

# Chronos Age-Depth Plot: A Java application for stratigraphic data analysis

Geoffrey C. Bohling

Kansas Geological Survey, 1930 Constant Avenue, Lawrence, Kansas 66047, USA

## ABSTRACT

One of the most important steps in the analysis and correlation of stratigraphic data is establishing an age model or line of correlation (LOC) for an individual core hole or stratigraphic section. Chronos Age-Depth Plot (ADP) is a Java re-implementation of an earlier Macintosh program for plotting ages of stratigraphic events versus depth and interactively fitting an LOC to those data. ADP is delivered from the Chronos Web site via Java Web Start and runs locally on the user's machine, allowing access to local data files as well as data from the Neptune database via a geographic-information-system interface. ADP allows the user to develop an LOC using a palette of tools for adding, moving, and deleting LOC control points. In addition, the user is able to zoom in to a more detailed view of different plot regions, toggle the display of different plot groups, and edit plot and axis titles and labels. Output options include saving the LOC control point data, the event data projected to the LOC, and the plot itself in portable-network-graphics (PNG), Joint-Photographic-Experts-Graphic (JPEG), or scalable-vector-graphics (SVG) format. The SVG output file is designed for convenient editing of various plot characteristics, with font and line characteristics for various plot elements and the palette of plotting symbols defined at the top of the file. In addition to being a useful application in its own right, ADP serves as a prototype for other clients interfacing with the Web services provided by the Chronos portal.

**Keywords:** paleontology, geoinformatics, Java, databases.

## INTRODUCTION

The Neptune database, originally developed at the Geologisches Institut of the ETH

Zentrum, Zürich, and now hosted by the Chronos system, contains micropaleontological data derived from cores obtained by the Deep Sea Drilling Project (DSDP) and the Ocean Drilling Program (ODP) (Lazarus, 1994; Spencer-Cervato, 1999). Originally, Neptune contained only taxon occurrence data for Cenozoic sections and the original Age-Depth-Plot (ADP) read corresponding first and last appearance datum information from external text files provided by the user. Neptune now contains information on quality-controlled Cenozoic datums from a selected set of DSDP and ODP holes (Spencer-Cervato, 1999), and Mesozoic datums have recently been added as well. Lazarus (1992) developed the original ADP in True BASIC for the Macintosh as a tool for displaying age-depth data, interactively fitting a line of correlation (LOC), and assigning revised age estimates to the stratigraphic events. Neptune has served as a primary test bed for development during the initial proof-of-concept phase of Chronos. The new Java implementation of ADP is one of the primary products developed during this phase, demonstrating the ability to deliver effective end-user tools through the Web, providing easy access to a Chronos-hosted database, and enabling intuitive data exploration and analysis. The Java environment provides multi-platform compatibility, and ADP has been successfully tested under Windows, Mac OS X, Linux, and Solaris. In addition, Java's high-level Swing components and event-handling facilities greatly ease the development of a graphical, interactive application such as ADP. Java also provides a high level of support for network and database access, vital for a project of this nature.

As the Chronos system continues to integrate additional hosted and federated databases, ADP will be revised to seamlessly access data from these various sources and provide additional functionality appropriate to the add-

ed data types, playing a central role in an expanding Chronos data analysis toolkit.

## LAUNCHING ADP

ADP is launched from a link under the Tools and Services tab of the Chronos portal (<http://portal.chronos.org>) using Java Web Start, a technology that allows Web deployment of stand-alone Java applications (<http://java.sun.com/products/javawebstart/index.jsp>). The first time a user launches the application, Web Start caches the application on the user's hard drive. Thereafter, the cached local copy is used, allowing offline use with local data files. However, whenever the program is launched and a network connection is active, Web Start will automatically check for and download program updates. Thus, Web Start provides a best-of-both-worlds approach. Since the application is running locally, it has the responsiveness necessary for a graphically oriented program such as ADP. In addition, interaction with the local file system occurs in a much more native fashion than is possible from within a browser environment (although the user still needs to grant the application access to the file system on initial launch). At the same time, the automation of the installation and upgrade processes frees the user from the maintenance chores typical of traditional applications. Web Start is included with recent versions of the Java Runtime Environment (JRE). ADP requires JRE Version 1.4.2 or higher. The latest JRE may be downloaded from [www.java.com](http://www.java.com) and a link to this site is included on the ADP Web page.

The ADP page on the Chronos portal provides program operating instructions. Realizing that most users are reluctant to read instructions, we have attempted to keep the operating instructions concise. In addition, each button on the ADP palette is equipped with a tooltip, a brief explanation that appears when the cursor hovers over the button. Most

users should be able to master the program fairly quickly with a cursory glance at the operating instructions and initial guidance from the tool tips.

### GEOGRAPHIC-INFORMATION-SYSTEM INTERFACE TO NEPTUNE DATABASE

ADP provides a zoomable map interface for retrieving data from the ODP and DSDP core holes stored in the Neptune database (Fig. 1). The world relief–bathymetry image is obtained from ESRI’s ArcIMS MapServer via a URL-based request processed through a server at the Kansas Geological Survey. Upon zooming, the request is resubmitted to obtain a map at the appropriate resolution. The ODP and DSDP core hole location data are extracted from the Neptune database using a URL-based query submitted to the Chronos query description format (QDF) service. The Chronos QDF service implements a number of preprepared queries into the Chronos databases, along with the ability to submit free-form Structured Query Language (SQL) queries. The use of hypertext transfer protocol (HTTP)–based transactions, as opposed to Java database connectivity (JDBC) calls directly to the database, provides greater security on the database side, and allows us to distribute a significantly lighter application, without the need for additional database-specific JDBC libraries. In addition, the presence of this QDF service middle layer will ease the process of integrating additional databases as they are federated with the Chronos system.

When the user clicks on the map while the selection tool is activated, all holes within a certain distance of the click are highlighted and summary information for these holes is displayed in the data table at the bottom. This information is derived primarily from Neptune’s hole summary table and includes the hole identification string, latitude and longitude, the number of samples for that hole in Neptune’s sample table, water depth at the hole location, and meters penetrated and recovered. The user selects the particular hole to plot by highlighting the appropriate row in the data table and clicking the Retrieve Selected Data button.

We plan to add further data selection capabilities, including the ability to select a hole simply by typing in its name and adding a more general-purpose query tool for narrowing the search to holes that meet certain criteria.

### DATA EXTRACTION

For the selected core hole, ADP extracts taxon occurrence data by depth from the Neptune sample and sample taxa tables and corresponding first and last occurrence ages from the datum table and uses these to construct an age-depth plot of first and last occurrences. It is important to note that the connection between the number of samples and the number of datums that will actually appear on the plot is somewhat indirect. Each sample can contain any number of observed taxa and all the occurrences of a given taxon are searched to find the lowest and highest occurrences, which are assumed to be associated with the first appearance datum (FAD) and last appearance datum (LAD), respectively. If a corresponding age appears in the Neptune datum table, then a point appears on the age-depth plot. Because the datum table contains age calibration information only for a subset of the taxa in the database (Spencer-Cervato, 1999), as is typical in biostratigraphy (Bolli et al., 1985), it is possible that a number of the taxa observed in the hole will not appear on the plot. ADP also extracts and displays the established age model for the hole, if one is available in the database.

The automated selection of LAD and FAD information from Neptune’s occurrence and datum tables provides a significant improvement over the tedious manual process of selecting FADs and LADs from range charts. However, the program currently assigns the LAD and FAD depths based on the most extreme occurrences, ignoring the possibility of errors due to reworking, downhole contamination, or misidentification. In the future we will add code allowing the user to examine the occurrence data for a selected taxon and interactively revise the FAD and LAD picks. Working with data from local files, rather than from the database, gives the user complete control over the specification of FADs and LADs, as discussed below.

Prior to searching for LADs and FADs, ADP attempts to resolve any taxon names indicated as synonyms in Neptune’s taxonomy table to corresponding valid taxon names (Spencer-Cervato, 1999). However, the problem of resolving synonymies in paleontological data is vast and will require continued programming effort as the Chronos system continues to integrate more databases.

Depth error bars for the FAD and LAD are based on the depths of the next samples outside the observed range of a particular taxon. That is, the actual LAD depth is assumed to occur somewhere between the depth of the

sample with the highest observed occurrence and the depth of the next sample immediately above the highest occurrence (and similarly but oppositely for the FAD). The depth range for the LAD is drawn between the depths of these two samples, and the plotting symbol is drawn at the midpoint. However, it is important to realize that there is really no reason to favor the midpoint over any other depth in the range and that the line between the two depths is not an error bar in the usual statistical sense. Furthermore, the assumption that failure to observe a taxon outside a certain depth range corresponds to actual absence of that taxon outside that range is quite strong and will almost certainly be violated for a number of plotted datums. Both of these factors imply that developing a plausible line of correlation requires a considerable amount of expert judgment that would be difficult if not impossible to incorporate into automated fitting procedures, as pointed out by Lazarus (1992). Most datum ages in Neptune are specified as single values. However, some are specified as ranges and in this case an error bar will also be plotted along the age axis. Note that the Neptune sample table does not contain age ranges for individual samples and the new ADP does not display sample age ranges, a capability that was available in the original ADP.

Alternatively, the input data may be read from local files. The code has been developed to read tab-delimited text files in the same layout as those employed for the original Macintosh version of ADP (Lazarus, 1992). Reading data from local files allows offline use of the program (without an Internet connection) and gives the user direct control over the depths and ages of first and last occurrences, rather than relying on the program’s automated retrieval of this information from the Neptune occurrence and datum tables. The fundamental input file contains the stratigraphic event data, including LAD and FAD depths and ages, with the option of entering either a single value or two values representing a range along either axis. Datum depths may be specified either directly in meters below seafloor (mbsf) or in the DSDP/ODP “core-section-interval” format. Depths in the latter format are translated into mbsf based on information read from an auxiliary core-depth data file which describes the depth range penetrated by each core obtained in the hole. In the absence of a core-depth data file, depths in the core-section-interval format in the stratigraphic event data file will be ignored and only depths in mbsf will be plotted. Finally, data points representing a previously de-

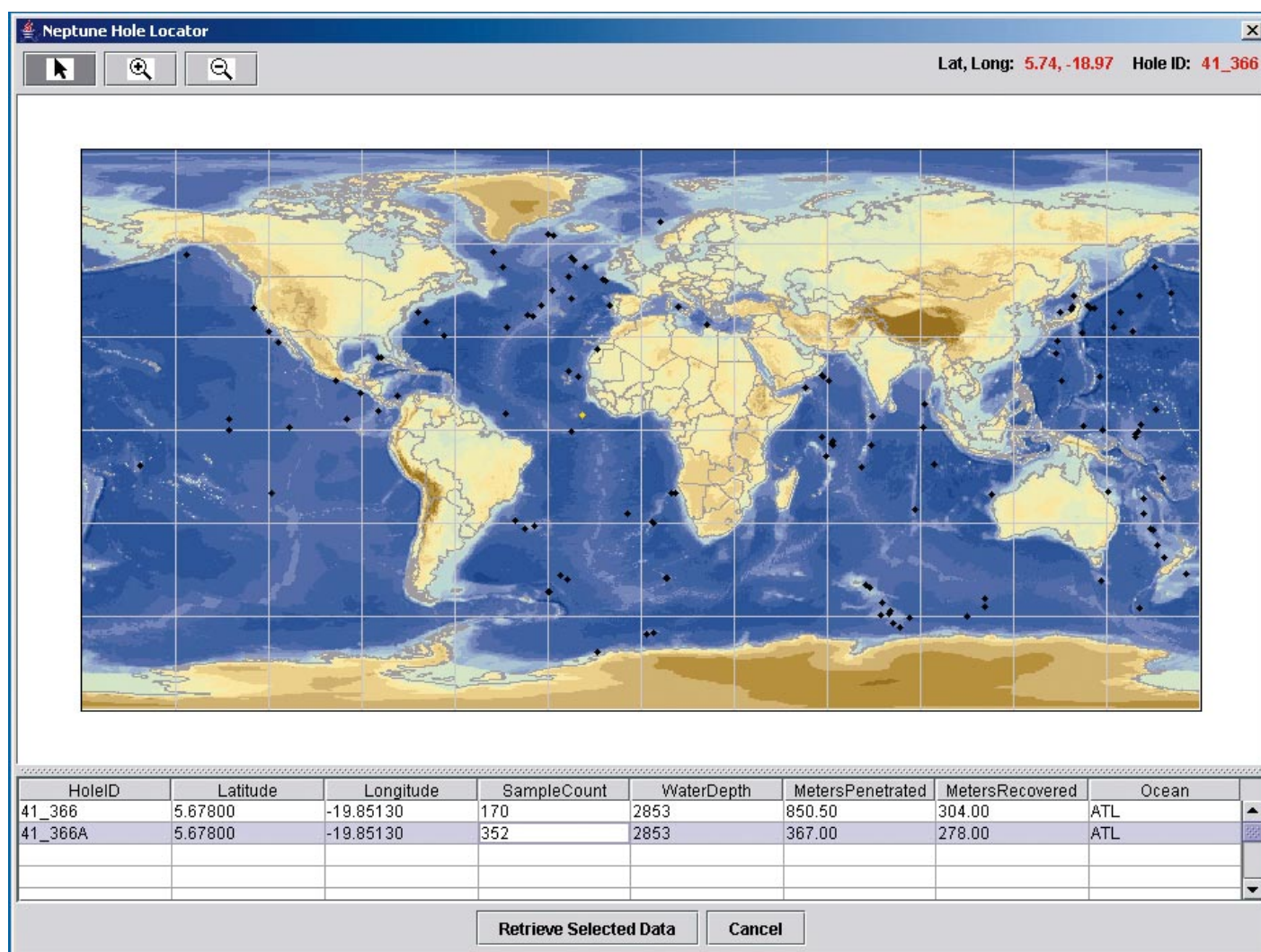


Figure 1. Geographic-information-system interface to Deep Sea Drilling Project and Ocean Drilling Program core hole data stored in the Neptune database.

veloped LOC may also be read from a local file.

The listing below shows the first ten lines of a stratigraphic event data file, with sample depths in core-section-interval format.

In the future, the input options will be expanded, both in terms of connecting seamlessly with other Chronos databases and in terms of handling a variety of local input file formats.

#### THE AGE-DEPTH PLOT

Figure 2 shows the plot of the Neptune data for Hole 41-366A. The Strat File Info line near the top of the application window shows both the hole's identification string and its location in decimal latitude and longitude. The LOC File Info line displays information about the LOC extracted from the database including the revision number, quality, and revision

574 2 DBL Most F datums deleted to keep size <200 pts.

Grp	Event	plotcode	Y.Age	OldAge	TopDepth	BotDepth
D	top R.praebergonii	tRp	1.93		1-CC	2-CC
D	top A.ellepticus	tAel	3.35		''3-2,75''	''3-4,75''
D	base T.convexa	bToc	3.83		''3-4,75''	3-CC
D	base A.elegans	bAe	4.2		3-CC	''4-4,75''
D	top N.cylindrica	tNc	4.69		''4-4,75''	4-CC
D	base T.oestrupii	bTo	5.61		4-CC	''5-2,75''
D	top A.acutiloba	tAa	5.89		''5-5,75''	5-CC
D	top N.miocenica	tNm	6.15		''5-5,75''	5-CC

date for the LOC. The storage and extraction of this LOC metadata in the Neptune database, not available through the original ADP, is crucial to documenting the development of each age model. Different plotting symbols are used for the different fossil groups represented in the database, as shown in the legend (F—foraminifera; N—nannoplankton; R—radiolarians; and D—diatoms). These four group codes (F, N, R, and D) are assigned a fixed set of plotting symbols, for consistency between plots, and any additional plot codes are dynamically assigned one of five other symbols based on the order they are encountered in the input data. The display of each plotting group can be toggled on and off by clicking on the box to the left of the corresponding legend entry, allowing the user to easily fit the LOC to a more trusted fossil group, for example,

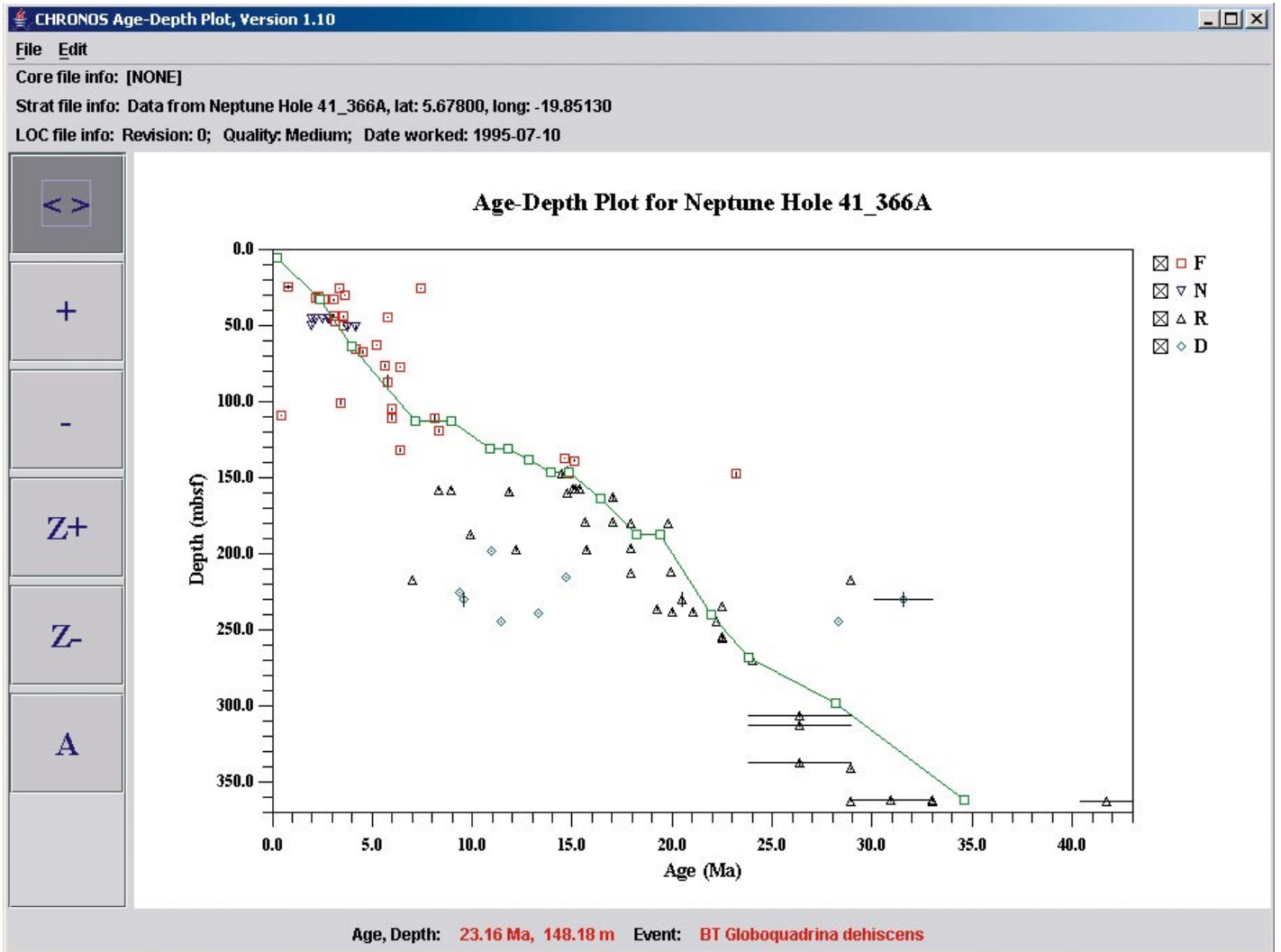


Figure 2. Age-depth plot developed from Neptune sample occurrence data for Hole 41-366A. F—foraminifera; N—nannoplankton; R—radiolarians; and D—diatoms; mbsf—meters below seafloor.

and redisplay other groups later for comparison.

The Age, Depth and Event labels at the bottom of the application window continuously update as the cursor moves over the plot region, showing the user the age-depth coordinates of the current cursor location and the name of the nearest datum within five pixels of the cursor. The screen shot in Figure 2 was taken when the cursor was hovering over the datum BT [bottom or base] *Globoquadrina dehiscens*, at ~23 Ma and 148 mbsf. These labels remain active while the user is moving or adding LOC control points (next section), providing crucial feedback for the line-fitting process. In addition, short plotting labels, read from the Neptune datum table or from the stratigraphic event data file, may be displayed by selecting Show Event Labels from the Edit

menu. The plotting labels tend to be of limited utility when the data are dense, but zooming in on a region of interest will reduce the amount of overplotting and make the labels more legible.

Zooming is accomplished by clicking the Z+ button on the palette along the left side of the application window and either dragging out a rectangle over the region of interest or clicking on a single point. In the latter case both axes will be zoomed by a factor of two, centered on the selected point. Clicking on the zoom-out button, Z-, will zoom the plot back out to the default axis limits. The user may also set the axis limits and tick intervals by clicking on the axis parameters button, A, and typing numbers in the resulting dialog box. Items on the Edit menu allow the user to edit the main plot title and axis titles.

The continuous feedback on cursor location and identities of nearby datums and the ease of zooming are both significant improvements over the original ADP, which required additional steps to accomplish these tasks.

#### DEVELOPING THE LINE OF CORRELATION

The solid green line on the plot (Fig. 2) represents the LOC, which consists of a sequence of straight line segments between control points, represented by the open green squares. The top three buttons on the palette are used to modify the LOC by moving (<>), adding (+), and deleting (-) LOC control points. When the <> button is activated, the user may move a control point by clicking on it, dragging it to a new location, and then re-

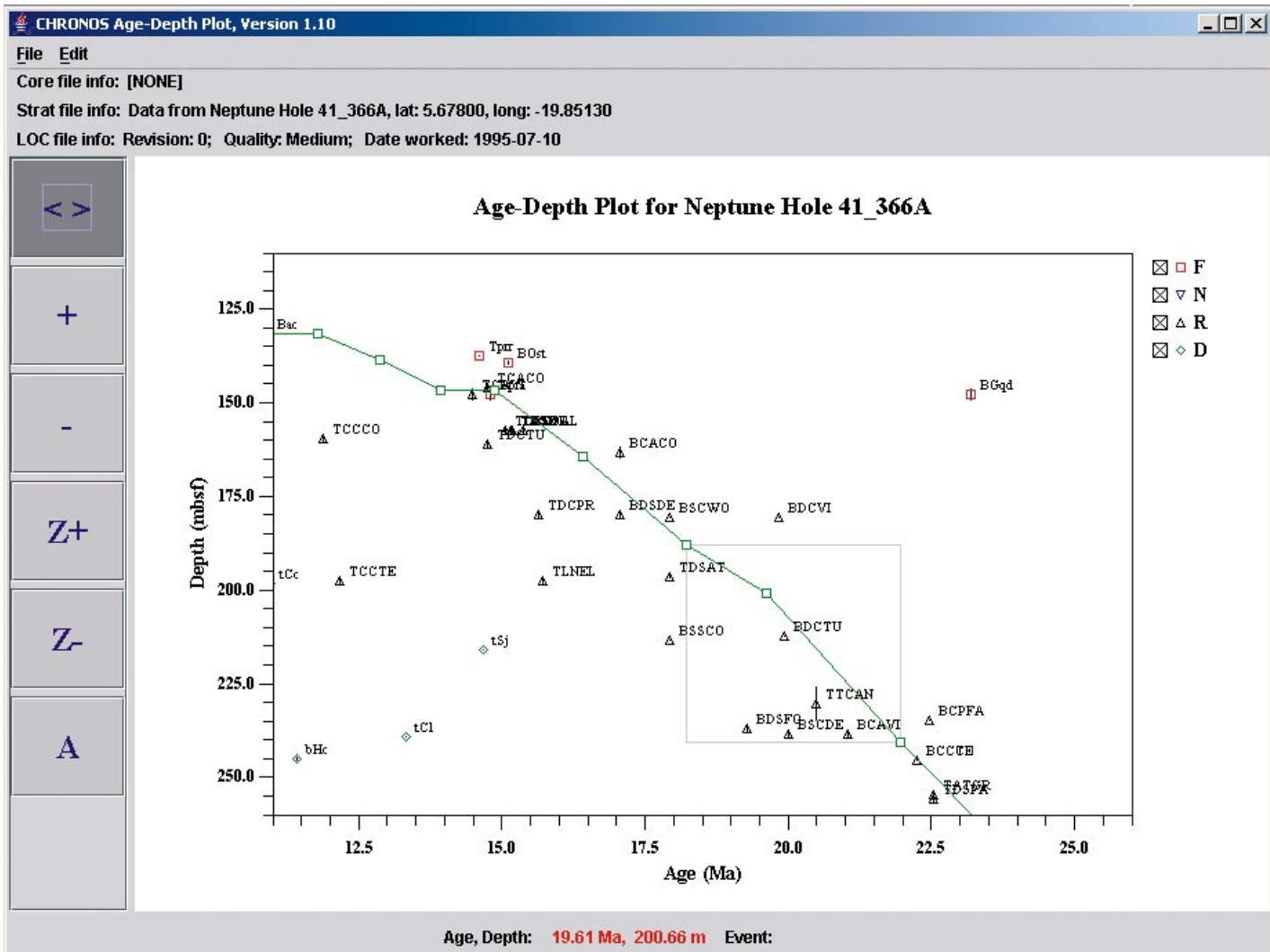


Figure 3. Age-depth plot shown while user is moving a line of correlation (LOC) control point. Age-Depth Plot enforces a model representing nondecreasing age with increasing depth by limiting movement to the region bounded by the gray rectangle. Short labels for events are also displayed. See Figure 2 for definitions.

leasing the mouse button. While the move is occurring, ADP displays a gray rectangle representing the ages and depths of the adjacent control points and does not allow the user to move the control point outside this box (Fig. 3). Similarly, ADP does not allow new LOC control points to be added outside the bounding rectangles defined by subsequent pairs of existing points. Thus, ADP enforces the development of a proper LOC (nondecreasing age with increasing depth) by limiting the regions in which points may be added or moved. A hiatus is added to the LOC by moving a control point to the upper or lower limit of the bounding box for its move. Users have found the LOC updating process to be quite simple and intuitive, with the Age, Depth and Event labels at the bottom of the application provid-

ing valuable feedback for the placement of control points.

If no LOC is available for a given hole in the Neptune database or if the stratigraphic event data have been read from a local file, then the initial LOC displayed is a single straight line connecting the extreme age and depth values in the data. Starting from this point the user may add and move control points to develop a LOC or read an existing LOC from a local file.

#### DATA OUTPUT

ADP's data output options, available through the File menu, include saving the stratigraphic data projected to the LOC to a local file, saving the LOC itself to a local file,

and submitting the LOC to Neptune. Under the first option, the stratigraphic events are projected to the LOC by replacing the observed ages with the age predicted at that depth by the LOC. Minimum and maximum depths for an event will both be projected, producing a corresponding age range, even if the event had only a single age value to begin with. The depths and projected ages are saved to a local file in the same format as a normal stratigraphic event data file, meaning the projected data can be read back and displayed on the plot.

The option to submit the LOC to Neptune is available only if the stratigraphic data have been read from Neptune, so that the submitted LOC can be associated with an existing hole ID (and control points can be optionally as-

Age	Depth	KeyDatum	Comment
0.275	5.465	[None]	
2.403	33.006	[None]	
3.965	64.244	[None]	
7.17	113.163	[None]	
8.932	113.163	[None]	
10.854	131.434	[None]	
11.776	131.434	[None]	
12.857	138.507	[None]	
13.939	146.758	[None]	
14.86	146.758	[None]	
16.422	164.44	[None]	
18.224	188.016	[None]	
19.426	188.016	[None]	
21.949	240.471	[None]	
24.018	270.442	[24.03 Ma, 270.08 mbsf]: B Lychnocanoma elongata	A somewhat arbitrary choice
28.178	298.864	[None]	
34.606	362.273	[None]	

**Figure 4.** Line of correlation (LOC) control point data table. User has associated the control point (at 24 Ma, 270 m below seafloor [mbsf]) with a datum at that location and has added a comment on the selection.

sociated with existing datum IDs). The user may associate any LOC control point with a particular datum by displaying the LOC data window (through an item on the Edit menu) and using the dropdown list in the KeyDatum field to select the associated datum (Fig. 4). A comment for each control point may be entered in the Comment field. When the LOC is submitted to Neptune, the control age and depth values along with any key datum IDs and comments are submitted to a user-age-model table in Neptune. In addition, the user is prompted to enter his or her name and an overall comment on the age model and this information is stored in a table of age model metadata. Work is currently under way on a Chronos age model database, not solely tied to the Neptune database, and age models submitted through ADP will eventually go to this more general-purpose database. Development of a protocol for submitting, attributing, and sharing age models derived from multiple data sources will represent a major step in creating a forum for cooperative investigation and refinement of the geologic time scale, a central thrust of Chronos.

## GRAPHICAL OUTPUT

The plot itself may be saved to a local file using the Save Plot option on the file menu. The available output formats are portable network graphics (PNG), Joint Photographic Experts Graphic (JPEG), or scalable vector

graphics (SVG). The first two options employ built-in Java functionality to create an automatic screen dump of the plot region and save the resulting raster image in the selected format. SVG is an Extensible Markup Language (XML) specification for describing the placement of various graphical elements (lines, shapes, text, etc.), recommended by the World Wide Web Consortium (Eisenberg, 2002). That is, SVG is a graphics description language, much like PostScript, formulated as an XML extension. SVG files can be viewed in a Web browser using a free plug-in (<http://www.adobe.com/svg/viewer/install/main.html>) or read and edited by recent versions of certain drafting programs, such as Adobe Illustrator CS. However, it should be noted that different SVG interpreters seem to have their own quirks. A considerable amount of effort was spent in working around certain oddities in Illustrator's interpretation of the output SVG files, perhaps to the detriment of results under other interpreters.

The SVG file is a plain text file and the characteristics of a number of plot elements (plotting symbols for different groups, line style for the LOC, fonts for the main title and axis title, etc.) are specified in a "defs" block at the top of the file. These component styles and plotting symbols can be easily changed by opening the SVG file in a text editor and modifying the appropriate styles and symbol definitions. The meanings of most style specifi-

cations as well as the commands for drawing the symbols are self-evident.

## CONCLUDING REMARKS

Scientists at the Mesozoic and Cenozoic Paleooceanography and Paleogeography Workshop in October 2004 ([http://www.chronos.org/pdfs/MzCz\\_wkshp.pdf](http://www.chronos.org/pdfs/MzCz_wkshp.pdf)) subjected ADP and other aspects of Chronos to extensive usability testing. Along with a general endorsement of ADP's utility and ease of use—once they cleared the initial hurdle of figuring out the fundamentals of program operation—the workshop participants provided a list of suggested improvements. Work on incorporating these suggestions, along with other improvements, is under way. One of the most significant revisions will be to incorporate the paleomagnetic data that have recently been added to the Neptune database. The original ADP plotted observed paleomagnetic patterns for a hole along the depth axis and established magnetic polarity records along the age axis, greatly aiding in the development of the LOC. A similar capability will be added to the Chronos ADP tool. Additional planned revisions include allowing more flexible management of plotting symbols, adding additional capabilities for fitting the LOC (such as snapping a LOC segment through a selected pair of points), and adding code for managing the selection of different geography-specific datum scales, among others. We anticipate that

Chronos tool development will continue to involve a positive feedback loop, with an expanding array of accessible tools triggering more user interest and suggestions for further improvement.

#### ACKNOWLEDGMENTS

This work was supported by a Kansas Geological Survey subcontract to the National Science Foundation grant funding the Chronos project, NSF-EAR 0315216. I would like to thank Dave Lazarus and

Brian Huber for their helpful reviews of this manuscript.

#### REFERENCES CITED

- Bolli, H., Saunders, J., and Perch-Nielsen, K., eds., 1985, *Plankton Stratigraphy*: Cambridge, UK, Cambridge University Press, 1040 p.
- Eisenberg, J., 2002, *SVG essentials*: Sebastopol, California, O'Reilly & Associates, Inc., 335 p.
- Lazarus, D., 1992, Age depth plot and age maker: Age modeling of stratigraphic sections on the Macintosh series of computers: *Geobyte*, v. 7, no. 1, p. 7–13.

Lazarus, D., 1994, The Neptune Project—Developing a large relational database of marine microfossil data on a personal computer: *Mathematical Geology*, v. 26, p. 817–832.

Spencer-Cervato, C., 1999, The Cenozoic deep sea microfossil record: Explorations of the DSDP/ODP sample set using the Neptune database: *Palaeontologia Electronica*, v. 2, no. 2, <http://palaeo-electronica.org/1999.2/neptune/issue2.99.htm> (June 2005).

MANUSCRIPT RECEIVED BY THE SOCIETY 7 FEBRUARY 2005

REVISED MANUSCRIPT RECEIVED 9 MAY 2005

MANUSCRIPT ACCEPTED 20 MAY 2005

Printed in the USA