The Strategic Environmental Archaeology Database (SEAD)

An International Research Cyber-Infrastructure for Studying Past Changes in Climate, Environment and Human Activities

Introduction

In 2008 the Environmental Archaeology Lab (*Miljöarkeologiska laboratoriet:* MAL) and HUMlab (The Humanities Computing Lab) at Umeå University, Sweden, were awarded a research and development grant for the construction of an international standard research database for environmental archaeology. The funding was provided by the Swedish Research Council's Database InfraStructures Committee (DISC) and forms part of a national and European level drive towards increasing open access to scientific data and results. It was a requirement of the grant that the project be outside the scope of any single research group, and the SEAD team are working to achieve their goals in collaboration with a number of international database and research groups working with scientific archaeological and palaeoenvironmental data.

A large proportion of environmental archaeology data are buried in archaeological file reports (grey literature), the majority of which never receive wide-scale readership and are rarely used in scientific research or method development. SEAD will provide a mechanism for providing free online access to these data. The primary aim is to be able to provide a high quality resource for scientists, an online research infrastructure for environmental archaeology, by the end of 2010, and then build on this over the subsequent two years by adding more datasets and tools.

Need for SEAD

Environmental archaeology is the science of the past interaction of people and their environments. This encompasses a wide spectrum of concepts, including climate change, human impact, sustainable land use and agriculture, cultural landscapes, resource management and availability, household activities, food cultures and many more areas where biological or physical/chemical remains can help us to understand the past. There is an inclination in the work of many environmental archaeologists towards focussing on humans as the central, unifying theme, although it is understood by the majority that an awareness of the natural (or otherwise) surroundings of any group of people is essential when interpreting data. Without an understanding of background signals, or "normal" processes of change, it is difficult to resolve and understand the anomalies caused by human activities. Innovations in agricultural technologies, for example, have had massive impacts on the landscape and have helped to transform Holocene landforms into the largely anthropogenic European countryside of today.

The evidence for these events and their environmental settings is contained in buried sediments and the fossil plants, animals and other organisms which they sometimes contain. When they were alive, the organisms naturally lived within their ecological tolerances, that is they were able to live at a particular place because the environment and climate supported them. Humans have a habit of modifying environments and creating microclimates which support organisms beyond their "natural" geographical ranges, and a number of species have adapted to these, relatively new, possibilities. Geochemical and physical properties of sediments can also give us useful information on the nature of events that affected them prior to and post deposition. By reading these archaeo- and palaeoenvironmental records, and determining their ages, we can therefore interpret the preserved remains in terms of the environment(s) that existed at particular places and times. Geochemistry and soil properties, along with other geoarchaeological techniques, also provide methods for mapping variations which can help us to understand the spatial extent and nature of past land management and archaeological sites. In other words, the fossils and geoarchaeology provide proxies for past changes.

The amount of data collectively provided by the endless hours of field and laboratory work resulting from these methods is considerable. The fossil insect database BugsCEP, for example, which is now part of SEAD, contains over 250,000 records, each of which includes several fields of data. These correspond to something in the order of 2.5 million data elements which represent the large part of the published European fossil beetle record (Buckland & Buckland 2006; Buckland 2007). Even a collection of small samples from a poorly preserved Bronze Age well may result in over 750 records of data for insects, snails, plant macrofossils, pollen and geochemistry. Whilst such relatively small datasets, from a single site, may feasibly be managed in a series of spreadsheets, there are limits to how efficiently analyses and interpretations can be made without solutions for easily cross comparing proxies. Multi-proxy analyses are generally the most informative, as the various data sources complement each other in terms of the different components of the past that they are able to reflect. The usual method is to represent the data in separate diagrams and tables of abundance and magnitude, and then visually compare them. The more proxies, archaeological features, bog and lake profiles in an investigation, the more difficult this becomes. Storing the data in a relational database and providing user-friendly interfaces for management, retrieval, analysis and visualisation greatly improves the power and efficiency with which this process can be undertaken. This is especially so if the system can provide facilities for visualising data from multiple sites on the same diagram.

The true power of databases becomes evident when using data from multiple sites, the amount of data increasing geometrically with each site included in the analyses. Through SEAD, data can be queried and aggregated from any number of sites to help provide answers to complex spatio-chronological questions such as: "How did winter and summer temperatures change across Europe during deglaciation?" (see Buckland 2007), and, "How did agriculture initially spread across Scandinavia?" The system provides the appropriate data and visualisations with which further analyses, statistics and interpretations can be made. Fig. 1, by way of example, shows a set of, originally database linked, maps which were produced when providing data for an English Nature investigation into the evidence for the nature of mid-Holocene woodlands in the British Isles (Buckland *et al.* 2005).

For each and every species found, ecology, distribution and to an extent, ethnographic data are needed in order to understand the meaning of the finds, as is an understanding of the taphonomies and methodologies involved. An extremely time-saving function of SEAD/BugsCEP is the ability to provide a report of the ecology/habitat and distribution of every insect species recorded from a site, or an aggregate collection of samples from multiple sites. Hours, if not days of archive, Internet and literature searching can be saved in this way, and the feature will eventually be in-

UK sites with fossil insects for three time-slices

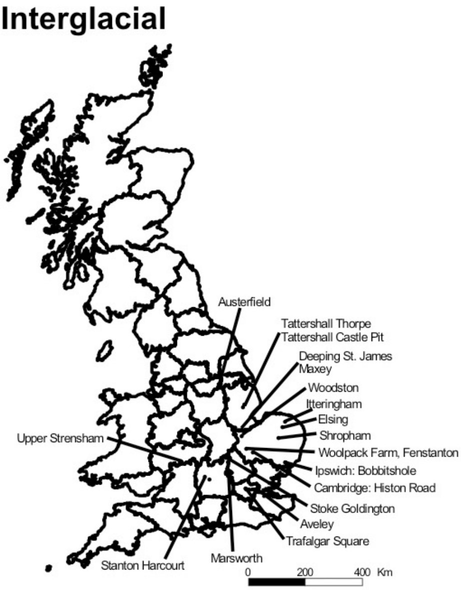


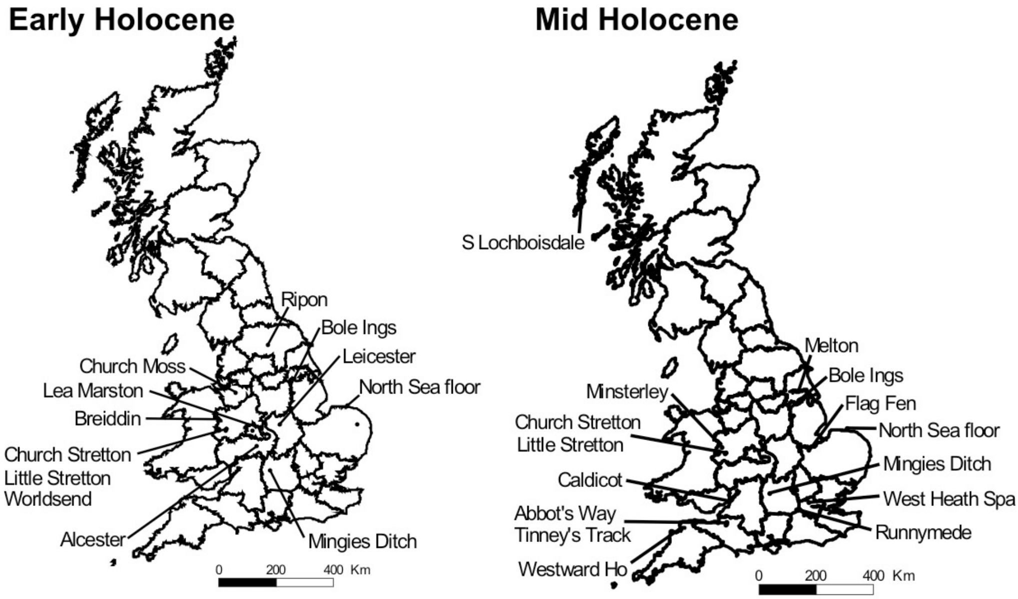
Fig. 1. The results of GIS analysis of fossil insect sites from the UK, looking at C14 dated finds of species indicating woodland environments. SEAD will allow the plant macro and pollen data to be easily added to these analyses for a more complete picture of Holocene woodlands. Modified with permission of the authors (Buckland *et al*. 2005).

troduced for plants and other proxies. SEAD also stores descriptions of the methods used to create all of the data held within it along with a linked bibliography of all original data sources and relevant publications.

Database and Tools Integrated

SEAD is a client-server based research database infrastructure system with both Internet based online components and downloadable, synchronisable data entry and quality assurance facilities. Centralisation of the master database ensures that users will always have access to the latest data. Downloadable components also allow users to securely analyse their unpublished data without having to upload it to the public database. The SEAD system concept includes web-based tools for data retrieval and visualisation, which will be expanded to include additional analysis tools and enhancements with time. All development tools, libraries and software components used in producing SEAD are freely available and the project is part of a series of ongoing innovative ICT solution and method development projects at HUMlab and MAL.

SEAD is stored in a single PostgreSQL 8.3 relational database and was developed by Philip I. Buckland, Erik J. Eriksson and Johan Olofsson (MAL) in consultation with Fredrik Palm (HUMlab), Eric Grimm (Neotoma/Tilia, Illinois State Museum, USA), Brian Bills (Center for Environmental Informatics, Penn State University, USA) and Paul Buckland (Bugs, Sheffield, UK). The structure is highly normalised, in that tables store data in their smallest divisible form where detailed data retrieval



requests are anticipated, with 128 tables. The database also has the capacity to hold data from any type of investigation producing proxy data in continuous, integer or relative values, and is highly expandable, should the storage of new methods be required. The system can also register "master datasets" which ensures that data collections, or ingested (virtual) databases, are assigned appropriate credit and can be extracted independently.

Whilst there is essentially no limit to the potential geographical scope of SEAD, the focus is primarily European, the initial data scope reflecting the spatial extent of the original constituent databases (BugsCEP and MAL; Fig. 2). The potential practical problems associated with variations between the species naming systems used in different countries are reduced by allowing for multiple taxonomic systems and synonyms for each of the ca 18,000 taxa currently stored. With the import of all of the data currently held at MAL, in addition to the BugsCEP virtual database, SEAD will become the largest repository of environmental archaeology data in the World, with the possible exception of the Neotoma database with which the project cooperates (Neotoma 2009).

The developers of SEAD firmly believe in the integration of database and tools for its use from an early stage. This ensures that the database structure can be adapted to any late occurring user needs and if necessary adapt to opportunities afforded by new software and interface related technologies. It also ensures that the project will produce the necessary tools for non-expert users to access the data themselves, rather than have to rely on standard database management software or the assistance

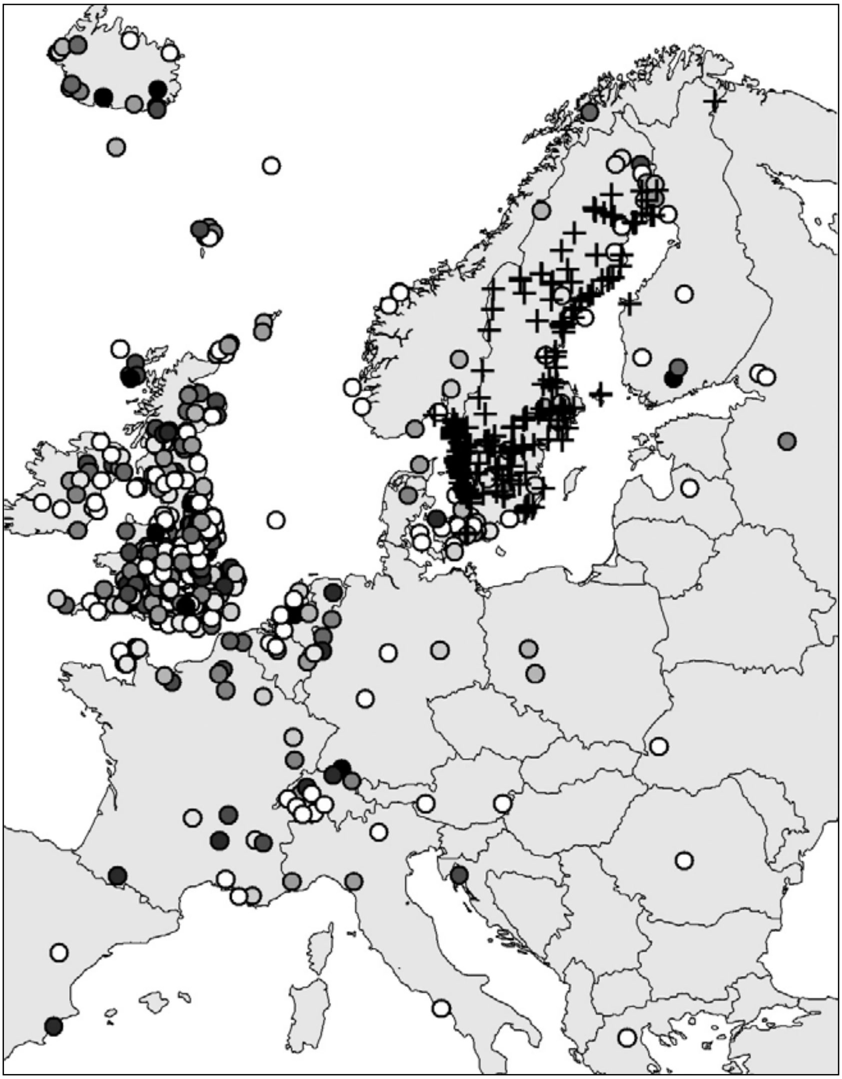


Fig. 2. Map of sites currently stored or scheduled for storage in SEAD, representing the initial datasets extracted and digitised from BugsCEP (circles) and the MAL archive (plusses). Circle point shade is proportional to the ratio *total number of individuals/total number of taxa* per site (quantile classified), a simple site diversity overview made easy with a large database.

of IT staff. Environmental archaeology is implicitly spatially and chronologically aware, and SEAD embraces this concept through a combination of browser, analysis tools and GIS (Geographical Information Systems) functionalities. It is hoped that by centralising this resource, ensuring the quality of data through external clearing houses, and providing free, unrestricted access, an international and multidisciplinary audience will be reached.

Concluding Points and Future Plans

By the summer of 2010 the SEAD project will have released public beta versions of data retrieval interfaces using the faceted browser, map and chronology query systems. The database will have been populated with the BugsCEP data and initial datasets from MAL. It is hoped that these will provide a taste of what the future can hold in terms of integrated, multidisciplinary scientific research infrastructure resources and that the international research community will use them to produce new knowledge on the nature of past climates, environments and human interactions with these. The system will also be integrated into teaching at a number of institutions, and specific teaching and online learning interfaces will be developed in the near future. Future developments will also be coordinated in collaboration with the Neotoma (2009) international consortium for palaeoecology databases.

SEAD CORE TEAM

Environmental Archaeology Lab (MAL), Dept. of Historical, Philosophical & Religious Studies, Umeå University:
Philip I. Buckland, Erik J. Eriksson, Johan Linderholm & Karin Viklund
HUMlab, Umeå University, Sweden:
Fredrik Palm & Patrik Svensson
The SEAD website: http://www.sead.se

ACKNOWLEDGEMENTS

Credit should be given to Paul Buckland (Sheffield, UK), Roger Engelmark (Umeå) and Eric Grimm (Illinois State Museum, USA) for acting as external advisors in various aspects of the project. Russ Graham (Penn State University, USA) has also provided essential feedback on the educational outreach aspects of SEAD. A number of other issues have been resolved by attendees of the "development of analysis and visualisation tools for database orientated environmental archaeological, palaeoenvironmental and palaeoclimatic studies" workshop in Umeå, 2009.

REFERENCES

Buckland, P. C., Buckland, P. I. & Hughes, D. (2005). "Palaeoecological evidence for the Vera hypothesis?" in *Large Herbivores in the Wildwood and Modern Naturalistic Grazing Systems*, K. H. Hodder, J. M. Bullock, P. C. Buckland & K. J. Kirby (English Nature Research Reports 648), Peterborough: English Nature, pp. 62–166.

Buckland, P. I. (2007). *The Development and Implementation of Software for Palaeoenvironmental and Palaeoclimatological Research. The Bugs Coleopteran Ecology Package (BugsCEP)* (Archaeology and Environment 23), diss. Environmental Archaeology Lab., Dept. of Archaeology & Sámi Studies, Umeå: Umeå University.

Buckland, P. I. & Buckland, P. C. (2006). *BugsCEP Coleopteran Ecology Package* (IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series 2006: 116), NOAA/NCDC Paleoclimatology Program, Boulder CO, USA; http://www.ncdc.noaa.gov/paleo/insect.html (Computer software).

Neotoma (2009). "The Neotoma multiproxy Palaeoecology Database;" http://www.neotomadb.org/; access date 28 April 2010.

Philip I. Buckland
Environmental Archaeology Lab (MAL)
Dept. of Historical, Philosophical & Religious Studies
Umeå University
Sweden
phil.buckland@arke.umu.se