European Pollen Database
Newsletter Number 1
Arles - January 1991

Introduction

As we all know, pollen analysis is a time-consuming activity. Pollen-analytical data are thus a valuable scientific resource that should be permanently archived for future generations of researchers. We are living in the 'information revolution' with an ever-increasing availability of powerful personal computers, with the development of computer-based databases and data-retrieval systems, and with the enormous expansion of the primary scientific literature. It is against this background that a European Pollen Database (EPD) is now being developed to provide for all palynologists a permanent archive of the basic data generated by pollen analysts in Europe, a tool for further research on palaeoecological and biogeographical problems at a variety of temporal and spatial scales, and a primary data-source for furthering our understanding of past environmental history at a time when research on Global Change is becoming increasingly important.

The aims of this Newsletter are 1) to outline the history, rationale, financing and computer systems of the EPD, 2) to clarify the relationships between different pollen database projects within Europe, 3) to outline the data-submission and distribution procedures for the EPD, and 4) to present the protocols that will govern the operation of the EPD. There is also a brief questionnaire at the end of the Newsletter that we would ask all recipients to complete and return. We would also request that you bring this Newsletter to the attention of any of your colleagues whom we may have overlooked in our mailing.

History, rationale and financing of the Database

During the closing session of the IGCP 158B project in Cracow during June 1988, a project in which integration of European palaeoecological research had played a prominent role, the establishment of a EPD was discussed as a logical development. A result of this meeting was a consensus that a well-organised database was needed for ecological synthesis at the European scale and as a contribution to the Global Change program. Following these preliminary discussions, Björn Berglund (Lund) and George Jacobson (Orono) agreed to coordinate a workshop to discuss the establishment of a European database. A similar conclusion as to the desirability of establishing such a database was reached independently by the participants in a meeting at Le Puy during September 1987 of those involved in the European Commission Palaeoclimate Program. Joel Guiot (Marseille), Brian Huntley (Durham) and Colin Prentice (Uppsala) were requested to initiate the process. A parallel initiative in North America began at about this time with Eric Grimm (Springfield) as coordinator.

Following a meeting between Berglund, Grimm, Guiot, Huntley and Jacobson during May 1989, a meeting was organised that took place in Frostavallen in Sweden during August 1989. Palynologists from 18 European countries gathered there. Although they agreed upon the importance of a database, several researchers mentioned a variety of practical and ethical problems which had to be solved. In addition, similar problems were discussed at the 12th European Quaternary Botanists Meeting in Czechoslovakia during June 1989.

The Frostavallen meeting also discussed problems of a practical and financial nature concerned with the housing and support of such a database. The proposal by Armand Pons (Marseille) of a series of rooms in the monastery of St. Trophime in Arles, near Marseille was adopted. The meeting appointed an Executive Committee who were to be primarily responsible for seeking funds; this comprised Brigitta Ammann (Bern), Armand Pons and Bill Watts (Dublin). An Advisory Board was also appointed to help with regional

and taxonomic questions and also to consider the ethics of the use of data in the database.

Subsequently, Brigitta Ammann proposed that a questionnaire be sent to the 120 European palynologists listed on the INQUA Holocene Euro-Siberian and Mediterranean Sub-Commissions mailing lists seeking their opinions about certain aspects of the proposed database. This questionnaire was sent during the spring of 1990. General information was included about the proposed EPD and a series of questions were posed.

The 68 responses to the questionnaire that were received before April 1st 1990 were compiled by Yrjo Vasari (Helsinki) and are summarised below.

91%	would provide raw pollen data
57%	want to use the database
22%	preferred a central database only
76%	preferred regional databases
69%	preferred a relational database
19%	had experience with database software

Concerning data accessibility:

28%	thought that all data should have unrestricted access
18% to 63%	thought that data should be available to contributors only / thought that data should be available to non-contributors if the contributors were contacted

During this same period, Armand Pons succeeded in obtaining financial support for a database from the EPOCH Program of the European Commission; the project supported by the EC focussed upon global climate change during the last 30,000 years. The funding period is July 1990 to July 1993 and a sum of 105,000 ECU was obtained by the University of Marseille to develop a centre in its annexe in Arles. The Marseille group also obtained from the Conseil Regional de Provence-Cote d'Azur approximately 250,000 FFr for the purchase of computer and office equipment.

The Advisory Board and Executive Committee gathered for the first time in Wilhelmshaven on 21st and 22nd September 1990. The aims of this meeting were 1) to resolve some of the practical problems involved in starting the database, 2) to further discuss the ethical problems raised at Frostavallen and subsequently, and 3) to propose an organisational structure and protocols for the database.

Eric Grimm, coordinator of the North American Pollen Database, reported that funding of 52,000 US$ had been obtained for the first year of this project from the U.S. National Geophysical Data Center (NGDC), an agency in the National Oceanic and Atmospheric Administration (NOAA). The pollen data centre will be the Illinois State Museum in Springfield, Illinois. The project began on August 15th 1990 and is employing a computer programmer (John Keltner) to develop suitable database applications software.

Brian Huntley reported that grants had been received from the U.K. Natural Environment Research Council that would support data compilations for the late-glacial in Durham and for the last interglacial in Cambridge. These compilations were intended to be contributions towards the central database and would be carried out in close collaboration with the database centre in Arles. Additional funds from the EPOCH program would also facilitate this collaboration.

Brigitta Ammann reported upon the organisation of an Alpine Pollen Database as part of a Swiss NSF-funded project on long-term vegetation dynamics in the Alps.

It was agreed that high data quality must be guaranteed, and that high taxonomic standards and sound radiocarbon chronologies were essential; it was envisaged that this would mean that the database would contain primarily, but not solely, data generated since 1960. It was also agreed that the database must contain the data in their entirety for all sites contributed, but that data from isolated samples of materials such as buried soils, cave sediments and archaeological materials would not be includeed. The database coordinators must cooperate closely with the contributors, with regional advisors and groups and with the Advisory Board in order to resolve pollen-taxonomic questions. Appropriate protocols must be drawn up for these purposes. It was agreed that the nomenclatural conventions used should be those presented by Birks

(1973)* pp. 225-226.

Discussion took place about the structure of the database and especially regarding the issues of synonymy, taxonomy, evaluation of radiocarbon dates, linguistic difficulties and the motivation for the database development. It was agreed that the structure should not be rigid, but that sufficient flexibility must be provided to suit those desiring to cooperate in the development of other databases as well as those wishing to contribute data to the central database. The participants emphasised the special need of some countries for the acquisition of computer hardware and software. A number of suggestions were made of possible sources of financial support for computer acquisition by these laboratories and it was agreed that this matter would be pursued.

A proposed set of protocols for the database were outlined by John Birks. Following discussion these were adopted in a modified form. These protocols, which are reproduced below, attempt to establish the credibility and integrity of the database and to set standards for acceptable use.

Following discussion, it was decided to use the relational database package Paradox 3.5 to build the database. This software is widely available and runs on IBM PC-compatible computers under DOS; it is also being adopted by the North American Pollen Database, thus application software developments can be shared.

At the end of the meeting Armand Pons and Bill Watts indicated their desire to step down from the Executive Committee and nominated Jacques-Louis de Beaulieu (Marseille) and Brian Huntley respectively as their successors; these nominations were accepted by the Advisory Board. Björn Berglund also indicated that he wished to step down from the Advisory Board and nominated Marie-José Gaillard-Lemdahl (Lund) as his successor; this nomination was also accepted.

The Executive Committee met, along with Eric Grimm and several technical advisors and coordinators (John Birks, Sytze Bottema, Mike Field, Annabel Gear, Joël Guiot, Steve Juggins and John Keltner) in the St. Trophime monastery between the 5th and 8th January 1991. A proposed technical structure for the Paradox database was presented by Eric Grimm. After modification following discussion, this was adopted for the EPD. A number of other practical matters were resolved and the participants prepared this first Newsletter.

Relationships beween different pollen-database projects

The European Pollen Database project is centralised in Arles, France. Its personnel consist of a post-doctoral scientist (to be appointed soon) working closely with Joël Guiot and Jacques-Louis de Beaulieu. The EPD has an Executive Committee responsible for overseeing the development of the project and for finding further financial support for the project. There is also an Advisory Board responsible for protocols, assisting with pollen-taxonomic and related problems, and arbitration in any disputes beween contributors, users, or coordinators. The composition of the Executive Committee is currently Brigitta Ammann, Jacques-Louis de Beaulieu and Brian Huntley. The Advisory Board consists of Karl-Ernst Behre (Wilhelmshaven), John Birks (Bergen), Elizaveta Bozilova (Sofia), Maria Follieri (Rome), Marie-José Gaillard-Lemdahl, George Jacobson Jr., Roel Janssen (Utrecht), Meilute Kabailiene (Vilnius), Magdalena Ralska-Jasiewiczowa (Cracow) and Jim Ritchie (Toronto and Taunton).

In addition there are at present three directly related database subprojects and one indirectly related database project. The three subprojects are concerned with specific geographical areas or temporal aspects and were, at their outset, designed to be integral components of the EPD. They consist of:

1. Eastern Mediterranean database coordinated by Sytze Bottema (Groningen) as part of the EPOCH project on Global climate change of the last 30,000 years.

2. 9000 - 15,000 B.P. ('late-glacial') of Europe coordinated by Annabel Gear (Durham) and Brian Huntley as part of their project on European Palaeoclimate

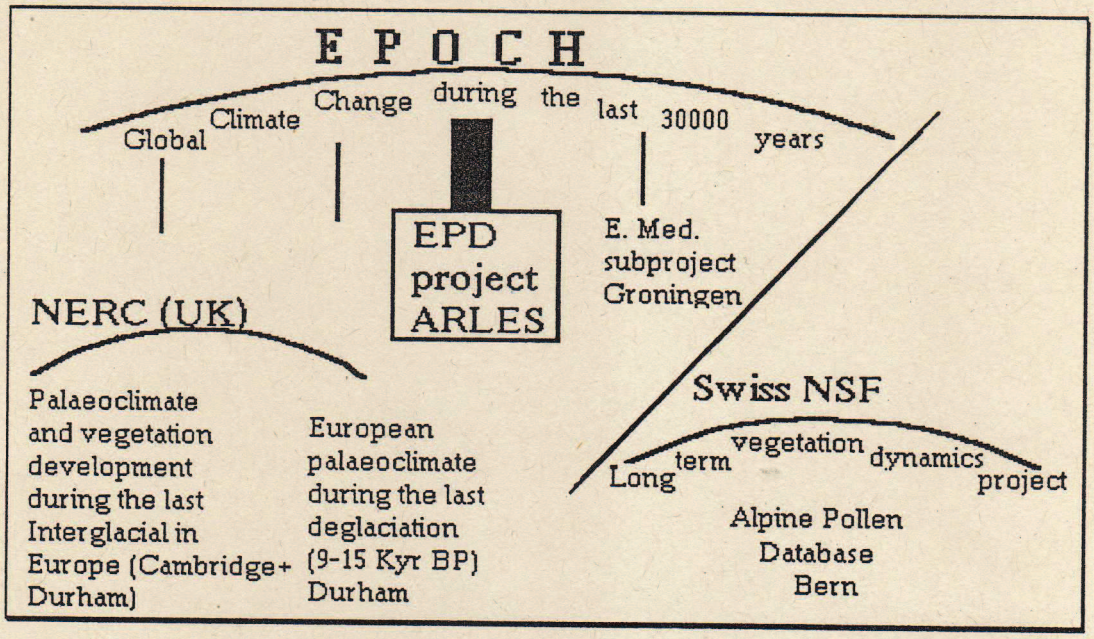
* Birks, H.J.B. (1973) *Past and present vegetation of the Isle of Skye - a palaeoecological study.* Cambridge University Press, London.

during the last deglaciation.

3. Last interglacial of Europe coordinated by Mike Field (Cambridge), Phil Gibbard (Cambridge) and Brian Huntley, as part of their project on Palaeoclimate and vegetation development during the last interglacial in Europe.

The indirectly related project is the Alpine database. This is part of a larger project specifically focussed on long-term vegetational dynamics in the Alps and immediate surroundings coordinated by Brigitta Ammann, John Birks, Otto Hegg (Bern), Steve Juggins (Bergen and London) and Felix Kienast (Zurich). It was not originally designed to be an integral component of the EPD. However, it is attempting to achieve the maximum possible compatibility with the EPD in terms of database design, software, etc.

The current relationships between these various activities can be summarised diagramatically:



Because of the structure that has evolved, it is clearly possible for a scientist to be approached by more than one project or subproject. The extent and nature of collaboration with these different activities are the decisions of individual scientists and may vary from one person to another.

The basis of all collaboration is open communication. The EPD and its three subprojects are using this and subsequent Newsletters as their primary means of communication to individual scientists. Because of the small number of laboratories involved, the Alpine project holds meetings and pollen-taxonomic and computing workshops about every six months as a means of communication with individual scientists.

The EPD closely collaborates with the NOAA NGDC American pollen database project coordinated by Eric Grimm and John Keltner, particularly in connection with computer software development.

Procedures for data submission and distribution

Enclosed with this Newsletter is a Questionnaire that should be returned to the EPD centre at Arles. Amongst the questions posed is one that asks whether you are willing to contribute data, and if so to indicate for which site(s). You are also asked to indicate for each site its location and the approximate time span that it represents. When your reply is received, any sites that you are willing to contribute will be sorted geographically and stratigraphically and the relevant data compiler(s) will be notified. The compiler responsible for any given site will then, using the published work that is available to them, begin to complete a data sheet that requests the information required for entry to the database. This sheet will then be sent to you for completion and checking, and should be

returned along with the pollen and/or macrofossil data for the site. Upon receipt by the compiler your data become subject to the EPD protocols. These data may be sent either on paper (e.g. photocopies of count sheets or data tables) or by electronic means (e.g. on diskette). Whichever compiler handles any given site, the data will be incorporated into the central database in Arles and you will receive a printout of the database entry.

It is intended that the database and the associated applications software will be available for distribution to users as provided for by the protocols (see below). This release of the database will not take place piecemeal, however. The first release is estimated to be available in about two years from now; subsequent updates will be released at intervals of 6 months to 1 year. Users who request a copy of the database will be charged only for the cost of the magnetic media (diskettes).

The database compilers and coordinators have agreed that they will regard all data that are added to the database as restricted (see protocols) until they appear in a version of the database that has been released to other users.

Protocols for the European Pollen Database

22nd September 1990 - Wilhelmshaven

The following protocols for the database were discussed and agreed by the Advisory Board at its meeting in Wilhelmshaven. In the event that it becomes necessary to modify them in the future, all database contributors and users will be notified and the changes will be published in the Newsletter.

A. Data

1. Data must consist of the original counts, not percentages or digitized data.

2. Database must contain the original taxonomic identifications, with exceptions of exact nomenclatural synonymy. Taxa will not be lumped into higher taxonomic groups in the database.

3. Data will be classified as unrestricted or restricted. All data will be available in the database. In other words, the central database will distribute all data, restricted and unrestricted. Thus, restricted data can be viewed by a user, but cannot be used except as provided below.

4. Unrestricted data are available for all uses.

5. Restricted data may be used only by permission of the data originator. Appropriate and ethical use of restricted data is the responsibility of the data user.

B. Contributors

1. Can declare data unrestricted or restricted.

2. Can ask to verify that data in the database are correct. As a matter of general policy, the central database should routinely return to the data originator a hardcopy printout of the data as they are entered in the database for optional verification by the originator.

3. May use any unrestricted data.

4. Can obtain copies of application software and the database itself for use on his/her own computer.

5. Should receive a periodic newsletter or report concerning the database.

6. Can ask at any time that his/her data be withdrawn from the database or that their status (unrestricted or restricted) be changed.

7. In the case of a dispute regarding inappropriate use of restricted data, the Advisory Board will serve as arbitrator.

C. Users

1. Must ask permission from the data originator for use of restricted data.

2. Should, as a matter of courtesy, inform data originators of the uses being made of their data.

3. If the contributor wishes, should show the contributor results of analyses and manuscripts for publication for critical comment.

4. Should cite, in any publication using data from the database, the contributors' original publications describing their data.

5. Should send contributors reprints of publications that use their data.

6. Should acknowledge contributors for use of unpublished data and for any advice they may have provided.

7. No user can pass data on to another party. All users must obtain data from the central database.

8. Normal ethics apply to co-authorship of publications. The contributor should be invited to be a co-author if a user makes significant use of a single contributor's data, or a single contributor's data comprise a substantial portion of a larger dataset analyzed, or a contributor makes a significant contribution to the analysis of the data or to the interpretation of the results. This guideline applies to unrestricted as well as to restricted data.

9. The data are available only to non-profit-making organizations and for research. Profit-making organizations may use the data, even for legitimate uses, only with the written consent of the Advisory Board, who will determine or negotiate the payment of any fee required.

D. Executive Committee and Database Coordinators

1. Should prepare a periodic newsletter or report about every six months for contributors, users, the Advisory Board, etc.

2. Must follow the same protocols that apply to all other users concerning the use of data.

3. Must closely cooperate with the data originator and/or relevant Advisory Board member(s), regional correspondents or taxonomic advisors when making taxonomic decisions.

4. Should assemble a mailing list of Quaternary palynologists in Europe and others associated with European data, and should inform them of the opportunity to contribute to and participate in the database development. In addition, should announce the development of the database in appropriate newsletters and publications.

5. Should incorporate all data into the database, subject to certain minimum requirements, without assignment of quality.

6. Should organize workshops on matters related to the database and should work to facilitate acquisitions of hardware and software by laboratories not having access to these.

7. Should send the protocols to all potential contributors and users.

In addition to abiding by these protocols, the database compilers and coordinators have agreed that they will regard all data that are added to the database as restricted (see above) until they appear in a version of the database that has been released to other users.

Questionnaire

1. Are you willing to cooperate with the Database Centre?
Yes/No

If you are willing to cooperate, then in what way?

as a data contributor Yes/No
as a national/regional correspondent Yes/No
as a member of a national/regional working group
Yes/No
as a pollen taxonomic advisor Yes/No
as a software developer Yes/No
as a database user Yes/No
in some other way (please specify).....

2. If you are willing to contribute data to the Database then please provide an indication of the number of sites that are available and their localities and age spans.

3. If you wish to be a database user then please provide an indication of the uses that you would envisage.

4. Do you use, or have access to, an IBM PC-compatible computer with a hard disc? Yes/No

If YES then do you wish to receive a copy of the data-entry software? Yes/No

(If YES then please enclose a formatted DOS PC diskette (5.25 or 3.5 inch) with your reply.)

If NO then do you use, or have access to, some other type of computer? Yes/No
If YES then please indicate the type of computer.....

5. Please enter below (or upon a separate sheet) any comments that you wish to make with respect to points in the protocols.

6. Please enter below (or upon a separate sheet) any general comments that you wish to make with respect to the database.

7. Do you wish to receive future issues of the Database Newsletter?
Yes/No

Please enter below your name and address in the form in which you prefer it to be used.

Name:
Address:

Telephone:
Fax:
E-Mail:

Please photocopy this questionnaire as necessary if more than one reply is needed.

SEND YOUR REPLY AT THE ADDRESS OF THE EDP:
European Pollen Database
Centre Universitaire d'Arles, Espace Van Gogh
13637 Arles, France