How can a database full of Bugs help reconstruct the climate?

Phil Buckland
Environmental Archaeology Lab.
Department of Archaeology & Sami Studies
University of Umeå,
901 87 Umeå
Sweden
Phone: +46 90 786 9792 - Fax: +46 90 786 7663 - E-mail: phil.buckland@arke.umu.se

Paul Buckland
Department of Archaeology & Prehistory
University of Sheffield
Northgate House, West Street
Sheffield, S1 4ET, UK
Phone: +44 114 22 22913 - Fax: +44 114 27 22 563 - E-mail p.buckland@sheffield.ac.uk

Abstract: The BUGS Insect Ecology Package was originally constructed (using Dbase and Clipper) to compile Coleoptera (beetle) habitat and distribution data from a myriad of sources into one, easy to use, and publicly available database. Its primary users were researchers and teachers within the palaeoentomology field. The present system, five versions and many revisions later, is built in MS Access 2000, and covers some 5300 species, 2000 references, and 240 sites (archaeological and Quaternary), and is of value to archaeologists, ecologists, and conservationists alike.

BUGS is essentially a relational database management system constructed around three components:
- the species data (modern ecology and distribution)
- the bibliography
- the site data with species lists

Its implementation in several institutions has greatly accelerated the efficiency with which palaeoentomological investigations can be carried out, and greatly improved the teaching of the subject.

Palaeoenvironmental reconstructions are performed by the superimposition of the ecology and distribution of modern insect populations over fossil assemblages. At the moment, this is essentially performed semi-quantitatively by cross-reference of the data (which BUGS collates for a species list and then exports as an RTF file to any word processing package). BUGS contains a wealth of ecological data which can be employed in the interpretation of archaeological sites and contexts. In natural deposits, away from the artificial heat islands created by human activity, insect distributions are essentially constrained by climatic parameters. Tim Atkinson (UEA) and Dave Perry (formerly at Birmingham University) digitally encoded the temperature range data for over 400 species into a program for the calculation of palaeoclimates through the MCR (Mutual Climatic Range) method, and this has been extensively used in the modelling of Quaternary climates from beetle remains. The aim of our present phase of BUGS development is to implement MCR functionality into the BUGS database system. From this point it should be possible to move on to other ecological variables such as habitat and vegetation types, and increase the precision of modern climatic data, thus enhancing the value of insects in archaeological interpretation and the modelling of past climates.

Key words: Bugs, beetles, Coleoptera, database, ecology, palaeoecology, climate change, environmental archaeology, palaeoentomology

Introduction

The phylum Insects (= Hexapoda) makes up approximately 98% of the planet’s animal biodiversity. Some 75% of these species are beetles, most of which have very specific habitat requirements, within which temperature is a very important variable. In addition, they are perhaps the most common—and useful—macrofossil the archaeologist is likely to encounter both on site and in suitable adjacent localities from which local palaeoenvironments may be reconstructed (cf. Buckland 2000; Ashworth et al. 1997; Buckland & Coope 1991; Buckland & Sadler 2000; Elias 1994). Insect remains are abundant in all anaerobic, permanently dry or permanently frozen contexts, from Eskimo middens (Böcher & Fredskild 1993) through Norse Greenland farms (Buckland et al. 1996), urban York (Kenward & Hall 1995) to the deserts of Egypt (Panagiotakopulu 2000). The reconstruction of these past environments relies upon the ability to project modern ecological data over fossil assemblages, allowing models to be built, often of very intimate places (cf. Kenward 1999; Panagiotakopulu & Buckland 1999). The close
climatic constraints on the distribution of many species also means that insects are sensitive indicators of climate change, and it was study of fossil beetle faunas which first indicated the precipitate nature of the end of the last glaciation (Ashworth 1972; Osborne 1972; 1980; Bishop & Coope 1977; Coope & Brophy 1973), long before its confirmation in the Greenland ice core data (Mayewski et al. 1996; White 1993). The large sample size (> 5 litres preferred) and the slow process of identification by comparison with modern reference collections, however, has made this a field in which there are few practitioners. Interpretation also has had problems (cf. Kenward 1976), and habitat data on individual taxa are not only widely dispersed through an often obscure literature but also sometimes of doubtful quality – readers of Donisthorpe's (1939) Coleoptera of Windsor Forest for example would be convinced that most beetles just sat around on flowers! The BUGS database collates habitat (including environmental dependency variables) and distribution (both modern and fossil) for over 5000 species of beetle – the complete UK fauna, and a significant part of the North European fauna. With its twin on the Egyptian beetle fauna (Buckland et al. in press), they contain in excess of 100 mb of information, of use not only to palaeoecologists but also to applied entomologists and those involved in species' conservation. As part of a Leverhulme Trust funded project developing the use of flies in archaeological interpretation, a similar database is being constructed for the Palaeartic Diptera, and a mollusc database is also in process of development. The basic engine developed in Microsoft Access © can be applied to any biological group.

A detailed description of the Bugs Coleopteran Ecology Package, albeit an older version, has been presented elsewhere (Buckland et al. 1997), and so only a brief outline will be provided here. Instead this paper concentrate on the enhancements planned in the current expansion phase of the project, centring on the development of a semi-automated system for producing proxy climate data from fossil insect remains.

From a palaeoclimatological point of view, many groups of insects have the useful trait of being specific to particular habitats (stenotopic) or ranges of environmental variables. Despite earlier ideas that insects represent rapidly evolving lineages, it has become evident though work on the Quaternary fossil record that, during periods of climate change, the more stenothermal species of insect tend to migrate as a population, tracking the changing environment, rather than evolving (Coope 1990; 1995). Potentially therefore, they may provide at least two million years of proxy climatic data. Taken with their propensity to occupy very specific microhabitats, it allows us to reconstruct fossil environments, both human and natural, to a very detailed level, confined only by the limits of sample resolution (cf. Buckland & Perry 1989). In the case of climate reconstruction, the most significant recent work has involved the larvae of non-biting midges, the Chironomidae (cf. Brooks & Birks 2000), where sample resolution has allowed significant refinement on earlier models of Lateglacial Holocene transition, but the group is essentially restricted to lacustrine sediments and do not allow the parallel reconstruction of regional environmental changes. In this the most useful group is the Coleoptera (beetles), which include taxa, which are dependent on plant foods, as well as highly mobile predators, such as many of the species of carabid (ground beetles). These latter species, independent of specific plant or animal hosts, have the ability to migrate quickly as populations to take advantage of climate changes. They may, for example, be among the first species to arrive at newly exposed glacial tills during a retreat phase after rapid climatic amelioration. By collating the temperature-tolerance data for species, essentially via the proxy of modern distributional data, found fossil in a sample it is possible to begin to construct snapshots of past environments and the changes that occurred in them, constrained only by sample resolution. This technique, called Mutual Climatic Range Analysis (MCR) (Atkinson et al. 1986; 1987) relies upon the definition of individual climatic envelopes for each species, and then overlaying these to define the range in which they may mutually occur. It is the intention to expand the BUGS program to include a facility for calculating palaeo-temperatures.

BUGS in Brief

BUGS, the Coleopteran Ecology Package, was designed as a research aid for both palaeoecologists and entomologists. Its roots lie in a DOS database (DBase IV & Clipper) database created by Jon Sadler and Mike Rains (Sadler et al. 1992), and the present form, built in Microsoft Access ©, BUGS 2000v5, consists of a primary database, of ecology, distribution and fossil record data for over 5,300 species, continually updated (with images and key identification characters being added slowly); the associated bibliography lists over 2500 references (Fig 1). A separate listing Qbib includes all known papers on fossil insects from the Quaternary, including the archaeological record. In addition, to this, a database of about 240 sites, archaeological and Quaternary, contains species spread sheets in taxonomic order, which can be accessed through BUGS or Microsoft Excel © (Fig.2). In the process of data entry, BUGS automatically orders any entered species list into taxonomic order. Only basic site data is included, such as latitude and longitude, elevation, interpretation, references, and dating evidence, to avoid the duplication of parallel site databases.

Basic query functions allow the user to retrieve a subset of the species database, which is restricted to their search criteria. This could be used, for example, to find all the species, which are attracted to fire, and found in Britain at the present day (Bug's names 54 species). This feature is particularly useful when a sample produces species that are no longer found in the existing environment of the sample, or even the country under research. Similarly, the Red Data Book (RDB) field allows a researcher to see whether a species is present/rare at the present day.

Built in reports provide simple, but very useful, printable outputs for: species data; search results (just the names, or full details); site species lists; ecology and distribution for species at a site; and references/bibliography. These reports, which can be printed directly, or saved as rtf files, effectively reduce the time necessary to produce meaningful results from an insect assemblage. Habitat codes, based upon those used by Koch (1989a & b; 1992) and incorporating the ideas of both Kenward (1978) and Robinson (1983; 1991), are being developed to help in the process of dimensioning a faunal assemblage across taxonomic boundaries to refine interpretation.
Where beetles like to be – the Mutual Climatic Range

Any species of beetle (or animal/plant for that matter) can only survive within a certain temperature range. So, when finding a fossil population of a certain species in a sample it is possible to say that the temperature of the environment represented by the sample must have been within the range of tolerance for that species. This may sound perfectly logical, but there always questions of interpretation to consider – such as what does the sample actually represent. Depositional conditions, as well as post depositional disturbances, redeposition or bioturbation, have to be worked out with the archaeologists/geologists during both primary sampling and interpretation; the interdisciplinary nature of both Quaternary studies and archaeoentomology cannot be over stressed. It must always be considered whether the individuals found actually lived in the sample environment – possibly as a breeding population, or were just passing through, so to speak. This is nowhere more problematic than in the archaeological context, where the artificially warmed habitats created by both human and domestic animal waste products, as well as house, barn and byre interiors, provide heat islands well north of natural distribution areas for many species (Barlow et al. 1998). A wide range of species, termed synanthropic, largely accidentally transported by Man, occur in these contexts, from the house fly, Musca domestica, probably of Mediterranean or Old World subtropical origin (Skidmore 1985) to the grain weevil, Sitophilus granarius, an early fellow traveller (Büchner & Wolf 1997; Buckland 1990; Panagiotakopulu 2000) and not now known from a natural habitat. In addition, the accidental import of insects in dunnage and ballast on ships may lead to either temporary establishment of local populations or an enigmatic fossil find in an archaeological context. The single find of the rather southern scarabaeid Oxyomus sylvestris at medieval Bessastaðir in Iceland for example can easily be dismissed as an accidental traveller (Amorosi et al. 1992), and the ants from post-medieval Reykholt, Iceland (Buckland et al. 1992), probably arrived in straw around the priest’s wine bottles, but the several specimens of the small rove beetle Metopsis retusa/clypeata, now only found in the very south of Norway and Sweden (Lundberg 1995), in the midden of the medieval fishing station at Langenes in the Norwegian Arctic is more problematic. Was the medieval thermal optimum this much warmer (Buckland & Wagner 2001), or is this again a ballast and dunnage accidental traveller? Only further refinement of the climate model, and more fossil faunas from the site can provide the answers.

Setting aside these essentially taphonomic problems, and taking as given that a sample accurately represents the environment in question, then it is possible to go on to refine the climatic interpretation by overlaying the individual climatic envelopes derived from modern distributional data to find the range of temperatures in which all species present could survive (Fig.3). This is where choice of species is fundamentally important. When considering transient environments, or where a chronology is required, it is necessary to employ species that are mobile (but not necessarily winged), omnivorous, and not plant host specific. In this way, it should be possible to eliminate the effects of any other variables from the equation (although these variables will of no doubt be important for the overall site interpretation). Where only a snapshot reconstruction is required, then the migration abilities of the species used in the calculations is of less importance, but it is apparent that succession is important in terms of what arrives first. Those species which form part of what Crowson (1981) termed the aerial plankton appear first onto any tabula rasa created for example by deglaciation. Thus chironomids and caddis flies (Trichoptera) may arrive before any beetles, and incidentally their food amongst the diatoms and algae must also be the first colonisers. In most contexts, however, sample resolution is insufficient to define this process.

Existing Mutual Climatic Range (MCR) Software

BUGS runs in the MS Windows® environment, and is built around the Access 2000© database engine. Having been developed in DOS® during the 1980s, MCR has yet to undergo an updating process, and at present two pieces of software exist for calculating MCR envelopes from species lists. Both run in MSDOS and are referred to as MCR Birmingham, and MCR UEA (University of East Anglia) after the institutions where they were created (Atkinson et al. 1986). Although both systems were excellent tools at their time of creation, and Sinka and Atkinson (1999) have recently developed a version which utilises plant distribution, their interfaces are difficult to use and their data storage system cumbersome in terms of the present day Windows environment. Both systems used a binary temperature matrix (Fig.4) to store species thermal envelopes, which allows a resolution of only 1°C, although it should be noted that further refinement is ultimately constrained by the coarse nature of modern distribution data, often conflating records of over a hundred years of insect collecting. The use of the present form of MCR to reconstruct climate from archaeological assemblages from medieval sites is therefore questionable (cf. Hellqvist & Lemdahl 1996; Buckland & Wagner 2001).

Implementing MCR in BUGS (Figs 5-6)

As a first step, the existing digital MCR data has been ported over into BUGS as a working base. However, it is generally agreed that these datasets, initially obtained by digitising small scale distribution maps of species' Palaeartic range (Atkinson et al. 1986; Perry 1986), are no longer considered to be of high enough resolution, and will eventually be superseded by new datasets. The old form of MCR was very much restricted by the technology of the 1980's and the availability of computers. The new system is intended as an upgradeable expert system, which can be asked to recalculate thermal envelopes and MCR overlaps in the light of new data. As with the previous version, however, data entry remains the main problem, since each record requires location, elevation, and year of capture to be collated against the relevant interpolated climate data, something which would have been impossible with the equipment, and datasets available twenty years ago.

Data storage (fig.5)

Base distribution data for species comes from a variety of
Issues yet to be Resolved

The final nature of calculation of the TMAX/TRANGE values from the temperature/lat/long data has yet to be resolved, for example, whether we will use a basic 6°C per 1000 m altitude lapse rate, or cubic splining of the temperature trend surfaces. It is possible that a better approach would be to use remotely sensed data rather than interpolations from weather stations, although the constraints here are likely to be financial, and the data is only available for the last couple of decades. The availability of weather station data varies from country to country, whereas satellite data is more uniformly available, albeit more expensive. On the other hand binary raster data take more processing power/time and could prove impractical for the purposes of interpolation, although one can always limit the processed area within ArcView. The mathematics involved in the calculation of MCR overlaps are significantly complicated by the switch from binary temperature matrices (fig.4) to linear data tables (Fig.5b), even if it does have great advantages in terms of expandability and accuracy. Further consideration needs to go into the statistical implications, and associated error sources, of the various methods available.

ArcView does have some database manipulation capabilities, a query engine, and the Avenue scripting language, and it is possible that some calculations will be more easily performed in ArcView than Access where spatial data are concerned. The practical desire to keep as much calculating as possible within Access (for reasons of portability) may have to be compromised for the sake of simplicity. However, with the incorporation of the VisualBASIC language into the latest version of ArcView, this may become less of a problem, with both engines using different programming languages. The desire to keep BUGS as a freeware product, which can be used by those who do not have expensive software such as ArcView (or even Access, with the runtime installation), may lead to Access as the motor for most of the calculations.

In a project of this type, data entry is always a very time consuming task, and makes up a considerable part of the workload. It has therefore been decided to limit initially the study area to Scandinavia and the UK. The primary dataset will also only include the Carabidae. The geographic areas have good histories of insect recording, and have easily available distribution and weather data, whilst the carabids have been extensively collected in both regions (e.g. Lindroth 1985; 1986; Luff 1998). Even though this is a compromise, and excludes species which occur in Quaternary fossil deposits with either Mediterranean or Siberian distributions (Angus 1973; Coope & Luff 1998), it still allows a 3000 km north-south transect, and the system is designed with expandability firmly in mind. It goes without saying, the more detailed and extensive the base data are, the better the accuracy of the system. The Icelandic and Greenlandic fossil insect data would suggest that the coldest period (Buckland & Wagner 2001) throughout the Holocene, part of the Holocene was the Little Ice Age of the post-medieval period (Lindroth 1961; Lundberg 1995). Greenlandic fossil insect data would suggest that the coldest part of the Holocene was the Little Ice Age of the post-medieval period (Buckland & Wagner 2001). Throughout the Holocene, however, the fossil insect evidence, as Blytt and Semander (1908) long ago realised with fossil plant distribution, implies

Asking Questions

Correlation coefficients have frequently been used to compare sample assemblages, and Kenward (1978) has consistently championed Fisher’s α in the comparison of archaeological assemblages. BUGS is being upgraded to include a series of common biological comparative methods (e.g. Jaccard, modified Sorensen (Southwood & Henderson 2000), so that similarities between samples and sites can be sought. The querying form is also customised to allow comparison within thermal groups, or to take thermal envelopes into consideration when comparing assemblages. This should give opportunities to think of new questions to ask the datasets, which were previously difficult to handle.
oscillation between relative oceanicity and continentality, and the east-west transect provided by the dataset is at least as important as the north-south; the difficulty, as always, is taking Man, and human impact on available habitats, out of the equation (Wagner 1997).

Future Applications of MCR Bugs

As Perry showed nearly twenty years ago in terms of the Icelandic beetle fauna (Buckland et al. 1983), MCR can be used for predictions, in both the past and the future. Given a particular climatic envelope, it is possible to provide a list of what species should be present given unimpeded access. Simulation is an extremely important area within meteorology, geography, and ecology, and BUGS is designed with this in mind. Useful questions might be, for example, if the temperature were to decrease by 1°C what would happen to the distribution of species X, which is a predator on crop pest Y? In the archaeological context, there are opportunities for re-examining the Elm decline here (Girling & Greig 1985; Robinson 2000), utilising fine climate data derived from contemporary insect faunas – does the expansion of the elm bark beetle, Scolytus scolytus coincide with a period of climatic stress? The fine resolution of beetle chronologies has been instrumental in illustrating the rapid nature of climatic change, and has complemented other research, which stresses the often precipitate nature of glacial/interglacial switches (e.g. Taylor et al. 1993), and there is much to be gained from further study.

Modern insect distribution maps are often full of gaps where collecting, for one reason or another, has been scarce. This may leave holes in distributions where the species quite probably could be found, if there was anybody interested enough to look for it there. MCR BUGS provides a database, which can be used as a source to direct collectors to particular areas. Andersen (1993; 1996) has queried whether fossil beetles can be used to retrodict climate, and although his arguments were effectively countered by Coope and Lemdahl (1996), there is no doubt that MCR is currently a very coarse though effective tool. Its development within the BUGS framework should help to dispel any remaining doubts.

Conclusion

Temperature is just one many environmental variables that influence insect distribution. Rainfall and number of daylight hours/day length are perhaps two of the more obvious others that could be considered in palaeoenvironmental reconstructions. Both being essential when crop yields are considered, and even more so in the potentially extreme limits of farming systems (cf. Barlow et al. 1998). BUGS is designed to be expandable, and extra columns fields can be added to the MCR tables to store as many variables as required. In the long term, its various facets should add much to archaeological and palaeoecological interpretation, although it should perhaps be added finally that all palaeoecological reconstruction is only as good as the initial field sampling, and no amount of computer application can replace good field techniques.

BUGS is distributed for free over the net from a website (http://www.umu.se/envarchlab), or by post to those who request it.

Acknowledgements

The two elements currently being combined into BUGS began with Jon Sadler, at the University of Birmingham, and Tim Atkinson at University of East Anglia in Norwich. Both owe much to the primary stimulus of Russell Coope, then at Birmingham, later at Royal Holloway, University of London. All are gratefully acknowledged. Development of the initial access version of BUGS was by Yuan Zhuo Don, whilst an MSc student at the University of Sheffield, and later revisions came from the Genisys development group there. The authors are grateful to Roger Engelmark, University of Umeå for continued interest in the project, and for funding to the Northern Crossroads project devised by Noel Broadbent, also at Umeå. Various parts of the work have been funded by an NSF grant to Tom McGovern, City University of New York, as part of NABO, the North Atlantic Biocultural Organisation, and by the Leverhulme Trust, whom are hereby thanked.

References


Coope, G. R. & Brophy, J. A., 1972. Late Glacial environmental changes indicated by a coleopteran succession from North Wales. Boreas, 1, 97-142.


Kenward, H. K., 1976. Reconstructing ancient ecological conditions from insects remains: some problems and an experi-
mental approach. *Ecological Entomology*, 1, 7-17.


Kenward, H. K., 1999. Public lice (*Pthirus pubis*) were present in Roman and medieval Britain. *Antiquity*, 73, 908-911.


Osborne, P. J., 1980. The Late Devensian-Flandrian transition depicted by serial insect faunas from West Bromwich, Staffordshire, England. *Boreas*, 9, 139-147.


Figure 1. Screen shots of Bugs, showing: (a) the main data screen; (b) bibliography for the selected species; and (c) fossil record for the selected species.

Figure 2. The Site database with the species count sheet for "Reykholt" selected (Buckland et al. 1992). Species names are listed down the left of the count sheet, and samples displayed in columns thereafter. Double clicking opens the sheet in MS Excel ©.

Figure 3. Schematic of a Mutual Climatic Range Diagram, in this case showing the thermal overlap of three species. Tmax is the mean temperature of the warmest month of the year, which the species will tolerate, and TRange the difference between this and Tmin, the mean of the coldest month. The Mutual Climatic Range itself is the area of most

Figure 4. Binary temperature envelope data from the MCR Birmingham program. Each digit represents 1°C, of a grid with the dimensions 0 to 38°C for Tmax and 5 to 65°C for T Range.

Figure 5. Storage of MCR data within Bugs. Distribution data is (a) digitised/entered into (b) Bugs MCR tables, and then (c) ArcView's Spatial Analyst used to calculate TMAX.
(a) Create MCR subsets for each sample

![Site count sheet](image)

![Filter sample for species on MCR list](image)

![Sample MCR subset](image)

Species in sample MCR list

(b) Calculate thermal overlap of species in the sample

![Sample MCR subset](image)

![Bugs MCR tables](image)

![MCR Diagram](image)

(c) Repeat a & b for all relevant samples and then construct time series diagram

![Temperature changes at site over time](image)

Figure 6. How Bugs calculates MCR data for a site. (a) Individual samples are extracted from the count sheets and parsed against the list of MCR viable species in Bugs, to produce a Sample MCR subset. (b) Each species in the MCR subset has an associated MCR table which is used to calculate thermal envelopes, which are then overlaid to produce the MCR diagram. (c) The calculations are performed for every sample in the count sheet, it is up to the researcher to judge the relevance of the results. If the count sheet represents a chronological/stratigraphic sequence then a composite time series diagram can be produced.