

THE UNIVERSITY OF CHICAGO

LATE CRETACEOUS TO PLEISTOCENE CLIMATES: NATURE OF THE  
TRANSITION FROM A 'HOT-HOUSE' TO AN 'ICE-HOUSE' WORLD

VOLUME ONE

A DISSERTATION SUBMITTED TO  
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

DEPARTMENT OF THE GEOPHYSICAL SCIENCES

BY  
PAUL J. MARKWICK

CHICAGO, ILLINOIS

JUNE, 1996

Copyright © 1996 by Paul J. Markwick

All rights reserved

## CHAPTER III

# THE VERTEBRATE DATABASE, v.3.5.

*"Statistics and probability help clarify one's thinking and improve one's capacity to deal with practical problems and to understand the world..."*

*...On the other hand, important research has been done by people with absolutely no formal knowledge of statistics. And a little study of statistics sometimes befuddles students into thinking that statistical principles are guides to research design and analysis. This mistaken belief only inhibits the exercise of sound research thinking...*

*...If the researcher finds that she/he must use refined statistical tests to reveal whether there are differences, the differences do not matter much...*

*Therefore, we should constantly direct our attention to ensuring that the data upon which we base our calculations is good data."*

Simon, 1992

## III.1. INTRODUCTION

Large datasets, such as that used in this study, present workers with logistical as well as scientific problems. Data must be stored in a readily accessible form in order to facilitate manipulation and analysis. But, while today computerized databases make this process easier, decisions concerning what information needs to be entered and how the database must be structured remain. The Paleogeographic Atlas Project at the University of Chicago has had some 15 years of experience of making these decisions, and this has proved invaluable in designing the database used here, a database that is informally referred to as the "Vertebrate Database," although it houses a variety of geological and

paleoecological information. In addition, the designer of the Vertebrate Database was also the principal data collector, data inputter and data user, and this provided valuable insights into all aspects of the design process, insights that few other database workers have enjoyed.

The "Vertebrate Database" was originally conceived in 1989 as a tool for a study of fossil crocodilian distributions through the last 100 million years of geologic history (Donn, 1987; Habicht, 1979; Markwick, 1992, 1993a, 1993b; Romer, 1961). The database utilizes the Helix Express software for the Macintosh (originally called Double Helix; NorthCon Technologies, 1992). This software allows customized fully relational databases to be easily constructed, without compromising power, and has been used extensively by the Paleogeographic Atlas Project. The original design of this database was based on one constructed for floral studies by Michael Hulver and Jiping Yao, graduate students in the Paleogeographic Atlas Project, University of Chicago. However, the present Vertebrate Database is a product of six years of subsequent development and evolution by the present author. Versions of the present database have now been successfully adapted for other projects and are used by a number of other groups.

### **III.1.1. A Successful Database must be "simple enough that it can be used, but comprehensive enough that it will be useful."**

A central tenet for designing a successful database is that the database be simple enough that it can be used, but comprehensive enough that it will be useful.<sup>1</sup>

---

1. Although the development of the Vertebrate Database has followed this paradigm, the number of fields in each relation is considerable. This reflects the need for qualifying the data. For example consideration of present day vertebrates indicates that size may reflect some climate parameter (Bergmans' Rule). As a consequence a size information field was added to the Taxa by Locality relation. But what size?



Unfortunately, reaching this balance is a question of trial and error, hence the evolution of the database over the last six years. After going over the database for a fifth time adding additional information as I had found needed, I decided that the best solution was to enter all vertebrate and pertinent geological data for each locality, and this is presently being completed. One consequence of this is that the database now comprises the largest dataset of vertebrate information of which I am aware (5700 localities, 17167 taxa, 26576 entries, 1060 present day climate stations with complete information with 51558 taxa by stations, 1024 references; memory size  $\approx$ 170 MB). Its comprehensiveness allows additional issues to be addressed, issues that impinge upon paleoclimatic questions including questions of paleobiology, taphonomy, paleoenvironments, paleoecology and paleobiodiversity. Inclusion of modern day data means that contemporary macroecological questions can also be addressed.

A further important qualification for any comprehensive database is that it have the flexibility to change with time as required. Although initially set up for fossil vertebrates, versions of the database are also now being used for the compilation of both macro- and micro-floral data.

In order to facilitate data entry and utilization, the database has been designed to be as 'user-friendly' as possible. To this end "Help" buttons are available on every entry form, which give a field by field description of what information should be entered and what format it should have. Where possible the database is designed to make calculations automatically and fill appropriate fields; for instance when a stratigraphic name (e.g. "Eocene") is entered in the specified field in the "Main Locality" relation, the database

---

the average? the maximum? Thus it was necessary to provide fields for minimum, maximum and average size (the average is calculated).

automatically calculates the absolute ages represented; similarly, on entry of a new genus name in the "Main Taxonomy" relation the database automatically appends the appropriate higher taxonomy to it. This is done through a series of peripheral relations that remain unseen by most users (see section III.2.4). In addition all entry forms include copious "comments" fields which provide the user with further freedom to qualify entries or include notes for future reference. Helix Express also allows for fields to be qualified, such that entries that fail specified criteria result in an error message (for instance latitude can never exceed +90°; if a number greater than 90 is entered the database is designed to refuse to enter the error until it is corrected).

But, a database is only as good as the data that it contains, and one recurring concern has been that the data taken from the literature are often unreliable. In such a broad database, which here includes not only vertebrate paleontology but also stratigraphy, sedimentology, paleoecology, paleoclimatology and tectonics, it is often difficult to grasp fully the intricacies of every entry. To this end all information is entered into the database unless blatantly false. Entry forms are designed to distinguish between well or poorly resolved data, whether it be geographic or temporal. Thus a locality described as the "Eocene of India" is only of moderate use, but if that is the only record for India for that time interval, it is important. In this case, sorts of the database can be run using the "Geographic Precision" code as a means of discriminating such coarse data: this is a 5 tier code with 1 being the best resolved site (position within 1 km) and 5 is the least resolved (country scale, >500 km). As better resolved data are entered, such coarse records can be replaced. A further problem is that few complete faunal lists are ever published, and one is constantly dependent upon partial lists (particularly biased towards mammal or dinosaurs lists) or systematic descriptions. In this instance the user can designate the nature of the

faunal list using the "Nature of faunal data field," and again, sorts can be designed to list only "complete" faunal lists. Such codes are replete throughout the database as the only way of dealing with a literature that is inherently nonstandardized and problematic and a record that is at best incomplete.

One last point concerning the database is that interpretative fields are kept to a minimum. Interpretation should not be muddled with the data itself. Some interpretation is of course unavoidable, such as for designating the environment. In such cases I have resorted to a broad brush approach, i.e. "fluvial" or "marine," classifications that would be hard to dispute, as opposed to a finer classification such as "ox-bow lake" which may be less easy to demonstrate and which may in any case be beyond the resolution of most locality records used in the database. To this end the database is designed to allow the user to pull back to the resolution at which more confidence may be applied. For instance the distribution of crocodilians may be displayed with more confidence than say the distribution of alligators which may be susceptible to a greater degree of misidentification. However, fields for greater detail are provided, because throughout the design process I have felt it important that the range of resolution obtainable using the database not be limited by the database itself.

Only a brief summary of the basic structure and nature of the database is presented here. A complete description is given in Appendix D.<sup>2</sup> In the following sections screen pictures of each entry form (with an example in each case) are used to show the types of data included. Most fields are self-explanatory (viz., latitude, longitude, age). Nonetheless

---

2. This is based on an unpublished manuscript, "*The Vertebrate Database User's Manual, v.3.0*" (Markwick 1993, unpublished).

a short text discussing why certain fields are included and, the decision process responsible is provided for each relation.

## **III.2. THE DATABASE**

### **III.2.1. Basic Structure**

As with any relational database, this one is divided into individual tables called "Relations." Each of these stores related information. The present database comprises six principal relations: one each for taxonomic information (fossil and modern--in the database referred to as the "Main Taxa" relation), locality data (fossil only--"Main Locality") and Climate station data (modern only--"Climate Stations"), and two that provide the links between them, taxa by locality (fossil occurrences--"Main Taxa By Locality") and taxa by climate station (modern occurrences--"Taxa By Climate Stations"). Reference information is stored in the "Main References" relation, from which it is related to all other relations. This then provides a comprehensive bibliography for all entries. In addition there are peripheral relations that contain information used to make the database more 'user friendly' - these include one for Standard Taxonomy (this stores the higher taxonomy used throughout the database such that whenever a new taxon is entered in the main taxonomic relation ("Main Taxonomy") the database accesses this standard taxonomic relation to see what higher taxonomy to use -- the taxonomy used throughout the database is that of Carroll (1988); this relation is referred to as "Standard Taxonomy"); Timescale (absolute ages assigned to stratigraphic ages: two relations are involved here, "Timescale," which stores the international stratigraphic names, and "Timescale II," which stores local names, including Mammal ages), Geography (using the area codes used by the Paleographic Atlas

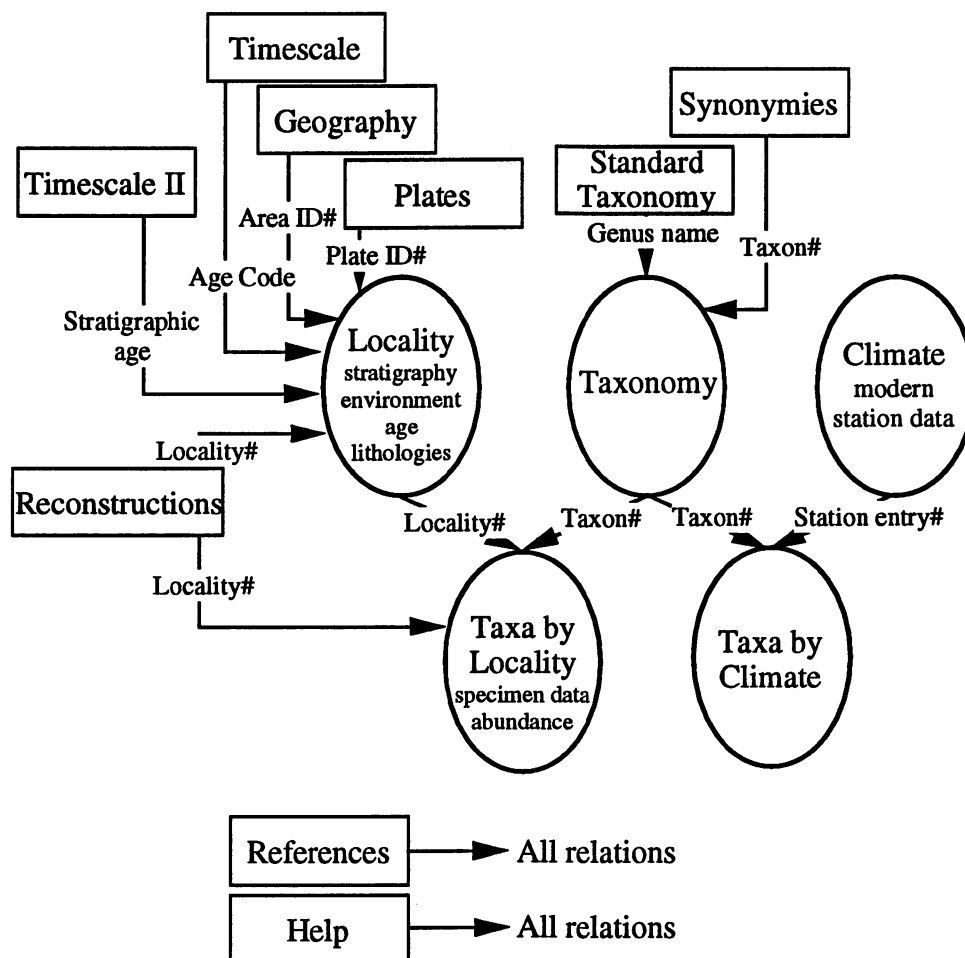
Project in their databases--"Geography"), Plates (also using Paleogeographic Atlas Project codes used for reconstructing the paleoposition of localities--"Plate ID's"), and journal information ("Journal Lookup").

There are two further important design qualifications that all databases should follow. First, each record should contain a unique identifier that is true for that record, but no other. This facilitates linkages between relations. Although it is often tempting to ascribe some additional meaning to this identifier (inclusion of an age code, for instance), experience has shown that a wholly arbitrary identifier is to be preferred and to this end such identifiers in this database are sequential numbers generated by the database itself.<sup>3</sup> Second, there should be only one point of data entry (entry form) in each relation. This avoids potential duplication of data, as well as potential confusion when linking relations. The basic structure of these individual relations is shown in Figure III.1.

Once data have been entered, they can be viewed in a variety of different ways. As individual records (viz., each entry form in turn), as simple listings within a single relation (for instance a list of taxon names with associated taxonomy; Figure III.2.), or as subforms linking multiple relations (for instance, a specified genus name with all its occurrences; Figure III.3.). In this regard, Helix Express provides the user with a great deal of freedom. What fields are listed and what multirelational summaries are made is entirely at the user's discretion, as are the structure and choice of each index and query used. In the "Vertebrate Database" the basic query for each relation is based on the entry form and consequently any field or combination of fields can be queried. This now also includes boolean searches. The power to search the database for any combination of factors in any combination of relations

---

3. One problem with ascribing meaning to the unique identifier occurs when that meaning, say the age, changes.



**FIGURE III.1. Basic Database Structure.**

This diagram shows the principal and peripheral relations in the present database together with the main linking fields. A full description of the database is given in Appendix D.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other

List Main Taxa

25 taxa **MAIN TAXON LISTING**

7495	†Alligator chineasis	7492	†Alligator curvieri	7494	†Alligator helois
3206	Alligator lucius	3006	Alligator mcgrewi	3008	Alligator mefferdi
3005	Alligator mississippiensis	3010	Alligator olseni	3011	Alligator parabyensis
3009	Alligator prenasalis	7496	†Alligator sinense	3256	Alligator sinensis
1	Alligator sp.	3007	Alligator thomsoni	3914	Alligator visleri
4040	Alligatorellus beaumonti	4039	Alligatorellus beaumonti bavaricus	4037	Alligatorellus sp.
3065	Alligatoridae indet.	329	Alligatoridae indet.	4043	Alligatorium depereti
4044	Alligatorium franciscanum	4042	Alligatorium maireri	4045	Alligatorium nainianense
4041					

List Main Taxa with taxonomy

25 taxa **MAIN TAXON LISTING**

Taxon#	Occur.	Taxon Name	Family	Order	Class
7495	0	†Alligator chineasis	Alligatoridae	Crocodylia	Reptilia
7492	0	†Alligator curvieri	Alligatoridae	Crocodylia	Reptilia
7494	0	†Alligator helois	Alligatoridae	Crocodylia	Reptilia
3206	1	Alligator lucius	Alligatoridae	Crocodylia	Reptilia
3006	1	Alligator mcgrewi	Alligatoridae	Crocodylia	Reptilia
3008	1	Alligator mefferdi	Alligatoridae	Crocodylia	Reptilia
3005	4	Alligator mississippiensis	Alligatoridae	Crocodylia	Reptilia
3010	11	Alligator olseni	Alligatoridae	Crocodylia	Reptilia
3011	0	Alligator parabyensis	Alligatoridae	Crocodylia	Reptilia

Taxo

3206	+	1	0	Alligator lucius	Amphibious				
3006	+	1	0	Alligator mcgrewi	Amphibious				
3008	+	1	0	Alligator mefferdi	Amphibious				
3005	⊕	4	9	Alligator mississippiensis	Amphibious	aquatic	carnivore	snakes, turtles, snails, insect	
3010	+	11	0	Alligator olseni	Amphibious				

FIGURE III.2. Screen picture of typical listings.

In order to illustrate the types of listings used in the database three are shown in this screen picture, all from the "Main Taxonomy" relation. The top listing shows a list of all taxa in the database that have the genus name 'Alligator' together with the unique Taxon number that applies to each; the listing is read from left to right and then top to bottom. Double-clicking on any of these entries brings up the appropriate entry record (see Figure III.5). The Helix Express software gives the database user *carte blanche* in designing each listing, the choice of fields included, their format, how they are arranged and the order in which they are shown. A query is assigned to each listing. For example the listing in the middle of the figure shows the same taxa, but this time with associated higher taxonomy--the abbreviation "Occur." relates to the number of records in the "Main Taxa by Locality" relation which relate to this taxon. The bottom listing shows the taxon names with associated ecological information.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other

Quick Taxa with localities

Enter taxon# and tab: **3010** **3010**

*Alligator olseni*

11 Records Age data, Ma: 28.3 to 16.3 longevity, Ma: 12.0

Enter Taxon# and Tab: **3010**  
If you can't remember the Taxon: *Alligator olseni*

Family: **Alligatoridae**

Locality Name

Loc #	Locality
265	Switchyard B of Seaboard Air Line Railroad Company, Tallahassee, Leon County, Florida, USA, NAM
1834	Griseom Plantation, Leon County, Florida, USA, NAM
1837	Midway, Gadsden County, Florida, USA, NAM
1838	Thomas Farm, Gilchrist County, Florida, USA, NAM
2218	Unit 5, Thomas Farm, Gilchrist County, Florida, USA, NAM
2219	Unit 6, Thomas Farm, Gilchrist County, Florida, USA, NAM
2220	Unit 7, Thomas Farm, Gilchrist County, Florida, USA, NAM
2221	Unit 8, Thomas Farm, Gilchrist County, Florida, USA, NAM

Locality Name	Age	Environment	Notes
Switchyard B of Seaboard Air Line Railroad Company, Tallahassee, Leon County, Florida, USA, NAM	L	5	Terrestrial in det.
Griseom Plantation, Leon County, Florida, USA, NAM	L	5	Terrestrial in det.
Midway, Gadsden County, Florida, USA, NAM	L	5	Terrestrial in det.
Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 5, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 6, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 7, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 8, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 11, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Unit 15, Thomas Farm, Gilchrist County, Florida, USA, NAM	L	5	Sink hole
Toledo Bend Dam (near), Newton County, Texas, USA, NAM	L	5	Fluvial

FIGURE III.3. Screen picture of typical subform listings.

As a truly relational database this database provides links between each relation. In the example shown in this figure the user has entered the taxon number appropriate to the taxon of interest, in this case *Alligator olseni*, and the database has accordingly constructed a listing of all occurrences recorded in the "Main Taxa by Locality" relation with this taxon. As with the simple listings shown in Figure III.2., double-clicking on entries in these subform listings provides access to the appropriate record (in this case displayed using the main entry form in the "Main Taxa by Locality" relation --see Figure III.8). Subforms consequently provide the ability to access all information from one part of the database applicable to entries in another part. Not only this, but the link between relations can be based on any entry field. Thus the user might specify a basin name in the "Main Locality" relation and the database access all taxa, or genera or even skulls (if so desired), etc., that occur in that basin. In the case of this figure the database is used to access the paleoenvironmental information applicable to all localities in which the taxon *Alligator olseni* occurs.



means that the dataset can be readily used for addressing a multitude of questions that would otherwise be precluded. For more information, readers are referred to the Helix Express User's manuals (Harrington, 1992).

### **III.2.2. Data Sources**

The dataset used in this study is derived from the published literature. Initially data were derived from Steel's (1973) review of fossil crocodilians and an unpublished term paper written by Arthur Busbey, while a graduate student at the University of Chicago. As the project developed a more comprehensive survey became necessary, utilizing, as far as possible, the original literature. This involved a comprehensive survey through the geological literature housed in the libraries of the University of Chicago and the Field Museum, Chicago. Additional references were obtained through Interlibrary Loans, courtesy of the John Crerar Library. I gathered additional Chinese literature from the library of the Nanjing Institute of Paleontology during field work in southern China in 1991. To maximize the scope of the collection, these literature surveys were implemented in the following way: first all of the principal paleontological journals were checked, beginning in each case with volume 1 and proceeding to the present. This was followed by searches of more general geological journals, followed by national, regional and local museum and geological survey reports from around the world. The benefits of this approach are that obscure manuscripts that may not be accessible by computerized bibliographies, such as GEOREF, or the biases inherent in most publications. The initial concentration on crocodilian data was rapidly expanded to include turtles, as the need to qualify absences became important (the use of control groups is discussed below). Ultimately it became important to survey all vertebrate localities in order to examine the potential biases inherent

in collection, and also to address other problems such as dating and environmental considerations. In order to expedite this as quickly as possible, these latter data were based on reviews of mammals and dinosaurs. I am at present augmenting the dataset with complete faunal lists, where available in the primary literature. The net result of this compilation is that any individual taxon or fossil group can be traced back through time. In addition broader questions of paleoecology and paleodiversity can also be addressed.

### **III.2.3. Principal Relations**

The following subsections outline the information and structure of each of the six primary relations. In each case, a screen picture of the entry form (with representative data) is presented. As already stated above, a detailed description of each field and abacus (calculations made by the database, including lookups for specified parameters) is given in Appendix D.

#### **III.2.3.1. Main References**

This relation contains all reference information (Figure III.4.). These data are then linked to all other parts of the database using a unique, but arbitrary, record number. Most of the fields in this relation are self-explanatory, author name and so forth. All others are included to facilitate 'housekeeping' efforts, for instance the name of the initial compiler (the person who first entered data into this record), the location of the reference (whether it is stored in the present authors library, or is still on the library shelf<sup>4</sup>) and PAPeRS record number, which refers to the bibliographical database of the Paleogeographic Atlas Project.<sup>5</sup>

---

4. All references used in the database have been photocopied and are stored in the present author's library. This follows the example of Alfred Ziegler and the Paleogeographic Atlas Project.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other

ENTER References

### REFERENCE DATA ENTRY SHEET

This sheet is for the entry of reference information. If you have an queries click on the HELP button below or contact the Paleogeographic Atlas Project. Tel. (312) 702 8146

Reference# 1  PAPeRS Ref#

Compiler  Date 1/26/91 Reference type:   
☐ P.J.M. Collected? ☒ Yes ☐ No Category:   
Location of Reference?  Library  Call#

Author 1, Last Name:  Other Author(Full Names)   
Author 1, First Names:   
Author 2, Last Name:   
Author 2, First Names:   
Year of Publication:   
Reference Abbreviation:

Title (Paper, book etc.):

Journal Name (Full):   
Volume:  Number:  Pages / No. of pages:

File Edit User View DATA ENTRY Listings Dumps Display Codes Other

ENTER References

Volume:  Number:  Pages / No. of pages:   
Volume Title (if a special Volume, collection of papers etc):

Editor 1, Last Name:  Other Editors   
Editor 2, First Names:   
Editor 2, Last Name:   
Editor 2, First Names:   
Edition:   
Language Used:   
Abstract Language:   
Translator 1, Last name:  Translator 1, First names:   
Translator 2, Last name:  Translator 2, First names:   
Publisher:   
Place of Publication:

Comments / Notes:

Abstract

FIGURE III.4. Screen picture of principal entry form in the "Main References" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.


### III.2.3.2. Main Taxonomy

This relation stores all taxonomic information used throughout the database (Figure III.5). In order to maintain consistency, the higher taxonomy attributed to each genus is that given in Carroll's (1988) review of vertebrate paleontology. This information is stored in a separate relation--"Standard Taxonomy" (see section III.2.4.1. and comments therein). On entry of a name into the "Genus" field, the database automatically accesses this subsidiary relation and enters the appropriate higher taxonomy in the appropriate fields in the "Main Taxonomy" relation. A taxon is here defined as the finest identifiable taxonomic level for specimens at a locality. Thus a species is a taxon, but so is a family where that is the best level of resolution defined. In cases where specimens at a single locality are identifiable to both a species and its family (e.g. *Alligator mississippiensis* and Alligatoridae indet. are defined at the same locality), the database will treat these as two separate taxa present at that locality. This then carries the uncertainty in the Alligatoridae indet. designation, since although the specimens may indeed represent *Alligator mississippiensis*, they could also represent another alligatorid. The total number of taxa at a locality thus represents the maximum number of taxa given the specimens present, and is different from the total number of genera or species, which are derived based on only defined genera and species, respectively.


Because this relation stores all taxonomic information, including that of extant taxa, a number of fields are included here which are only appropriate to studies of the Recent.

---

5. Most of the codes used throughout the "Vertebrate Database, v.3.5" are those of the Paleogeographic Atlas Project (PAP). They are included in order to provide consistency between the existing databases of the PAP, most of which have also been constructed using the Helix Express software. Most of this is the work of Michael Hulver, who has been involved with PAP since the early 1980's.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other 



ENTER MAIN Taxonomy

4 occurrences  MAIN TAXON DATA ENTRY 9 stations  
 Click for occurrences Click for stations

Taxon# 3005 *Alligator mississippiensis* (Daudin, 1802)

Synonymized to:

**TAXONOMIC DATA**





GENUS *Alligator*  
 Species *mississippiensis*  
 Plant Organ   
 Describer(s) Daudin  
 Year described 1802 ( )? ( )   
 Common American Alligator  
 Synonymy# Ref. p.

To enter 'minor' synonyms click this button


Subfamily  
 Family Alligatoridae  
 Superfamily  
 Infraorder  
 Suborder Eusuchia  
 Order Crocodylia  
 Superorder Archosauria  
 Subdivision  
 Division  
 Infraclass Archosauromorpha  
 Subclass Diapsida  
 Class Reptilia  
 Superclass

**'Major' Synonyms:**  
 #7491 - *Crocodilus lucius* CUVIER, 1807  
 #7492 - *Alligator curvieri* BORY DE ST VINCENT, 1824  
 #7493 - *Champsia lucia* WAGLER, 1830  
 #7494 - *Alligator holois* COPE, 1865

To access list of 'minor' synonyms click this button

Environmental affinity: Amphibious   
 Habitat:   
  
 Elevation. Min.: Max.: Ave.:  
 Nocturnal/diurnal?   
 Size information: Minimum Maximum Average  
 Total length, mm:

Superclass  
 Phylum Vertebrata  
 Kingdom Animalia

Diet: carnivore   
 snakes, turtles, snails, insects, fish, small mammals, birds

**Size information:**

	Minimum	Maximum	Average
Total length†, mm:			
Tail length, mm:			
Head & Body*, mm:	1800.0	5000.0	3400.0
Skull length, mm:			
Weight, g:			

†skull length for turtles; \*snout to vent

REF# Reference Abbrev.  
 2 Steel (1973)  
 371 Conant and Collins (1991)

pages  
 73  
 39

1991 P.J.M.

COMMENTS, TAXONOMIC

Picture:

FIGURE III.5. Screen picture of principal entry form in the "Main Taxonomy" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

These include information on general diet, habitat and size. While such information can be interpreted for fossil taxa, a precept of this dataset is that such interpretation be kept to a minimum.

Synonymies can be specified in this relation, where known. If defined, the major synonym will be used throughout the rest of the database.<sup>6</sup>

### III.2.3.3. Main Localities

The "Main Localities" relation stores all data pertaining to localities, including stratigraphic, lithologic and paleoenvironmental information (Figure III.6). Each record represents a 'locality,' as described in the literature. However, a 'locality' *per se* has no geographic or temporal definition. In order to facilitate comparisons, temporal and spatial resolution is defined within the database using qualifiers, including the "Geographic Precision" code described in section III.1.1.<sup>7</sup> A qualifier is also used to describe the type of faunal information associated with each locality, viz., "Complete Faunal List," "Composite Faunal List," etc. In this case the user is given only a limited selection of choices by the

---

6. Synonymies are a perennial source of concern in any dataset of palaeontological information. In the "Vertebrate Database" synonymies are dealt with in three separate places. "Major" synonymies (those that always apply to the taxon concerned) are entered in the "Main Taxonomy" relation. Whenever the original taxon is entered as an occurrence in the "Main Taxa by Locality" relation, the database automatically searches to see if a "major" synonymy is defined. If it is, then the database subsequently defaults to the synonymy. In contrast, synonymies may only apply to a specific specimen and this therefore cannot be treated in the same way as synonymies in the "Main Taxonomy" relation. Instead such synonymies are entered in the "Main Taxa by Locality" relation. Alternatively, the "Minor Synonymies" relation provides a means of reviewing the history of synonymies for each taxon, including the name of the reviewer.

7. The Geographic Precision code (GP) is represented by an interger between 1 and 5: 1 = <1km equivalent to site by site resolution; 2 = <10km, equivalent to nearest town; 3 = <100km, equivalent to US county; 4 = <500km equivalent to US state; 5 = >500km equivalent to country scale.


[File](#)
[Edit](#)
[User](#)
[View](#)
[DATA ENTRY](#)
[Listings](#)
[Dumps](#)
[Display](#)
[Codes](#)
[Other](#)
[Wired](#)


ENTER Localities									
Fauna bot:					top.				
Zone bot:					top.				
Local Stage:					top.				
Age Bot	41220	Eocene, Lutetian			}	Age Bot (Ma):		50.000000	
Age Top	41220	Eocene, Lutetian				Age Top (Ma):		42.100000	
Summary: {Eoc2, Lut}						Time Span (Ma):		7.900000	
					Interval Mean Age (Ma):		46.050000		
age given as Lutetian REF#197 p.15 <div style="float: right;">↑ ↓</div>									
<b>COMMENTS, STRATIGRAPHY</b> <div style="float: right;">↑ ↓</div>									
Group:		▼	thickness:		ft	m		m/Ma	
Formation:	Messel Shale	▼	thickness:		ft	190.0	m	24.1	m/Ma
Member:		▼	thickness:		ft		m		m/Ma
Bed:		▼	thickness:		ft		m		m/Ma
Fossil horizon lithologies: shale									
Associated	black laminated oil shale, sands, clays				↑	Horiz PAP Lith M			
Rock Type						Horiz Summary 4			
						PAP Lith MS			
					↓	Summary 3			
HORIZON: Mudstone dominant; ASSOC.LITH: Mudstone with sandstone									
<b>COMMENTS, FORMATION / ROCK TYPE</b> <div style="float: right;">↑ ↓</div>									
<b>ENVIRONMENT / BIOGEOGRAPHIC / CLIMATE</b>									
Environment	Lacustrine ▼			REF#1304			TOTALP <div style="float: right;">↑ ↓</div>		

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other U

### ENTER Localities

**ENVIRONMENT**

PAP Environment Code:  L Land(L), Shelf(S), or Ocean(O) Implied palaeoelevation (m)  
 PAP Environment: Coastal plains, lower river systems, delta tops min 0 max 200  
 Geological recognition: Alluvial complexes; Major floodplain complexes; Swamps and channel sands

Tectonic setting:  **COMMENTS, CLIMATE**  
 Floral data present?  REF#1304 p.23  
 Megaflorea? ☐ wood? ☐ palynoflorea? ☐

**GENERAL COMMENTS**

**REFERENCES QUOTED**

Ref#	Reference	Pages, Fig, Plate, etc.
197	Frey et al (1987). Neues Jahrbuch fuer Geologie und Palaeontologie	15
2	Steel (1973).	59
884	Berg (1964). Berichte der Naturforschenden Gesellschaft zu Freiburg, 58 :	45
1127	Savage & Russell (1983).	91
1304	Schaal and Ziegler (1992).	

Editors	Dates Edited	Editors	Dates Edited	Editors	Dates Edited
P.J.Markwick	Feb 11, 1991				
P.J.Markwick	Jan 14, 1994				

1991 PJM

FIGURE III.6.continued.

database ("Nature of Faunal List"<sup>8</sup>). This ability to restrict possible entries in a field, especially for qualifiers, both simplifies and standardizes the entry process.

Present day distributional data are stored elsewhere in the database using the modern climate stations in the "Climate Station" relation as 'localities' (see section III.2.3.4). However, in order to facilitate range information, which is automatically calculated by the database using occurrences recorded in the "Main Taxa By Locality"

8. This comprises the following choices using a "pop-up menu" (see Table III.1): Complete faunal list; Complete faunal list (composite); Faunal list (Completeness ?); Faunal list (composite - herps only); Faunal list (crocodilians only - review); Faunal list (herps only); Faunal list (systematics); Faunal list (turtles only - review); Faunal list (mammals only - review); Incomplete Faunal list; Incomplete Faunal list (composite); No faunal list given



relation (see section III.2.3.5), the "Present Day" is also represented by a single locality in the "Main Locality" relation (Locality number 0 in the database).

Each locality is assigned a reconstructed paleolatitude and paleolongitude based on the continental plate upon which it is sited and the age assigned to it. In addition a paleoposition is calculated for each locality for each time interval used in the analyses given in Chapters V and VIII (Early Eocene, Middle Eocene, Eocene, etc.), where appropriate. These latter paleolatitudes and longitudes are stored in a separate relation, "Reconstructions" (see section III.2.4.7) and cross-linked with both the "Main Locality" and "Main Taxa by Locality" relations. All reconstructions are derived using the latest rotation parameters calculated by David Rowley, Paleogeographic Atlas Project, University of Chicago.

The age of each locality is given as the stratigraphic age. On entry, the database assigns absolute ages, in millions of years, to these stratigraphic intervals, based on the Harland timescale (Harland et al., 1990). The database also stores absolute age assignments from the timescale of Berggren et al. (1985; also referred to as the DNAG--Decade of North American Geology--timescale). Consequently the database can be run using the Berggren rather than the Harland timescale if required. This information is stored in two separate relations: "Timescale" (standard international terms, which is linked to the "Main Locality" relation using the 'Age Code' assignment--see section III.2.4.2) and "Timescale II" (local stratigraphic names, including mammal ages, which is linked to the "Main Locality" relation by the stratigraphic name itself--see section III.2.4.3). The use of stratigraphic ages reflects the fact that absolute ages bounding stratigraphic intervals are subject to change depending on the latest timescale. By keeping absolute ages in a separate

relation provides for future changes, such that new ages can be entered in the "Timescale" and "Timescale II" relations and then posted<sup>9</sup> throughout the database as appropriate.

One drawback with many computerized databases is the need to standardize entries in order to facilitate searches. For example a fluvial paleoenvironment may be referred to by different data enterers as "river," "riverine," "stream," "fluvial," "channel" and so forth. This is becoming less of a problem as personal computers and database software become more powerful. Indeed in Helix Express, boolean searches for all possible terms satisfying a specified search parameter are facilitated by a special query, called a "power query." Nonetheless, adoption of a standardized terminology streamlines searches. But it is always important to keep track of the more detailed information from which the generality is derived. To this end the "Vertebrate Database" provides fields for both levels of description, and this is exemplified best by the entry fields for paleoenvironmental and lithological data in the "Main Locality" relation (Figure III.6). Hence a pop-up menu provides a list of choices for the general environmental assignment appropriate to each record,<sup>10</sup> but space is provided for a fuller description (see Appendix D). Unlike DOS-based databases, Macintosh databases do not have a pre-specified field size and consequently large amounts of data can be entered in each field; for example, in the case of lithological information the "Associated Rock Type" field may contain a complete lithological summary of an entire section. In order to facilitate searches of such fields, they are designated as "keyword" fields, which using the Helix Express software means that the specified field is individually indexed. Experience has shown that searches of such fields

---

9. Posting refers to the means of disseminating data from one part of the database to another, on command. This can also be done using "Lookup" abaci.

10. The list of choices provided in this menu are as follows: Aeolian; Archaeological Site; Cave; Deltaic; Fissure Fill; Fluvial; Fluvio-lacustrine; Lacustrine; Littoral; Marine; Near shore marine; Paralic; Sinkhole; Tar pits; Terrestrial indet.; Transitional; Unspecified.


are completed almost instantaneously (even when used on a Macintosh Powerbook 180, which is presently the machine on which the database is housed<sup>11</sup>). As with other parts of database, the appropriate Paleogeographic Atlas Project (PAP) code is also included in order to provide compatibility with existing databases of this project and this includes a code for the paleogeographic setting of each locality. (See Ziegler et al., 1985, for the rationale for the use of summary codes by the Paleogeographic Atlas Project; a complete listing of all codes used in the database is given in Appendix D). Environmental data are also coded according to its paleogeographic position using the summary codes given in Ziegler et al., (1985).

#### **III.2.3.4. Main Climate Stations**

This relation stores climate information for 1060 globally distributed climate stations as compiled by Müller (1982; Figure III.7). These stations are used as localities to which are assigned Recent faunal lists (see section III.2.3.6, "Main Taxa by Climate Stations"). In addition to the monthly climate data the dataset now also includes the monthly Normalized Difference Vegetation Index (NDVI) for each climate station (taken from the gridded data arrays provided by the NOAA-EPA Global Ecosystems Database Project, 1992). This NDVI is derived from observations of the Advanced Very High Resolution Radiometer (AVHRR) housed on the NOAA-9 (1985-1988) and NOAA-11 (1989) satellites and gives a measure of the density of vegetation coverage based on the nature of light reflected back from the ground to the satellite. A systematic relationship between this index and Net Primary Productivity (NPP) has been shown in a number of

---

11. All screen pictures are taken from that computer.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other 


ENTER Climate data

### CLIMATE STATION ENTRY FORM

Entry#: 597 Locality (from Müller): 77 PAP area code: 175  
 Continent: North America NAM Country: USA Province: Illinois  
 Station: Chicago *Chicago, Illinois, USA, NAM*  
 Latitude deg: 41 min: 47 Latitude: 41.8 Elevation (m) 185.0 Climate symbol, Köppen: D f a Df  
 Longitude deg: -87 min: -47 Longitude: -87.8 Climate symbol, Troll: III 7 III  
 BIOME (Ziegler, 1990, after Walter, 1985): ▼  
 FAUNAL/FLORAL DATA: Complete amphibia? ☒ reptilia? ☒ mammals? ☒ birds? ☒ plants? ☒

**SUMMARY** MAT, °C: 10.5 MART, °C: 27.6 CMM, °C: -3.3 WMM, °C: 24.3 MMDC, °C: -7.7 MMDW, °C: 27.7  
 ACT, °C: -30.6 AWT, °C: 40.6 Months MDT ≤ 0°C: 3 MDT ≥ 5°C & P ≥ 40mm: 7 MDT ≥ 10°C & P ≥ 40mm: 6  
 MEANS: Winter T: -2.4 P: 45 Spring, T: 9.2 P: 81 Summer, T: 23.1 P: 90 Autumn, T: 12.1 P: 65

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	YEAR	p	%
Mean daily temperature, °C	-3.3	-2.3	2.4	9.5	15.6	21.5	24.3	23.6	19.1	13.0	4.4	-1.6	10.5	30	
Mean daily max. temperature, °C	0.0	1.1	6.1	12.8	18.3	23.9	27.2	26.1	22.8	16.1	8.3	2.2	19.7	75	-0
Mean daily min. temperature, °C	-7.7	-6.7	-1.7	4.4	10.0	15.6	18.9	18.3	14.4	8.3	1.1	-5.0	5.8	75	0
Absolute max. temperature, °C	18.3	20.0	27.8	32.8	36.7	38.9	40.6	38.9	37.8	31.1	25.6	20.0	40.6	77	
Absolute min. temperature, °C	-28.9	-29.4	-24.4	-8.3	-2.8	1.7	9.4	8.3	-1.7	-10.0	-18.9	-30.6	-30.6	77	
Mean relative humidity, %	75	76	71	68	66	69	67	70	70	69	72	76	71	12	
Mean precipitation, mm	47	41	70	77	95	103	86	80	69	71	56	48	843	30	
Maximum precipitation, mm	96	85	127	212	193	163	228	150	153	306	95	169		16	se
Minimum precipitation, mm	10	8	8	33	20	20	34	25	12	8	23	9		16	not
Maximum ppta in 24 hours, mm	73	39	64	104	74	116	159	79	65	143	47	60	159	18	
Mean # days w/ppt ≥ 0.25 mm	10	10	12	13	12	11	9	8	7	10	10	10	120	18	
Mean duration of sunshine, hours	126	142	199	221	274	300	339	299	247	216	136	118	2611	30	
Mean quantity of radiation, ly/day	96	147	227	331	424	458	473	403	313	207	120	76	273		
Mean potential evaporation, mm	0	0	6	37	79	122	146	192	90	51	12	0	675	51	
Mean windspeed, m/sec	5.0	5.0	5.0	5.0	5.0	4.0	3.0	3.0	4.0	4.0	5.0	5.0	4.4	18	

File Edit User View DATA ENTRY Listings Dumps Display Codes Other 

ENTER Climate data

	W	W	W	W	SSW	SSW	SW	SW	S	S	W	W		16
Mean predominant wind direction														
(Mean Ppt - Mean Pot. Evap), mm	47	41	64	40	16	-19	-60	-52	-21	20	44	48	168	(P-E)
Moisture Index (MI)	Inf	Inf	106	108	20	-16	-41	-39	-23	39	367	Inf	25	
"Productivity" NDVI**	107	95	109	123	139	138	136	134	135	129	110	107		
NDVI Summary: Mean annual: <u>122</u> Max: <u>139</u> Min: <u>95</u> Annual Range: <u>44</u> Median: <u>126</u> 1 SD: <u></u>														
REF: <u>1302</u> Müller (1982)														
PP: <u></u>														

**NOTES**

\* some of the yearly values given in Müller, 1982 do not agree entirely with the monthly values, where this is the case the difference between calculated and Müller's total (calculated - Müller) is put in the column on the right titled with the "?" (started after entry#90).

† the yearly values for maximum and minimum precipitation are not sums of the monthly values, but represent the maximum and minimum annual precipitation values; the monthly values need not all be for the same year.

‡ Mean potential evaporation (PE) is derived using the equations of Thornthwaite, 1948.

\*\* "Productivity" data is actually the calculated value of the Normalized Difference Vegetation Index (NDVI), derived from satellite observations of the "greenness" of the Earth's surface. See NOAA-EPA Global Ecosystems Database Project, 1990 Global Ecosystems Database Version 1.0 User's Guide, Documentation, Reprints, and Digital Data on CD-ROM.

USDOC/NOAA National Geophysical Data Center, Boulder, CO.

One further note should be made. Values of <1 (or <0.1) are represented in Müller by <1 (or <0.1) or tr (trace). In this datab these are entered as the values 1 or 0.1 depending upon the decimal places used in adjacent values. In these cases the total ye values need not add up to the value given in Müller, which in general is the value used here.

**ABBREVIATIONS**

MAT = Mean Annual Temperature, °C; MART = Mean Annual Range of Temperature, °C; CMM = Coldest Month Mean Temperature, °C; WMM = Warmest Month Mean Temperature, °C; MMDC = Minimum Mean Daily Minimum Temperature, °C; MMDW = Maximum Mean Daily Maximum Temperature, °C; ACT = Absolute Minimum Temperature, °C; AWT = Absolute Maximum Temperature, °C; MDT = Mean Daily Temperature, °C; MMP = Mean monthly ppta, mm; MDMAT = Mean Daily Minimum Temperature, °C; P = Mean precipitation; p = period (in general the 30 refers to the period 1931-1960).

**Comments**

Figure III.7. Screen picture of principal entry form in the "Climate Station" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

studies (Goward and Dye, 1987; Goward et al., 1985).<sup>12</sup> Abacii within the database are used to calculate additional metrics, including Mean Annual Temperature (MAT), Mean Annual Range of Temperature (MART), maximum annual NDVI, and so forth. Further calculations are made as required.

The propensity for meteorological stations to occur in lowland sites reflects Müller's original requirements: acceptable stations must contain data for a large array of climate parameters representing time series on the order of 30 years (typical of "climate"). Consequently most stations are located in large towns or cities that historically occupy lowland sites.<sup>13</sup> This bias closely mimics the biases in the geological record that tend to overrepresent low elevation environments (see also Chapter IV).

### III.2.3.5. Main Taxa by Locality

The "Main Taxa by Locality" relation links the data stored in the "Main Taxonomy" and "Main Locality" relations (Figure III.8). Each record represents an occurrence of a

---

12. The data included in the climate station dataset are obtained by taking the monthly NDVI gridded datasets for the years 1986-1990, producing average monthly values for this period, and then extracting values that occur at the geographic sites of each climate station. Applications of the NDVI dataset have been discussed extensively in the remote sensing literature, especially the use of the vegetation index to examine Net Primary Productivity (Cihlar et al., 1991; Goward and Dye, 1987; Goward et al., 1985; Hielkema et al., 1987; Running and Nemani, 1988; Sellers, 1987). Comparisons between satellite derived data and ground-based observations of NPP have shown a systematic underestimate in predicted compared with actual values (Goward and Dye, 1987; Goward et al., 1985). Recently, Lottes and Ziegler (1994) have used the NDVI data to examine the relation between climate, productivity and peat formation. Regardless, the intergration of this dataset with the global diversity and climate dataset, described in this Chapter, provides the potential for examining many present issues in ecology, while also providing further power for examining the past. This will be examined after completion of this dissertation.

13. Large urban areas tend to influence local climate. Peterson (1973) notes that mean annual temperatures (MAT) of cities are generally  $\approx 0.6$ - $0.8$  °C higher than surrounding rural areas, with minimum winter temperatures being  $\approx 1.1$ - $1.7$  °C warmer. This disparity depends on the size of the urban area, for Chicago the difference in MAT is  $0.6$  °C, while for New York City it is  $1.1$  °C. With few exceptions most of the cities and towns used by Müller (1982) are smaller than these extremes and urban effects on climate are minor compared with the overall patterns being investigated.

[illegible]

ENTER Taxa by Locality														
E B R A T E S		N	N	N	N	N	N	N	N	N	N	N	N	N
Mainly Y	O	O	O	O	O	O	O	O	O	O	O	O	O	O
Complete? N	O	O	O	O	O	O	O	O	O	O	O	O	O	O
MNS each category:	1	2	2	0	0	0	0	0	0	0	0	0	0	0
Trace fossils present?:	<input type="radio"/> Yes <input type="radio"/> No <input checked="" type="radio"/> Undefined    Nature of trace fossil: _____													
Wood Palynomorphs Leaves Fruits Seeds Stem/fibre MATERIAL (DETAILS)														
F L O R A	Present ?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mainly complete? N	O	O	O	O	O	O	O	O	O	O	O	O	O	O
MNS each category	0	0	0	0	0	0	0	0	0	0	0	0	0	0
microscopic characters: <input type="radio"/> Y <input checked="" type="radio"/> N														
(T) ant. portion of a snout & lower jaw; (RS) skull & jaw material														
Abundance, Relative _____														
Minimum Number of Individuals (MNI) 2														
Minimum Number of Specimens (MNS) 5														
Is the material mainly Articulated? <input checked="" type="radio"/> YES <input type="radio"/> NO <input type="radio"/> UNDEFINED														
Specimen matrix: shale														
Mode of fossilization: Original material														

**FIGURE III.8. Screen picture of principal entry form in the "Main Taxa by Locality" relation.**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other

ENTER Taxa by Locality

Mode of fossilization:

**SIZE & PHYSIOGNOMIC INFORMATION** HELP

VERTEBRATES (in mm)				FLORA (in mm)			
	Largest	Smallest	estimate?	n=?			
Skull Length:	700	700	observed	1	Margin?	<input type="radio"/> Entire	<input type="radio"/> Non-entire
Skull Width:					Drip tip present?	<input type="radio"/> Y	<input type="radio"/> N
Body Length:	4000	4000		1	Long axis (mm)		
					Short axis, maximum (mm)		
Shell Length:					Wood: length of log (mm)		
Shell Width:					Wood: diameter of log (mm)		
Shell Height:							

*Comments*  
REF#2, p.59; size data from REF#1304 p.111

**REFERENCES**

Ref#	Ref. Abbrev.	Page#s
Ref. 1: 2	Steel (1973)	59
Ref. 2: 884	Berg (1964)	45
Ref. 3: 1304	Schaal and Ziegler (1992)	111
Ref. 4:		
Ref. 5:		

FIGURE III.8. continued.

taxon in a specified locality. Once the link is made, by entering the unique record numbers for the appropriate taxon and locality, additional information specific to that occurrence at that place, can be entered. This includes qualification of the taxon identification ("cf.", "aff.", "?", etc.) but also specimen, size and abundance information. The inclusion of this information (as well as basic taphonomic observations, such as degree of articulation and completeness) allows for some measure of confidence to be assigned to taxonomic identifications. The database can therefore be sorted for only those occurrences that comprise well preserved specimens. In addition, gross taphonomic questions can be addressed.<sup>14</sup>

14. As with so much of this study, this is work in progress. One consequence of such a large, comprehensive dataset is that every question begets additional questions, which can be readily answered. This tends to promote excursions at tangents to the original study, and this may explain the time taken to complete this dissertation.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other **UCED**

Enter Taxa by Climate Station

## ENTER TAXA BY CLIMATE STATION

Entry#: 3005/646      Date: December 14, 1995 20:57      List stations by taxon#  
 Taxon# 3005      Syn#      Climate station Entry#: 646      Area data      List taxa by Station#

Genus: Alligator      Species: mississippiensis      (Daudin, 1802)  
 synonym:  
 Subfamily:      Family: Alligatoridae      Superfamily:  
 Infraclass: Archosauria      SubOrder: Eusuchia      Order: Crocodylia  
 Superorder: Archosauria      Subdivision:      Division:  
 Infraclass: Archosauriformes      Subclass: Diapsida      Class: Reptilia  
 Superclass:      Phylum: Vertebrata      Kingdom: Animalia

*Alligator mississippiensis*

Principle Biogeographic Reference:      pages      Geographic region  
 Ref# 363      Grenard (1991)      74      Florida  
 Additional Biogeographic References:  
 Ref#  
 Ref#  
 Ref#

**Apalachicola, USA, NAM**

Lat: 29.7 Long: -85.0 Elev: 4 Station: Apalachicola Province Florida  
 Country USA Continent North America NAM Köppen Symbol: Cfa Troll Symbol: IV.7  
 MAT: 20.5 CMM: 12.8 WMM: 27.5 MART: 14.7 MMDC: 8.9 MMDW 30.8  
 Mo ≤0°C: 0 mo MDMint≤0°C: 0 AWT 39.0 ACT -8.0 Mean radiation: 444  
 Mo ≥20mm ppt: 12 ≥40: 12 Mo P-E >0mm: 9 ≤0mm: 3 P 1429 P-E 349  
 Months MDT≥5°C: 12 Mo MDT≥10°C: 12 Mo MDT≥5°C, ppt≥40mm: 12 MDT≥10°C, ppt≥40mm: 12  
 Mean Winter TC: 13.3 Mean Spring TC: 19.9 Mean Summer TC: 27.3 Mean Autumn TC: 21.4

**FIGURE III.9. Screen picture of principal entry form in the "Taxa by Climate Station" relation.**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

### III.2.3.6. Main Taxa by Climate Stations

This relation links the appropriate taxon (from "Main Taxonomy") to the recent locality--climate station ("Main Climate Stations") in which it occurs (Figure III.9). An occurrence is based on the intersection of a taxon's present distribution with a 50 km radius circle drawn around each station.



### III.2.4. Subsidiary Relations

Subsidiary relations store information that is used to facilitate data entry and avoid repetitious tasks in the six primary relations described above. By isolating this information into separate relations, subsequent changes (for instance to the timescale) can be effected without going to each individual record in the primary entry relations. This makes updating easier and therefore more likely to be enacted.

#### III.2.4.1. Standard Taxonomy

In order to ensure a consistent taxonomic hierarchy throughout the database,<sup>15</sup> all entries in the "Main Taxonomy" relation derive their higher taxonomies (above generic level) from a 'standardized' taxonomy contained within the "Standard Taxonomy" relation (Figure III.10). At present this standard taxonomy is that given in Carroll (1988). Genera described subsequent to 1988 are placed in Carroll's scheme according to their subfamily or family assignments. Other absences are compared with Romer's (1971) classification, which forms the basis of Carroll's work. In the case of spelling (or any other) inconsistencies, experience has suggested deference to Romer.

The use of Carroll (1988) should not be taken as this author's wholehearted acceptance of his comprehensive classification. It is more a reflection of expediency, since this study is not intended to be a phylogenetic one, but requires some level of consistency

---

15. As discussed earlier, consistency within a database ensures the success of sorts and queries of the data. However, it must be remembered that consistency is not necessarily a guarantee of truth. A detailed survey of crocodilian classification is given in Clark (1986), which also provides the most recent cladistic analysis (Benton and Clark, 1988); readers are also directed to Norell's (1989) analysis of extant taxa.

**ENTRY SHEET FOR STANDARD TAXONOMY**

Genus:  Source:

Synonymized to:  Synonymies:

Subfamily:

Family:

Superfamily:

Infraorder:

Suborder:

Order:

Superorder:

Subdivision:

Division:

Infraclass:

Subclass:

Class:

Phylum:

Age Btm:  Age Top:

Age Top (Ma): 0.0  
Age Btm (Ma): 35.4  
Duration: 35.40 Ma

*Oligocene to Holocene*

Ref:  PP:

Ref:  PP:

FIGURE III.10. Screen picture of principal entry form in the "Standard Taxonomy" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

in the classification used. Carroll's scheme is thus a compromise, providing a recent, readily accessible and comprehensive source for all vertebrates.<sup>16</sup>

#### III.2.4.2. Timescale

As described in section III.2.3.3, the ages ascribed to each locality are generally the stratigraphic ages. Absolute ages applied to these are derived from two relations,

16. Discussions with a number of vertebrate palaeontologists has suggested that this choice, although not without its problems, is presently an acceptable compromise and the least likely to cause offense.

File Edit User View DATA ENTRY Listings Dumps Display Codes Other Help

Paul's PB180

et 250-Q

Enter TimeScale

	#	Abbr.	Name	Comments
Era	4	Cz	Cenozoic	
Period	2	Q	Quaternary	
Epoch	1	Ple	Pleistocene	
Stage	2	mPle	middle Pleistocene	
Div	3	lmPle	Late middle	

Stratigraphic Code = 42123

	DNAG 83	Int?	Harland 89	Int?
Top Age:	0.01 Ma		0.01 Ma	
Bottom Age:	0.15 Ma	Y	0.15 Ma	

Trash

**FIGURE III.11. Screen picture of principal entry form in the "Timescale" relation.**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

"Timescale" and "Timescale II." The "Timescale" relation contains the ages for the international stratigraphic names given in Harland et al. (1990) and is accessed through a 5-digit code, representing the Era, Period, Epoch, Stage and stage division respectively (Figure III.11). On entry of this code in the appropriate entry field in the "Main Locality" relation, the absolute ages are automatically entered in the locality record. This code was originally constructed by Michael Hulver, a graduate student at the University of Chicago, for use in the Lithologic Database of the Paleogeographic Atlas Project. This database was used extensively by the group's oil company sponsors. Using a numeric code greatly accelerated access times using the software then available.

**File Edit User View DATA ENTRY Listings Dumps Display Codes Other**

**Enter Local Time Units**

**ENTER LOCAL TIME UNITS**

Time Unit Name:

Nature of Unit:

Geographic Area:

	Harland 89	Berggren 85	DNAG 83
Age Top (Ma)	7.31	7.00	
Age Btm (Ma)	8.28	8.20	

If the Unit given is defined in time click yes; if ESNE is used click No  
☒ YES ☐ NO

**EUROPEAN STAGE NAME EQUIVALENTS (ESNE)**

Btm Code:	41453	Mio3, ITor	7.93 Ma
Top Code:	41453	Mio3, ITor	6.70 Ma

**Comments**  
 related to magnetostratigraphy following Berggren et al 1985; the Harland values are obtained by fitting the chron boundaries to the Harland Time scale and interpolating the Local Time unit boundaries from this derived curve

Initial Compiler:  Date:

Ref#:  Woodburne (1987) pp

Def#:

Trash

**FIGURE III.12. Screen picture of principal entry form in the "Timescale II" relation.**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

### III.2.4.3. Timescale II

This relation contains the absolute ages for local stratigraphic names and others not included in Harland (1990) (Figure III.12). This includes local stage names, mammal stages, and zones. This relation provides a greater resolution of dating to localities than is provided by using the established epochs and European marine stage names. Unlike the "Timescale" relation, this relation is not accessed by a code, but on entry of the appropriate

**File Edit User View DATA ENTRY Listings Dumps Display Codes Other Help**

**Enter Areas**

**ENTER AREA CODES**

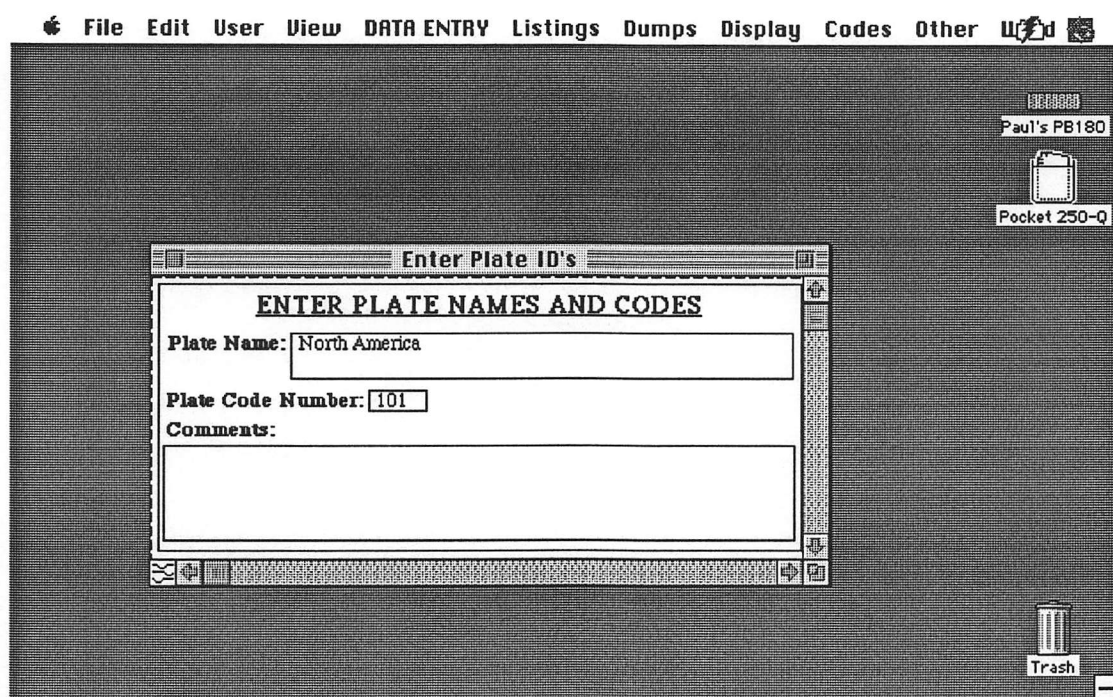
Area Code: 312  
 Area Abbr.: England  
 Area Name: England  
 State/Prov.:  
 Country: England (United Kingdom)  
 Continent: Europe  
 Former country names:  
 Area (km2): 130360  
 Ref#: 1125 Times Atlas of the World (1988) P.  
 Population: 46846000 Census year: 1983  
 Ref#: 1125 Times Atlas of the World (1988) P.  
 Comments

**FIGURE III.13. Screen picture of principal entry form in the "Geography" relation.**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

stratigraphic name in the appropriate entry field in the "Main Locality" relation.<sup>17</sup> For each record the "Time Unit" is tied to chronostratigraphic ages where they are provided. Harland equivalents are generally obtained using magnetostratigraphy where available and appropriate. Each "Time Unit" is also related to a European Marine Stage equivalent in order to provide some continuity with the general timescale used throughout the database.

17. Given the increased speed of presently available software and computers, compared with when the database was originally designed, the use of the stratigraphic name rather than a code is now easily facilitated. For new users, this will probably be the preferred means of data entry rather than learning or looking up codes.



**FIGURE III.14.** Screen picture of principal entry form in the "Plate ID's" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

#### III.2.4.4. Geography

In all of the Paleogeographic Atlas Project databases, each area is assigned a 3-digit code number (Figure III.13). As with the 5-digit timescale code described above (see section III.2.4.3), the PAP area code was originally designed to make searches more efficient. Although no longer as necessary, it is maintained in this database again to retain compatibility with existing PAP datasets. All of these codes are given in Appendix D.



**File Edit User View DATA ENTRY Listings Dumps Display Codes Other**

**Enter Journals**

**ENTRY SHEET FOR JOURNAL INFORMATION**

Journal #: 7

Journal Name (Full): American Museum Novitates

Journal Abbreviation: Amer.Mus.Novit.

Publisher:

Library Held: Crerar \*ILL details:

Library Code: QH1.A532

Complete set? ☒ Yes ☐ No Details if No:

Comments:

**FIGURE III.15. Screen picture of principle entry form in the "Journal Lookup" relation**

An example record is illustrated to show the type of information entered in each field. See Appendix D for details.

#### III.2.4.5. Plate ID's

Each tectonic plate recognized by the Paleogeographic Atlas Project for their plate reconstructions is assigned a 3-digit number (Figure III.14). This number is then assigned to each locality that is sited on it, and this, together with the present coordinates of the locality and its age, are then used to reconstruct the locality's paleoposition. A listing of all plate codes and the names of the plates is given in Appendix D, which also provides a series of maps that show the geographic extent of each plate.

#### **III.2.4.6. Journal Lookup**

This relation stores all journal information including the library in which the journal was located (Figure III.15). A direct link with the "Enter References" entry form ("Main References" relation) is made through a dynamic pop-up menu using the "Journal Name" entry field.

#### **III.2.4.7. Reconstructions**

Reconstructed coordinates are stored in two places in the database: in the "Main Locality" relation, where paleocoordinates are calculated based on the actual age of the locality, and in the "Reconstructions" relation, where paleocoordinates are calculated for one designated age for each interval in which the locality 'occurs' (Figure III.16). This is the source of information for the maps shown in Chapter V, and ensures that all of the reconstructed coordinates for each map are appropriate to plate reconstructions shown on that map. Thus, for example, all localities that occur within, or include, the Early Eocene have reconstructed paleocoordinates for 55 million years, despite the fact that in the "Main Locality" relation they may represent the earliest part of the Early Eocene and therefore they have paleocoordinates representing, say, 56.7 million years. Again this is a question of standardization.

Entry into this relation is accomplished through loads from external text files, which are derived as output from PAP computer programs used to calculate rotated coordinates. Data in this relation are accessed by the "Main Locality" and the "Main Taxa by Locality" relations using the locality record number as the link.



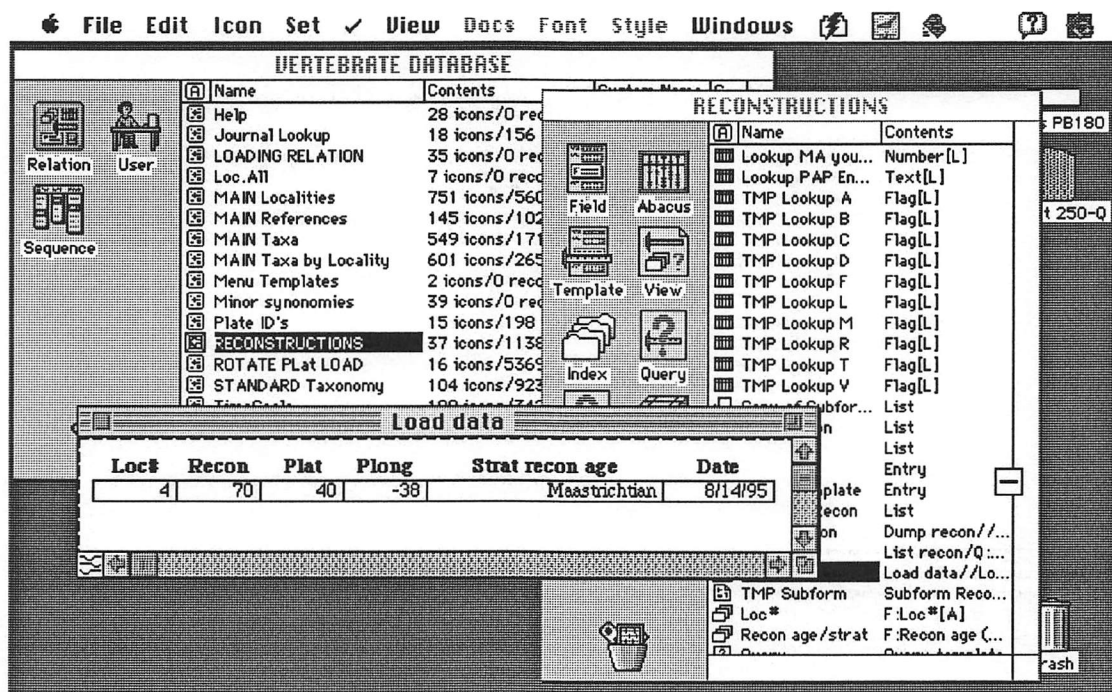


FIGURE III.16. Screen picture of principal entry form in the "Reconstructions" relation.

An example record is illustrated to show the type of information entered in each field. See Appendix D for details. In this figure the background to the entry form comprises lists of the relations in the database and an open window showing some of the abacii and fields in the "Reconstructions" relation.

#### III.2.4.8. Minor Synonymies

This relation does not form an interactive part of the database, in the sense that entries here do not impart upon any other part of the database. Instead, it is designed as a reference source of past taxonomic reviews for any given taxon. To this date, however, no data have been entered into this relation (Figure III.17). It is included here because the principle accession point to these data is through a button in the "Main Taxonomy" entry form, which is illustrated in section III.2.3.2.

**ENTER Minor Synonyms**

This provides an opportunity to enter all the synonyms which various workers have assigned to each taxon in the database.

Taxon#:

SynTaxon#:

Genus name  species name

Source the reference for the synonymized taxon name referred to by the reviewer  
 REF#  Page#

Reviewer the reference of the worker(s) reviewing the synonymies  
 REF#  Page#

Comments  Year of review:

**HELP**

**FIGURE III.17.** Screen picture of principal entry form in the "Minor Synonymies" relation.

See Appendix D for details.

### III.2.3. Aims of the Database, Future Versions, and Potential Users

Although the present version of the database includes a considerable amount of lithological, stratigraphic and paleoenvironmental data, the major emphasis has been on the inclusion of vertebrate data covering the last 100 million years of Earth History. In future versions this will be expanded to incorporate the entire history of terrestrial life on Earth, including the floral record,<sup>18</sup> which has long been recognized as one of the most sensitive

18. Some 1000 megafloral references have already been collected for this purpose.

indicators of past global change (Lyell, 1830). The database has been designed with these additions in mind. A further planned development is to link the "Taxa by Locality" and "Taxa by Climate Station" relations, such that the climate space defined by the modern distribution of each taxonomic group ("Taxa by Climate Station") can be directly applied to each fossil occurrence, and hence used automatically to reconstruct the climate of each locality.

While Helix Express provides a powerful platform for the storage and manipulation of such a large dataset, it is intended ultimately to run the database through a Geographic Information System (GIS). Throughout this study and dissertation, I have continually emphasized the importance of visualization and inductive thought on directing how questions of geologic global change should be addressed. A GIS based database will greatly facilitate this.

For the time being the database is only being used by the present author, although versions of the database structure have been adapted for use by other workers. Nonetheless, the database may be useful to other scientists including paleontologists, paleoclimatologists and geologists, whether or not they are interested in vertebrates, *per se*, or wish to use the database structure for their own datasets. In addition the picture capability of Helix Express means that scanned images of the various taxa can be entered if desired. Such information thus gives the database potential as a teaching tool.