Pollen Database Manual

Global Pollen Database
African Pollen Database
Alpine Pollen Database
Base de données polliniques et macrofossiles de Québec
Chinese Pollen Database
Czech Pollen Database
Eastern Mediterranean Pollen Database
Eastern Siberian and Far East Pollen Database
European Pollen Database
Indo-Pacific Pollen Database
Last Interglacial Pollen Database
Late-glacial Pollen Database
Latin American Pollen Database
Moscow Pollen Database
North American Pollen Database

Pollen Database Manual

08 August 1996

Pollen Database Manual

1. Introduction

1.1. Overview

This manual is a guide to the Global Pollen Database (GPD) as well as its immediate precursors such as the European and North American Pollen Databases (EPD and NAPD). These databases are intended to be both archives of pollen and associated data as well as important research tools for studies in paleoecology and paleoclimatology.

Initiatives to develop comprehensive, archival pollen databases in Europe and North America began about the same time, in the late 1980's. Although separate databases, with separate internal identification systems, were developed, the organizers intended for the design and implementation of the databases to be highly compatible, such that any software developed would work with either database.

The establishment of a World Data Center-A for Paleoclimatology in coordination with the Past Global Changes (PAGES) project (IGBP 1992) combined with the power of the Internet as a means of distributing data furnished the momentum for the creation of a Global Pollen Database as the superset of regional efforts. Regional databases or database cooperatives do the hard work of preparing the data that are then passed on to the GPD for insertion into the database tables and distribution to the world's scientific community. These final steps in the process are performed by the World Data Center-A for Paleoclimatology located at the National Geophysical Data Center in Boulder, Colorado, USA.

The GPD and the WDC-A for Paleoclimatology provide users with a uniformly organized set of data products that are available free and without restriction, 24 hours a day over extremely fast data connections. The GPD already contains data from North America, Central and South America, and northern Asia. Work in progress will add data from the Indo-Pacific region, central Asia, and Europe.

The database itself is PC-based, currently running under *Paradox*® (version 4.5 for DOS/ Windows) from Borland International. Users of the database, however, are not required to have *Paradox*. *Paradox* tables can also be read and manipulated by other database products such as Microsoft *Access*®. Planning has begun to make the data available in *Oracle*® as well.

1.2. History

Before the initiation of the NAPD/EPD, Thompson Webb III at Brown University had already assembled a substantial database of pollen counts from North America as part of the Cooperative Holocene Mapping Project (COHMAP). In Europe, Brian Huntley and John Birks had assembled pollen-percentage data at given time intervals, which they used for their atlas (Huntley and Birks 1983). These databases have been used for many valuable paleoclimatic and paleoecological studies, culminating in the COHMAP paper in *Science* (COHMAP Members 1988; Wright et al 1993). Such studies have demonstrated not only the extreme value of the pollen databases, but also the need for greater accessibility and completeness. Also of concern are the long-term archiving and curation of pollen data, which are usually published in only summary or graphical form and whose collection is highly labor intensive. These needs and concerns fostered the creation of the North American and European Pollen Databases.

The need for a European pollen database was discussed during the closing session of the International Geological Correlation Programme (IGCP) 158B project in Krakow, Poland, in June 1988. Following these preliminary discussions, Björn Berglund (Lund, Sweden) and George Jacobson (University of Maine, USA), who was on sabbatical in Lund, agreed to coordinate a workshop to discuss the establishment of a European database. Independently, participants in a meeting of the European Commission Palaeoclimate Program reached a similar conclusion in Le Puy, France, in September 1987. Joël Guiot (Marseille, France), Brian Huntley (Durham, England) and Colin Prentice (Uppsala, Sweden) agreed to initiate the process. In North America, Eric Grimm (Illinois State Museum) had been having conversations with Tom Webb concerning the curation and fate of the COHMAP pollen database, including preliminary discussions of a grant proposal, although no action had been taken.

In May 1989, Berglund, Grimm, Guiot, Huntley, and Jacobson met and organized a meeting for the following August in Frostavallen, Sweden, which palynologists from 18 European countries attended. Although the participants agreed on the importance of a database, several researchers brought forth a variety of practical and ethical problems, some of which had been discussed the previous June at the 12th European Quaternary Botanists meeting in Czechoslovakia. The Frostavallen meeting also discussed practical problems concerning the housing and financial support of such a database. Armand Pons and colleagues Joël Guiot and Jacques-Louis de Beaulieu (Marseille) presented a proposal, which was subsequently adopted, for housing the database at a new Centre Universitaire d'Arles in the Monastery of St. Trophime in Arles, near Marseille. The meeting appointed an Executive Committee, composed of Brigitta Ammann (Bern, Switzerland), Armand Pons, and W. A. Watts (Dublin, Ireland), who were primarily responsible for seeking funds. An Advisory Board was also appointed to help with regional and taxonomic questions and to consider the ethics of the use of data in the database.

Armand Pons succeeded in obtaining financial support for a database from the European Commission's EPOCH Program, which is focused on global climate change during the last 30,000 years. The original funding period was July 1990 to July 1993 and a sum of 105,000 ECU was obtained by the University of Marseille to develop a center at 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. Independently, Brian Huntley received a grant from the U. K. Natural Environment Research Council to support data compilations in Durham for the late-glacial and in Cambridge for the last interglacial. These compilations were intended to be contributions towards the central database and would be carried out in close collaboration with the database center in Arles. Additional funds from the EPOCH program would also facilitate this collaboration. In addition, Brigitta Ammann and colleagues from the Alpine region had organized an Alpine Pollen Database with funds from a Swiss NSF project on long-term vegetation dynamics in the Alps.

The EPD Advisory Board and Executive Committee met in Wilhelmshaven, Germany, in September 1990. The aims of this meeting were to resolve some of the practical problems involved in starting the database, to discuss further the ethical problems raised at Frostavallen, and to propose an organizational structure and protocols for the database. Discussions centered on the structure of the database, taxonomy, synonymy, evaluation of radiocarbon dates, linguistic difficulties, and the motivation for the database development. It was agreed that the central database would cooperate with those desiring to develop regional databases, and that these in turn would serve as conduits to the central database. At the end of the meeting, Armand Pons and W. A. Watts resigned from the Executive Committee and nominated Jacques-Louis de Beaulieu and Brian Huntley as their successors, which the Advisory Board accepted.

In the United States, the National Oceanic and Atmospheric Administration (NOAA) initiated a Paleoclimatology Program in its agency the National Geophysical Data Center (NGDC) as part of its Climate and Global Change Program. NOAA/NGDC appointed a Paleoclimate Advisory Panel, whose mission was to make recommendations for the future of the Paleoclimatology Program. The Panel met for the first time in June 1989 and recommended that acquisition of paleoclimate proxy data be a major priority. It further recommended that experts in the various fields should manage the data acquisition, rather than NGDC itself. The Panel identified the following data areas: (1) pollen/packrat midden/ macrofossil data, (2) tree-ring data, (3) marine and lacustrine sediments, (4) documentary and long instrumental data sets, (5) ice core and glacier records, (6) paleoclimate model output, and (7) hydrology/stream/lake-level data. Pursuant to these recommendations, NOAA/NGDC solicited proposals for development of databases. Eric Grimm submitted a successful proposal for the development of the North American Pollen Database. The project began in August 1990. John Keltner was employed as the Database designer/ programmer, with major responsibility for development of database software for both the NAPD and EPD. The NAPD also has an Advisory Board of palynologists representing different geographic regions and with a broad range of expertise.

Keltner and Grimm created an initial design for the database, which they presented to the first NAPD Advisory Board Meeting in November 1990. In January 1991, Grimm and Keltner met with several technical advisors and coordinators of the EPD (Brigitta Ammann, J.-L. de Beaulieu, John Birks, Sytze Bottema, Mike Field, Annabel Gear, Joël Guiot, Brian Huntley, Steve Juggins) in Arles. This workshop resulted in substantial modification of the database structure to accommodate both North American and European idiosyncrasies, as well as to accommodate other kinds of data, such as plant macrofossils, and diatoms. The participants agreed that any future modifications would be made to both databases after joint consultation.

The Global Pollen Database evolved directly from the North American Pollen Database as new data became available from Latin America, Asia, and the Indo-Pacific region and the World Data Center-A for Paleoclimatology was organized in Colorado. It became obvious to all involved that the work required to unify pollen data from around the globe (primarily in developing a global pollen-type hierarchy) would be less costly than maintaining distinct regional databases, particularly for its users.

1.3. Database Administration

1.3.1. European Pollen Database

Location	Centre Universitaire d'Arles, Arles, France
Coordinators	J.-L. de Beaulieu Joël Guiot
Postdoctoral Scientist	Rachid Cheddadi

Executive Committee

J.-L. de Beaulieu, Marseille, France
Brian Huntley, Durham, England
Brigitta Ammann, Bern, Switzerland

Advisory Board

Sheila Hicks, Oulu, Finland (Chairman)
H. J. B. Birks, Bergen, Norway
Sytze Bottema, Groningen, The Netherlands
Elizaveta Bozilova, Sofia, Bulgaria
George L. Jacobson Jr., Orono, Maine, USA
C. R. Janssen, Utrecht, Netherlands
Meilute Kabailiene, Vilnius, Lithuania
Henry Lamb, Aberystwyth, United Kingdom
Thomas Litt, Leipzig, Germany
M. Ralska-Jasiewiczowa, Kracow, Poland

1.3.2 North American Pollen Database

Location	Illinois State Museum Sprinfield, Illinois, USA
Administrator	Eric C. Grimm Illinois State Museum
Administrator	Stephen C. Porter Illinois State Museum

Advisory Board

Konrad J. Gajewski, University of Ottawa
George L. Jacobson Jr., University of Maine
Glen M. MacDonald, McMaster University
Louis J. Maher, University of Wisconsin
Vera Markgraf, INSTAAR
Pierre Richard, Université de Montréal
Thompson Webb III, Brown University
Cathy Whitlock, University of Oregon

1.3.3 Global Pollen Database

Location	World Data Center-A for Paleoclimatology, National Geophysical Data Center, Boulder, Colorado, USA
Coordinator	Eric C. Grimm Illinois State Museum
Designer/Programmer	John Keltner NOAA/Paleoclimatology Boulder, Colorado, USA

1.4. Cooperating Database Projects

Unless otherwise noted, the projects below are ongoing.

1. Base de données polliniques et macrofossiles de Québec (BDPMQ) developed by Pierre Richard and Alayn Larouche at the University of Montréal.

2. The Alpine Database, which is part of a larger project focused on long-term vegetational dynamics in the Alps and immediate surroundings coordinated by Brigitta Ammann, H. J. B. Birks,

Otto Hegg (Bern), Steve Juggins (Newcastle), Felix Kienast (Zurich), and Pim van der Knapp (Bern). This database was conceived before the EPD and was not originally designed to be an integral component of it. However, it is attempting to achieve compatibility with the EPD in terms of database design and software.

3. Eastern Mediterranean database coordinated by Sytze Bottema (Groningen) as part of the EPOCH project on Global climate change of the last 30,000 years. This project is complete and its data have been transferred to the European Pollen Database.

4. 9000-15,000 B.P. ('late-glacial') of Europe coordinated by Annabel Gear and Brian Huntley (Durham) as part of their project on European Paleoclimate during the last deglaciation. This project is complete and its data have been transferred to the European Pollen Database.

5. Last interglacial of Europe coordinated by Mike Field (Cambridge), Phil Gibbard (Cambridge) and Brian Huntley, as part of their project on paleoclimate and vegetation development during the last interglacial in Europe. This project is complete and its data have been transferred to the European Pollen Database.

6. Indo-Pacific Pollen Database coordinated by Geoff Hope at the Australian National University in Canberra, Australia.

7. Latin American Pollen Database coordinated by Vera Markgraf and Lysanna Anderson at the University of Colorado in Boulder, Colorado, USA.

8. The development of a Chinese Pollen Database is being coordinated by Kam-biu Liu at Louisiana State University in Baton Rouge, Louisiana, USA.

9. The development of an Eastern Siberian and Far East Pollen Database is being coordinated by Pat Anderson at the University of Washington in Seattle, Washington, USA.

2. The Database Management System: *Paradox*®

2.1. Overview of *Paradox*

Paradox from Borland International is a Relational Database Management System (RDBMS) for MS-DOS based microcomputers. An RDBMS is one that adheres (fully or partially) to the relational model of data management (see Codd 1970, 1985, 1990; Date 1990). Users of an RDBMS perceive their data as a set of tables, and only as a set of tables. Tables represent classes of real-world objects; they are the familiar two-dimensional structures consisting of columns (or fields) that contain exactly one type of information and rows (or records) that represent individual instances of the table class. For example, a table designed to store information about sites and their locations might contain columns for the site's name, its latitude, and its longitude. Individual rows in this table would then store the name, latitude, and longitude for a single site; each row is an instance of the site-location class.

Every row in an RDBMS table should be unique. Thus, some set of values, drawn from one or more columns, should distinguish each row from all other rows. The set of columns whose values uniquely identify each row in a table defines the primary key for that table. The combination of table name, column name, and primary key gives each cell in a table a unique address. The columns comprising the primary key are specified when a table is first created. With *Paradox*, this is done by listing the primary key columns in sequence at the top of the table and adding an asterisk to each column's data type specification (see below). If a primary key is subsequently modified, *Paradox* removes any duplicate rows (rows with identical values in each of their primary key columns) from the table with the modified key, and places them in a new table named, KEYVIOL.

This last point illustrates a fundamental property of an RDBMS, namely that all operations produce new tables from old ones. To create or modify a table, you specify its structure in a STRUCT table (the table names, such as

STRUCT, are specific to *Paradox*). Modifying the structure of a table may, as noted above, produce a KEYVIOL table. Viewing a table produces a read-only copy of that table. Editing a table produces a copy of that table that will replace the existing table when you make the changes final (when you DO_IT!, in *Paradox for DOS*; *Paradox for Windows* tables are in read/write mode by default when viewed). Querying a database produces an ANSWER table. Deleting rows from a table using a query form produces a new table with the deleted rows removed to a DELETED table. Any of these tables, except STRUCT, may be empty (contain zero rows).

2.1.1. Features

2.1.1.1. PC based, hardware requirements

Paradox for DOS version *4.5* runs on any 100% IBM-compatible, protected-mode capable 80286, 80386 (or higher processor) personal computer with one hard disk and a floppy drive. The system must have at least 2Mbytes of extended memory and be running DOS 3.0 (or Windows 3.1 or OS/2 2.0) or higher.

2.1.1.2. Forms, reports, scripts

Paradox provides tools for organizing the collection, retrieval, and display of data. These tools include report and forms generators, and a programming language that is an extension of the *Paradox* interface and whose scripts (programs) are played (run) from the interface.

Forms allow a user to see only the data needed for a particular task, whether that task is data entry, editing, or merely viewing. This keeps the actual complexity of the database out of sight and out of the way. A form may contain selected information from one or more tables, text for titles or prompts, display only or calculated fields. With forms, the presentation of the database becomes more consistent and reliable.

Reports are the means for producing printed output of selected data in the database. *Paradox* reports may be arranged in almost any manner. They may include titles and other text, calculated fields, groupings, and much more. Users can design custom reports or use defaults. Custom reports can be saved and reused or modified.

Scripts are executable programs written in PAL (the *Paradox A*pplication *L*anguage). Playing a script may invoke a complete, complex application or a simple, keystroke-saving macro. The former is created using a text editor and PAL commands, whereas the latter are usually created using *Paradox's* macro recording capabilities. Scripts have an important role in providing an interface between the user and the underlying database. Common tasks should always be placed under the control of tested and trusted scripts. Like forms, scripts provide the user with a consistent and helpful interface that restricts the view of the database to just those data required to complete a given task. With scripts, data can be filtered (for example, converted to all uppercase) where appropriate, saving the user from having to remember details of data formatting.

Scripts have a further role in database maintenance: they can help protect the database from inadvertent errors. Simple changes often have side-effects that are easily overlooked when directly manipulating the tables. Scripts are the best way to control undesirable side-effects. Lastly, although a poorly written script can corrupt the database as thoroughly as any user, repairing the damage caused by a script is often easier since the nature of the damage is more readily assessed. **All routine interactions with the database should be mediated with scripts.**

2.1.1.3. Table creation, naming tables and fields

A table is created by assigning it a name and then listing the names and data types of each of its columns. To *Paradox*, table names are one to eight, alphanumeric characters, and follow the rules for MS-DOS file names; case is ignored. Column names may be up to 25 alphanumeric characters in length, embedded spaces are allowed; again, names are not case sensitive.

2.1.2. Data types

The database structure is documented table by table in section 8, below. For each table the fields are listed, along with an indication of their data type and brief indication of what they contain.

Paradox data types (also called field types) must be one of the following:

S = a short integer (-32766 <= *x* <= 32767).

N = a floating point number (15 significant digits, $10^{-307} <= x <= 10^{+308}$)

A*n* = an alphanumeric string with a fixed length of *n* characters (1 <= *n* <= 255). The string can include letters, numbers, and special symbols.

M*n* = Memo, an alphanumeric string of practically unlimited size, of which up to *n* characters (1 <= *n* <= 240) are visible in table view.

D = a date between 1 Jan 100 and 31 Dec 9999.

$ = currency. These differ from floating point numbers only in their default display.

2.1.3. Keys

All tables in the database have a Primary key. Additionally, some tables have Alternate and Foreign keys. Primary keys uniquely identify each row in a table; an Alternate key is any other column (or combination of columns) that could also uniquely identify each row; it is an alternate primary key. A Foreign key is any column (or combination of columns) whose non-null values are constrained to be values in the primary key of some table (possibly the same table).

By way of example, the Pollen Database ENTITY table (an entity is a core, section, or surface sample collected from some site and from which samples are drawn) has a primary key made up of a single, integer-valued column (E#). *Paradox* therefore guarantees that no two rows in the ENTITY table will have the same value in the E# column - every row is uniquely identified. ENTITY has another column (Sigle) that stores a 3-8 character, short name for the entity whose values should also be unique; Sigle is an Alternate key. ENTITY has another column, Coll#, that stores an identifier for the collector of the entity. The non-null values in the Coll# column are constrained to be values that currently exist as the primary key of some row in table WORKERS (i.e. there must be some worker identified by the value in Coll# in table WORKERS). Column Coll# is designated a foreign key referencing table WORKERS. A table that meets all foreign key constraints is said to have referential integrity. Unfortunately, *Paradox for DOS* provides no direct support for maintaining alternate keys and referential integrity (but see section 2.1.5. for a description of the set of **system** tables that have been developed).

2.1.4. Configuration to International Sort Order

Paradox arranges or selects records in a certain order, which must be either: ASCII, Norwegian/Danish, Swedish/Finnish or International. By agreement, all pollen databases use International sort order. "Intl (International) sort order combines capital, lowercase and special characters into a unified sorting order so that records are arranged according to their alphabetical position, irrespective of case or diacritical marks. Dual characters, such as ü are sorted correctly; that is, ü is sorted as if it were "ue")" p. 680 *Paradox* 4.5 Users Guide. Sort order affects the primary key ordering of records in a table, the ordering of records in the ANSWER table, the result of using the SORT command, and the answers to queries that use selection operators (<,>, <=, >=).

It is essential that *Paradox for DOS* is configured to the International sort order (see p. 686 of the *Paradox* 4.5 Users Guide) and that it is not changed thereafter. *Paradox for Windows* should be configured to Paradox ‘intl’ sort order.

2.1.5. Limitations

Although *Paradox for DOS* has great advantages such as widespread availability, low-cost, and ease-of-use, it is not without its limitations as well. Chief among these are the lack of a central system catalogue for the database, no support for maintaining the consistency of alternate and foreign keys, and no provision for transaction logging or rollback. Unfortunately, at the time *Paradox* was first chosen these highly desirable features were only available in much more expensive and difficult to use database management systems. The alternative to having such features built into an RDBMS is to emulate them with external software and to use this software diligently to check the status of the database.

A system catalogue stores information about each table in the database. It stores information on the names, data types, and domains for each column. It stores information on a table's primary key as well as information on any alternate and foreign keys that a table has. A set of system tables (e.g. SYSCAT) have been created for the Pollen Database to stand-in for a DBMS-managed system catalogue. These tables, in conjunction with a set of system scripts provide Pollen Database administrators with the ability to check alternate and foreign key integrity, as well as the specifications of primary keys and column data types.

Transaction logging means keeping a record of all actions that affect the content of the database: every insertion, deletion, or update. The existence of such a log, obviously, greatly facilitates recovery from simple data entry errors or such catastrophes as power failures and file damage. A transaction is any series of actions on a database that must occur in its entirety or not at all. Should a transaction fail, for whatever reason, rollback means playing the log in reverse until the database is returned to its pre-transaction state. Implementing transaction logging and rollback in application software would be both complicated and time-consuming. Adding such a facility may be worth considering for the most routinely used or critical applications. Alternatively, frequent, incremental backups of the database may be a more appropriate and achievable solution - the worst catastrophes may then be limited to a few days lost work.

2.1.6. Common disk directory structure

The database uses a standard directory structure (the actual disk drive used is not mandated).

Database Directory Structure

```
C:\
|--...
|--PDB          parent directory for pollen database
   |--DB          database tables (e.g. SITELOC.DB)
   |--DBV         database views (e.g. P_VARIS.DB)
   |--DOC         documentation (e.g. PDM.DOC)
   |--EXE         DOS executables (e.g. PD2TIL.EXE)
   |--SC          parent directory for scripts
      |--CMN        script modules (e.g. DTYPE.SC)
      |--DB         base table scripts (e.g. PD2TIL.SC)
      |--...
      |--SYS        system table scripts (e.g.
   |--...       CHECK_DB.SC)
   |--SYS
   |--TMP         system tables (e.g. SYSCAT.DB)
                  temporary tables; error reports
```

3. The entry of pollen data using *Tilia*

3.1. Requirements, source, and cost

While the totality of pollen data are best managed using an RDBMS, the data for a single site are much more conveniently manipulated and displayed using the MS-DOS-based applications, *Tilia* and *Tilia·graph*, written by Dr. Eric Grimm. This software requires an IBM PC or compatible computer, MS-DOS 3.0 or higher, 640 kb of conventional memory, and a graphics display for *Tilia·graph*. *Tilia* is available free of charge; *Tilia·graph* costs 250 US$. Both are available from Eric Grimm at the Illinois State Museum, Research and Collections Center, 1011 East Ash Street, Springfield, IL 62703, USA.

3.2. The *Tilia Software Package*

The *Tilia Software Package* consists of a set of programs for manipulating, analyzing, and graphing stratigraphic data. The programs are particularly designed for pollen data, but can be used for a wide variety of stratigraphic, geologic, and ecological data. The programs are all menu driven.

Tilia is a spreadsheet program designed especially for stratigraphic data. Data can be entered, deleted, and easily rearranged. The user can define up to 26 different sums or groups, and percentages can be based on any sum. The program can read and write a variety of different ASCII formats, and the program will assign ages to levels by one of three options: linear interpolation, fitting a polynomial, or a spline curve. Radiocarbon or other ages can be entered. The program can also calculate concentrations and accumulation rates. A standard dictionary of variable names can be maintained. The dictionary can be loaded or linked, so that variable lists do not have to be repeatedly reentered.

Version 2.0 of *Tilia* introduced *Tilia Forms*. The *Forms* provide a means of recording such important information as site details (sitename, latitude, longitude, etc.) along with the pollen data themselves. A number of such forms have been developed for pollen data: site details, samples, ^{14}C dates, core drives and sections, lithology, loss-on-ignition, chronology, and publications. Additionally, users may easily develop their own forms.

CONISS carries out cluster analysis according to the incremental sum of squares method (also known as Ward's method, minimum variance, error sums of squares). The analysis can be stratigraphically constrained and, therefore, a method for numerical zonation. The program is accessible from *Tilia*.

CA performs correspondence analysis and detrended correspondence analysis. The program is accessible from *Tilia*.

Tilia·graph plots pollen diagrams or other stratigraphic data. The program has a default format, and a spreadsheet of pollen frequencies can be viewed as a pollen diagram with only a few keystrokes. The user has a high degree control over the diagram format. Graphs may be plotted as silhouette curves, histograms, presence/absence, or raw numbers. Curves can be filled with a variety of patterns or depth bars, and exaggerated curves can also be plotted (the exaggeration is defined by the user). The user can define the width of histogram bars, which can be filled with a pattern. Cumulative or overlaid graphs can also be plotted. The user can scale both the horizontal and vertical axes; define major, minor, and labeled tic marks; and enter axis labels. If the data are plotted against depth, the program will optionally plot an age axis correctly scaled, or vice versa. Graph names can be plotted at any angle from 0 to 90°. Zone boundaries and labels can be plotted. A variety of fonts are available.

Tilia is able to read files that are in any of the following formats: Brown, Cambridge (POLLDATA output), Marseille, Minnesota, Wisconsin, and WK1 or WKS.

3.3. Transferring data between the database and *Tilia*

Paradox scripts are available for moving data between the Pollen Database and *Tilia* and *Tilia Forms*. These scripts, titled PD2TIL and TIL2PD, directly read data to and from *Tilia* .TIL file format. Note that TIL2PD uses the data in the .TIL file to replace existing data in the database, and that once replaced, the existing information is not retrievable; no backups are made.

These scripts (and others such as CHECK_DB to check the structure and integrity of the Pollen Database Paradox tables) are available via anonymous *ftp* (login is anonymous, give your email address as password) from ftp.ngdc.noaa.gov (192.149.148.109) in the directory /paleo/pollen/pdb/dba-files.

4. Conventions adopted for the databases

4.1. Place names

Country names and codes follow the ISO standard (3166) for the representation of names of countries, dependencies and areas of special sovereignty and have been entered into table POLDIV1. In addition, GER has been assigned for Germany until a new code is officially assigned by the Swiss National Bureau of Standards. Codes for the new states in eastern Europe will be added when they become available.

The NAPD second political subdivisions are states or provinces. Two-character codes have been assigned from the U.S. Department of Commerce FIPS PUB 10-3. The NAPD third political subdivisions are counties or census areas. These codes can be found in lookup tables POLDIV2 and POLDIV3.

4.2. IGCP type regions (EPD)

The International Geological Correlation Programme Project 158B subdivided the following countries into "type regions, which are uniform areas as regards geology, morphology, climate and biota" (Magdalena Ralska-Jasiewiczowa, 1986 p. 1): Austria, Belgium, Bulgaria, Czechoslovakia, Denmark, Finland, France, Germany, Great Britain, Greenland, The Netherlands, Norway, Poland, Sweden, Switzerland, and some areas of the USSR including Kola, Karelian, Estonia, Latvia, Lithuania, Belarus, Ukraine and Moldavia.

The IGCP catalogue includes a series of maps which indicate the type region boundaries, and a set of codes which uniquely identify each region. These codes are cross-referenced to the look-up table IGCPTYPE.

4.3. Bibliographic conventions

Bibliographic references will be entered and formatted following the conventions used in the journal *Ecology*. The Council of Biology Editors Style Manual is followed for details of style and presents "rules, rationale, principles, definitions and examples to assist in formulating bibliographic references from a wide variety of materials" (p. 55).

4.4. Alphanumeric data storage

Alphanumeric codes are uppercase (e.g. "CAN" for Canada in field PolDiv1 of table SITELOC).

4.5. Typographical errors

Typographical errors in published data will be corrected when entered into the database. For example, a sampling device described as a "modified Livingston", will be entered as "modified Livingstone".

4.6. Special characters

It is sometimes necessary to use special characters when entering text or workers' names. A list of the ASCII special characters that are available for data entry can be found on p. 682 of the *Paradox for DOS* 4.5 Users Guide.

4.7. Dummy values

When data are not available the field may be temporarily flagged with a dummy or nominal value (e.g. "0" for latitude or longitude).

5. Differences between the European and Global Pollen databases

5.1. Internal code numbers

The database structure is the same for both the GPD and the EPD. However, internal code numbers for sites (Site#), entities (E#), pollen variables (Var#), worker names (Worker#), and publications (Publ#) are independently assigned from Arles for the EPD and from Boulder for the GPD.

5.2. Taxonomy, nomenclature, and synonymy

Three kinds of synonymy exist: *nomenclatural*, *syntactic*, and *morphological*. The Database has nomenclatural and syntactic conventions, and will substitute nomenclatural or syntactic synonyms for contributors' names that do no conform to these conventions. In general, the Database will not change names based on morphological synonymy.

Nomenclatural synonymy refers to botanical nomenclature. For North America, the GPD follows Cronquist (1981) and floras conforming to this authority, including the Flora of North America Editorial Committee (1993), Gleason and Cronquist (1991), Great Plains Flora Association (1986), and Hickman (1993). Nomenclature for bryophytes follows Crum and Anderson (1981), Crosby and Magill (1981), and Schuster (1992). The nomenclature of these floras has precedence over those used for other parts of the continent. Examples of nomenclatural synonymy in North America are Poaceae = Gramineae, Asteracae subfam. Asteroideae = Compositae subfam. Tubuliflorae, and *Alnus incana* = *Alnus rugosa*. The EPD follows *Flora Europaea* (Tutin et al. 1964-1980) for the Pteridophyta and Spermatophyta, with the exception of taxa that are not native to Europe, which follow Willis (1985), Smith (1980, 1990) for the Bryophyta, Prescott (1969) for the green and red algae, and Ainsworth et al. (1973, 1983) for the Eumycota.

Syntactic synonymy refers to nomenclatural syntax and applies particularly to the non-Latin parts of pollen-type names. The language of the database is English. Some examples of syntactic conventions are:

Ambrosia-type rather than *Ambrosia* type or *Ambrosia* (type)
Ostrya/*Carpinus* rather than *Ostrya-Carpinus*
Pinguicula cf. *P. villosa* rather than *Pinguicula* cf. *villosa*

Morphological synonymy refers to different but valid nomenclatural and syntactic names applied to the same morphological pollen type. Morphological synonyms often have a geographical ingredient. Examples in North America are *Fraxinus nigra*-type = *Fraxinus quadrangulata*-type and *Larix* = *Pseudotsuga*.

No attempt has been made to determine pollen-morphological synonyms. It is recognized that the association of a pollen-morphological taxon with a particular plant taxon, or group of taxa, is a matter of subjective interpretation, because it is unlikely that the present biogeography had prevailed in the past. Nevertheless, it is common to apply different nomenclature to morphologically indistinguishable pollen taxa when working in different geographical regions because the associated plant taxa had limited geographical ranges. As a matter of principle, as much information is being preserved as possible, even at the expense of long taxon lists that contain many pollen-morphological synonyms. In the interest of somewhat limiting the number of morphological synonyms, the database administrator may suggest possible morphological synonyms to data contributors for their approval. The ultimate decision on morphological taxonomy is that of the contributor, however. The development of a uniform taxonomy and nomenclature for pollen-morphological taxa is perhaps a desirable long-term goal, but in the short-term only botanical nomenclatural and syntactic conventions are being imposed.

Both the EPD and GPD have adopted the following syntactic conventions

Type
Type should always be hyphenated e.g. **-**type. Exceptions: *Armeria* (type A), *Armeria* (type B), *Artemisia norvegica* (type 1)

cf.
Form is Genus cf. Genus species, but Family cf. Genus. For example, *Betula* cf. *B. nana* and Oleaceae cf. *Phillyrea*

Genus/Genus-type is not acceptable
For example, *Malus*/*Mespilus*-type should be either *Malus*/*Mespilus* or *Malus*-type or *Mespilus*-type.

Genus-type/Genus-type is not acceptable

Genus indeterminable or Genus indet. is not acceptable
For example, *Plantago* indeterminable = *Plantago* undiff.

spp. is not acceptable
Genus undiff. or family name (having checked with the author where necessary)

s.l. (sensu lato)
Elymus s.l. = *Elymus*-type

5.3. Restrictions on the use of data

EPD data are classified as restricted or unrestricted. Although all data will be distributed, that marked as restricted can only be used with the permission of the originator.

GPD data are available without restrictions, although professional courtesy is expected when using data.

6. Overview of the Pollen Database Structure

6.1. Categories of Tables

The Pollen Database is ambitious in the scope of data it seeks to archive. Furthermore, good database design favors a larger number of tables each with a few fields, rather than a smaller number of tables each with many fields, and repeating information. Tables are linked together by like fields for data retrieval. The Pollen Database presented below comprises over 200 fields of data organized into approximately 40 tables, with each table containing a different kind of data. Although such numbers must at first seem daunting, understanding the database is simplified by an awareness of the relationships between the different categories of tables, and the realization that the core data are stored in a relatively small number of tables.

Five categories of table can be distinguished: Archival, Look-up, Research, System, and View. **Archival** tables store the data reported on each entity, data that are not expected to change once entered into the database, except to add missing information or to correct errors. SITELOC is an Archival table that stores the locational information on sites. **Look-up** tables save space and reduce errors in the database by collecting repeatedly used items in a single place. Other tables then reference these items using short numbers or codes. For example, the POLDIV1 table contains names and three-character codes for all countries. Redundancy is avoided and space is saved in the SITELOC and WORKERS tables by storing just the country code and not the full country name. **Research** tables store data that are derived by manipulation from the Archival tables or else are of an interpretive or subjective nature. Such data are expected to change or be supplemented. Users are likely to develop a variety of research tables, apart from those initially provided for, according to the requirements of their projects. It is envisaged that applications software be developed initially to allow a few basic research tables to be generated by the user, but that the database distributed by the EPD will not be accompanied by any centrally-developed research tables. P_AGEDPT is an example of a Research table, it stores the estimated ages for samples in a profile calculated from the radiocarbon or other age determinations. **System** tables are used by application programs to help maintain the integrity of the data in the database. They store a description of the structure of the Pollen Database, including the names of all tables and columns, and the composition of primary, alternate, and foreign keys. The SYSCAT table stores the name and category of each table in the Pollen Database. **Views** contain information that is derived from other existing tables or views, making it more convenient to access. Pollen Database archival tables such as P_VARS and DESCR store hierarchical information that is subsequently used to derive views for the quick retrieval of hierarchical relationships (P_VARIS and DESCRIS).

6.2. Pollen v. Macrofossil data

A set of parallel tables that store analogous data have been designed to allow for the recording of plant and macrofossil data. Where available, this information will be included in the EPD. These tables are distinguished by the prefixes "P_" for pollen data, and "M_" for macrofossil data. For example, both the P_ENTITY and M_ENTITY tables hold additional information about an entity (e.g. a core). Two tables are used rather than one because P_ENTITY stores pollen-specific data and M_ENTITY stores macrofossil-specific data (e.g. the year that the analysis was completed for a given entity, which might not be the same where both types of data were analyzed).

6.3. Fundamental Tables

The Archival tables form the core of the database. The most important and frequently used of these are listed below (A = archival; R = research).

Table	Stores
SITELOC (A)	Name and location of the site.
ENTITY (A)	Name and description of the entity.
GEOCHRON (A)	Depths and material dated for geochronological dates.
C14 (A)	Reported radiocarbon dates.
P_SAMPLE (A)	Sample depths and analyst.

P_VARS (A)	Names and codes of pollen variables.
P_COUNTS (A)	Non-zero pollen counts.
PUBL (A)	Bibliographic citations.
WORKERS (A)	Name and address of workers: contact persons, analysts, etc.
CHRON (R)	Name and date of preparation, age model used.
AGEBASIS (R)	Dating information used to produce the age-depths.
P_AGEDPT (R)	Age-depths and deposition times for each sample for each chronology.

7. References

Ainsworth, G. C., F. K. Sparrow, and A. S. Sussman. 1973. The fungi: an advanced treatise. Academic Press, London, England.

Ainsworth, G. C. 1983. Dictionary of fungi. Seventh edition. Commonwealth Mycological Institute, Kew.

Borland International. 1994. Paradox 4.5 User's Guide. Borland International, Scotts Valley, CA, USA.

Codd, E. F. 1970. A relational model of data for large shared data banks. Communications of the Association for Computing Machinery 13:377-387.

Codd, E. F. 1985. Is your DBMS really relational? Computerworld, October 14 and 21.

Codd, E. F. 1990. The Relational Model for Database Management. Version 2. Addison-Wesley, Reading, Massachusetts, USA.

COHMAP Members. 1988. Climatic changes of the last 18,000 years: observations and model simulations. Science 241:1043-1052.

Council of Biology Editors 1983. The CBE Style Manual. Fifth edition. American Institute of Biological Sciences, Washington, D.C., USA.

Cronquist, A. 1981. An integrated system of classification of flowering plants. Columbia University Press, New York, New York, USA.

Crosby, M.R., and R.E. Magill. 1981. A dictionary of Mosses. Missouri Botanical Garden, St. Louis, Missouri, USA.

Crum, H.A., and L.E. Anderson. 1981. Mosses of eastern North America. 2 volumes. Columbia University Press, New York, New York, USA.

Date, C. J. 1990. An Introduction to Database Systems. Fifth edition. Addison-Wesley, Reading, Massachusetts, USA.

Flora of North America Editorial Committee. 1993. Flora of North America North of Mexico. Volumes 1 and 2. Oxford University Press, New York, New York. USA.

Gleason, H. A., and A. Cronquist. 1991. Manual of vascular plants of northeastern United States and adjacent Canada. Second edition. New York Botanical Garden, Bronx, New York, USA.

Great Plains Flora Association. 1986. Flora of the Great Plains. University Press of Kansas, Lawrence, Kansas, USA.

Hickman, J.C., editor. 1993. The Jepson manual: higher plants of California. University of California Press, Berkeley, California, USA.

Huntley, B., and H. J. B. Birks. 1983. An atlas of past and present pollen maps for Europe 0-13,000 years ago. Cambridge University Press, Cambridge, England.

IGBP 1992. The PAGES Project: Proposed Implementation Plans for Research Activities, edited by John A. Eddy. IGBP Report No. 19. IGBP Secretariat, Stockholm, Sweden.

Prescott, G. W. 1969. The algae: a review. Nelson, London, England.

Ralska-Jasiewiczowa, M., editor. 1985. Palaeohydrological changes in the temperate zone in the last 15,000 years. Subproject B. Lake and Mire Environments. International Geological Correlation Programme Project 158B.

Schuster, R.M. 1992. The Hepaticae and Anthocerotae of North America east of the hundredth meridian. Volume 6. Field Museum of Natural History, Chicago, Illinois, USA.

Smith, A. G. 1980. The moss flora of Britain and Ireland. Cambridge University Press, Cambridge, England.

Smith, A. G. 1990. The liverworts of Britain and Ireland. Cambridge University Press, Cambridge, England.

Tutin, T.G. 1964-1980. Flora Europaea. Cambridge University Press, Cambridge, England.

Willis, J. C. 1985. A dictionary of the flowering plants and ferns. Cambridge University Press, Cambridge, England.

Wright, H.E., Jr., J.E. Kutzbach, T. Webb III, W.F. Ruddiman, F.A. Street-Perrott, and P.J. Bartlein, editors. 1993. Global climates since the last glacial maximum. University of Minnesota Press, Minneapolis, Minnesota, USA.

8. Database Tables and Field Descriptions
8.1. Archival Tables

TABLE WORKERS
Names of workers.

The **WORKERS** table is an Archival table that stores the names of all workers known to the Pollen Database. To the Pollen Database a worker may be a collector, contact person, analyst, or author.

```
FIELDS          ( Worker#       S
                  WorkerIs#     S
                  Status        A1
                  LastName      A30
                  Initials      A10
                  FirstName     A15
                  Suffix        A5
                  Title         A10
                  Country       A3
                  Phone         A30
                  Fax           A30
                  EMailAddr     A60
                  Address       M65 )
PRIMARY KEY     ( Worker# )
FOREIGN KEY     ( Country ) references POLDIV1.
                ( WorkerIs# ) references WORKERS.
```

Worker#	Unique identifier for the name of each worker.
WorkerIs#	Identifier for the name currently used by the worker. This number is the same as Worker#, except where the worker has changed his or her name.
Status	One character code for the status of worker. It must be one of the following: A=Active, I=Inactive, D=Deceased, U=Unknown.
LastName	Last name of the worker. Name prefixes, particles, articles, and prepositions will be retained and placed in Ecology style according to guidelines suggested by the 5th edition of the CBE Style Manual.
Initials	Full initials of the worker, without spacing (e.g. "E.C.").
FirstName	First name of the worker.
Suffix	Suffix for the worker's name (e.g. "III").
Title	Title of the worker (e.g. "Dr.").
Country	Three character code for the worker's country of residence.
Phone	Telephone number, including the international code (e.g. "44 091 3743741").
Fax	FAX number, including the international code (e.g. "33 90 939803").
EMailAddr	Electronic mail address.
Address	Mailing address of worker (65-character wide lines of text) including country.

TABLE SITELOC

Site location information.

The **SITELOC** table is an Archival table that stores information on sites, their names and locations.

FIELDS	(Site#	S
	SiteName	A40
	SiteCode	A14
	SiteExists	A5
	PolDiv1	A3
	PolDiv2	A2
	PolDiv3	A3
	LatDeg	S
	LatMin	S
	LatSec	S
	LatNS	A1
	LatDD	N
	LatDMS	A9
	LonDeg	S
	LonMin	S
	LonSec	S
	LonEW	A1
	LonDD	N
	LonDMS	A10
	Elevation	S
	AreaOfSite	N)
PRIMARY KEY	(Site#)	
ALTERNATE KEY	(SiteName)	
	(SiteCode)	
FOREIGN KEY	(PolDiv1) references POLDIV1.	
	(PolDiv1, PolDiv2) references POLDIV2.	
	(PolDiv1, PolDiv2, PolDiv3) references POLDIV3.	

Site# Unique identifier for the site.

SiteName Unique name by which the site is generally known (e.g. "La Grande Pile"). Modifiers are added in square brackets, if necessary, to make the site name unique. Political division names are preferred as modifiers (e.g. "Devils Lake [Wisconsin]"). Where an author's SiteName is general or nonspecific, the author's name and the year of publication, in square brackets, will preface the SiteName (e.g. "[Mott & Camfield 1969] Site 3").

SiteCode Coded name for site, in 3 divisions: country, 2nd and 3rd political subdivisions, and site (e.g. "USA-MN053-WOLS").

SiteExists Year that the site was last known to exist or not to exist (e.g. "Y1985" or "N1977").

PolDiv1 Code for the first political subdivision - country.

PolDiv2 Code for the second political subdivision (USA = state). The value "00" will be entered where the second political subdivision is unspecified or unknown.

PolDiv3 Code for the third political subdivision (USA = county). The value "000" will be entered where the third political subdivision is unspecified or unknown.

LatDeg Degrees of latitude of the site. This value is a positive integer in the range 0-90.

LatMin Minutes of latitude of the site. This value is a positive integer in the range 0-59.

LatSec Seconds of latitude of the site. This value is a positive integer in the range 0-59.

LatNS A single character (**N** or **S**) indicating whether the latitude is north or south of the equator.

LatDD Latitude expressed in decimal degrees.

LatDMS Latitude in degrees, minutes, and seconds formatted as a string (e.g. "60.58.44N").

LonDeg Degrees of longitude of the site. This value is a positive integer in the range 0-180.

LonMin Minutes of longitude of the site. This value is a positive integer in the range 0-59.

LonSec Seconds of longitude of the site. This value is a positive integer in the range 0-59.

LonEW A single character (**E** or **W**) indicating whether the longitude is east or west of the Greenwich meridian.
LonDD Longitude expressed in decimal degrees.
LonDMS Longitude in degrees, minutes, and seconds formatted as a string (e.g. "69.59.31W").
Elevation Elevation (altitude) of the site in meters above sea level.
AreaOfSite Approximate area of the present site in hectares (i.e. the open area of a lake).

TABLE SITEDESC

Site descriptions.

The **SITEDESC** table is an Archival table that stores textual descriptions of a site. This table is a subtype of table **SITELOC.**

```
FIELDS          ( Site#          S
                  SiteDescript   A40
                  Physiography   A40
                  SurroundVeg    A40
                  VegFormation   A40
                  IGCPtype       A8 )
PRIMARY KEY     ( Site# )
FOREIGN KEY     ( Site# ) references SITELOC.
                ( IGCPtype ) references IGCPTYPE.
```

Site# Identifier for the site.
SiteDescript Brief description of the present nature of the site (e.g. "lake with marginal fen").
Physiography Physiographic description (e.g. "rolling stagnation moraine").
SurroundVeg Vegetation surrounding the coring site.
VegFormation Vegetation formation (e.g. "boreal forest"). The vegetation formation in Europe is defined as the unit depicted on the Map of the Natural Vegetation of the member countries of the European Community and the Council of Europe (1987). If a site is not located within one of these countries then an appropriate vegetation formation is assigned (e.g. "*Quercus*/*Fraxinus* woods").
IGCPtype IGCP type region code (e.g. "F-x", for the Forez, Velay and Aubrac region).

TABLE SITEINFO

Sites and associated published information.

The **SITEINFO** table is an Archival table that stores information that cross-references sites and the published information on them.

```
FIELDS          ( Site#          S
                  ICode          A3
                  Publ#          S )
PRIMARY KEY     ( Site#, ICode, Publ# )
FOREIGN KEY     ( Site# ) references SITELOC.
                ( ICode ) references INFOTYPE.
                ( Publ# ) references PUBL.
```

Site# Identifier for the site.
ICode Three character identifier for the type of information present in the publication (e.g. ALG=algae, COL=coleopterans, POL=pollen). These identifiers and their descriptions are stored in table INFOTYPE.
Publ# Identifier for the publication.

TABLE ENTITY

Entity identification and description.

The **ENTITY** table is an Archival table that stores the description of an entity. An entity is either a core, a section, or a surface sample that is associated with exactly one site. The entity is the unit of organization of the database.

FIELDS	(E#	S
	Site#	S
	Sigle	A8
	Name	A30
	IsCore	A1
	IsSect	A1
	IsSSamp	A1
	Descriptor	A4
	HasAnLam	A1
	EntLoc	A40
	LocalVeg	A40
	Coll#	S
	SampDate	A10
	DepthAtLoc	N
	IceThickCM	S
	SampDevice	A30
	CoreDiamCM	N
	C14DepthAdj	N
	Notes	M60)
PRIMARY KEY	(E#)	
ALTERNATE KEY	(Sigle)	
FOREIGN KEY	(Site#) references SITELOC.	
	(Descriptor) references DESCR.	
	(Coll#) references WORKERS.	

E# Unique identifier for the base data entity (core, section, etc.).

Site# Identifier for the site.

Sigle Unique code name for the entity. The Sigle is any unique combination of uppercase letters, numbers, and certain special characters. These special characters include those that are legal DOS filename characters.

Name Any modifier of the site name as published on the pollen diagram (e.g. "core B" or "boring 16A"). For NAPD, where none is given, Name is left blank (null). An entity is a core, which may include a series of segments, a section, or a surface sample, or any combination thereof.

IsCore "Y" if the entity is a core, otherwise "N".

IsSect "Y" if the entity is a section, otherwise "N".

IsSSamp "Y" if the entity is a modern surface sample, otherwise "N".

Descriptor Identifier for the description of modern entity site. These identifiers and their descriptions are listed in table DESCR.

HasAnLam Code describing the type of analyzed annual laminations: T=annually laminated to top, P=laminated, but not to top, N=no laminations or laminations not analyzed. The term 'varve' is reserved for varves *sensu stricto* (i.e. glacial varves). [Applies to Cores and Sections only.]

EntLoc Location where the sample or core was taken (e.g. "north side of lake").

LocalVeg Local vegetation at the collection site.

Coll# Identifier of the worker who took the core or collected the surface sample.

SampDate Date that the sample or core was collected. Enter date as yyyy-mm-dd (e.g. "1977-08-25", or where the day is not reported, "1962-01-00", or if no date, "0000-00-00").
DepthAtLoc Water depth at the coring location (cm). [Applies to Cores and Surface Samples only.]
IceThickCM Thickness of the ice through which a sample was taken (cm).
SampDevice Name of the sampling or coring device (e.g. "Livingstone piston sampler").
CoreDiamCM Diameter of the core (cm). [Applies to cores only.]
C14DepthAdj Adjustment for ^{14}C depths needed to relate the depth of the dated samples to the pollen or other samples. This is necessary since depths cited for samples taken for radiocarbon and other geochronological dates may have been measured from a different datum than those for pollen and other samples (cm). (See also, table GEOCHRON.) [Applies to Cores and Sections only.]
Notes Additional information on the description of the entity.

TABLE P_ENTITY

Pollen-specific entity data.

The **P_ENTITY** and **M_ENTITY** tables are Archival tables that store additional information about the entity including the source and form of the data, and its use status. These tables are subtypes of table **ENTITY**.

```
FIELDS          ( E#            S
                  Contact#      S
                  DataSource    A80
                  DataForm      A2
                  UseStatus     A1 )
PRIMARY KEY     ( E# )
FOREIGN KEY     ( E# ) references ENTITY.
                ( Contact# ) references WORKERS.
```

E# Identifier for the entity.
Contact# Identifier for the worker responsible for the data, the contact person.
DataSource Source of the data (e.g. "publication", "originator", "COHMAP").
DataForm Two character code indicating the form of the data. It must be one of the following: RC=raw counts, RP=raw percentages, DC=digitized counts, DP=digitized percentages. The EPD protocols require that all data be of form RC.
UseStatus One character code indicating the use status of the data. It must be one of the following: USA: P=published, U=unpublished; EUR: U=unrestricted, R=restricted. The EPD will distribute all data, restricted and unrestricted. Restricted data can be viewed by the user, but can only be used with the permission of the originator.

TABLE M_ENTITY

Macrofossil-specific entity data.

```
FIELDS          ( E#            S
                  Contact#      S
                  DataSource    A80
                  UseStatus     A1 )
PRIMARY KEY     ( E# )
FOREIGN KEY     ( E# ) references ENTITY.
                ( Contact# ) references WORKERS.
```

E# Identifier for the entity.
Contact# Identifier for the worker responsible for the data, the contact person.
DataSource Source of the data (e.g. "publication", "originator", "COHMAP").
UseStatus One character code indicating the use status of the data. It must be one of the following: USA: P=published, U=unpublished; EUR: F=free, R=restricted.

TABLE COREDRIV

Core drive identification and description.

The **COREDRIV** table is an Archival table that stores information on core-drives.

```
FIELDS          ( E#            S
                  Drive#        S
                  DriveLabel    A10
                  DriveTopCM    N
                  DriveBotCM    N
                  InfTopCM      N
                  InfBotCM      N
                  RecoveryCM    N )
PRIMARY KEY     ( E#, Drive# )
FOREIGN KEY     ( E# ) references ENTITY.
```

E# Identifier for the entity.
Drive# Unique identifier for the drive.
DriveLabel Label given to the drive (e.g. "4B")
DriveTopCM Depth at which the drive for the core segment started (cm).
DriveBotCM Depth at which the drive for the core segment stopped or bottomed (cm).
InfTopCM Inferred depth of the top of the core segment, according to the data originator, when account has been made for any compaction or sediment disturbance (cm).
InfBotCM Inferred depth of the bottom of the core segment (cm).
RecoveryCM Length of the segment recovered (cm).

TABLE SECTION

Section identification and description.

The **SECTION** table is an Archival table that stores information related to the identification of the sections taken for an entity.

```
FIELDS          ( E#            S
                  Section#      S
                  SectionLabel  A60
                  SectionTopCM  N
                  SectionBotCM  N )
PRIMARY KEY     ( E#, Section# )
FOREIGN KEY     ( E# ) references ENTITY.
```

E#	Identifier for the entity.
Section#	Unique identifier for the section.
SectionLabel	Label given to the section.
SectionTopCM	Depth of the top of the section relative to datum (cm).
SectionBotCM	Depth of the bottom of the section relative to datum (cm).

TABLE GEOCHRON

Analytically determined dates.

The **GEOCHRON** table is an Archival table that stores geochronological information that is common to all of the different types of geochronological dates (e.g. radiocarbon or fission track dates). This common information includes depth, material dated, and a reference to the published description of the event. The **GEOCHRON** table is the super-type to tables that store information for a particular geochronological dating method (e.g. tables **C14** and **FT**).

```
FIELDS          ( E#            S
                  Sample#       S
                  Method        A1
                  DepthCM       N
                  Thickness     N
                  MaterialDated A30
                  Publ#         S )
PRIMARY KEY     ( E#, Sample#, Method)
FOREIGN KEY     ( E# ) references ENTITY.
                ( Publ# ) references PUBL.
```

E#	Identifier for the entity.
Sample#	Unique identifier for the dated sample.
Method	A one character code indicating the method used to date the material. It must be one of the following: A=AAR, C=^{14}C, E=ESR, F=FT, K=KAR, P=^{210}Pb, S=Si^{32}, T=TL, U=USERIES.
DepthCM	Reported depth of the midpoint of the sample (cm). **Subtract the value of C14DepthAdj (table ENTITY) from the reported depth to correlate geochronological with fossil depths.**
Thickness	Thickness of the sample (cm).
MaterialDated	Type of material and fraction dated. This should be defined as precisely as possible (e.g. "gyttja", "wood (*Pinus sylvestris*)", "shell (*Arctica islandica*)").
Publ#	Bibliographic citation for the dating information for this sample.

TABLE AAR

Amino Acid Racemization dating information.

The **AAR** table is an Archival table that stores Amino Acid Racemization dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	ErrorLimits	N
	TaxonDated	A30
	LabNumber	A10
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeBP Age determination of the sample in years BP.
ErrorLimits Error (±) in the reported age.
TaxonDated Taxonomic classification of the dated material (e.g. *Arctica islandica*).
LabNumber Laboratory number assigned to the sample.
Notes Additional notes on the sample.

TABLE C14

Carbon-14 dating information.

The **C14** table is an Archival table that stores Radiocarbon dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	AgeSDUp	N
	AgeSDLo	N
	GrThanAge	A1
	Basis	A1
	Enriched	A1
	LabNumber	A10
	DeltaC13	N
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
ALTERNATE KEY	(LabNumber)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeBP Age determination of the sample in years BP (= 1950 using the Libby half-life).
AgeSDUp Upper standard deviation of this age.
AgeSDLo Lower standard deviation of this age.
GrThanAge "Y" if the determination is stated as being 'greater than', otherwise "N" (the default).
Basis A one character code indicating the method used to make the radiocarbon determination. It must be one of the following: A=accelerator, G=gas decay count, L=decay count liquid scintillation, U=decay count method unknown.
Enriched "Y" if the sample was enriched, otherwise "N".

LabNumber	Laboratory number assigned to the sample (e.g. "St-1050").
DeltaC13	Value of delta ^{13}C.
Notes	Additional notes on the sample.

TABLE ESR — Electron Spin Resonance dating information.

The **ESR** table is an Archival table that stores Electron Spin Resonance dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	ErrorLimits	N
	LabNumber	A10
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E#	Identifier for the entity.
Sample#	Identifier for the dated sample.
AgeBP	Age determination of the sample in years BP.
ErrorLimits	Error (±) in the age determination.
LabNumber	Laboratory number assigned to the sample.
Notes	Additional notes on the sample.

TABLE FT — Fission Track dating information.

The **FT** table is an Archival table that stores Fission Track dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	ErrorLimits	N
	LabNumber	A10
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E#	Identifier for the entity.
Sample#	Identifier for the dated sample.
AgeBP	Age determination of the sample in years BP.
ErrorLimits	Error (±) in the age determination.
LabNumber	Laboratory number assigned to the sample.
Notes	Additional notes on the sample.

TABLE KAR — Potassium/Argon dating information.

The **KAR** table is an Archival table that stores Potassium/Argon dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	ErrorLimits	N
	LabNumber	A10
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeBP Age determination of the sample in years BP.
ErrorLimits Error (±) in the age determination.
LabNumber Laboratory number assigned to the sample.
Notes Additional notes on the sample.

TABLE PB210

Lead-210 dating information.

The **PB210** table is an Archival table that stores ^{210}Pb dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeAD	N
	AgeSDUp	N
	AgeSDLo	N
	GrThanAge	A1
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#, Sample#) references GEOCHRON.	

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeAD Age determination of the sample in years AD.
AgeSDUp Upper standard deviation of the age determination.
AgeSDLo Lower standard deviation of the age determination.
GrThanAge "Y" if the determination is stated as being 'greater than', otherwise "N" (the default).
Notes Additional notes on the sample.

TABLE SI32

Silicon-32 dating information.

The **SI32** table is an Archival table that stores Silicon-32 dating information. This table is a subtype of table **GEOCHRON**.

FIELDS	(E#	S
	Sample#	S
	AgeBP	N
	AgeSDUp	N

```
                          AgeSDLo       N
                          GrThanAge     A1
                          LabNumber     A10
                          Notes         M60)
PRIMARY KEY             ( E#, Sample# )
FOREIGN KEY             ( E#, Sample# ) references GEOCHRON.
```

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeBP Age determination of the sample in years BP (= 1950 using the Libby half-life).
AgeSDUp Upper standard deviation of this age.
AgeSDLo Lower standard deviation of this age.
GrThanAge "Y" if the determination is stated as being 'greater than', otherwise "N" (the default).
LabNumber Laboratory number assigned to the sample.
Notes Additional notes on the sample.

TABLE TL — Thermoluminescence dating information.

The **TL** table is an Archival table that stores Thermoluminescence dating information. This table is a subtype of table **GEOCHRON**.

```
FIELDS                  ( E#            S
                          Sample#       S
                          AgeBP         N
                          ErrorLimits   N
                          GrainSize     A10
                          LabNumber     A10
                          Notes         M60 )
PRIMARY KEY             ( E#, Sample# )
FOREIGN KEY             ( E#, Sample# ) references GEOCHRON.
```

E# Identifier for the entity.
Sample# Identifier for the dated sample.
AgeBP Age determination of the sample in years BP.
ErrorLimits Error (±) in the age determination.
GrainSize Range of grain size of the dated sediments in microns (e.g. "2-10").
LabNumber Laboratory number assigned to the sample.
Notes Additional notes on the sample.

TABLE USERIES — Uranium-series dating information.

The **USERIES** table is an Archival table that stores Uranium-series dating information. This table is a subtype of table **GEOCHRON**.

```
FIELDS                  ( E#            S
                          Sample#       S
                          AgeBP         N
                          ErrorLimits   N
                          LabNumber     A10
                          Notes         M60)
PRIMARY KEY             ( E#, Sample# )
FOREIGN KEY             ( E#, Sample# ) references GEOCHRON.
```

E#	Identifier for the entity.
Sample#	Identifier for the dated sample.
AgeBP	Age determination of the sample in years BP.
ErrorLimits	Error (±) in the age determination.
LabNumber	Laboratory number assigned to the sample.
Notes	Additional notes on the sample.

TABLE SYNEVENT

Dates from synchronous events.

The **SYNEVENT** table is an Archival table that stores information on ages associated with an entity based on evidence from synchronous events. The description of the event is stored in the **EVENT** table.

```
FIELDS          ( E#            S
                  Event#        S
                  DepthCM       N
                  Thickness     N )
PRIMARY KEY     ( E#, Event# )
FOREIGN KEY     ( E# ) references ENTITY.
                ( Event# ) references EVENT.
```

E#	Identifier for the entity.
Event#	Identifier for the event (event names are stored in table EVENT).
DepthCM	Depth of the midpoint of the event (cm).
Thickness	Thickness of the event (cm).

TABLE EVENT

Event names and dates.

The **EVENT** table is an Archival look-up table that stores the description of an event. The description includes its name, age, the uncertainty in the age, and a reference to the published description of the event.

```
FIELDS          ( Event#        S
                  Event         A1
                  Name          A30
                  AgeBP         N
                  AgeUncertUp   N
                  AgeUncertLo   N
                  Publ#         S )
PRIMARY KEY     ( Event# )
FOREIGN KEY     ( Publ# ) references PUBL.
```

Event#	Unique identifier for the event.
Event	One character code for the type of the event that is being used to date or correlate the core: T=tephra, C=^{137}Cs, etc.
Name	Name of the event (e.g. "China Hat" for a tephra; "Vedde ash bed").
AgeBP	Estimated age of the event in ^{14}C years BP.
AgeUncertUp	Upper bound of range of uncertainty in the age of the event.
AgeUncertLo	Lower bound of range of uncertainty in the age of the event.
Publ#	Bibliographic citation for the age of the event.

TABLE ALSEGS

Annual laminations segments description.

The **ALSEGS** table is an Archival table that stores information on annual lamination segments. A segment is defined as a continuous stratigraphic sequence of annual laminations. An entity can have any number of segments. Depths and counts are stored in the **P_ANLDPT** or **M_ANLDPT** tables.

FIELDS	(E#	S
	Seg#	S
	DepthTopCM	N
	DepthBotCM	N)
PRIMARY KEY	(E#, Seg#)	
FOREIGN KEY	(E#) references ENTITY.	

E# Identifier for the entity.
Seg# Unique identifier for the segment.
DepthTopCM Depth of the top of the annually laminated segment (cm).
DepthBotCM Depth of the bottom of the segment (cm).

TABLE P_ANLDPT

Pollen annual laminations depths information.

The **P_ANLDPT** and **M_ANLDPT** tables are Archival tables that store information on annual laminations for each of an entity's segments (segment descriptions are stored in table **ALSEGS**). These tables store the lamination counts and depths of samples within a segment.

FIELDS	(E#	S
	Seg#	S
	Sample#	S
	DepthCM	N
	Thickness	N
	CountTop	N
	CountBot	N)
PRIMARY KEY	(E#, Seg#, Sample#)	
FOREIGN KEY	(E#, Seg#) references ALSEGS.	

E# Identifier for the entity.
Seg# Identifier for the annual lamination segment.
Sample# Unique identifier for the sample.
DepthCM Depth of the midpoint of the sample (cm).
Thickness Thickness of the sample (cm).
CountTop Annual laminations count at the top of the sample.
CountBot Annual laminations count at the bottom of the sample.

TABLE M_ANLDPT

Macrofossil annual laminations depths information.

FIELDS	(E#	S
	Seg#	S
	Sample#	S
	DepthCM	N
	Thickness	N
	CountTop	N

	CountBot	N)
PRIMARY KEY	(E#, Seg#, Sample#)	
FOREIGN KEY	(E#, Seg#) references ALSEGS.	

E# Identifier for the entity.
Seg# Identifier for the annual lamination segment.
Sample# Unique identifier for the sample.
DepthCM Depth of midpoint of the sample (cm).
Thickness Thickness of the sample (cm).
CountTop Annual laminations count at the top of the sample.
CountBot Annual laminations count at the bottom of the sample.

TABLE P_VARS

Look-up table for pollen variable names.

The **P_VARS** and **M_VARS** tables are Archival tables that store the names, codes and hierarchy information for each pollen variable recognized by the Pollen Database, including synonyms.

FIELDS	(Var#	S
	AccVar#	S
	SynType	A1
	VarCode	A8
	VarName	A60
	HVar#	S
	MHVar#	S
	Auth#	S
	Notes	M60)
PRIMARY KEY	(Var#)	
ALTERNATE KEY	(VarCode)	
	(VarName)	
FOREIGN KEY	(AccVar#) references P_VARS.	
	(HVar#) references P_VARS.	
	(MHVar#) references P_VARS.	
	(Auth#) references PUBL.	

Var# Unique identifier for the variable.

AccVar# For EPD: Identifier for the *accepted* name for this variable within a European, as opposed to local or regional, context; both the EPD name and the Author's name are valid. For NAPD: Identifier of the accepted Var# for this variable. AccVar# = Var# where VarName is the senior synonym.

SynType One-character code indicating the type of synonym. It must be one of the following: M=morphological; N=nomenclatural; 1=name of a globally-monospecific genus (e.g. *Zea*); 2=name of a genus that is monospecific within a continent, but not globally (e.g. *Tilia americana* in North America); or <*null*> for accepted variables. For EPD: not used.

VarCode Eight character code for the pollen variable name.

VarName Name of the pollen type or variable, including Linnean and non-Linnean names (*Acer saccharum*-type, Cerealea, *Avena*/*Triticum*), or non-pollen (microspheres).

HVar# Var# of the next higher taxon of which the pollen type is a member. (Special cases: if HVar# = Var#, then name is valid and it is the highest taxon recognized. HVar# is **null** for non-accepted variables.) For EPD: not defined, see MHVar#.

MHVar# Like HVar# except that in this hierarchy pollen morphological criteria take precedence over taxonomic criteria. For example, the taxonomic higher variable for *Dryopteris* is

Dryopteridaceae whereas the morphological higher variable is *Dryopteris*-type (which includes non-Dryopteridaceae).

Auth# Identifier for the publication cited as authority for a pollen variable name. For NAPD this will be the authority for the Latin portion of the pollen variable's name (e.g. Gleason and Cronquist 1991 is the authority for *Dalea purpurea* in *Dalea purpureatype)*.

Notes Notes on the variable.

TABLE M_VARS

Look-up table for macrofossil variable names.

FIELDS	(Var#	S
	AccVar#	S
	SynType	A1
	VarCode	A8
	VarName	A60
	HVar#	S
	Auth#	S
	Notes	M60)
PRIMARY KEY	(Var#)	
ALTERNATE KEY	(VarCode)	
	(VarName)	
FOREIGN KEY	(AccVar#) references M_VARS.	
	(HVar#) references M_VARS.	
	(Auth#) references PUBL.	

Var# Unique identifier for the variable.

AccVar# Identifier of the accepted Var# for this variable. AccVar# = Var# where VarName is the senior synonym.

SynType One-character code indicating the type of synonym. It must be one of the following: M=morphological; N=nomenclatural; 1=name of a globally-monospecific genus (e.g. *Zea*); 2=name of a genus that is monospecific within a continent, but not globally (e.g. *Tilia americana* in North America); or <*null*> for accepted variables.

VarCode Eight character code for the macrofossil variable named.

VarName Name of the macrofossil type or variable, including Linnean and non-Linnean names (*Acer saccharum*-type, Cerealea, *Avena/Triticum*).

HVar# Var# of the next higher taxon of which the variable is a member. (Special cases: if HVar# = Var#, then name is valid and it is the highest taxon recognized. HVar# is **null** for non-accepted variables.)

Auth# Identifier for the publication cited as authority for a macrofossil variable name.

Notes Notes on the variable.

TABLE P_SAMPLE

Pollen sample depths and thickness.

The **P_SAMPLE** and **M_SAMPLE** tables are Archival tables that store the sample depths for an entity. This table also stores an identifier for the worker who analyzed the sample.

FIELDS	(E#	S
	Sample#	S
	DepthCM	N
	Thickness	N
	Analyst#	S
	AnalyDate	A10
	Notes	M60)

PRIMARY KEY	(E#, Sample#)
FOREIGN KEY	(E#) references ENTITY.
	(Analyst#) references WORKERS.

E# Identifier for the entity.
Sample# Unique identifier for the sample.
DepthCM Depth of the midpoint of the sample (cm).
Thickness Thickness of the sample (cm).
Analyst# Identifier of the worker who counted the pollen sample.
AnalyDate Date that the pollen analysis was completed. Enter date as yyyy-mm-dd (e.g. "1977-08-25", or where the day is not reported, "1962-01-00", or if no date, "0000-00-00").
Notes Additional notes on the sample.

TABLE M_SAMPLE

Macrofossil sample depths and thickness.

FIELDS	(E#	S
	Sample#	S
	DepthCM	N
	Thickness	N
	Analyst#	S
	AnalyDate	A10
	Notes	M60)
PRIMARY KEY	(E#, Sample#)	
FOREIGN KEY	(E#) references ENTITY.	
	(Analyst#) references WORKERS.	

E# Identifier for the entity.
Sample# Unique identifier for the sample.
DepthCM Depth of the midpoint of the sample (cm).
Thickness Thickness of the sample (cm).
Analyst# Identifier of the worker who counted the macrofossil sample.
AnalyDate Date that the macrofossil analysis was completed. Enter date as yyyy-mm-dd (e.g. "1977-08-25", or where the day is not reported, "1962-01-00", or if no date, "0000-00-00").
Notes Additional notes on the sample.

TABLE P_COUNTS

Pollen counts.

The **P_COUNTS** and **M_COUNTS** tables are Archival tables that store the pollen counts for an entity. This table stores the non-zero counts for each combination of variable and sample depth (the depths values are stored in table **P_SAMPLE**).

FIELDS	(E#	S
	Sample#	S
	Var#	S
	Count	N)
PRIMARY KEY	(E#, Sample#, Var#)	
FOREIGN KEY	(E#, Sample#) references P_SAMPLE.	
	(Var#) references P_VARS.	

E# Identifier for the entity.

Sample#	Identifier for the sample.
Var#	Identifier for the variable.
Count	Pollen count. It must be a positive number.

TABLE M_COUNTS

Macrofossil counts.

```
FIELDS          ( E#            S
                  Sample#       S
                  Var#          S
                  Count         N )
PRIMARY KEY     ( E#, Sample#, Var# )
FOREIGN KEY     ( E#, Sample# ) references M_SAMPLE.
                ( Var# ) references M_VARS.
```

E#	Identifier for the entity.
Sample#	Identifier for the sample.
Var#	Identifier for the variable.
Count	Macrofossil count. It must be a positive number.

TABLE LITHOLGY

Sediment lithology.

The **LITHOLGY** table is an Archival table that stores sediment lithology information for an entity.

```
FIELDS          ( E#            S
                  Lith#         S
                  Descript      A40
                  DepthTopCM    N
                  DepthBotCM    N
                  LoBoundary    A40 )
PRIMARY KEY     ( E#, Lith# )
FOREIGN KEY     ( E# ) references ENTITY.
```

E#	Identifier for the entity.
Lith#	Unique identifier for the lithological unit.
Descript	Description of the lithological unit.
DepthTopCM	Depth of the top of the lithological unit (cm).
DepthBotCM	Depth of the bottom of the lithological unit (cm).
LoBoundary	Description of the nature of the lower boundary.

TABLE LOI

Loss-on-ignition information.

The **LOI** table is an Archival table that stores Loss-on-ignition information for an entity.

```
FIELDS          ( E#            S
                  Sample#       S
                  DepthCM       N
                  Thickness     N
                  TempLo        S
                  LOILo         N
                  TempHi        S
                  LOIHi         N
```

```
                    BulkDens        N )
PRIMARY KEY        ( E#, Sample# )
FOREIGN KEY        ( E# ) references ENTITY.
```

E#	Identifier for the entity.
Sample#	Unique identifier for the sample.
DepthCM	Depth of the midpoint of the sample (cm).
Thickness	Thickness of the sample (cm).
TempLo	Low loss-on-ignition temperature in °C (e.g. 500).
LOILo	Loss-on-ignition value, low temperature (organics), as a percentage of dry mass.
TempHi	High loss-on-ignition temperature in °C (e.g. 1000).
LOIHi	Loss-on-ignition value (carbonates), as a percentage of dry mass.
BulkDens	Dry mass of the sample in g/cm^3.

TABLE PUBL — Publications and their library accession numbers.

The **PUBL** table is an Archival table that stores publication information and the complete citation.

```
FIELDS             ( Publ#          S
                     Acc#           S
                     YearOfPubl     A4
                     Citation       M60 )
PRIMARY KEY        ( Publ# )
ALTERNATE KEY      ( Acc# )
```

Publ#	Unique identifier for the publication.
Acc#	A local library accession number (these are assigned as consecutive short integers). The absence of a value implies no local reprint of the publication.
YearOfPubl	Year of publication.
Citation	Complete bibliographic citation for a publication, in the style of the journal *Ecology*.

TABLE PUBLENT — Publications cross-referenced by entity.

The **PUBLENT** table is an Archival table that cross-references publications and the entities on which they report. This table may be used to find the entities associated with any publication, or the publications associated with any entity.

```
FIELDS             ( Publ#          S
                     E#             S )
PRIMARY KEY        ( Publ#, E# )
FOREIGN KEY        ( Publ# ) references PUBL.
                   ( E# ) references ENTITY.
```

Publ#	Identifier for the publication.
E#	Identifier for the entity.

TABLE AUTHORS

Publications cross-referenced by authors.

The **AUTHORS** table is an Archival table that cross-references authors to their publications. This table may be used to find the publications authored by a worker.

FIELDS	(Author#	S
	Publ#	S
	Order	S)
PRIMARY KEY	(Author#, Publ#)	
FOREIGN KEY	(Author#) references WORKERS.	
	(Publ#) references PUBL.	

Author# Identifier for the name of the author.
Publ# Identifier for the publication.
Order An integer that identifies the ordering of authorship (1,2,3,...,n) in a publication.

8.2. Lookup and Referral Tables

TABLE POLDIV1

Look-up table of country names.

The **POLDIV1** table is a Look-up table of 1st political subdivision names. These correspond to the names of countries.

```
FIELDS          ( PolDiv1     A3
                  Name        A40 )
PRIMARY KEY     ( PolDiv1 )
ALTERNATE KEY   ( Name )
```

PolDiv1	ISO/ANSI three-letter country codes (source: U.S. Department of Commerce FIPS PUB 1041). Note that *GER* will be used for the unified Germany until an ISO standard code is announced.
Name	Name of the country (spellings used are those specified by the ISO 3166 standard).

TABLE POLDIV2

Look-up table of 2nd political subdivision names.

The **POLDIV2** table is a Look-up table of 2nd political subdivision names. The 2nd political division in the United States is the state, it is the province or territory in Canada.

```
FIELDS          ( PolDiv1     A3
                  PolDiv2     A2
                  PostCode    A2
                  Name        A25 )
PRIMARY KEY     ( PolDiv1, PolDiv2 )
FOREIGN KEY     ( PolDiv1 ) references POLDIV1.
```

PolDiv1	Three letter country codes.
PolDiv2	Two digit codes for 2nd political subdivision. In North America these are states or provinces (source: U.S. Department of Commerce FIPS PUB 10-3). In Europe a list is currently being compiled of all the major political subdivisions of countries (i.e. Vice Counties in the U.K., départments in France, etc.). A map will eventually be prepared that will depict the boundaries between each of these areas.
PostCode	Two letter postal code.
Name	Name of the political subdivision.

TABLE POLDIV3

Look-up table of 3rd political subdivision names.

The **POLDIV3** table is a Look-up table of 3rd political subdivisions names. These correspond, generally, to county names in the United States.

```
FIELDS          ( PolDiv1     A3
                  PolDiv2     A2
                  PolDiv3     A3
                  Name        A40 )
PRIMARY KEY     ( PolDiv1, PolDiv2, PolDiv3 )
FOREIGN KEY     ( PolDiv1, PolDiv2 ) references POLDIV2.
```

PolDiv1	Three letter country codes.

PolDiv2 Two digit 2nd political subdivision code.

PolDiv3 Three digit codes for 3rd political subdivision. In North America these are counties or census divisions (source: U.S. Department of Commerce FIPS PUB 6-4 and Statistics Canada, Geography Division, Geography Attributes File 1986). In Europe, the minor political subdivisions (parishes, townships, etc.) are so numerous that they are only being entered into the database as necessary.

Name Name of the political subdivision.

TABLE SYNTYPE

Look-up table for description of synonym types.

The **SYNTYPE** table is a Look-up table that stores the full description of the synonym type codes used in **P_VARS** and **M_VARS**.

```
FIELDS          ( SynType       A1
                  Description   A40 )
PRIMARY KEY     ( SynType )
ALTERNATE KEY   ( Description )
```

SynType Unique one character identifier for the synonym type.

Description Description of the type of synonym (e.g. "globally monospecific genus").

SynType	Description
-------	------------------------
1	globally monospecific genus
F	folded into another family
N	nomenclatural synonym
S	subspecific synonym

TABLE INFOTYPE

Look-up table for information type names.

The **INFOTYPE** table is a Look-up table that stores the full names for the three character information type codes.

```
FIELDS          ( ICode         A3
                  InfoType      A30 )
PRIMARY KEY     ( ICode )
ALTERNATE KEY   ( InfoType )
```

ICode Unique three character identifier for the information type.

InfoType Name for the type of information (e.g. "algae", "dendrochronology").

ICode	InfoType
-------	------------------------
AAR	amino acid racemization
ALG	algae
ARC	archaeology
C14	radiocarbon
CES	cesium-134, cesium-137
CHA	charcoal
CHI	chironomids
CLA	cladocerans
COL	coleopterans
DEN	dendrochronology
DIA	diatoms
ESR	electron spin resonance
FIS	fission track
FOR	foraminiferans
GEO	geochemistry
GRA	granulometry
INS	other insects

ISO	other stable isotopes
KAR	potassium/argon
LED	lead-210
LOI	loss on ignition
LUM	thermoluminescence
MAC	macrofossils
MAG	paleomagnetism
MOL	molluscs
OST	ostracods
OTH	other
OXY	stable oxygen isotopes
POL	pollen
RAD	radon-226
RHI	rhizopods
TEF	tephra
URA	u-series

TABLE DESCR

Look-up table for entity site-descriptions.

The **DESCR** table is a Look-up table that stores the descriptions associated with each of the four character descriptor codes.

```
FIELDS          ( Descriptor    A4
                  HigherDescr   A4
                  Description   A40 )
PRIMARY KEY     ( Descriptor )
ALTERNATE KEY   ( Description )
```

Descriptor Unique four character identifier for the entity site-descriptions.

HigherDescr Identifier of the next higher descriptor in the hierarchy. HigherDescr=Descriptor at the top level of the hierarchy.

Description Description of the modern entity site (e.g. "lacustrine", "mire").

Descr	**Description (indentation indicates hierarchy)**
--------	--
AERI	aerial
AAIR	..............air pollen sampler
ARCH	archaeological
COAS	coastal
CEST	..............estuarine
CSLT	..............salt marsh
FLUV	fluvial
HOLL	small hollow
HVER	..............vernal pool
LACU	lacustrine
LART	..............artificial open-water
LOTH	....other (e.g. Norfolk Broads)
LRES	....reservoir
LSTK	....stock pond
LDRA	..............drained lake
LNAT	..............natural open water
LFLU	....fluvial origin (e.g. oxbow lake)
LGLA	....glacial origin
LCIR	....cirque lake
LGMC	....glacial meltwater channel
LGSC	....glacial scour lake
LKET	....kettle lake
LMOR	....morainally dammed lake
LPER	....periglacial origin
LTHK	....thermokarst lake
LSOL	....solution hollow
LTEC	....tectonic lake
LUNK	....origin unknown
LPLY	..............playa
LPRO	..............pro-glacial lake
LVOL	..............volcanic origin
LEXP	....explosion crater
MARI	marine
MIDD	animal midden
MPAC	..............packrat (Neotoma) midden
OTHR	other
PTRP	pollen trap (terrestrial)
PTLK	..............pollen trap (lacustrine)

TERR	terrestrial
TCAV	..............cave sediment
TCDU	....animal dung
TDET	..............organic detritus
TLUS	..............loess
TMIR	..............mire (= peatland, > 30cm peat)
TBOG	....bog
TBLA	....blanket bog
TRAI	....raised bog
TFEN	....fen
TFLO	....floating mire
TFLP	....flood plain mire
TOWT	....open-water transition mire
TSOL	....soligenous (mineraltrophic) mire
TVAL	....valley mire
TSWA	....swamp (= forested peatland = forested wetland)
TMOS	..............moss polster
TMSH	..............marsh
TSOI	..............soil
TMOR	....mor humus

TABLE RATIONAL — Look-up table for dating rationales.

The **RATIONAL** table is a Look-up table that stores the descriptions of the dating rationales (e.g. Tsuga decline) associated with each 3 character rationale code.

FIELDS	(RCode	A3
	Rationale	A60)
PRIMARY KEY	(RCode)	
ALTERNATE KEY	(Rational)	

RCode Unique three character identifier for the dating rationale.
Rationale Description of the rationale for the date assignment.

RCode	**Rationale**
AMR	Ambrosia rise
ANL	annual laminations
BOT	bottom of core
C13	corrected for isotopic fractionation with C13
C14	uncorrected radiocarbon date
CAJ	adjusted for ancient carbon
CAV	averaged radiocarbon date
CCA	calibrated radiocarbon date
CGT	radiocarbon greater than date
COM	corrected for compaction
DEN	C14 date with dendrochronological correction
DGL	date of deglaciation
ESH	European settlement horizon
EXT	extrapolated date
GUE	guess
INT	interpolated date
OTH	other
POL	pollen stratigraphic date
SED	sediment stratigraphic date
TEF	tephra date
TOP	top of core
TSD	Tsuga decline

TABLE IGCPTYPE — Look-up table for IGCP type region codes.

The **IGCPTYPE** table is a Look-up table that stores the names and codes for the IGCP type regions. The IGCP catalogue includes a series of maps that indicate the type region boundaries, and a set of codes that uniquely identify each region.

FIELDS	(IGCPtype	A8
	RegionName	A80)
PRIMARY KEY	(IGCPtype)	

ALTERNATE KEY (RegionName)

IGCPtype IGCP type region code.
RegionName Name of the IGCP type region.

8.3. Research Tables

TABLE P_GROUP

Major group to which a pollen variable is assigned.

Tables **P_GROUP** and **M_GROUP** are Research tables that store the group to which each pollen variable is assigned. These tables enable the user to define taxa into groups based on the life form and habitat of each taxon (e.g. trees, aquatics). The organization of these tables allows for multiple sets of groupings.

```
FIELDS          ( Set#          S
                  Var#          S
                  GroupId       A4 )
PRIMARY KEY     ( Set#, Var# )
FOREIGN KEY     ( Var# ) references P_VARS.
                ( GroupId ) references GROUPS.
```

Set# Unique identifier for the group set.
Var# Identifier for the pollen variable.
GroupId Four-character identifier of the major group to which the pollen variable belongs (e.g. "AQUA" for aquatics).

TABLE M_GROUP

Major group to which a macrofossil variable is assigned.

```
FIELDS          ( Set#          S
                  Var#          S
                  GroupId       A4 )
PRIMARY KEY     ( Set#, Var# )
FOREIGN KEY     ( Var# ) references M_VARS.
                ( GroupId ) references GROUPS.
```

Set# Unique identifier for the group set.
Var# Identifier for the macrofossil variable.
GroupId Four-character identifier of the major group to which the macrofossil variable belongs.

TABLE GROUPS

Names for major groups.

The **GROUPS** table is a Research table that acts as a look-up table for the names and codes for variable groups.

```
FIELDS          ( GroupId       A4
                  GroupCode     A1
                  GroupName     A60 )
PRIMARY KEY     ( GroupId )
ALTERNATE KEY   ( GroupCode )
ALTERNATE KEY   ( GroupName )
```

GroupId Unique four-character identifier for the major group (e.g. "UPHE" for upland herbs).
GroupCode One character code for the major group.
GroupName Name of the major group (e.g. "trees", "aquatics", "nonpollen").

TABLE CHRON

Registers the chronologies developed for each entity.

The **CHRON** table is a Research table that registers each chronology developed for an entity. This table also stores the name of the model used to derive the sample ages, the name of the person who prepared the chronology, and the date the chronology was last modified.

FIELDS	(E#	S
	Chron#	S
	Default	A1
	Name	A20
	PreparedBy	A30
	DatePrepared	A10
	Model	A60
	Notes	M60)
PRIMARY KEY	(E#, Chron#)	
FOREIGN KEY	(E#) references ENTITY.	

E#	Identifier for the entity.
Chron#	Unique identifier for the developed chronology.
Default	"Y" if the chronology is designated the default chronology for that entity, otherwise "N".
Name	Name given to the chronology (e.g. "author's preferred").
PreparedBy	Name of the person who prepared the chronology.
DatePrepared	Date the chronology was prepared or last modified. Enter date as yyyy-mm-dd (e.g. "1977-08-25", or where the day is not reported, "1962-01-00", or if no date, "0000-00-00").
Model	Name of the model used to derive the chronology (e.g. "cubic spline curve").
Notes	Additional notes on the chronology.

TABLE AGEBASIS

Dates used to establish age-depth.

The **AGEBASIS** table is a Research table that stores the data used to derive the age-depth curve for each chronology created for an entity (see table **P_AGEDPT** and **M_AGEDPT**).

FIELDS	(E#	S
	Chron#	S
	Sample#	S
	DepthCM	N
	Thickness	N
	Age	N
	AgeUp	N
	AgeLo	N
	RCode	A3)
PRIMARY KEY	(E#, Chron#, Sample#)	
FOREIGN KEY	(E#, Chron#) references CHRON.	
	(RCode) references RATIONAL.	

E#	Identifier for the entity.
Chron#	Identifier for the chronology.
Sample#	Unique identifier for the dated sample.
DepthCM	Depth of the sample (cm).
Thickness	Thickness of the sample (cm).
Age	Estimated age of the sample.
AgeUp	Upper bound for the estimated age.

AgeLo	Lower bound for the estimated age.
RCode	Three character code for the description of the rationale for the age.

TABLE P_AGEDPT

Pollen age-depths derived from age models.

The **P_AGEDPT** and **M_AGEDPT** tables are Research tables that store fitted age estimates for the sample depths for each chronology specified for an entity. Names for each chronology are stored in the **CHRON** table. The age-depth data are generated by applying the age-model to the data in the **AGEBASIS** table.

FIELDS	(E#	S
	Chron#	S
	Sample#	S
	AgeBP	N
	AgeUp	N
	AgeLo	N
	DepTime	N)
PRIMARY KEY	(E#, Chron#, Sample#)	
FOREIGN KEY	(E#, Chron#) references CHRON.	
	(E#, Sample#) references P_SAMPLE.	

E#	Identifier for the entity.
Chron#	Identifier for the chronology.
Sample#	Identifier for the depth of the sample.
AgeBP	Age of the sample in years BP.
AgeUp	Upper bound of the age.
AgeLo	Lower bound of the age.
DepTime	Deposition time (yr/cm).

TABLE M_AGEDPT

Macrofossil age-depths derived from age models.

FIELDS	(E#	S
	Chron#	S
	Sample#	S
	AgeBP	N
	AgeUp	N
	AgeLo	N
	DepTime	N)
PRIMARY KEY	(E#, Chron#, Sample#)	
FOREIGN KEY	(E#, Chron#) references CHRON.	
	(E#, Sample#) references M_SAMPLE.	

E#	Identifier for the entity.
Chron#	Identifier for the chronology.
Sample#	Identifier for the depth of the sample.
AgeBP	Age of the sample in years BP.
AgeUp	Upper bound of the age.
AgeLo	Lower bound of the age.
DepTime	Deposition time (yr/cm).

TABLE AGEBOUND
Reliable age bounds for dated entities.

The **AGEBOUND** table is a Research table that stores a best estimate for top and bottom ages for an entity. This table provides a simple mechanism for evaluating extrapolated sample ages.

```
FIELDS          ( E#            S
                  Chron#        S
                  Top           N
                  Bottom        N )
PRIMARY KEY     ( E#, Chron# )
FOREIGN KEY     ( E#, Chron# ) references CHRON.
```

E#	Identifier for the entity.
Chron#	Identifier for the chronology.
Top	Uppermost reliable age for this entity (years B.P.).
Bottom	Lowermost reliable age for this entity (years B.P.).

TABLE P_VTRANS
Pollen variable name translation table.

The **P_VTRANS** and **M_VTRANS** tables are Research tables that store variable names in alternate or preferred style to that stored in **P_VARS**. This alternate style is then available for data output.

```
FIELDS          ( Var#          S
                  TranslatesTo  A60 )
PRIMARY KEY     ( Var# )
FOREIGN KEY     ( Var# ) references P_VARS.
```

Var#	Identifier for the pollen variable.
TranslatesTo	Alternate or preferred name for this pollen variable (e.g. "Betula (arbustif)").

TABLE M_VTRANS
Macrofossil variable name translation table.

```
FIELDS          ( Var#          S
                  TranslatesTo  A60 )
PRIMARY KEY     ( Var# )
FOREIGN KEY     ( Var# ) references M_VARS.
```

Var#	Identifier for the macrofossil variable.
TranslatesTo	Alternate or preferred name for this macrofossil variable.

8.4. System Tables

TABLE SYSIDN

Database identification table.

The **SYSIDN** table is a System table that stores the identity of a database. There will be exactly one record in this table.

```
FIELDS          ( Id            A10
                  Name          A60
                  Place         A60
                  Type          A1 )
PRIMARY KEY     ( Id )
```

Id	Unique identifier for the database. It can be up to 10 characters in length (e.g. "EPD").
Name	Full name of the database.
Place	Place where the database is located (e.g. "Montreal, Canada").
Type	Type of database. It must be either C=Central or R=Regional.

TABLE SYSCAT

Names of database tables.

The **SYSCAT** table stores the name and category of every table in the Pollen Database. Table categories are System, Base, and View. The Base tables are further divided into Archival, Lookup, and Research subcategories. The purpose of this table is to enable programs to access information on the structure or keys of a table given its name or its category.

```
FIELDS          ( T#            S
                  TableName     A8
                  Category      A20 )
PRIMARY KEY     ( T# )
ALTERNATE KEY   ( TableName )
```

T#	Unique identifier for the table (T#'s less than 100 are reserved for the system tables).
TableName	Name of the table.
Category	Category of the table. It must be one of the following: "system", "base-archival", "base-lookup", "base-research", or "view".

TABLE SYSCOL

Names and data types of database columns.

The **SYSCOL** table stores the name, data type, and domain type for each column (field) of each table in the Pollen Database. The column name is the name of the *Paradox* field, and data type is the *Paradox* data type (e.g. A20). Domain is not currently implemented.

```
FIELDS          ( T#            S
                  C#            S
                  ColumnName    A20
                  DataType      A4
                  Domain        A20 )
PRIMARY KEY     ( T#, C# )
FOREIGN KEY     ( T# ) references SYSCAT.
```

T#	Identifier for the table.

C# Unique identifier for the column.
ColumnName Name of the column.
DataType Type of data in the column (S=short integer, N=floating point number, An=alphanumeric, where 1 <= n <= 255).
Domain Describes the range of valid values allowed in the column.

TABLE SYSIDX Identifies table columns that comprise primary, alternate, or foreign keys.

The **SYSIDX** table stores the keys that exist on each table in the Pollen Database. Key types that are recognized include, primary, alternate, and foreign. This table records each column comprising the key.

```
FIELDS          ( KeyType      A1
                  T#           S
                  C#           S )
PRIMARY KEY     ( KeyType, T#, C# )
FOREIGN KEY     ( T#, C# ) references SYSCOL.
```

KeyType Identifier for the type of the key (P=primary, A=alternate, F=foreign).
T# Identifier for the table.
C# Identifier for the column.

TABLE SYSFKS Identifies foreign keys - their attributes and the tables they reference.

The **SYSFKS** table stores the columns comprising each foreign key in the Pollen Database and the table that is referenced by each foreign key.

```
FIELDS          ( FK#          S
                  T#           S
                  C#           S
                  RefT#        S )
PRIMARY KEY     ( FK#, T#, C# )
FOREIGN KEY     ( T#, C# ) references SYSCOL.
                ( RefT# ) references SYSCAT.
```

FK# Unique identifier for the foreign key.
T# Identifier for the table.
C# Identifier for the column.
RefT# Identifier for the table referenced by the foreign key.

TABLE DBLOG

Data compilation history.

Table **DBLOG** stores a record of the compilation history of the database: when data or the database structure were modified, how, and by whom.

```
FIELDS          ( Date          D
                  Action#       S
                  Compiler      A2 )
PRIMARY KEY     ( Date, Action# )
FOREIGN KEY     ( Action# ) references DBACT.
```

Date Date the compilation or modification was performed.
Action# Identifier for the action performed.
Compiler Initials of the data compiler.

TABLE DBACT

Data compilation history - actions taken.

The **DBACT** table stores a record of the actions performed on the database.

```
FIELDS          ( Action#       S
                  Action        A255 )
PRIMARY KEY     ( Action# )
ALTERNATE KEY   ( Action )
```

Action# Unique identifier for the action performed.
Action Description of the action performed.

TABLE DBENT

Data compilation history - entities affected.

The **DBENT** table stores a record of the entities affected by a given action.

```
FIELDS          ( Date          D
                  Action#       S
                  E#            S )
PRIMARY KEY     ( Date, Action#, E# )
FOREIGN KEY     ( Action# ) references DBACT.
                ( E# ) references ENTITY.
```

Date Date the compilation or modification was performed.
Action# Identifier for the action performed.
E# Identifier for the entity affected.

8.5. Views

TABLE P_VARIS
View on P_VARS - pollen variable taxonomic hierarchy.

The **P_VARIS** and **M_VARIS** tables are Views that store the taxonomic hierarchy of variables in the Pollen Database. These views can be used to query at any level in the hierarchy (e.g. *Juglans* includes *J. cinerea*, *J. nigra*, and *J. undiff.*).

FIELDS	(Var#	S
	IsAlso#	S)
PRIMARY KEY	(Var#, IsAlso#)	
FOREIGN KEY	(Var#) references P_VARS.	
	(IsAlso#) references P_VARS.	

Var# Identifier for the variable.
IsAlso# Identifier for the variable that Var# is also one of (e.g. "*Acer rubrum*" is also an "*Acer*").

TABLE M_VARIS
View on M_VARS - macrofossil variable taxonomic hierarchy.

FIELDS	(Var#	S
	IsAlso#	S)
PRIMARY KEY	(Var#, IsAlso#)	
FOREIGN KEY	(Var#) references M_VARS.	
	(IsAlso#) references M_VARS.	

Var# Identifier for the variable.
IsAlso# Identifier for the variable that Var# is also one of.

TABLE DESCRIS
Description-Is-Also hierarchy: view on table DESCR.

The **DESCRIS** table is a Look-up table that stores the hierarchy of modern entity site-descriptors. This table can be used to query any level in the descriptor hierarchy (e.g. to query for bogs, or to query for mires, which include bogs, fens, and swamps).

FIELDS	(Descriptor	A4
	IsAlsoA	A4)
PRIMARY KEY	(Descriptor, IsAlsoA)	
FOREIGN KEY	(Descriptor) references DESCR.	
	(IsAlsoA) references DESCR.	

Descriptor Identifier for the entity site-description.
IsAlsoA Identifier that the Descriptor is also one of (e.g. a "fen" is also a "mire").

TABLE DATATYPS

Data types: view on tables SYSCAT and SYSCOL.

The **DATATYPS** table is a view based on tables **SYSCAT** and **SYSCOL** that contains the table name, column name, and data type for each column in the Pollen Database. This view is used by software to generate or check a data type, making software less susceptible to changes in Pollen Database structure.

FIELDS	(TblName	A8
	ColName	A20
	DataType	A4)
PRIMARY KEY	(TblName, ColName)	

TblName	Name of a Pollen Database table.
ColName	Name of a column in that table.
DataType	Data type of that column.