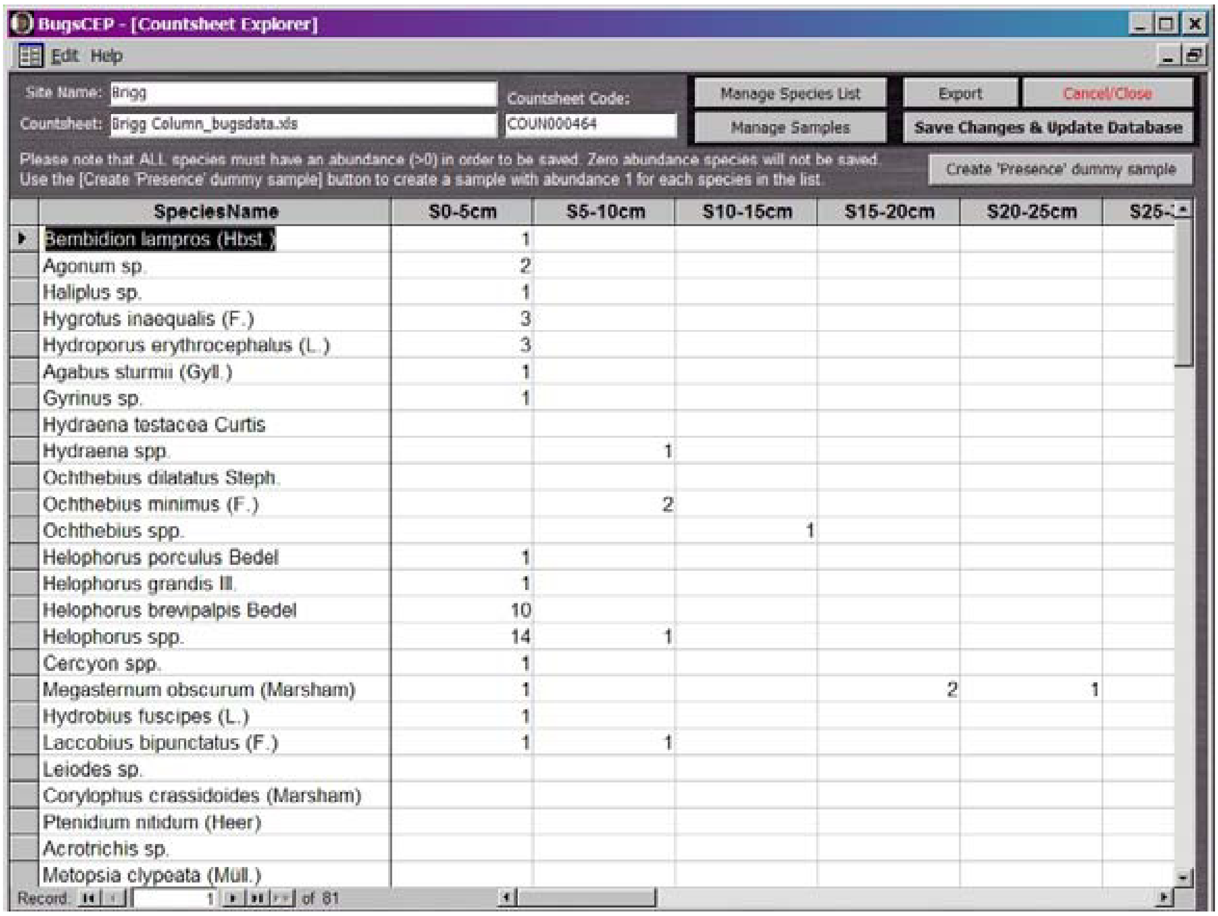


Figure 3.25. Screenshot: The Countsheet Explorer spreadsheet display, showing part of the abundance data for the countsheet 'Brigg Column_bugsdata.xls'. Taxa names fill the first column and sample names form the remaining column headers.

Taxa are added to or deleted from the list by clicking the [Manage Species List] button, which opens the **Species List Editor**, shown in Figure 3.26. This form consists of a taxonomic selection system

[xxvii] Note that the standard BugsCEP menu and toolbar are not available during a countsheet editing session.

identical to that on the navigation panel of the main screen, which can be used to locate the required taxon. Double clicking the taxon name, or pressing the [Add to Species List] inserts the selected taxon into the current species list. Species are inserted in their correct taxonomic order, and attempts to inset duplicates are ignored. A previously entered taxon can be deleted by selecting it in the current list, and clicking the [Delete Selected] button – but note that all abundance data for that taxon will also be deleted. Alternatively, a taxon in the current list can be replaced by selecting it, selecting the new replacement taxon in the navigation box, and then clicking the [Replace Selected] button. For example, in Figure 3.26 clicking [Replace Selected] would replace *Agonum* sp. with *Agonum thoreyi* Dej. This is essential where re-examination of a specimen has lead to increased taxonomic precision. Clicking [Apply Changes] saves any alterations, closes the editor, and displays an updated species list in the Countsheet Explorer.

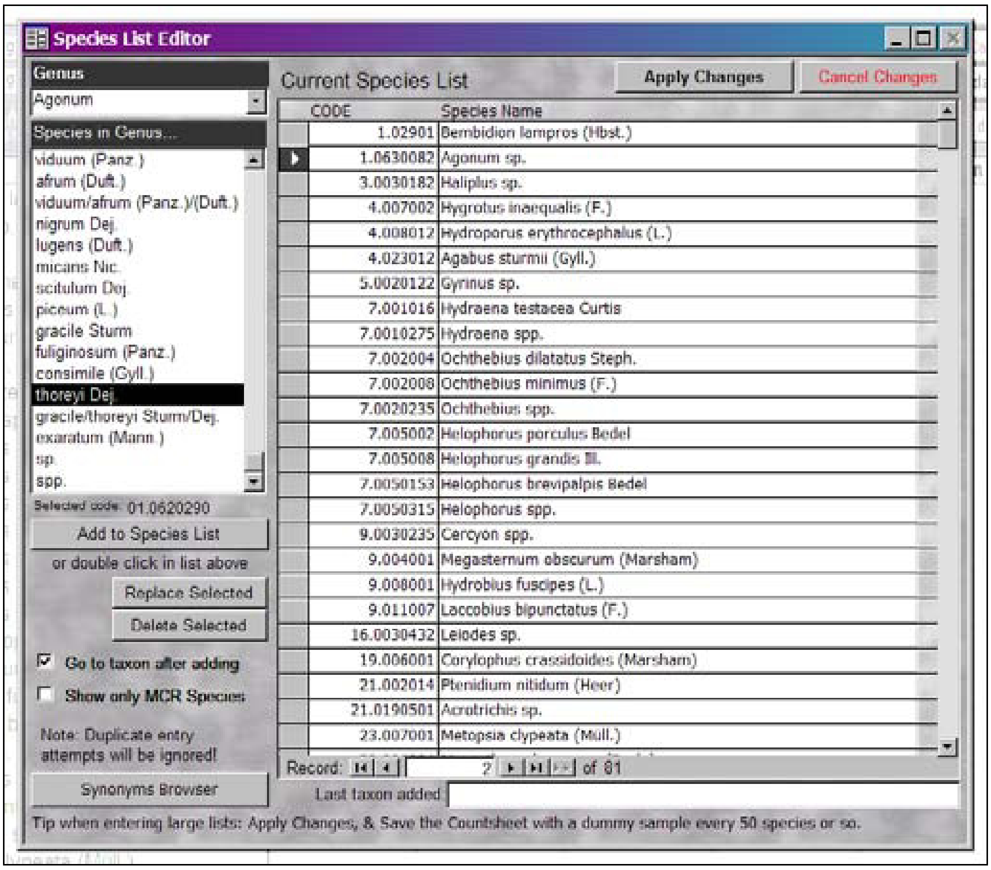


Figure 3.26. Screenshot: Species List Editor.

Samples may be viewed, added, deleted, renamed or given metadata by clicking the [Manage Samples] button and opening the **Samples Manager** (Figure 3.27). Sample depths and coordinates may be assigned to samples, and will be available in future versions for sorting samples in outputs.

The Countsheet Explorer includes a number of validation features, to prevent the input of non-numerical data, and the storage of empty species or sample rows, warning users when such attempts are made. It also includes the facility to create a 'dummy' sample with an abundance=1 for every species that has been entered. This feature, activated by clicking the [Create 'Presence' dummy sample] button, is useful when creating presence only species lists, eliminating the need to laboriously enter '1' in every cell. It is also useful when a long species list has been entered, but abundance data entry must be temporarily aborted. The Countsheet Manager will not allow a countsheet to be saved if there are species without any abundance values, and the creation of a dummy sample, which can be deleted when the full dataset has been entered, allows the incomplete countsheet to be saved. The

[Export] button allows the user to export the current countsheet as an MS Excel file, for use in other software or import into another copy of BugsCEP.

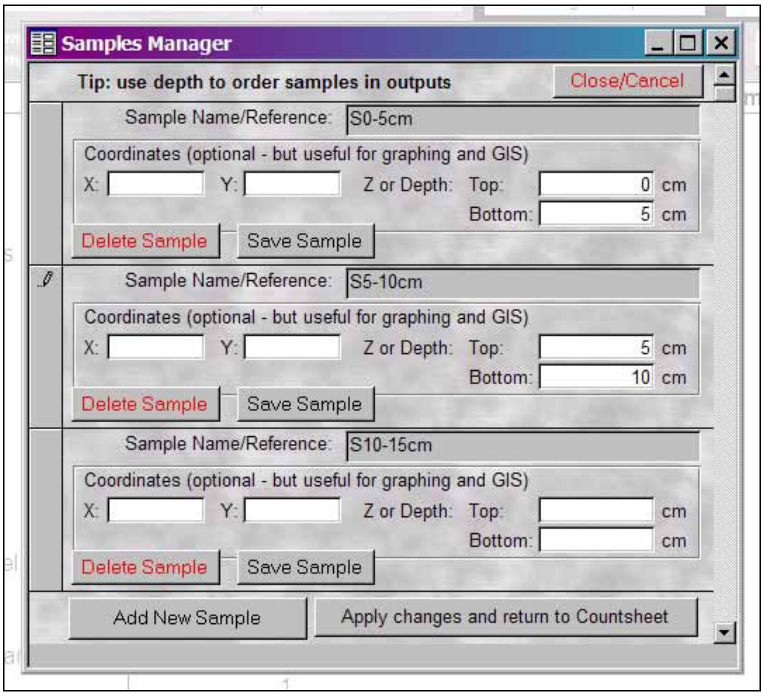


Figure 3.27. Screenshot: Samples Manager showing the first three samples of the stratigraphic sequence from the countsheet 'Brigg Column_bugsdata.xls'. Sample depths have been entered for the first two samples.

3.4.2.4 Importing MS Excel spreadsheets

Countsheets created in other software may be imported into BugsCEP under the following conditions[xxviii]:

1. The data are in an MS Excel file.
2. The data are in the BugsCEP format, or an importable format where the first column (A) contains Genus and species (and optionally authority), the first row (1) contains sample names, and the data cells only contain integer numbers (as in Figure 3.28).
3. A site has been created to store the countsheet.

Once a site has been created, click the [Manage Countsheets & Create Reports] button. As there is no data currently stored for the site the user will be given the option of creating an empty countsheet or importing data from an MS Excel file. If the user chooses to import a countsheet, then the **BugsCEP Import Wizard** will open, through which seven steps guide the import of external data. The seven steps, each of which incorporate validation checks on the data, must be followed in order, and give the user the chance to add additional metadata describing the samples and countsheet. Step one involves clicking the [Select countsheet for import] button, and then locating the file on the computer. If the data file is in the BugsCEP format, i.e. was created by BugsCEP, then import is quite straight forward, and Figure 3.29 shows the successful completion of the first two steps of selecting the file for import, and extracting the sample names. Step three only becomes available at this point.

[xxviii] There are also a number of additional important considerations in terms of the format of the MS Excel file, which are explained on the website http://www.bugscep.com/help/help_enteringdata.html#importnotes and on the equivalent page of the built in help files.

Step three opens the 'Define Samples...' dialog box in which users may add extra sample metadata, such as depth (top and bottom) and coordinates. The sample names cannot be changed during import, as this would make it impossible to match them with the source data columns.

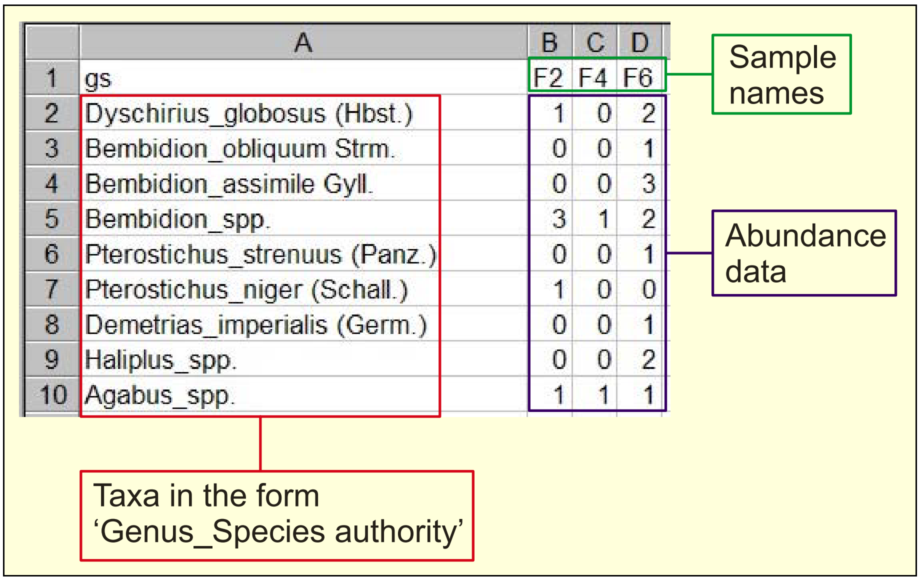


Figure 3.28. Spreadsheet format for import of MS Excel files into BugsCEP. Note that the divider between Genus and species could also be a space.

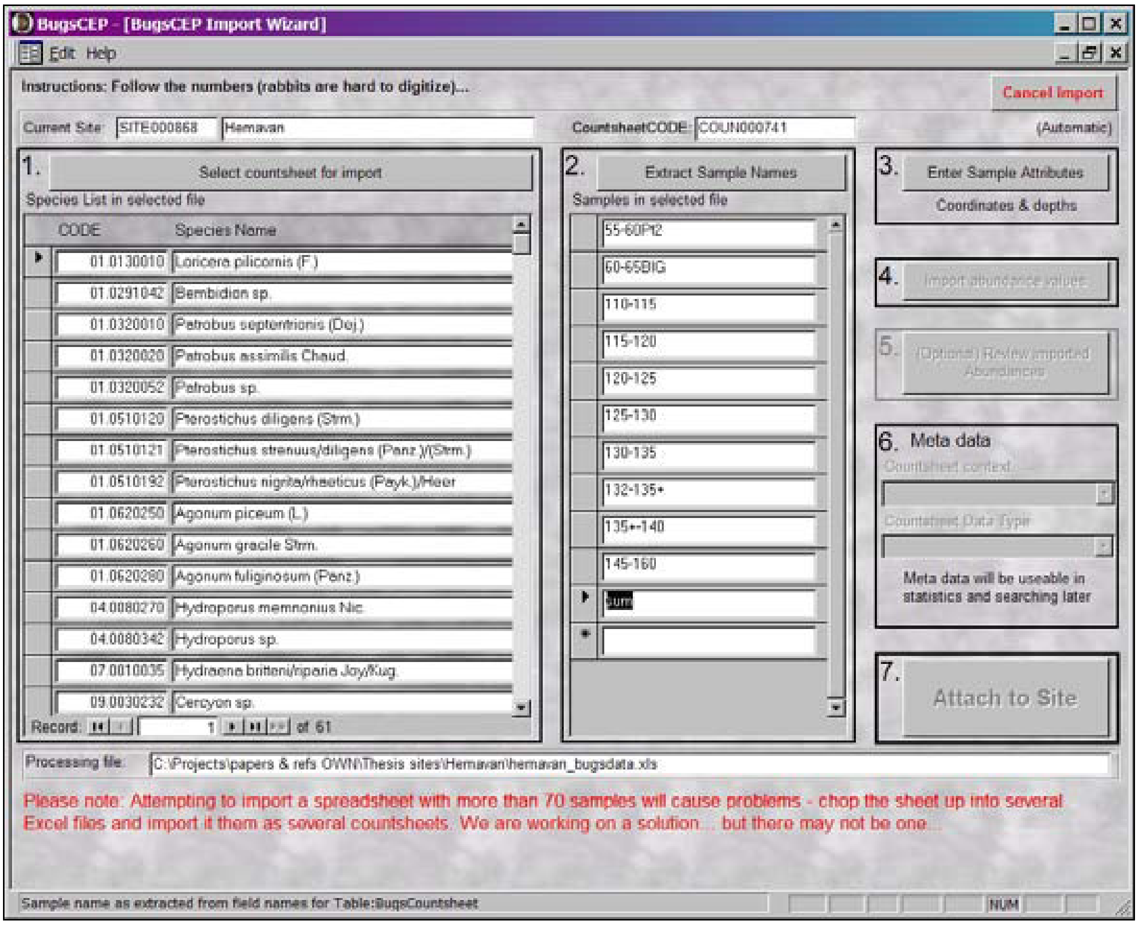


Figure 3.29. Screenshot: BugsCEP Import Wizard showing stages one and two completed in the import of the hemavan_bugsdata.xls MS Excel file.

Stage four imports the abundance data from the MS Excel file, and stage five allows the user to review these data. Stage six allows the countsheet to be assigned context related metadata. Stage seven consists of simply pressing the [Attach to Site] button, after which a 'completed' message is displayed, and the imported data are visible as a countsheet in the Countsheet Manager. From here it may be edited as any other countsheet.

If, during stage one, the data file is discovered not to be in the BugsCEP format, then the user will be informed and asked whether they wish to attempt to convert it. If [Yes] is chosen, then the **BugsCEP File Converter** will run, and attempt to match the species names in the import file with those stored in BugsCEP's master index (Figure 3.30). It does this by first trying to find the genus component of the text in each cell in the first column. When it has identified a known genus, it then moves on to the species component and tries to match that. If a match is made then the taxon is stored in the New Countsheet, otherwise the cell is left blank. If a match is made, but the full name, with authority, is not the same length in the import file and BugsCEP, then it is marked with an asterisk in the New Countsheet for correction.

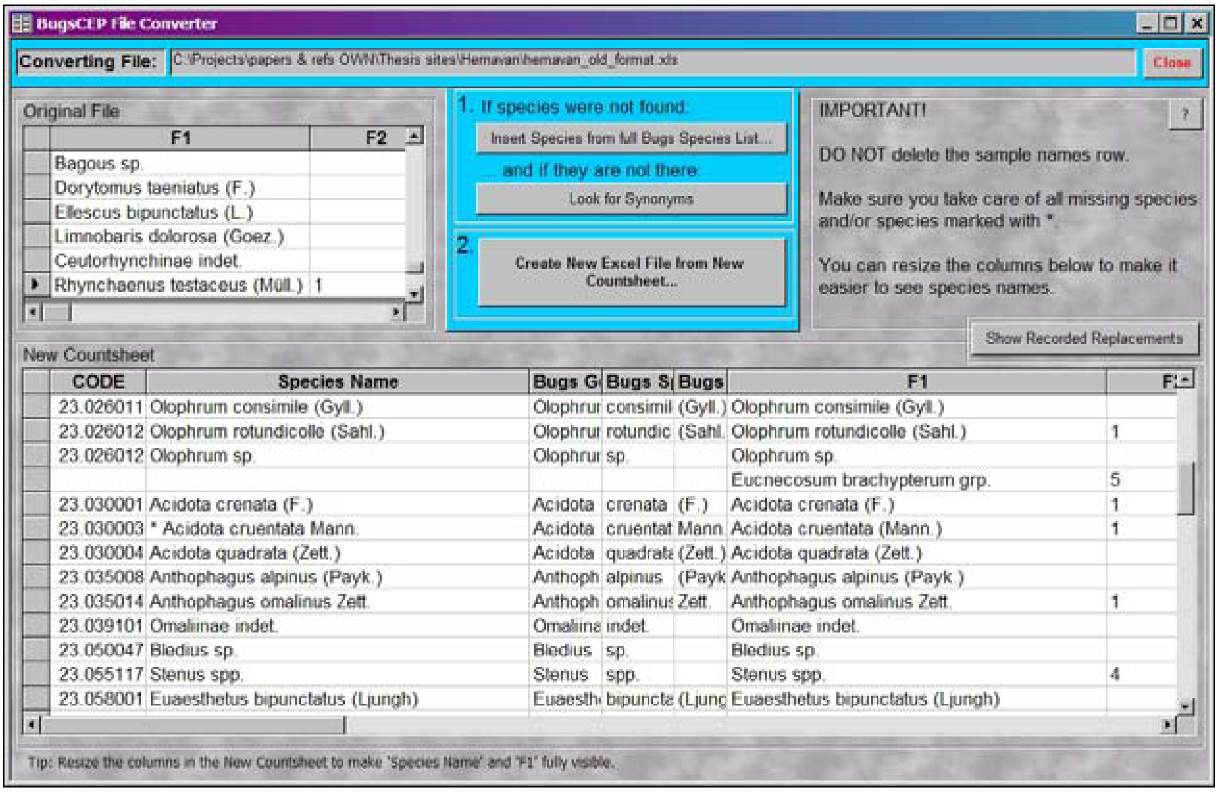


Figure 3.30. Screenshot: BugsCEP File Converter, having attempted to automatically convert the file hemavan_old_format.xls. Blank cells in the first few columns of the New Countsheet (bottom) indicate a failed match – the original name being shown in column 'F1'. Names marked with an asterisk are uncertain matches. Column widths have been adjusted for clarity, by drag the right hand edge of the column headers.

When the automatic conversion routine is complete, a message is displayed explaining the degree of its success. In the example here (Figure 3.31), eight taxa could not be matched, and three were matched, but the converter is unsure due to differences in length.

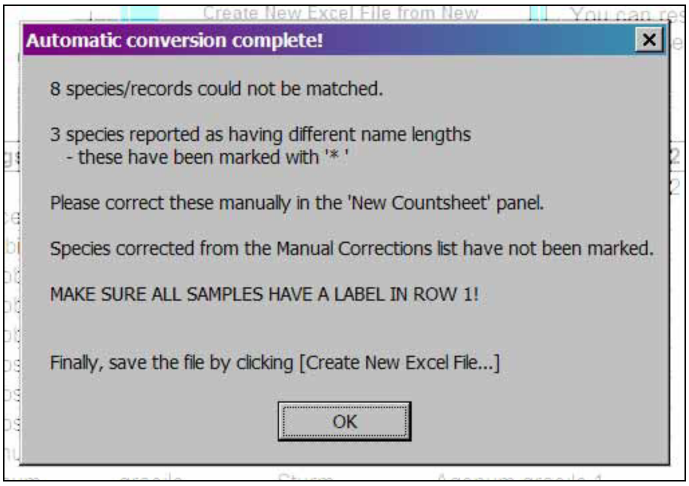


Figure 3.31. Screenshot: Message displayed on completion of the automatic phase of imported file conversion. All problems indicated here must be corrected manually in the File Converter before import can continue.

The 'New Countsheet' (bottom of Figure 3.30) is used to check asterisked names, and to insert the correct taxa where no match could be made. The latter is undertaken by selecting the relevant record and searching either the BugsCEP species index or the synonyms list and inserting the correct taxon. These replacements can be stored so that the File Converter knows what to use if it encounters the same problem during a later import[xxix]. Taxa should never be typed in manually, as this risks the insertion of an incorrect taxonomic code, and the abundance data for that taxon will be lost as the file is processed. The conversion routine is not perfect, and will miss what may seem obvious on many occasions, so patient diligence is required for its use.

3.4.2.5 The Date Explorer

The **Date Explorer**, which can be accessed from both the Site Manager and Site Information screens, is divided into two areas, the samples in the current countsheet being shown in a list on the left, and the dates for the selected sample to the right. If a site has more than one countsheet, they can be switched between using the drop down [Select Countsheet] box above the list of samples. It is also possible to toggle between the display of all samples, and only samples with dates using the [Show only samples with dates/Show all samples] button below the samples list. Tab controls are used to switch between the display of the three different date types. Figure 3.32 shows the two calendar dates stored for sample S783 at the Norse Greenlandic site of Sandnes (Sadler, 1987), indicating that the fauna in the sample most likely represent an environment present between 1000 and 1350 AD, according to historical records.

[xxix] Full instructions are to be found in the help files or on the website at http://www.bugscep.com/help/help_enteringdata.html#import

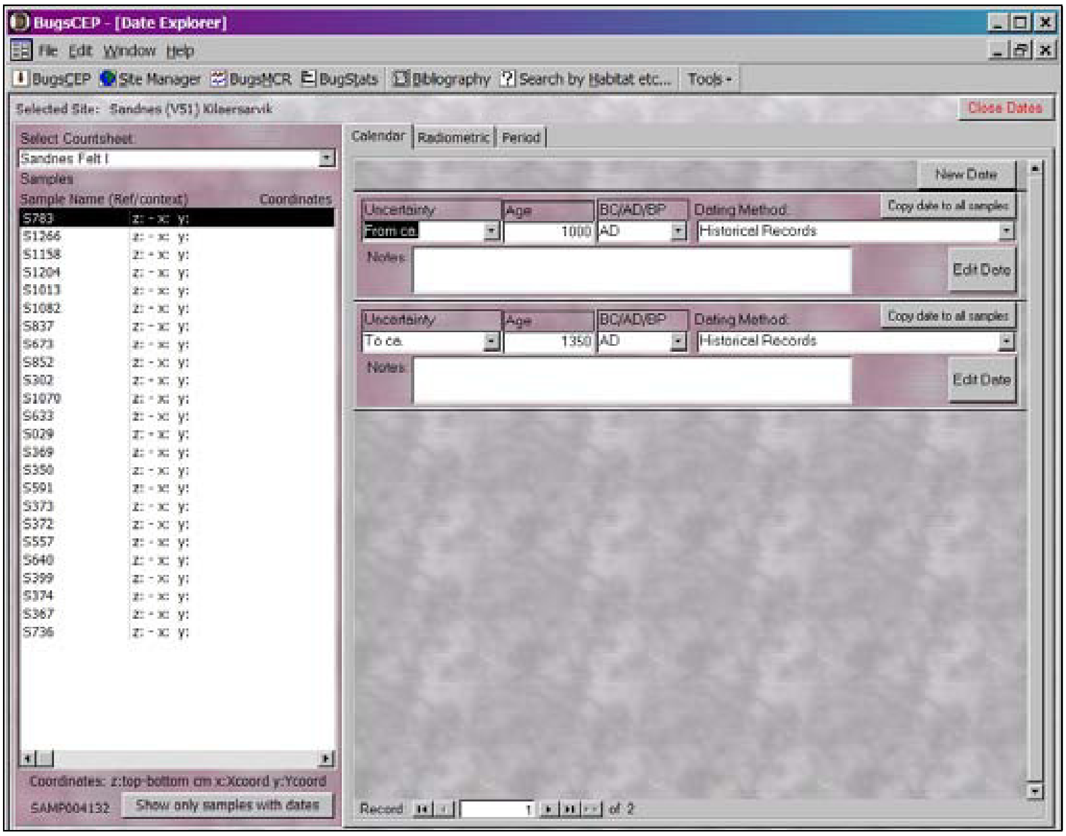


Figure 3.32. Screenshot: Date Explorer showing calendar dates for first sample in the countsheet Sandnes Felt I, Greenland.

The Lateglacial site at Messingham, Lincolnshire in the UK (Buckland, 1982), has two radiocarbon dated samples, the details of one of which are shown in Figure 3.33. A variety of dating methods are provided for, and 'other' categories are provided to allow for the storage of dates from uncommon methods. The facility to record a degree of uncertainty associated with any date is provided for all dating types, by way of drop down boxes with a limited set of choices for calendar and period dates, and as a free text box for radiometric dates. Any date may also have notes attached to it to explain further or describe the age.

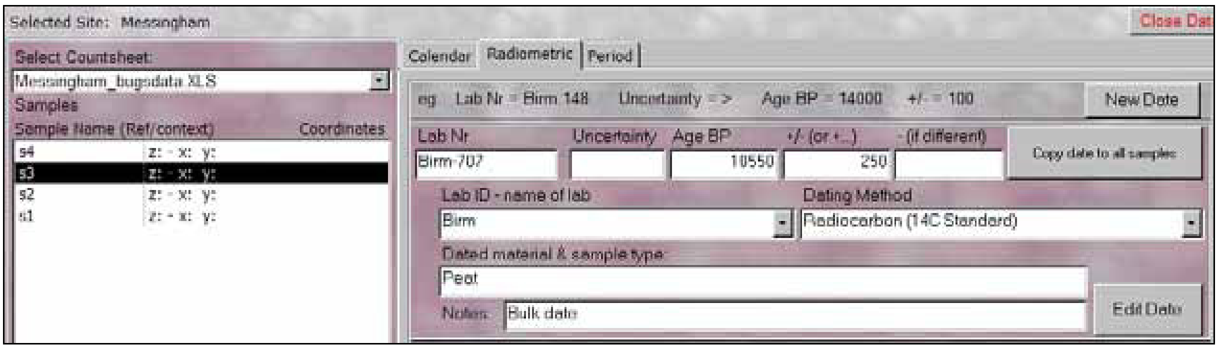


Figure 3.33. Screenshot: Date Explorer showing a radiocarbon dated sample from Messingham, UK.

3.4.3 Environmental reconstruction - BugStats

BugStats is an easy to use system for producing summaries and diagrams of the environmental implications of, or habitats represented by, the beetle taxa found at sites. It provides a number of tools and functions to assist in palaeoenvironmental reconstruction, and can help researchers to objectively compare faunas between sites and samples. This section explains the use of BugStats, whereas the mathematics and implications of the methods used are described in detail in Chapter 4. The implications of the various calculation options are also discussed in some of the case studies in Chapter 6.

BugStats is activated by clicking the [BugStats] component button under the main menu. Sites are selected through the [Site] drop down box towards the top left of the screen. The number of countsheets associated with the selected site will then be displayed just below the name, and the [Select Countsheet] box can be used to select any of these. Note that only sites with abundance or presence/absence data are available for selection. Figure 3.34 shows the BugStats main screen with the single countsheet for the site Stóraborg selected, a complex site, which has previously been treated to a degree of numerical analysis (Perry *et al.,*1985).

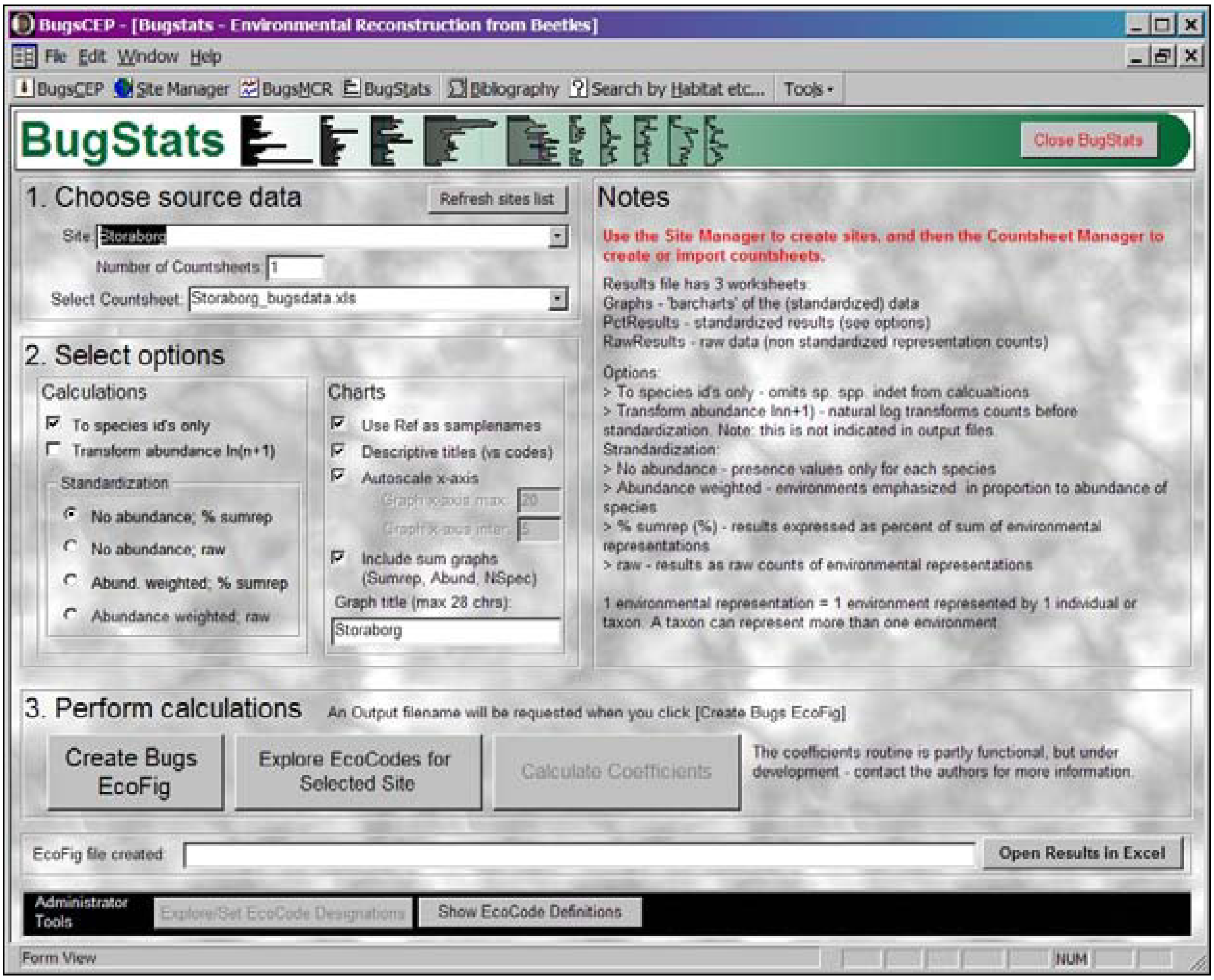


Figure 3.34. Screenshot: BugStats main screen with the site Stóraborg and countsheet 'Storaborg_bugsdata.xls' selected. The 'Graph title' has been entered manually.

A number of calculation and output options are available, represented by check boxes in the 'Select options' part of the screen, and as the implications of a number of these are important for interpreting the results, they are briefly explained on the right hand half of the screen. A key concept is that of *environmental representation counts*, which are the individual indications of a habitat class given by either an individual or a taxon, depending on the options chosen. These are the raw values upon which any further calculations are based.

3.4.3.1 Calculations options

To species id's only

When this option is checked, BugStats will ignore all sp., spp. and indet. taxa, that is to say those not identified to species level. Generic identifications are common in palaeoentomology due to the variability of preservation and identifiably of beetle fragments. These taxa have been assigned to the full range of habitats represented by their component species, and as such their inclusion in calculations creates a much broader, diluted picture of the environment represented by a sample. This is not always a bad thing, and users should use their discretion and knowledge of entomology when deciding whether to use this function or not. The use of this option probably produces the most reliable output, in that only habitats specifically represented by species are displayed, whereas switching it off may increase the probability that the environments reconstructed include those that actually existed at the time of deposition. The latter is especially true for poorly resolved or small faunas, where the removal of generic taxa may result in the loss of a significant proportion of the information available in the faunas.

Transform abundance ln(n+1)

This facility applies a natural logarithm transformation[xxx] to the raw reconstruction counts before standardization (see below). Some authors believe this compensates for inter-sample variation in a number of ways, including log/normal rank abundance distributions. This may be so with certain faunas, but ideally rank abundance curves should always be produced before assuming such a translation is necessary, and the facility to do this may be built into BugsCEP at a later date. The current recommendation is *not* to transform, as this will smooth the resulting diagram, reducing peaks and raising troughs, and thus may make it difficult to observe relative changes in habitat. See Chapter 4 for more details.

Standardization

By using raw values the results reflect the absolute numbers of species and/or individuals in each sample, and samples with higher numbers of species or individuals are likely to display stronger habitat indications. Although useful, this may not be ideal where abundances vary considerably between samples, or where more than one site is to be compared. Results may be standardized so as to allow for easier inter-sample and inter-site comparisons, the method chosen here to recalculate sample values as proportions of their respective sample totals. By doing so the results will always be comparable, irrespective of site or sample. The process is explained in more detail in Chapter 4, but the four options available are outlined in Table 3.12 in the order that they are shown in the BugStats interface (see Figure 3.34).

[xxx] The +1 prevents an error from occurring where abundance=0, and is an accepted workaround (e.g. Jongman *et al.* 1995).

Table 3.12. Outline of standardization options in BugStats.

Option	Summary
No abundance, % sumrep	Taxa occurrence only (presence/absence). Environmental representation sum for habitat class expressed as percent of sum of environmental representations for sample (sumrep).
No abundance, raw	Taxa occurrence only (presence/absence). Environmental representation counts for habitat class only.
Abund. weighted, % sumrep	Taxa environmental representation multiplied by its abundance in sample. Environmental representation sum for habitat class expressed as percent of sum of Environmental representations for sample (sumrep).
Abundance weighted, raw	Taxa environmental representation multiplied by its abundance in sample. Environmental representation counts for habitat class only.

3.4.3.2 Chart options

These settings are more or less self explanatory, and so will only be outlined here. Users should experiment in order to find the settings most appropriate for their needs.

- *Use Ref as sample names*
  Use the sample Reference/Context field to label samples (recommended) rather than the internal database sample code.
- *Descriptive titles*
  Use real words for habitat titles, rather than Bugs EcoCodes (see Chapter 4).
- *Autoscale x-axis*
  BugStats will scale diagrams to fit all data. Uncheck this to specify a maximum and interval yourself, which can be useful if either the autoscale function proves inadequate, or when wishing to magnify the graphs.
- *Include sum graphs*
  Output summary graphs showing:
  Sumrep - sample environmental representation sums
  Abundance - sample abundance sums (number of individuals per sample)
  NSpec - sample taxa sums (number of taxa per sample)
- *Graph title*
  Use this box to give the graph a meaningful title (BugStats will append the chosen standardization option to this name on output).

3.4.3.3 Creating EcoCode outputs

Two forms of output can be created for habitat reconstructions: a standard EcoFig; and a sample by sample EcoCode breakdown.

Standard EcoFig

Clicking the [Create Bugs EcoFig] button will open the standard 'Save As' dialog box, with which the output file can be given a name and saved to a particular location. After [Save] is clicked BugStats will begin creating the output file and running calculations. This will take only a few seconds on modern computers[xxxi], during which time the screen will flash and the diagram will be constructed onscreen. *Do not interrupt this process in any way!* Doing so may have unexpected consequences and could lead to the computer crashing and loss of data.

BugStats will indicate when calculations are complete, and display the path and filename of the output file in the [EcoFig file created:] box at the bottom of the screen. Clicking the [Open Results in Excel] will open the file and display the first few habitat diagrams, an example is shown in Figure 3.35. The rest of the diagram can be examined by scrolling and zooming out in MS Excel, copying all the diagram components into a graphics package such as CorelDraw, or by printing it. The output file is automatically set up to print on two landscape orientated A4 pages.

As can be seen in Figure 3.35 the habitat groups (or environments) are listed along the top of the diagram, samples are listed in a column to the left, and the individual chart scales are evident across the bottom of each diagram component. The actual meaning of the bars within the charts will depend on the options used, and Chapter 4, along with the case studies in Chapter 6 should be consulted for more details.

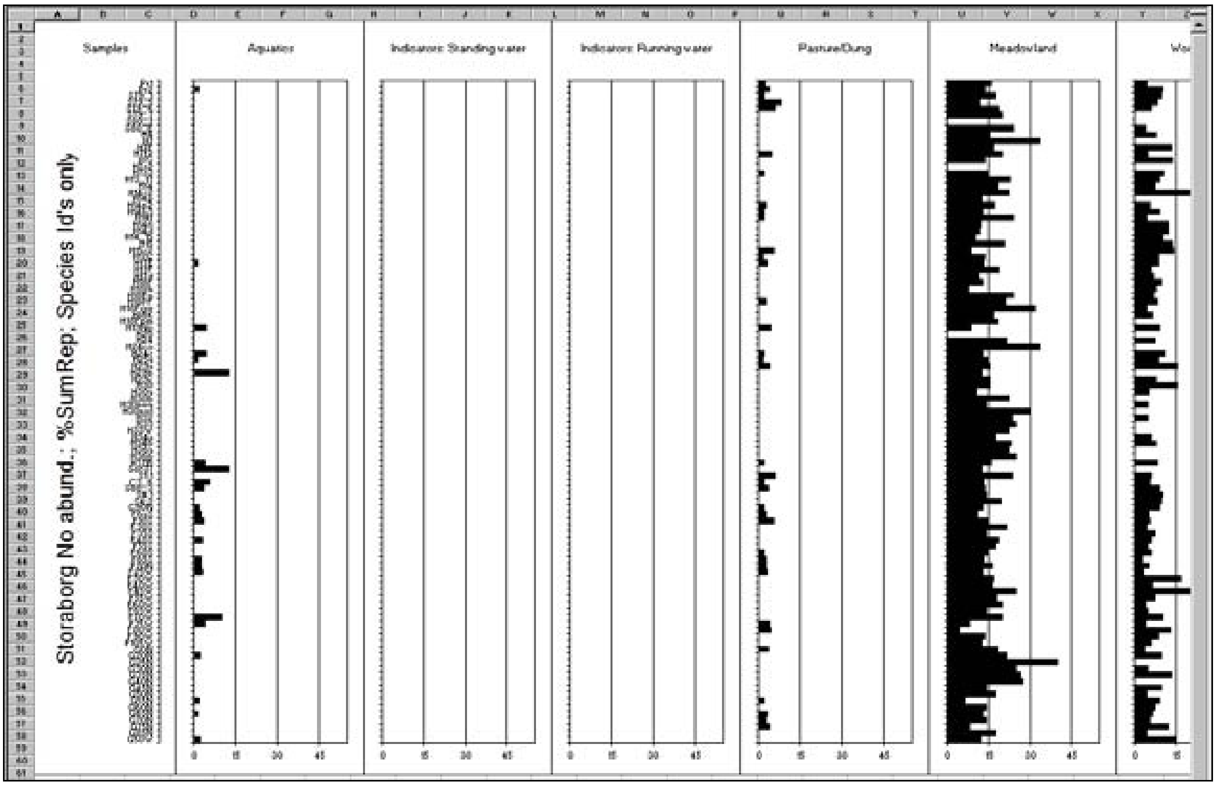


Figure 3.35. Screenshot: the first few charts of a Bugs EcoFig for Stóraborg, as opened in MS Excel for the first time.

Sample by sample EcoCode breakdown

In addition to producing the EcoFig, it may be useful to see which taxa lie behind the environmental representation counts. Clicking the [Explore EcoCodes for Selected Site] will create a report listing the taxa, their abundances and Bugs EcoCode classifications in each sample, as shown in Figure 3.36. This can help the user to understand why a particular sample indicates a particular habitat, and may also help when deciding which calculation options to choose.

xxxi Faster than 2 GHz with 512 MB RAM.

This report can be printed, or exported to MS Word or Excel by clicking the appropriate button on the toolbar. The report is created as a temporary file in the currently active directory, and should be saved from MS Word or Excel to a new location.

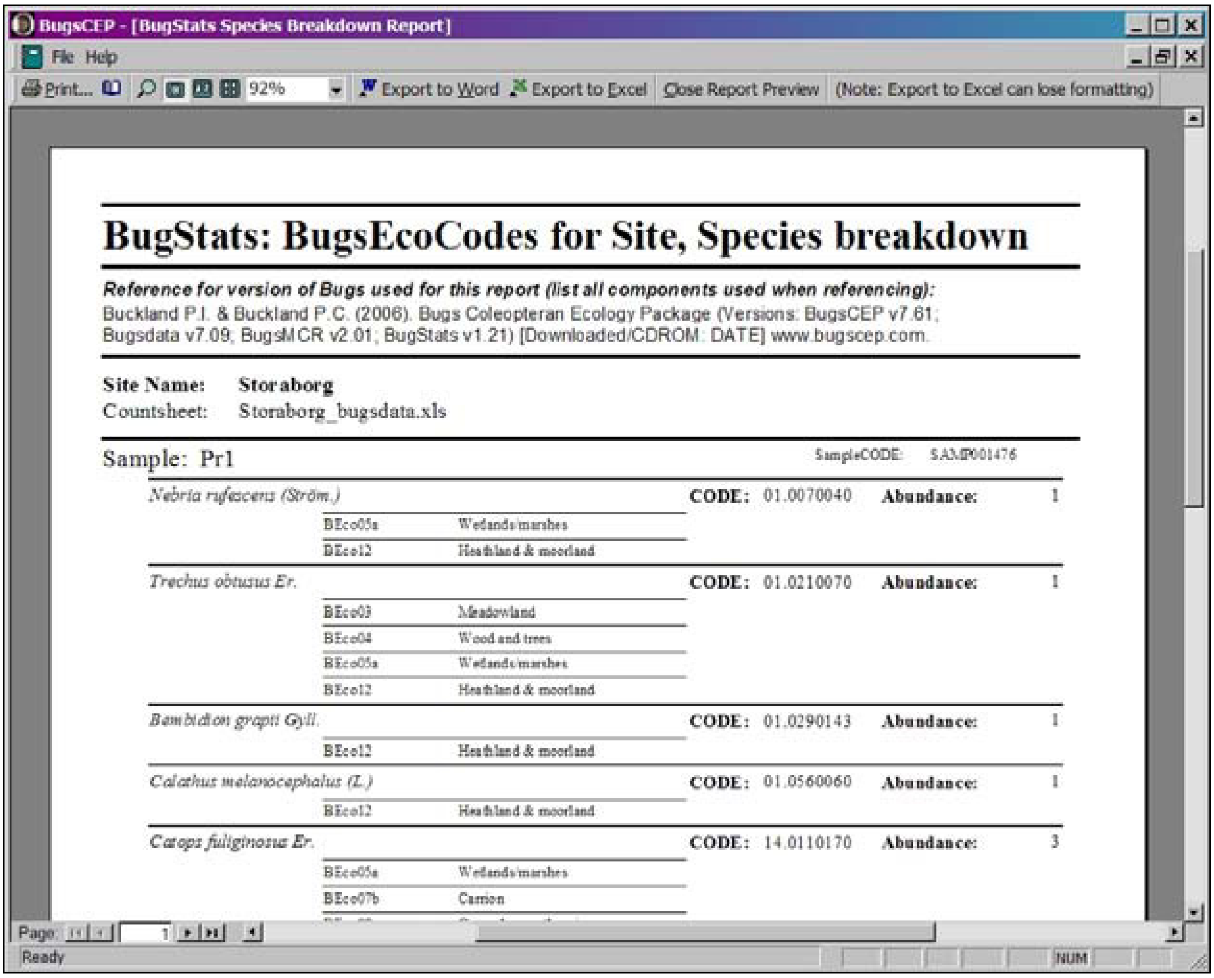


Figure 3.36. Screenshot: Start of report showing a sample by sample breakdown of the taxa found at Stóraborg, including their abundances and Bugs EcoCode designations.

3.4.3.4 Seeing the Bugs EcoCode definitions

Clicking the [Show EcoCode Definitions] button will open a form in which the formal definitions of all 22 Bugs EcoCodes can be seen. These are described in detail in Chapter 4.

3.4.3.5 Calculating coefficients of similarity

Correlation coefficients can be used to compare the similarity/dissimilarity of the species composition of samples. This is especially useful when trying to identify samples that represent similar environments in Quaternary geology and environmental archaeology, as shown in a number of the case studies in Chapter 6. BugsCEP currently only supports the modified Sørensen's coefficient of similarity (Southwood, 1978), which is the inverse (1-B) of the Bray-Curtis coefficient of dissimilarity (Krebs, 1999). The results of coefficient calculations can be used as the basis for cluster analysis in other software.

Clicking the [Calculate Coefficients] button will activate the coefficients interface (Figure 3.37), which is currently password protected and awaiting comprehensive stability testing[xxxii]. Site and countsheet selection is by drop down boxes as usual, and the option to (natural) log transform the raw count data is available through a check box. Log transformation has a tendency to increase the observed degree of similarity when using the modified Sørensen's coefficient, which is sensitive to sample abundances (see Chapter 4 for the formula).

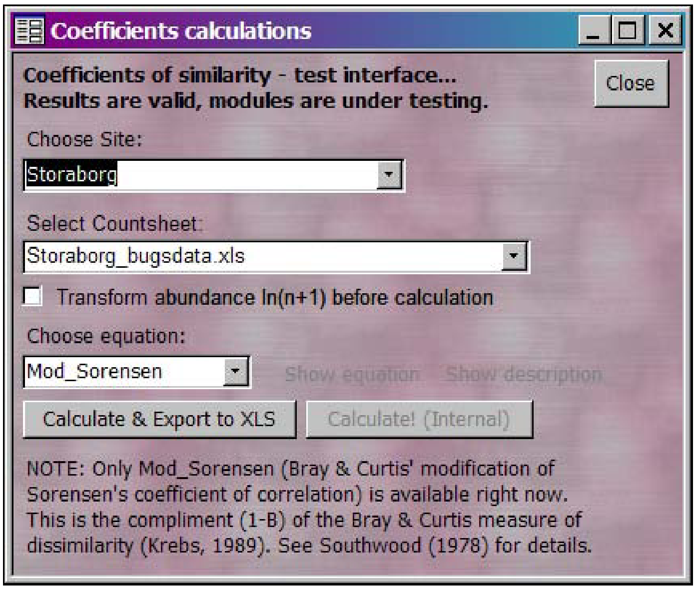


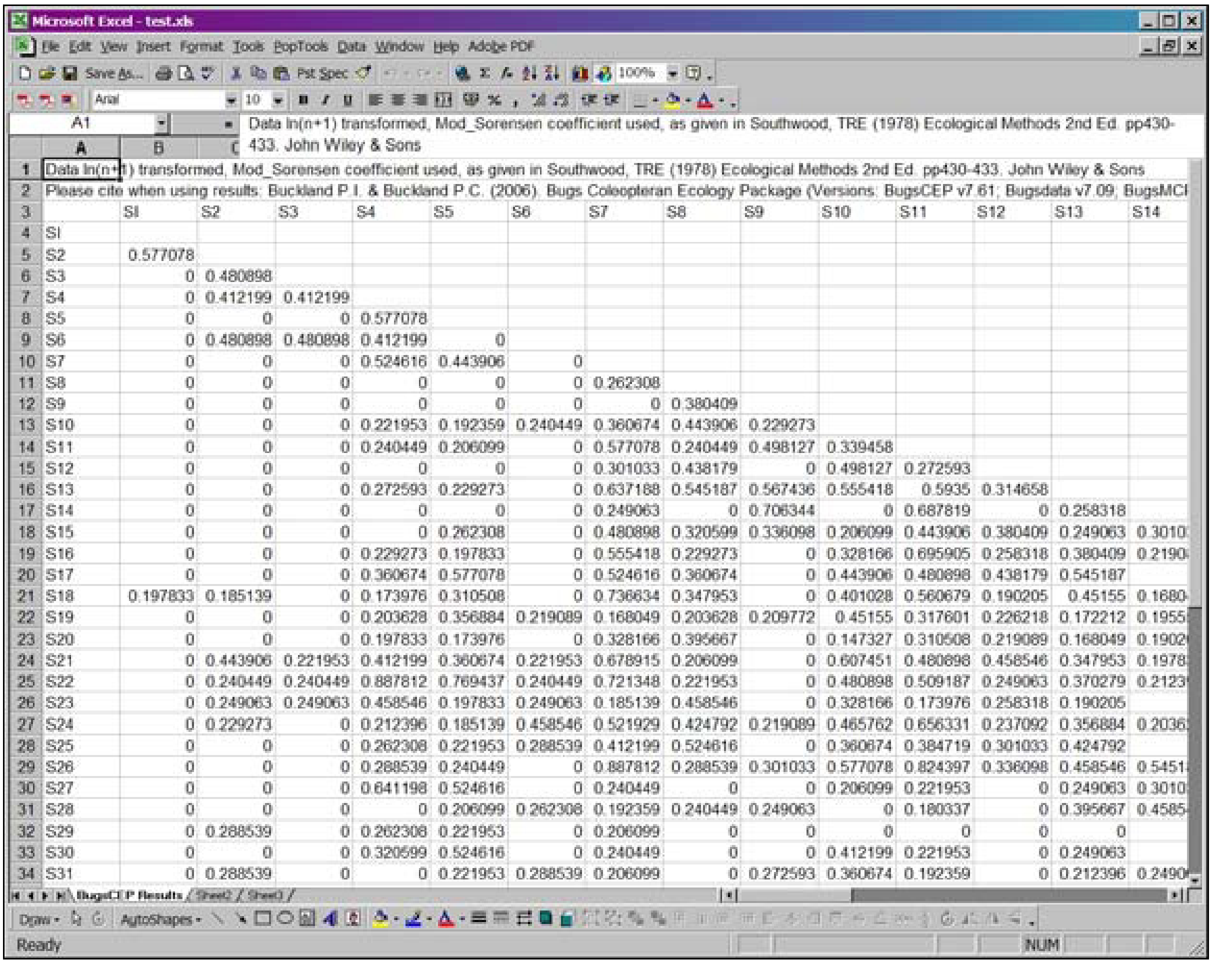
Figure 3.37. Screenshot: The coefficients calculation interface, with Stóraborg selected.

Clicking the [Calculate & Export to XLS] button runs the calculations, a 'busy' message being displayed until they are complete, after which a file name is requested through the standard 'Save as' dialog box. The screen will flash for a while as the results are exported to MS Excel, until a dialog box is displayed asking whether the user wishes to open the resultant file in MS Excel.

The results file is not formatted in any way, and users may wish to reformat the results cells as numbers with two decimal places to improve clarity. Underneath the two explanation and reference rows, the results are presented as a trellis (or matrix) diagram, sample names are listed in the first row and column, with the values in between indicating the correlation between the sample indicated by the row and column headers of each cell (Figure 3.38). Coefficient values are between zero and one, with zero indicating no similarity whatsoever between the two samples, and one indicating identical species lists in the two samples. The PopTools (Hood, 2003) 'Colour scale' function is recommended for shading the results by value.

Users should always remember that correlation is no guarantee of cause and effect, and that two samples may have similar faunas but for different reasons. It is recommended that users read standard ecological methodology texts (e.g. Southwood & Henderson, 2000; Krebs, 1999) for tips on interpreting correlations.

[xxxii] The password may be obtained from the author, and the protection will be removed when testing is complete.



Data ln(n+1) transformed, Mod_Sorensen coefficient used, as given in Southwood, TRE (1978) Ecological Methods 2nd Ed. pp430-433. John Wiley & Sons

Please cite when using results: Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCI

	SI	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
SI														
S2	0.577078													
S3	0	0.480898												
S4	0	0.412199	0.412199											
S5	0	0	0	0.577078										
S6	0	0.480898	0.480898	0.412199	0									
S7	0	0	0	0.524616	0.443906	0								
S8	0	0	0	0	0	0	0.262308							
S9	0	0	0	0	0	0	0	0.380409						
S10	0	0	0	0.221953	0.192359	0.240449	0.360674	0.443906	0.229273					
S11	0	0	0	0.240449	0.206099	0	0.577078	0.240449	0.498127	0.339458				
S12	0	0	0	0	0	0	0.301033	0.438179	0	0.498127	0.272593			
S13	0	0	0	0.272593	0.229273	0	0.637188	0.545187	0.567436	0.555418	0.5935	0.314658		
S14	0	0	0	0	0	0	0.249063	0	0.706344	0	0.687819	0	0.258318	
S15	0	0	0	0	0.262308	0	0.480898	0.320599	0.336098	0.206099	0.443906	0.380409	0.249063	0.3010
S16	0	0	0	0.229273	0.197833	0	0.555418	0.229273	0	0.328166	0.695905	0.258318	0.380409	0.2190
S17	0	0	0	0.360674	0.577078	0	0.524616	0.360674	0	0.443906	0.480898	0.438179	0.545187	
S18	0.197833	0.185139	0	0.173976	0.310508	0	0.736634	0.347953	0	0.401028	0.560679	0.190205	0.45155	0.1680
S19	0	0	0	0.203628	0.356884	0.219089	0.168049	0.203628	0.209772	0.45155	0.317601	0.226218	0.172212	0.1955
S20	0	0	0	0.197833	0.173976	0	0.328166	0.395667	0	0.147327	0.310508	0.219089	0.168049	0.1902
S21	0	0.443906	0.221953	0.412199	0.360674	0.221953	0.678915	0.206099	0	0.607451	0.480898	0.458546	0.347953	0.1978
S22	0	0.240449	0.240449	0.887812	0.769437	0.240449	0.721348	0.221953	0	0.480898	0.509187	0.249063	0.370279	0.2123
S23	0	0.249063	0.249063	0.458546	0.197833	0.249063	0.185139	0.458546	0	0.328166	0.173976	0.258318	0.190205	
S24	0	0.229273	0	0.212396	0.185139	0.458546	0.521929	0.424792	0.219089	0.465762	0.656331	0.237092	0.356884	0.2036
S25	0	0	0	0.262308	0.221953	0.288539	0.412199	0.524616	0	0.360674	0.384719	0.301033	0.424792	
S26	0	0	0	0.288539	0.240449	0	0.887812	0.288539	0.301033	0.577078	0.824397	0.336098	0.458546	0.5451
S27	0	0	0	0.641198	0.524616	0	0.240449	0	0	0.206099	0.221953	0	0.249063	0.3010
S28	0	0	0	0	0.206099	0.262308	0.192359	0.240449	0.249063	0	0.180337	0	0.395667	0.4585
S29	0	0.288539	0	0.262308	0.221953	0	0.206099	0	0	0	0	0	0	
S30	0	0	0	0.320599	0.524616	0	0.240449	0	0	0.412199	0.221953	0	0.249063	
S31	0	0.288539	0	0	0.221953	0.288539	0.206099	0	0.272593	0.360674	0.192359	0	0.212396	0.2490

Figure 3.38. Screenshot: Coefficients output opened in MS Excel for first time.

3.4.4 Climate reconstruction - BugsMCR

The Mutual Climatic Range (MCR) method is a technique for deriving quantitative temperature reconstruction from fossil beetle remains which was developed during the 1980s (Atkinson *et al.*, 1986). It is a Modern Analogue Technique (MAT) in that it uses the modern thermal distribution of taxa as the basis for 'retrodicting' palaeotemperatures (see Birks, 1995). MCR uses the area of mutual overlap of the thermal envelopes of the species in a sample as the most probable thermal range, rather than looking for similar species assemblages from the present day, and thus circumvents the non-analogue community problems that assemblage based methods can have (Pross & Klotz, 2002). Although a number of authors have used the calibration/correction method of Atkinson *et al.* (1987), this facility is omitted here due to its unsound mathematical foundations (see Chapter 5).

BugsCEP includes the facility for running MCR calculations and outputting the results in table and graph form, along with the sample thermal envelopes (climate space maps), to MS Excel. BugsMCR uses the Birmingham 'beetle.dat' dataset (see section 2.8.5.1, and Chapter 5), and is described in detail in Chapter 5. The BugsMCR component is activated by clicking the [BugsMCR] button on the main toolbar, and all temperatures are given in degrees Celsius

BugsMCR is operated in a similar fashion to BugStats, in that the source data (site and countsheet) are selected, calculations and output options chosen, and then the calculations run. A site is selected using the drop down [Site:] box at the top left, the number of countsheets stored for the site being displayed underneath on selection (Figure 3.39). The [Select Countsheet:] drop down box can then be used to select between different countsheets if required. In a slight difference from BugStats, the selected countsheet must then be activated by pressing the [Activate this countsheet] button. This loads the full site species list into the 'Active Species List', and allows the user to see how many MCR species[xxxiii] are present by pressing the [Show only MCR species] button. In Figure 3.39 the species list for the Saint Bees Lateglacial site in Cumbria (Coope & Joachim, 1980) has been activated, and shows a total of 282 taxa. Reducing this to only MCR species reveals that only 114 of them have climate data at the current time. This is to be expected, as there are only 436 taxa in the MCR database, a full list of which, along with their RECON equivalents (see Chapter 5) can be seen by clicking the Show All MCR Species link in the Tools box. The currently active countsheet data, in their reduced MCR form if the [Show only MCR species] button has been pressed, can also be seen by clicking the Show Current Countsheet link just below this.

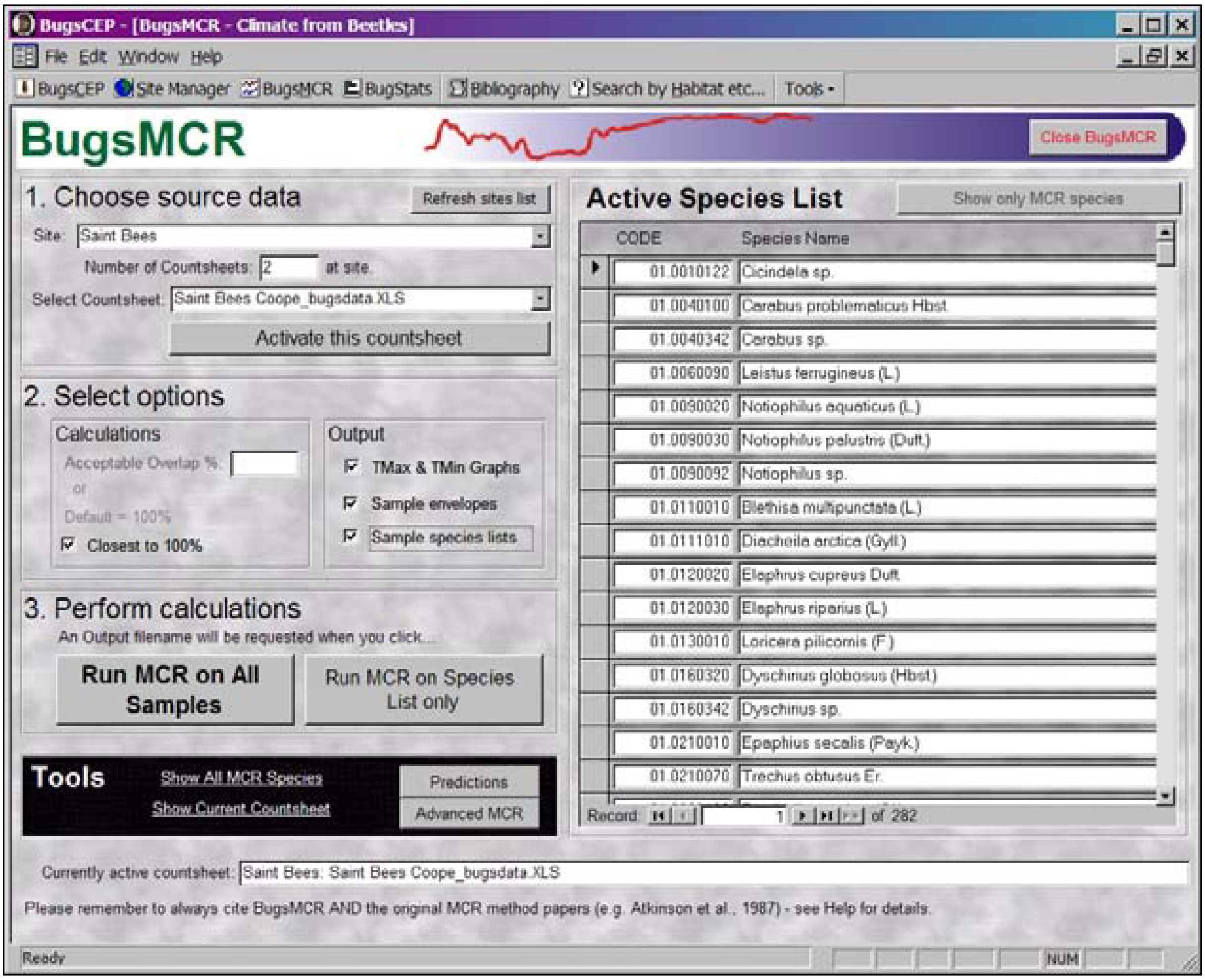


Figure 3.39. Screenshot: The BugsMCR interface with the Saint Bees site and 'Saint Bees Coope_bugsdata.XLS' countsheet activated.

Two calculation setting options are available in BugsMCR, the implications of which will be explained in Chapter 5, and only the usage is described here. If the [Closest to 100%] box is checked then BugsMCR will attempt to calculate a thermal reconstruction from the maximum extent of the area of highest overlap for each sample, allowing a greater scope for reconstruction than the traditional >90 % method. This does, however, come with the warning that reconstructions for the samples with less than 100 % overlaps are probably less reliable. Unchecking this box, and entering a number in the [Acceptable Overlap %:] box allows the user to specify the minimum level of overlap, in percent, on which a reconstruction should be undertaken. For the most reliable calculations this should be set to 100 %. Should the area of maximum overlap be non-contiguous, BugsMCR calculates using the

---

[xxxiii] Taxa that have thermal envelopes stored for them.

extreme limits of the combined areas. Note that the default setting for BugsMCR is 'closest to 100%', as this allows the user to examine the results from all viable samples, and decide themselves whether to recalculate with further restrictions. The user may also delete results from the output files manually, and the graphs will be updated automatically.

The output options, to the right of the calculation options (see Figure 3.39), control the creation of two possible results files. Checking the [TMax & TMin Graphs] option, which is selected by default, graphs the results in standard MCR diagram form, although at this point in time the X, or sample axis, can unfortunately not be scaled by time or depth (Figure 3.40). Only the numerical results are exported if this box is unchecked (Table 3.13, with an explanation of headers in Table 3.14).

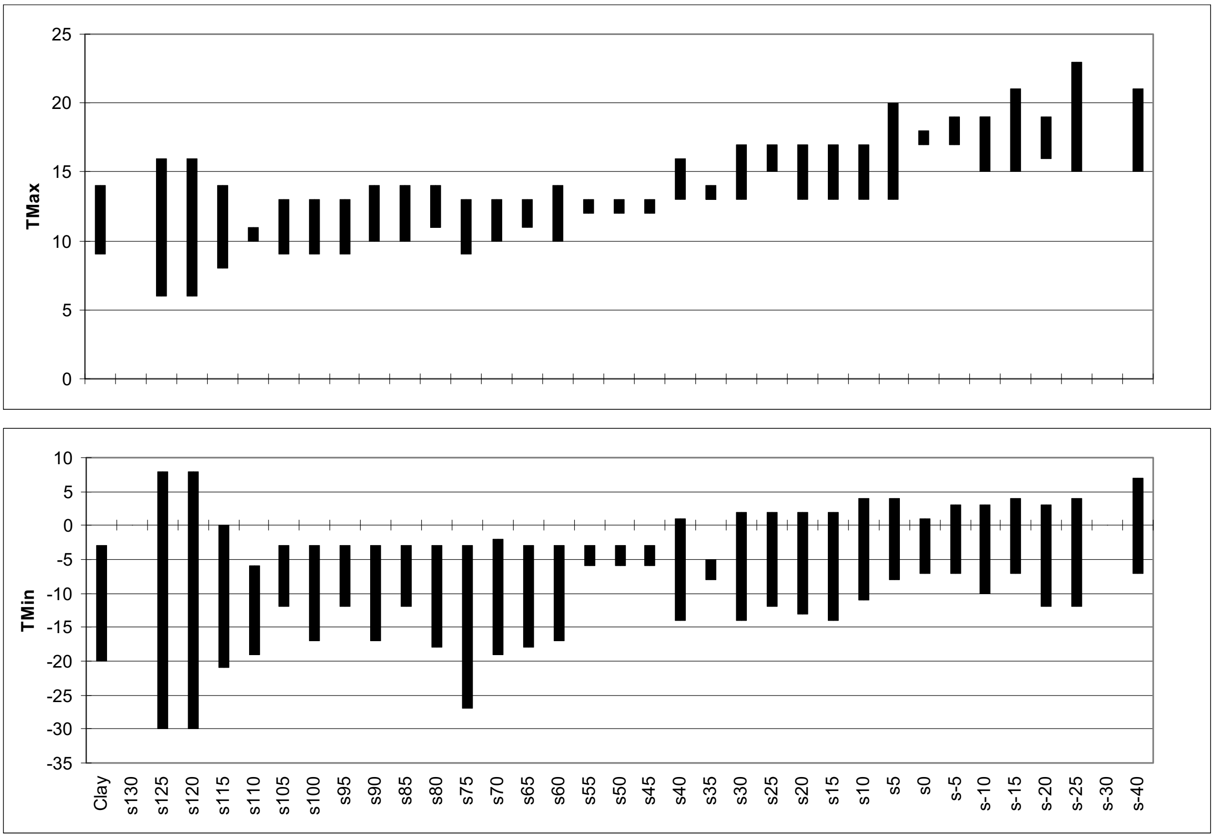


Figure 3.40. Thermal reconstruction from the Saint Bees site, using the 'Closest to 100%' overlap option. Note that the temperature axes may differ on output, and must be scaled manually. All results are in degrees Celcius, and the warming of both summer and winter temperatures through the sequence can clearly be seen from left to right.

Table 3.13. Extract of the numerical results output for the thermal reconstruction from the Saint Bees site, using the 'Closest to 100%' option. See Table 3.14 for an explanation of column headers.

Sample	TMaxLo	TMaxHi	TMinLo	TMinHi	TRangeLo	TRangeHi	NSPEC	Overlap
Clay	9	14	-20	-3	16	29	10	90
s130							0	
s125	6	16	-30	8	8	36	1	100
s120	6	16	-30	8	8	36	1	100
s115	8	14	-21	0	14	29	3	100
s110	10	11	-19	-6	17	29	11	100
s105	9	13	-12	-3	16	22	11	90.90909
s100	9	13	-17	-3	16	29	13	92.30769
s95	9	13	-12	-3	16	22	21	95.2381
s90	10	14	-17	-3	16	29	24	83.33334
s85	10	14	-12	-3	16	23	20	90
s80	11	14	-18	-3	16	29	18	88.88889
s75	9	13	-27	-3	16	36	15	100
s70	10	13	-19	-2	15	29	17	94.11765
s65	11	13	-18	-3	16	29	17	94.11765
s60	10	14	-17	-3	16	29	23	91.30434
s55	12	13	-6	-3	16	18	11	100

Table 3.14. Explanation of MCR results terms, see Chapter 5 for more details on the method. All results are in degree Celcius.

Column	Explanation
Sample	Sample name from database
TMaxLo	Lower limit of reconstructed mean temperature of warmest month
TMaxHi	Upper limit of reconstructed mean temperature of warmest month
TMinLo	Lower limit of the reconstructed mean temperature of the coldest month
TMinHi	Upper limit of the reconstructed mean temperature of the coldest month
TRangeLo	Lower limit of the reconstructed mean temperature ranges (TMax – TMin)
TRangeHi	Upper limit of the reconstructed mean temperature ranges (TMax – TMin)
NSPEC	Number of taxa used in reconstruction
Overlap	Percentage of sample taxa in the area of maximum overlap, used to calculate the temperature values

Checking the [Sample envelopes] box activates the creation of the second results file, which is automatically named (see below). In this file the climate space maps for each sample are exported, on individual worksheets, showing the percentage of sample species that could be found in each one degree climate cell (Figure 3.41). If the [Sample species lists] box is also checked, then the species list for each sample will be inserted under each map. These allow the user to gain a greater understanding of the thermal regime indicated by the faunas, and is particularly useful when teaching MCR theory.

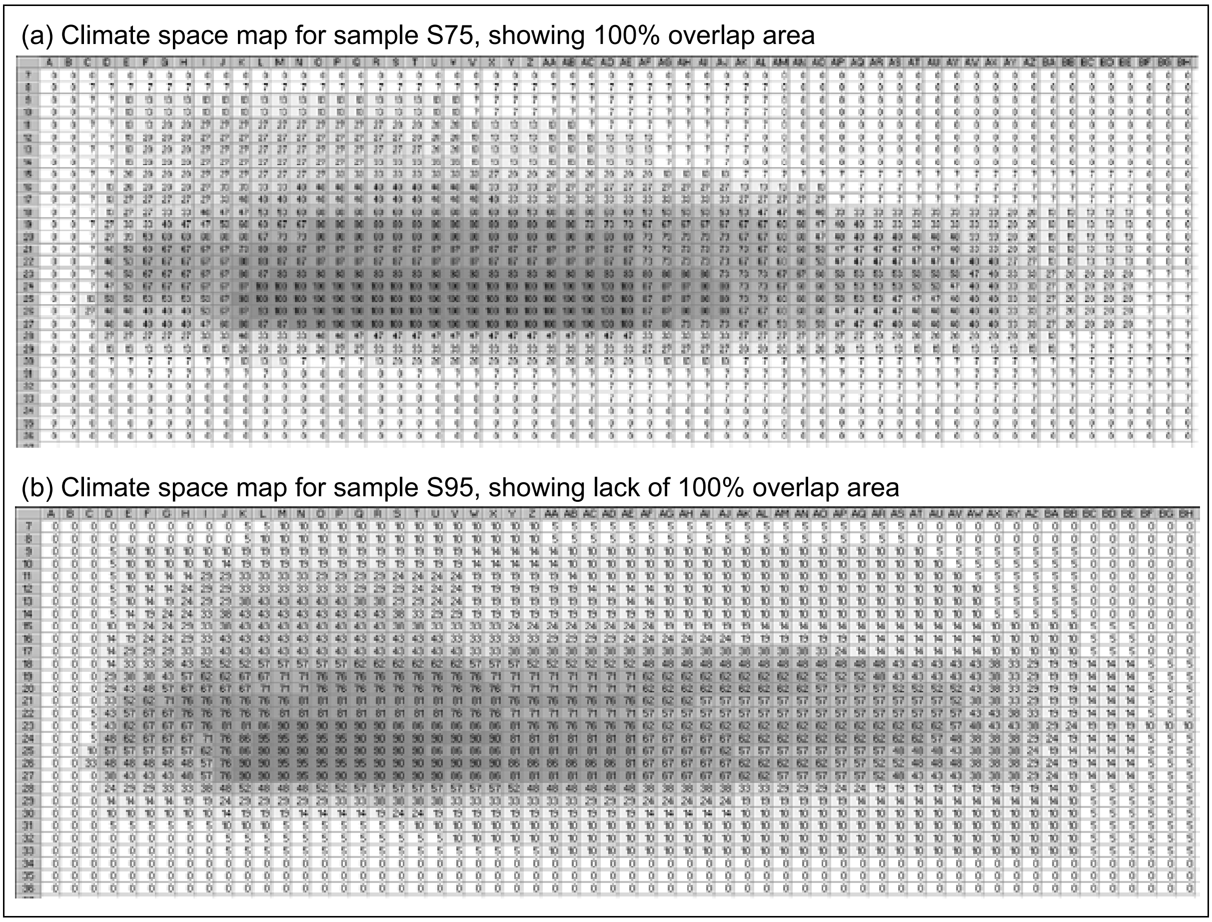


Figure 3.41. Climate space maps for two samples from Saint Bees, as exported by BugsMCR, showing (a) a 100 % overlap scenario and (b) a non 100 % overlap scenario with a complex area of maximum overlap. See Chapter 5 for further a more detailed explanation of envelopes and overlaps.

Two buttons start calculations: [Run MCR on All Samples] and [Run MCR on Species List only]. The latter of these ignores all samples and calculates temperature values for the entire species list as if it was one sample. This is primarily useful in teaching, using hypothetical data sets. The former button calculates temperature values for each sample that has at least one MCR taxon. Clicking on either button will open the standard 'Save as' dialog box, with which a save location and file name can be specified. A pop-up window (Figure 3.42) then indicates the progress of calculations (by way of a moving penguin), which may take some time for large datasets. If the export of sample envelopes is chosen, then this process will be repeated as the second output file is created. This second file is automatically given the same name as that specified by the user, but with the suffix "_matrices" added.

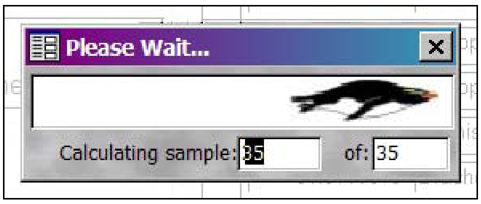


Figure 3.42. Screenshot: MCR calculation progress indicator, showing that sample 35 out of 35 is being calculated.

A message is displayed on completion of the calculations, from which the user can open the results files in MS Excel. This message must be closed in order to continue using BugsCEP.

BugsMCR also contains a number of advanced features, including species predictions and jackknifing, which are presented in Chapter 5.

3.4.5 The Search Interface

The BugsCEP search system is activated by clicking the [Search by Habitat etc...] component button on the main toolbar. Search settings are remembered between sessions, and if a previous search is found then the user will be asked whether they wish to “Continue from previous search session?” Clicking [No] leads to a request for confirmation with the question “Reset all search criteria, results and logs?”, and answering [Yes] to this opens the **Bugs Search Explorer** in its primary state (Figure 3.43).

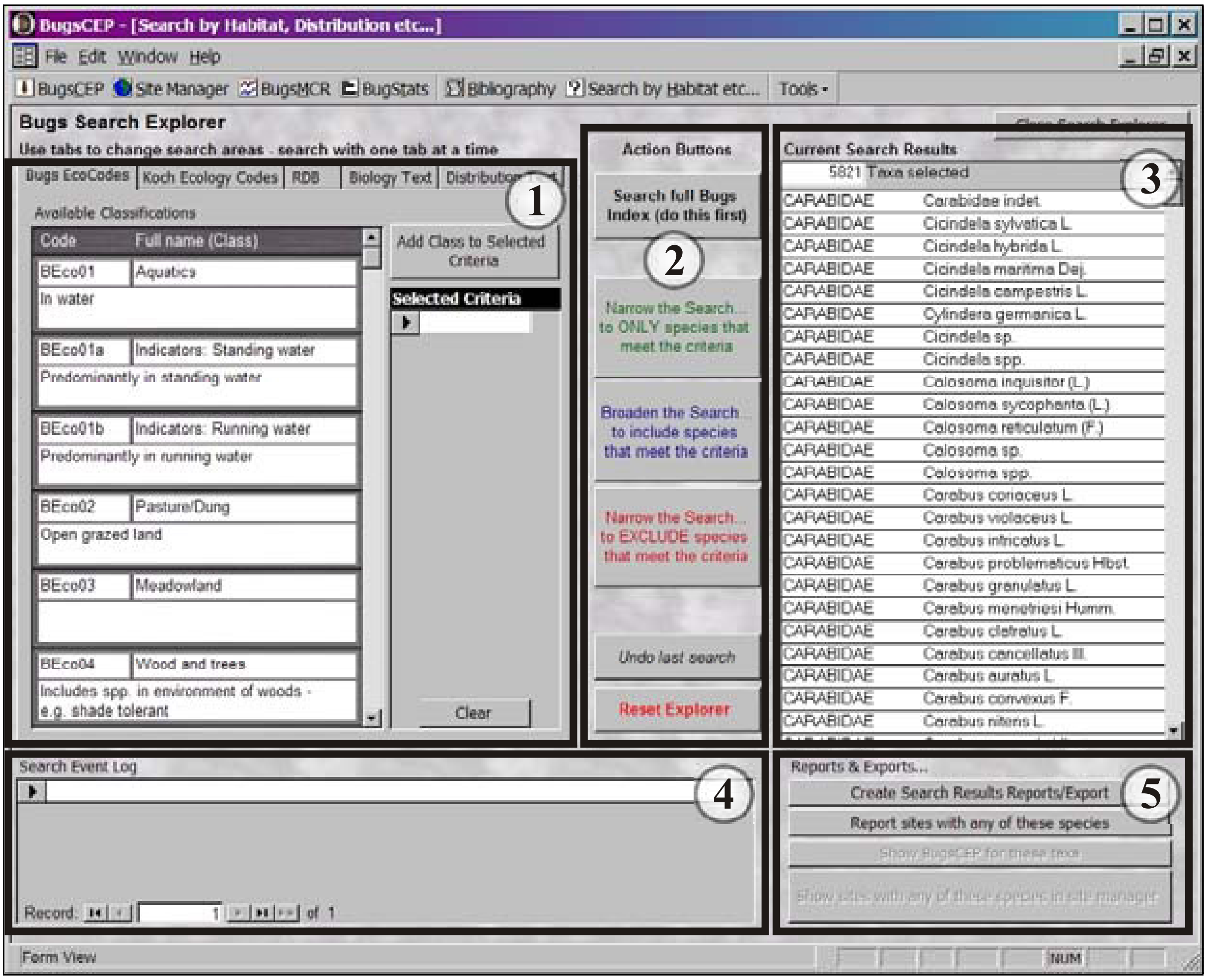


Figure 3.43. Bugs Search Explorer, with interface parts as follows: (1) Criteria type tabs (top) and search criteria selection for the active tab; (2) Action buttons; (3) Current species list (search results); (4) Current search session log; (5) Report/export buttons.

The Search Explorer is a powerful interface for retrieving species subsets through the successive application of search criteria to lists of species. The criteria type tabs, Figure 3.43(1), are used to select the area of the database to be queried, and then search criteria are either selected or entered, and then applied. The searchable data areas, along with the method by which criteria must be specified, are shown in Table 3.15.

Table 3.15. Data areas currently searchable with the Bugs Search Explorer, along with the method of criteria specification.

Search area	Method of criteria specification
Bugs EcoCodes	Code/term selection
Koch Ecology Codes	Code group and then code/term selection
RDB (Red Data Book)	RDB system and then classification selection
Biology Text	Free text entry (search terms can be saved)
Distribution Text	Free text entry (search terms can be saved)

Criteria must be selected/entered and applied using one tab at a time, although criteria are retained when the active tab is changed. On each search area tab the [Add Class to Selected Criteria] or [Add entered term to Selected Criteria] button adds the selected/entered term to a list, which is then used as the search terms when querying the current species list. Search terms are applied to the current species list, which may be the full BugsCEP species list or a result of a previous search, by clicking the action buttons, Figure 3.43(2). If multiple terms are specified, then all of them are searched for independently (i.e. logical 'OR' between terms). The resulting species are then shown in the 'Current Search Results', Figure 3.43(3), at the top of which can always be seen the number of taxa in the currently active list. Figure 3.44 illustrates the process of performing searches in the form of a flow diagram, and may help explain the logic of the process.

On each click of an action button a *Search Event Log* is created and displayed in the bottom left of the screen Figure 3.43(4), which includes the search terms used, the logic applied to the query, and the number of taxa that resulted from the search action. This log is essential for understanding how the currently displayed species list has been arrived at, and is exported with the reports available through the report/export buttons Figure 3.43(5).

It should be noted that the effectiveness of the search system is entirely dependent on the quality of the data, and the logic of the search sequence applied by the user. The biology and distribution fields are qualitative data in the form of extracted texts (see 3.1.2.1), and as such may very well contain text which could create misleading results. Lindroth (1985), for example, writes "Europe: not UK or Iberia" in his distribution notes for *Carabus coriaceus* L. The search system, on being asked to return species where the distribution data mentions 'UK' or 'Iberia' will return this species, ignoring the 'not'. An attempt has been made to design an 'intelligent' searching system which would be able to handle the 'not UK' case, but the 'not ... Iberia' is a more complex problem. This system has not been fully implemented in the release associated with this thesis, and users are recommended to always thoroughly examine the raw data for any species lists derived in this way.

Although the search system may appear daunting, it allow complex queries to be performed by linearly adding search terms and logically applying them to the previous search results. One of the more advanced features of the search system is that it is able to retrieve a list of sites from which any of the resulting species are known. Clicking the [Report sites with any of these species] button performs this action, and after a few seconds open a preview of a report which lists the summary data for each site, followed by a list of the species matched between the site list and the search results (see section 3.5.4). In principle, this allows researchers to define a specific environment by the application of search terms, and then retrieve a list of sites where that environment may have been present. This is an invaluable feature for comparative studies, and in general site interpretation. As always, this report may be exported to MS Word or Excel and saved, or printed directly.

Search result species lists can also be exported as Bugs Countsheet formatted MS Excel files, with a single presence-only sample (the final reporting option in the Report Generator, Figure 3.49a). This file can then be re-imported into a BugsCEP site and treated in the same way as any other countsheet.

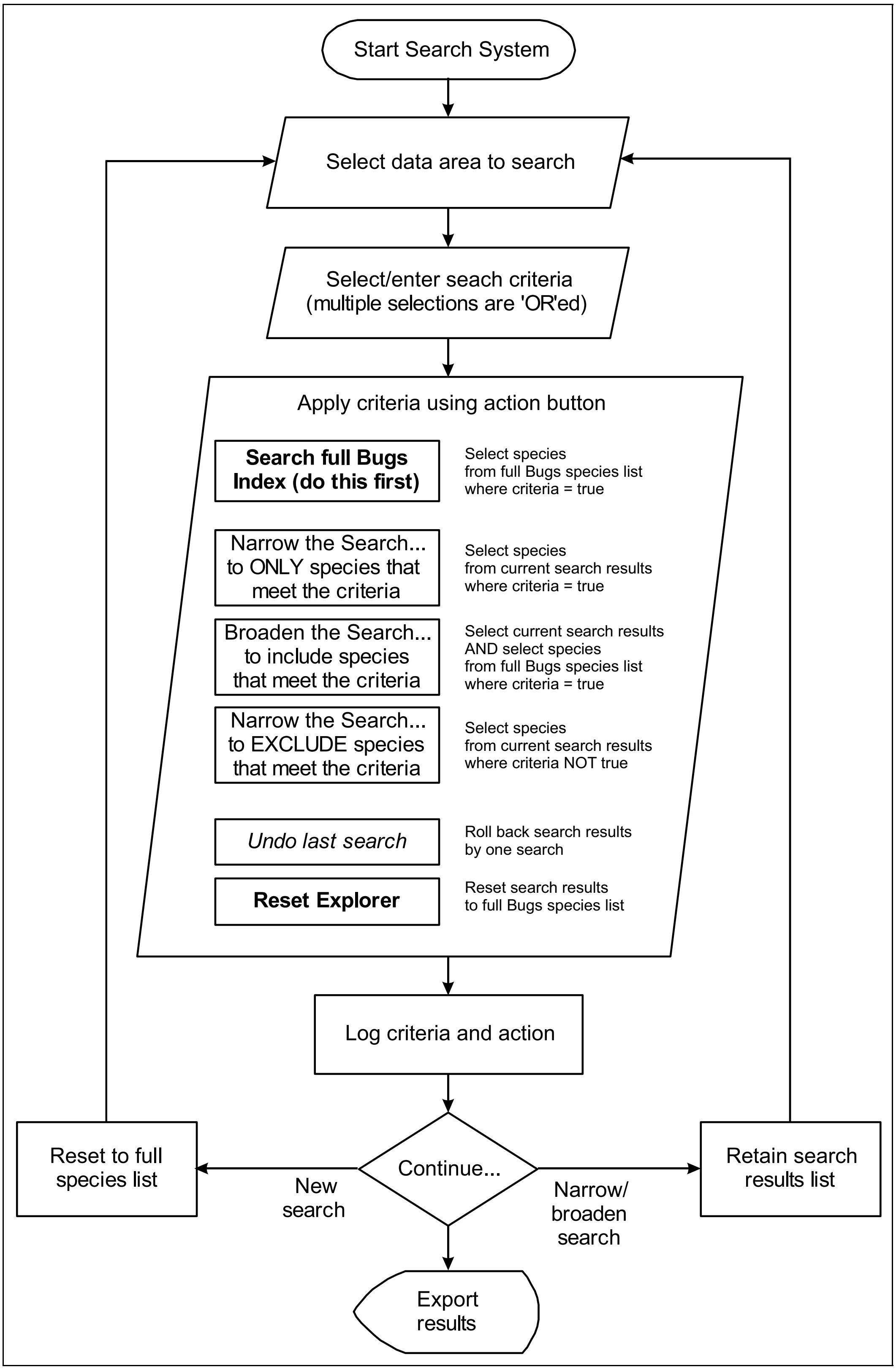


Figure 3.44. Flow diagram illustrating the sequence of events when using the Bugs Search Explorer.

3.4.5.1 A worked example

Imagine the scenario where a researcher wishes to retrieve a list of aquatic beetle species, which are extinct in the UK, but can be found still living in Sweden. The following steps describe how to answer this using the Bugs Search Explorer, with *italics* denoting the start of stages in criteria specification.

1. Reset the explorer, if a previous search is still active, by clicking the [Reset Explorer] button.
2. *Select the [RDB] criteria type tab.*
3. Select the 'United Kingdom: UKRDB' item from the [Select Country & RDB System] drop down box.
4. Click the 'X: Extinct in UK' item from the [Select Category] list, and click the [Add Class to Selected Criteria] button to add an 'X' to the [Selected Criteria] list.
5. Click the [Search full Bugs Index...] button. The Current Search Results are updated, with only 97 taxa now being displayed – these are all extinct in the UK, but more than just aquatic species are included. An item is also written to the Search Event Log.
6. *Select the [Bugs EcoCodes] criteria type tab.*
7. Click the 'Aquatics' class, and press the [Add Class to Selected Criteria] button to add 'BEco01'[xxxiv] to the [Selected Criteria] list.
8. Click the [Narrow the Search... to ONLY species that meet the criteria] button. This results in a list of only ten taxa that are extinct in the UK and classified as aquatic according to the BugsCEP classification system.
9. *Select the [Distribution Text] criteria type tab.*
10. Type the word 'Sweden' into the [Enter single search term:] box, and click [Add entered term to Selected Criteria].
11. Click the [Narrow the Search... to ONLY species that meet the criteria] button. This results in seven taxa, the names of which can be exported by pressing the [Create Search Results/Export] button and choosing the [Just the names] report (Figure 3.45). Page two of the report shows the search event log which lead to the species list (Figure 3.46), allowing the search to be easily replicated. The date of creation of the report is also given, as this could be important when comparing results from different versions of BugsCEP.

The accuracy of this list is dependant on a number of factors (in addition to those mentioned above), including the accuracy and completeness of the data, the reliability of the cited authors, and the reliability and extent of the Bugs EcoCode classifications. As a result the list should in no way be considered a fact, but only a suggestion, and an aid to further research. Future developments of the BugsCEP system will inevitably improve the search system, in particular by the classification of more species, the addition of real geographical distribution ranges and the ability to search fossil records and climate data through this interface.

Bugs Search Explorer Results

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Code	Family	Genus & species
4.0270032	DYTISCIDAE	Colymbetes striatus (L.)
4.027004	DYTISCIDAE	Colymbetes dolabratus (Payk.)
4.028002	DYTISCIDAE	Hydaticus continentalis Balfour-Browne
4.029001	DYTISCIDAE	Graphoderus bilineatus (Deg.)
4.032001	DYTISCIDAE	Cybister lateralimarginalis (Deg.)
5.002007	GYRINIDAE	Gyrinus natator (L.)
8.001001	SPERCHEIDAE	Spercheus emarginatus (Schall.)

Figure 3.45. Search results report showing 'Just the names' of the taxa which are RDB classified as extinct in the UK, classified as aquatic and known from Sweden.

[xxxiv] This is the Bugs EcoCode designation for aquatic taxa, as described in section 4.3.1.

Search log for exported session — Created: 30 January 2007
Full Bugs species list selected
Narrowed to only species where RDB = 10 (97 taxa)
Narrowed to only species where Bugs EcoCodes = BEco01 (10 taxa)
Narrowed to only species where Distribution Text = Sweden (7 taxa)

Figure 3.46. Final page of the 'Just the names' search results report, showing the sequence of events that lead to the final species list (the search log).

3.4.6 The help files

A simple, example based html help system is built into BugsCEP, the contents page of which (Figure 3.47) is accessible through the [BugsCEP Help Contents] item of the Help menu. Context sensitive help, that is to say specific help for the BugsCEP screen that is currently open, is available by pressing the <F1> function key on the keyboard. This opens the appropriate help page for the currently viewed area of the software, but, unfortunately does not yet move to the exact description of the currently selected button or feature. The help pages will be improved with time, and appropriate tool tips, indices and wizards will be added if users request them.

The entire help system may also be navigated through a series of buttons that are shown at the top of every help page (Figure 3.47). The system, which closely resembles a website, is essentially organised by program component, as described above, with additional introductory and overview pages to complement the information. Clicking on the title bar will attempt to open the BugsCEP online website through the help system.

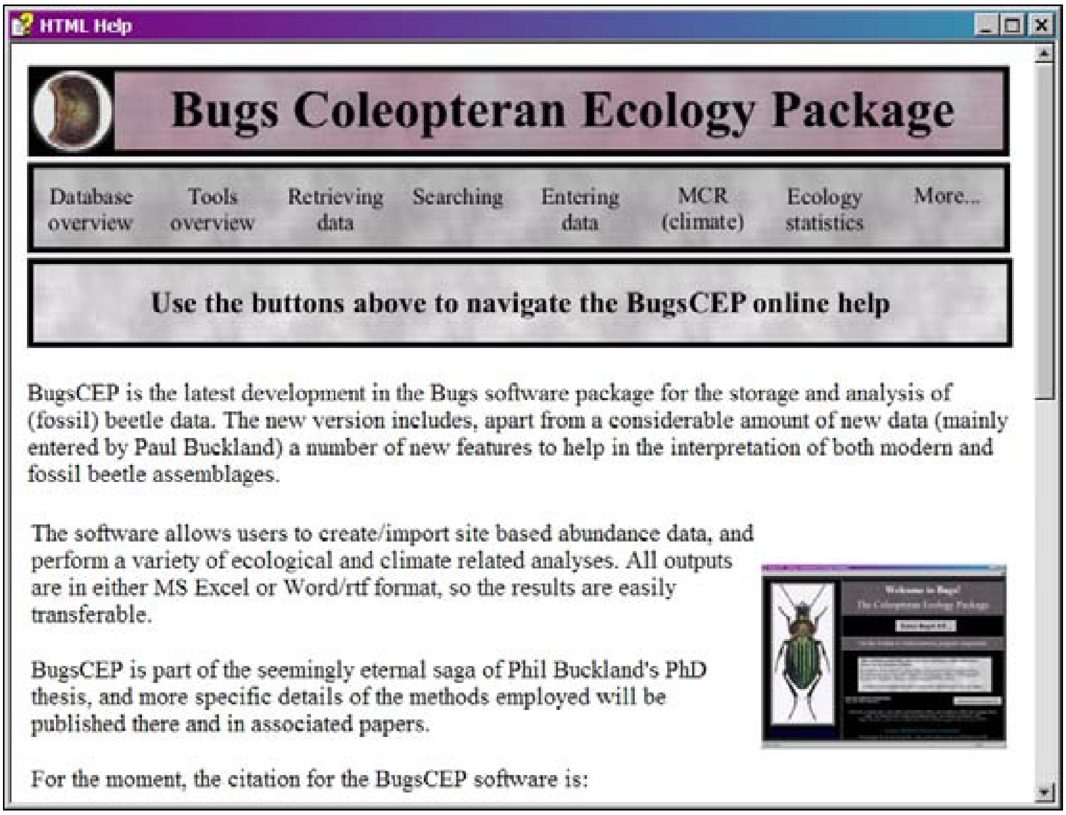


Figure 3.47. Screenshot: The BugsCEP help contents page, the links below the title bar providing navigation.

The help files are routinely duplicated on the website at http://www.bugscep.com/help/ where they are accessible to anyone with an Internet connection. It often proves useful to be able to direct users to instructions using a web address rather than trying to describe what to do over the phone or via email. The exposure of the help pages on the Internet also increases the visibility of the website through search engines, which are a valuable form of passive marketing and help to increase awareness of the software.

3.5 Reporting & Exporting Functions

BugsCEP includes a number of export and report functions that are spread throughout the program, which are presented here by program area.

Figure 3.48. Screenshot: The report toolbar, displayed at the top of all report previews.

Whenever a report is created in BugsCEP, it is first displayed in report preview mode, and the standard toolbar and menu is replaced with those shown in Figure 3.48. The first two buttons activate standard Print and Page Setup functions, to the right of which are a number of zoom controls which allow the report to be previewed as up to six pages at once. Following these are the [Export to Word] and [Export to Excel] buttons, both of which export the report to a temporary file (rtf format for MS Word, and xls for MS Excel) and then open it in the indicated program[xxxv]. The user can then save this file for future use.

A number of functions in BugStats and BugsMCR export complex data or results directly to MS Excel, and these files are described above (sections 3.4.3.3 and 3.4.4) and in Chapters 4 and 5.

3.5.1 Single species data

It is possible to output the information available for a single species by clicking the [Export] button on the BugsCEP main screen whilst viewing the species information. A confirmation request will be displayed, which can be prevented from appearing again, and then the BugsCEP **Report Generator** options pop-up window (Figure 3.49). This pop-up is used in more than one section of the program, and all the available options, which are described in section 3.5.4, may not be useful in every context.

After selecting the required report type, and clicking the [Preview/Export Report] button the appropriate report preview will be shown. The relevant references will be exported with reports (with the exception of XLS file output), and users are reminded to always cite the original data sources *as well* as BugsCEP. A file name will be requested if the [Export as Bugs Countsheet] option is chosen. The fossil record for a taxon cannot currently be exported in this way, and must be manually selected, copied and pasted into other software. A future update will provide this facility.

[xxxv] Note that the [Export to Excel] function of the Countsheet Manager, or [Export as Bugs Countsheet] report option (Figure 3.49), rather than text reports, are recommended when wishing to export species lists as they provide better formatting.

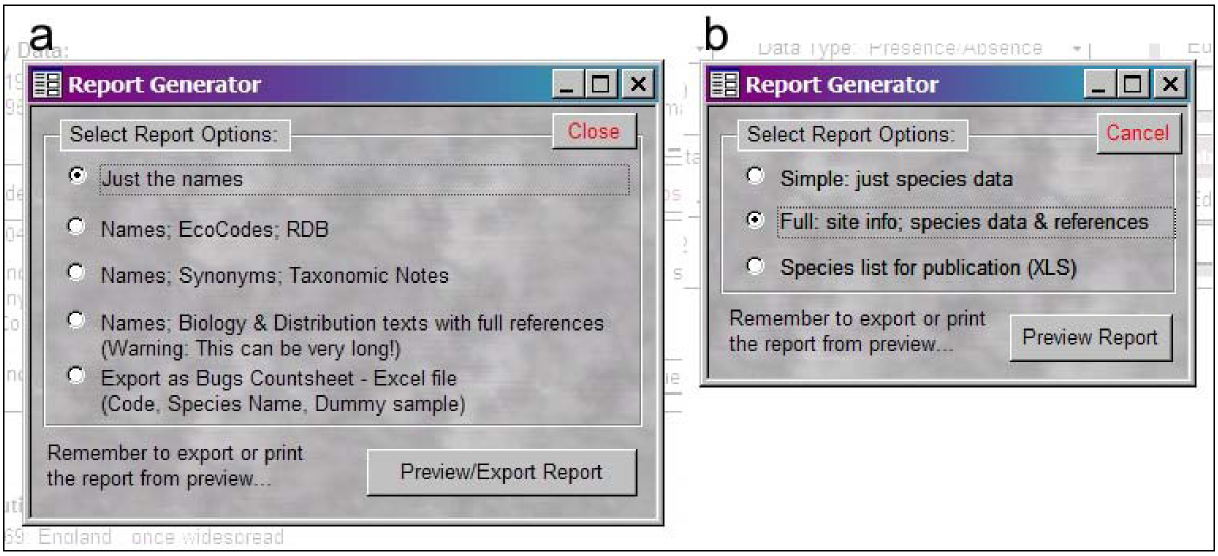


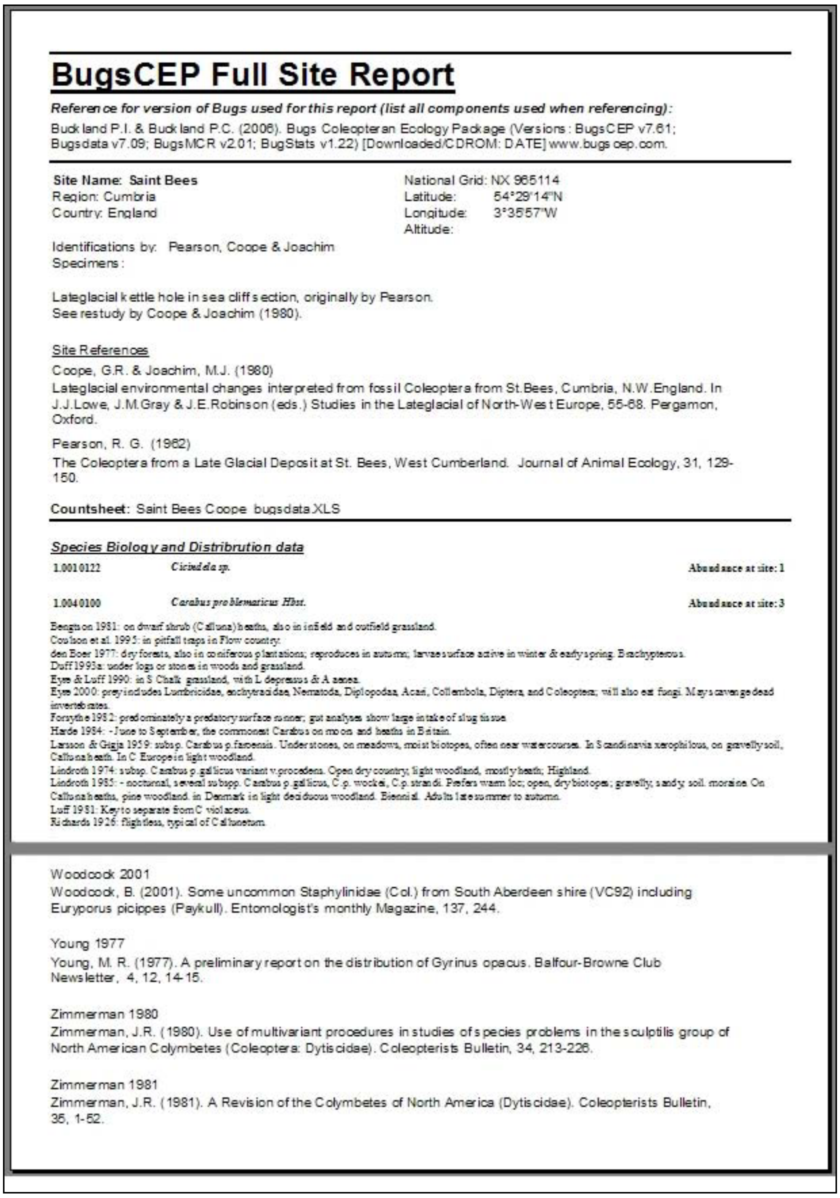
Figure 3.49. Screenshot: The Report Generator, which allows the user to specify which type of report to create. It is used in several areas of the program, and has a number of variations as shown by **a** and **b**.

3.5.2 Bibliographic data

The BugsCEP Bibliographic Browser includes the facility to output the references being displayed at any time to a standard report, by clicking the [Export current references] button. This report simply lists the active references by author.

3.5.3 Site reports

A compilation of the biology and distribution data for the species found at a site, along with the relevant references, can be created by clicking the [Create Report] button on the Countsheet Manager screen for any site. The Report Generator will open and display three report options (Figure 3.49b), the 'Full…' report, being the default. This report includes site information, including site references, followed by the biology and distribution information for each taxon found at the site, along with that taxon's total abundance (Figure 3.50). References to all works cited are exported at the end of the report.



BugsCEP Full Site Report

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Site Name: Saint Bees
Region: Cumbria
Country: England

National Grid: NX 965114
Latitude: 54°29'14"N
Longitude: 3°35'57"W
Altitude:

Identifications by: Pearson, Coope & Joachim
Specimens:

Lateglacial kettle hole in sea cliff section, originally by Pearson.
See restudy by Coope & Joachim (1980).

Site References

Coope, G.R. & Joachim, M.J. (1980)
Lateglacial environmental changes interpreted from fossil Coleoptera from St.Bees, Cumbria, N.W.England. In J.J.Lowe, J.M.Gray & J.E.Robinson (eds.) Studies in the Lateglacial of North-West Europe, 55-68. Pergamon, Oxford.

Pearson, R. G. (1962)
The Coleoptera from a Late Glacial Deposit at St. Bees, West Cumberland. Journal of Animal Ecology, 31, 129-150.

Countsheet: Saint Bees Coope bugsdata.XLS

Species Biology and Distribrution data

1.0010122 ***Cicindela sp.*** **Abundance at site: 1**

1.0040100 ***Carabus problematicus Hbst.*** **Abundance at site: 3**

Bengtson 1981: on dwarf shrub (Calluna) heaths, also in infield and outfield grassland.
Coulson et al. 1995: in pitfall traps in Flow country.
den Boer 1977: dry forests, also in coniferous plantations; reproduces in autumn; larvae surface active in winter & early spring. Brachypterous.
Duff 1993a: under logs or stones in woods and grassland.
Eyre & Luff 1990: in S Chalk grassland, with L depressus & A aenea.
Eyre 2000: prey includes Lumbricidae, enchytraeidae, Nematoda, Diplopodaa, Acari, Collembola, Diptera, and Coleoptera; will also eat fungi. May scavenge dead invertebrates.
Forsythe 1982: predominately a predatory surface runner; gut analyses show large intake of slug tissue.
Harde 1984: - June to September, the commonest Carabus on moors and heaths in Britain.
Larsson & Gigja 1959: subsp. Carabus p.faroensis. Under stones, on meadows, moist biotopes, often near watercourses. In Scandinavia xerophilous, on gravelly soil, Calluna heath. In C Europe in light woodland.
Lindroth 1974: subsp. Carabus p.gallicus variant v.procedens. Open dry country, light woodland, mostly heath; Highland.
Lindroth 1985: - nocturnal, several subspp. Carabus p.gallicus, C.p. wockei, C.p. strandi. Prefers warm loc; open, dry biotopes; gravelly, sandy soil. moraine. On Calluna heaths, pine woodland. in Denmark in light deciduous woodland. Biennial. Adults late summer to autumn.
Luff 1981: Key to separate from C violaceus.
Richards 1926: flightless, typical of Callunetum.

Woodcock 2001
Woodcock, B. (2001). Some uncommon Staphylinidae (Col.) from South Aberdeen shire (VC92) including Euryporus picipes (Paykull). Entomologist's monthly Magazine, 137, 244.

Young 1977
Young, M. R. (1977). A preliminary report on the distribution of Gyrinus opacus. Balfour-Browne Club Newsletter, 4, 12, 14-15.

Zimmerman 1980
Zimmerman, J.R. (1980). Use of multivariant procedures in studies of species problems in the sculptilis group of North American Colymbetes (Coleoptera: Dytiscidae). Coleopterists Bulletin, 34, 213-226.

Zimmerman 1981
Zimmerman, J.R. (1981). A Revision of the Colymbetes of North America (Dytiscidae). Coleopterists Bulletin, 35, 1-52.

Figure 3.50. Site report, full version, showing part of the first and last pages of the 82 page report generated for the Saint Bees site. The first page shows site summary information and the first taxa from the site, and the last page shows the last few references cited in the biology and distribution data.

3.5.4 Reporting search results

A number of reporting options (Figure 3.49a) are available when a list of species is arrived at through the Bugs Search Explorer, each of which can be created by pressing the [Create Search Results Reports/Export] button. The figures shown below are all produced from the results of the worked example given in section 3.4.5.1. The search event log (see Figure 3.46) is appended to all reports just before the list of references, and will not be duplicated below.

The simplest report form, 'Just the names', requires little explanation and is shown above (Figure 3.45 and Figure 3.46). The 'Names; EcoCodes; RDB' report lists the Bugs ecology classifications, and Red Data Book classifications for each species (Figure 3.51).

Bugs Search Explorer Results

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Note: Referencing the correct BugsStats (EcoCodes) version allows intercomparability of results.

Code	Family	Genus & species	
4.0270032	DYTISCIDAE	Colymbetes striatus (L.)	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB
4.027004	DYTISCIDAE	Colymbetes dolabratus (Payk.)	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB
4.028002	DYTISCIDAE	Hydaticus continentalis Balfour-Browne	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB
4.029001	DYTISCIDAE	Graphoderus bilineatus (Deg.)	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB
4.032001	DYTISCIDAE	Cybister lateralimarginalis (Deg.)	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB
5.002007	GYRINIDAE	Gyrinus natator (L.)	
	BugsEcoCODE: BEco01a: Standing water (ponds) indicators		
	BugsEcoCODE: BEco01: Aquatics		
	Rarity: United Kingdom	X: Extinct in UK	UKRDB

Figure 3.51. Search results report using the 'Names; EcoCodes; RDB' option. Each taxon name is followed by its BugsCEP habitat and Red Data Book classifications.

The 'Names; Synonyms; Taxonomic Notes' option may be more useful for taxonomic studies, or when identifying specimens, although these field contain a limited amount of information at the present time (Figure 3.52). For many users, the 'Names; Biology & Distribution texts with full references' option will be the most useful, as it exports the full extent of the databases extracted texts for each species, along with references to the cited works (Figure 3.53).

Bugs Search Explorer Results

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Code	Family	Genus & species
4.0270032	**DYTISCIDAE**	**Colymbetes striatus (L.)**
4.027004	**DYTISCIDAE**	**Colymbetes dolabratus (Payk.)**
4.028002	**DYTISCIDAE**	**Hydaticus continentalis Balfour-Browne**
	Synonym: Hydaticus stagnalis (F.) non (Fourcroy) Duff 2006	
	Taxonomy notes: Hodge & Jones 1995: very similar to H transversalis, but black area on hind margin of pronotum less extensive and elytra with longitudinal streaks; both characters variable.	
4.029001	**DYTISCIDAE**	**Graphoderus bilineatus (Deg.)**
4.032001	**DYTISCIDAE**	**Cybister lateralimarginalis (Deg.)**
5.002007	**GYRINIDAE**	**Gyrinus natator (L.)**
	Taxonomy notes: Foster 2006: Separated from Gyrinus substriatus on female genitalia.	
	Taxonomy notes: Angus & Carr 1982: Detailed separation from Gyrinus substriatus.	
8.001001	**SPERCHEIDAE**	**Spercheus emarginatus (Schall.)**

Figure 3.52. Search results report using the 'Names; Synonyms; Taxonomic Notes' option.

A list of the sites from which the species are known, along with site summaries is available by clicking the [Report sites with any of these species] button on the Search Explorer. The first few items in the report for the seven species listed above are shown in Figure 3.54.

Bugs Search Explorer Results

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Code	Family	Taxon
4.0270032	**DYTISCIDAE**	**Colymbetes striatus (L.)**

Dettner 1983: Key to C spp.
Koch 1989: in bog pools.
Nilsson & Holmen 1995: in more or less temporary ponds in open country; in S, more stable populations usually in smaller dystrophc ponds or clear water lakes or in ponds and ditches in bogs and fens. Frequently on flooded lake or river margins. Adults overwinter and larvae in late spring and early summer.

Dettner 1983: Europe: N & parts of C, Russia, Siberia.
Koch 1989: Europe : E.
Nilsson & Holmen 1995: Denmark : scattered rec in E. Norway :mainly in SE, but also in ST. Sweden & Finland : widespread & common, N to Arctic Circle. Baltic States. W Palaearctic : E France & Switzerland to Fennoscandia, E to W Siberia.

Code	Family	Taxon
4.027004	**DYTISCIDAE**	**Colymbetes dolabratus (Payk.)**

Böcher 1988: Greenland: typically in ponds with rich vegetation, highly eurytopic in almost any kind of stagnant water, except possibly temporary pools. In summer, adults and larvae found together, but in late summer and autumn the imagines migrate to hibernate in lakes which do not freeze to the bottom, i.e. > 3m. Flight records from Scoresby Sound. In Scandinavia, alpine and subalpine, found in marshes, sphagnum bogs and ponds devoid of vegetation. On Hardangervidda, Norway, almost confined to pools above 1000m, without vegetation, but with Carex spp, grasses and mosses along shore. Feeds probably on crustacea, Brachinecta, Cladocera. Adults throughout year. Pupa is terrestrial, found in small hollows under moss close to shore of ponds.
Larsson & Gigja 1959: - Iceland - throughout but rarer in S. and S.W.; prefers ponds rich in plants, but often in other biotopes, even clayey pools near glacier's edge.
Lindroth et al. 1973: in ponds, usually with rich vegetation. Wings full and functional. Icelandic form has been given the subsp name Thomsoni Sharp.
Nilsson & Holmen 1995: mainly in shallow ponds above the treeline, frequent on flooded lake and river margins, also in ponds in gravel workings. Larval development mainly late June-late July. Adults ready fliers, in Greenland migrate to deeper lakes for overwintering.
Zimmerman 1981: a highly variable sp. collected in pools of water on pack ice!

Böcher 1988: Greenland: most common and best known beetle, both coastal and inland from Kap Farvel in S to Upernavik on W and Clavering O on E. N limit is set by 10 month of freezing on lakes because of need to hibernate. Circumpolar, boreal to arctic, Europe, Iceland, N Scandinavia, N Russia, N Siberia, E to Kamchatka, Alaska, Canada: NW Territories to Labrador.
Larsson & Gigja 1959: Iceland.
Lindroth et al. 1973: Circumpolar N, Greenland. Iceland: throughout but rare in S.
Nilsson & Holmen 1995: Fennoscandia : mountains S to TE in Norway & DR to Sweden. Finland : N, S to Ob. A circumpolar sp, in N America along N edge of continent from Alaska to Quebec. Greenland N to 75N. Eurasia, through Arctic and N boreal regions from Iceland and

Figure 3.53. Search results report using the 'Names; Biology & Distribution texts with full references' option.

Bugs Search Explorer Results - Sites where selected taxa found

Reference for version of Bugs used for this report (list all components used when referencing):
Buckland P.I. & Buckland P.C. (2006). Bugs Coleopteran Ecology Package (Versions: BugsCEP v7.61; Bugsdata v7.09; BugsMCR v2.01; BugStats v1.22) [Downloaded/CDROM: DATE] www.bugscep.com.

Agerod V, Skåne, Sweden

Ageröd V. Specimens hand picked by the excavators of a mesolithic midden.
Lat: 55°55'60"N Long: 13°23'60"E Altitude: 39m
Identified by: Lemdahl

Search Result species found at site:
4.0320010 DYTISCIDAE Cybister lateralimarginalis (Deg.)

Björkeröds Mosse, Skåne, Sweden

Lateglacial to Early Holocene successions from profiles A6, B1 & 3b.
Dates from base of A6, the remainder interpolated from pollen data.
Note : Sample A6/1 is Early Holocene (EH).
Lat: 56°16'60"N Long: 12°32'0"E Altitude: 75m
Identified by: Lemdahl

Search Result species found at site:
4.0270040 DYTISCIDAE Colymbetes dolabratus (Payk.)

Brandon Upper Deposit, Warwickshire, England

Initially interpreted as Mid-Weichselian organic sediments, but see Maddy et al. (1994) for suggestion that this and Upper 2 are OS 6.
Lat: 52°27'0"N Long: 1°23'0"W
NGR: SP391753
Identified by: Coope
Specimens stored at University of Birmingham, UK.

Figure 3.54. Report showing the first few sites from which the seven species shown in Figure 3.45, which were retrieved through the Search Explorer, are known.

3.6 The Addition of Taxa, Biology/Distribution Data and References

Modern reference data entry is currently undertaken centrally by Paul Buckland, and the lack of update routines for the import of data entered by other users makes decentralization potentially difficult and time consuming. The current system also allows for more uniform implementation of data quality and validation criteria. Users are instructed to contact the authors with requests for the addition of specific data. Should additional taxa or reference data be required, then the authors would be only too pleased to add them on request. Users may, on the other hand, add as many countsheets and sites as desired. References for sites may be added by clicking the [Add Reference] button, on the Site Information screen. This becomes available either on the creation of a site, or when the [Edit Site] button is pressed and the correct 'Admin' level password entered. The reference entry system automatically inserts lettered suffices where duplicate author-date combinations occur (e.g. 'Buckland 2000a', 'Buckland 2000b', etc.), and allows the user to pick any existing reference for use in the current context.

A number of maintenance features have been omitted from this text as they would have been out of place, and will be documented at a later date. These allow data managers complete control over the taxonomic master list, ecology code and other lookup systems, as well as providing fail safes for when data become corrupt.