EPIC: AN OCEANOGRAPHIC DATA ARCHIVAL AND RETRIEVAL SYSTEM

N.N. Soreide
S.P. Hayes

NOAA/Pacific Marine Environmental Laboratory
7600 Sand Point Way N.E.
Seattle, WA 98115-0070

ABSTRACT

EPIC was developed to manage the large quantity of oceanographic data collected by PMEL oceanographers during climate study programs such as EPOCS and TOGA. It presently provides access to over 2000 data sets from CTD, XBT, TOPS profiling current meter systems and from other oceanographic instruments. Data sets are selected interactively by specifying geographic location, time and depth. Station locations of selected data can be displayed on a map. A user can generate data plots, contour plots and listings from the selected data sets with interactive and "user friendly" programs. On-line help describes EPIC programs and data formats. EPIC is written in FORTRAN 77 on a VAX 11/785 running under VMS.

I. INTRODUCTION

EPIC (Equatorial Pacific Information Collection) was designed to manage the large quantities of oceanographic data collected by oceanographers in the Ocean Climate Research Division of the Pacific Marine Environmental Laboratory (PMEL) during climate study programs such as EPOCS (Equatorial Pacific Ocean Climate Studies) and TOGA (Tropical Ocean Global Atmosphere). Between 1979 and 1983, more than 1000 data sets were obtained on 24 EPOCS cruises with the PMEL CTD and TOPS profiling current meter systems. More data sets are anticipated from the continuing EPOCS research and the onset of the TOGA program. The objective of EPIC was to provide easy access to these and other data sets on the PMEL VAX 11/785. EPIC consists of the data itself, in a standardized format, and a suite of programs which provide for selection and display of the data. It presently provides access to over 2000 CTD, XBT and TOPS data sets.

II. EPIC SOFTWARE

The software is a suite of interactive programs and command procedures written in FORTRAN 77 and VMS on a VAX 11/785 computer running under VMS. Use of these programs and the associated data formats is documented in an on-line help (HLP) program. The overall structure of EPIC is shown schematically in figure 1. The first program, EPIC, serves as an index to the data and is independent of data format or instrument type. Data sets are selected by specifying instrument type, geographic location, time, and depth in answer to questions. The format of responses to program questions is flexible. For example, a latitude of 1° 30'N can be given as 1 30 N or 1.5 N or 1.5. The program searches a relational data base containing descriptive information about available data to find those which meet the criteria. A list of these data sets is then provided in a file (called the EPIC file) which contains geographic location, time, depth range, cruise identification, and VAX file specifications for each data set. Data files can be copied to the users private disk space and station locations of selected data can be displayed on a map. For cruises where the TOPS profiling current meter was used simultaneously with the CTD, this association can be included in the EPIC file. These options are selected interactively by responding to questions.

The data index is based on RIM, a relational data base obtained for the PMEL VAX from Boeing Computer Services[1]. The information contained in this data index is detailed in Table 1. Each index entry in the data base tables contains more information than the standard selection criteria and file specifications which are used by the EPIC interactive program. This extended information can be obtained and sorted by directly using RIM's interactive command mode. For example, one can select all CTD data sets taken during the months of July and August, regardless of year or all the TOPS drops made from the NOAA ship OCEANOGRAPHER.

The data base tables are loaded by entering descriptive information for each data set into a FORTRAN program which utilizes RIM's program interface. Alternatively, data can be directly entered using RIM's interactive command mode. The program is the usual entry method and is less expensive. Correspondence between TOPS profiling current meter and CTD drops is entered into the data base tables using RIM's relational algebra capabilities.

U.S. Government Work. Not protected by U.S. copyright.

The major display programs in EPIC are STDLST, CONTOUR, and PLOTSETUP. These programs access the data sets indicated by the EPIC file. All are interactive and require no programming skills to operate. Program STDLST provides either a full or a summary listing of the data. The listing can be to the terminal or to a file.

CONTOUR produces contour plots of hydrographic parameters or velocity. The user names the EPIC file to be used, selects the variables to be contoured, specifies axis parameters (or uses defaults) and can either view the plots interactively or route them to a remote plotter. Once variables and axis parameters are selected, data sets indicated by other EPIC files can be contoured with the same specifications simply by naming the EPIC files. CONTOUR is written in PLOT5[2], which is an interactive and command driven plotting package. CONTOUR generates a PLOT5 command file, which the user can later edit with the VAX editor to change plotting parameters such as contour labels, intervals between contours and axis parameters. The user can also enter PLOT5, display a plot from this command file, enter interactive commands to change the plot appearance, and then redisplay the plot.

PLOTSETUP produces plots of selected variables from data sets indicated by an EPIC file. It has options for plots with up to 4 multiple axes (*e.g.*, Temperature, Salinity, Sigma-T and Oxygen vs. Pressure), and for over-plotting of several data sets with or without offsetting successive data sets on the plot. There is an option to fit selected data to a polynomial of order 1 through 9, and display the fit with the data on a plot. PLOTSETUP generates a set-up file containing plotting parameters. With this set-up file, a plot set-up can be repeated with data indicated by another EPIC file. The set-up file can also be edited, using program PLOTSETUP, to modify plotting parameters. Plots can be displayed interactively and routed to a plotter.

III. EPIC DATA FORMAT

The data format (Table 2) was designed for CTD, XBT or TOPS profiling current meter data. However, the format is flexible and can accommodate data from other instruments for which the dependent variable is not pressure. Data files are unformatted, have variable length records and are sequential. These characteristics conserve space on the disk, reduce read/write time, and simplify programming. Each data set (a CTD Cast or TOPS drop) is stored on an individual disk file consisting of a header of eight 80-character lines followed by the data scans.

The header (Table 2) contains data identification such as cruise and cast number, latitude, longitude, time, date, and other descriptive information. It includes the file creation date and a list of variable codes which define each variable in the data scan. There is ample room for comments. There is also a flag which identifies the header as a standard CTD header and another flag which indicates whether a second header (of the same length but probably a different format) follows this header. Subsequent headers can have any format except that these two flags must be present to indicate the header type and the presence or absence of subsequent headers.

The header format was designed so that many types of data can readily be stored in this format. In addition, the documentation of data processing and analysis by various programs can easily be incorporated with the processed data; the data files are self documenting.

The data scans are written sequentially. The number of scans in the file, the number of variables in a scan and a list of variable codes identifying each variable are all part of the header.

Data files are stored on a removable disk. The VAX computer has a hierarchical file structure which is well suited to organizing oceanographic data by cruise or experiment. All data resides in a directory [DATA] and its subdirectories. There is a subdirectory for each cruise, and each cruise directory has a subdirectory for each data type.

Each file has a name which uniquely identifies the data without a directory specification. Therefore, the file retains its identity by name even when it is copied to another directory. The file name consists of a 5-character cruise identifier, a 1-character data type identifier and a 3-digit cast number. The extension indicates data type. A typical file name complete with directory specification is DISK$HAYES: [DATA.EP38.CTD] EP381C001.CTD;1. This data was taken on EPOCS Cruise EP3-81 with the PMEL CTD.

IV. SUMMARY

EPIC is a method of storing, retrieving and displaying oceanographic data which requires no programming knowledge to use. It was designed for pressure-indexed data from CTD, XBT and TOPS profiling current meter system; however, planned future uses include time series data from instruments such as pressure gauges, current meters and moored thermistor chains. It is written in FORTRAN 77 on a VAX 11/785 running under VMS and uses the RIM data base to index files. The EPIC system provides a user friendly environment for rapid access and display of a variety of ocean data and is proving to be a useful tool in oceanographic research.

REFERENCES

[1] Boeing Computer Services Company, 1983: BCS RIM Relational Management System General Information Manual, The Boeing Company, Seattle, WA.

[2] Donald W. Denbo, JISAO/School of Oceanography, University of Washington, Seattle, WA, 98195, Private Communication.

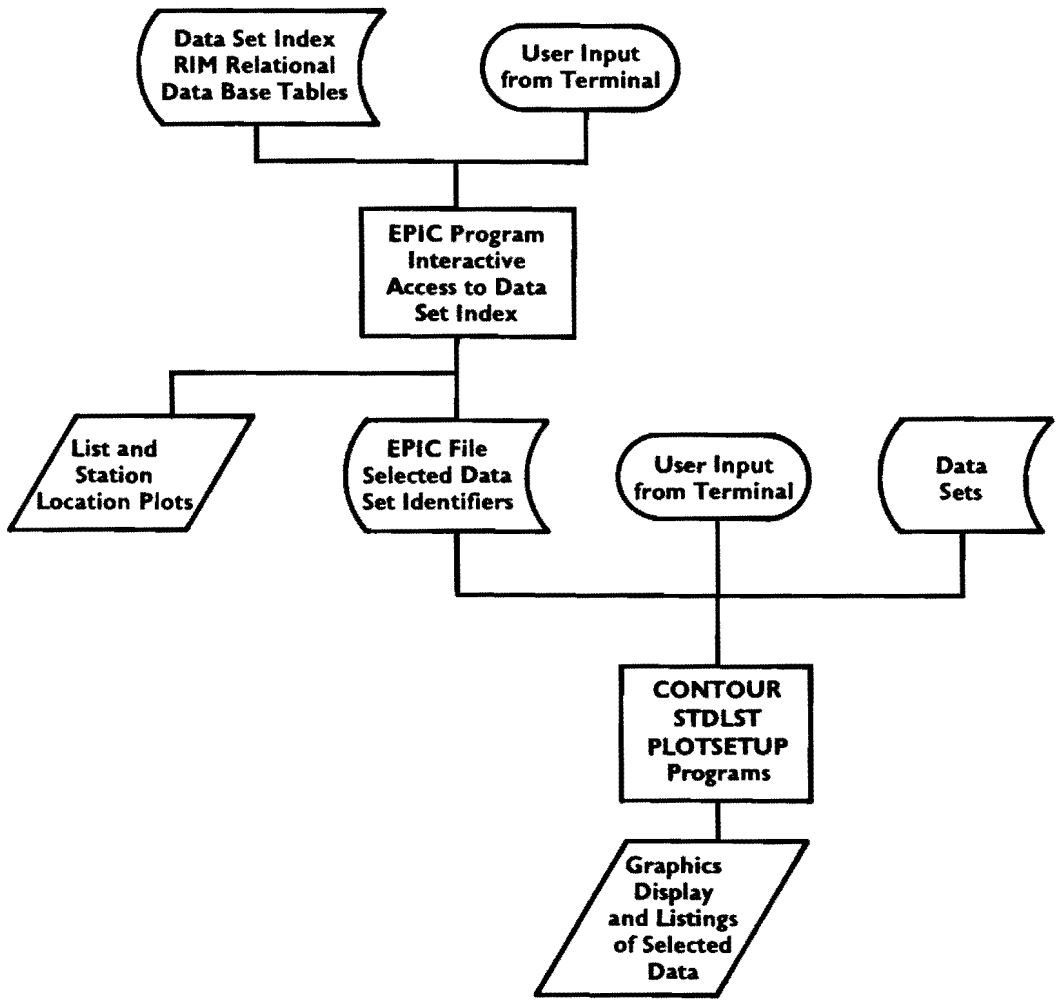

Figure 1. Flow chart of interactive access to data index and data sets. On-line help describes all programs and data formats.

TABLE 1. DESCRIPTIVE INFORMATION IN EPIC DATA BASE TABLES.

NAME	TYPE	DESCRIPTION
CRUISE	TEXT	Cruise ID, *e.g.*, EP3-81-OC
CAST	INTEGER	Cast or Drop Number
ZI	INTEGER	Initial depth (db)
ZF	INTEGER	Final depth (db)
DZ	INTEGER	Incremental depth (db)
LATI	REAL	Latitude North = positive, S = negative (degrees)
LONGI	REAL	Longitude West = positive, East = 360-longitude (degree)
TIMI	REAL	Time in hours since start of century
DATEXT	TEXT	Data type, *e.g.*, CTD
INSTYP	TEXT	Instrument type
DIRECT	TEXT	VAX directory name, *e.g.*, [DATA.EP381.CTD]
FILENAME	TEXT	VAX file name (without directory), *e.g.*, EP381C001.CTD;
DATORIG	TEXT	Data origin
TEMPFLG	TEXT	Flag to indicate if there is good temperature data
SALFLG	TEXT	Flag to indicate if there is good salinity data
UVFLG	TEXT	Flag to indicate if there is good velocity data
FILCREAT	TEXT	File creation date
NVAR	INTEGER	Number of variables in each scan
VAR1	INTEGER	Code for variable 1
VAR2	INTEGER	Code for variable 2
VAR3	INTEGER	Code for variable 3
VAR4	INTEGER	Code for variable 4
VAR5	INTEGER	Code for variable 5
VAR6	INTEGER	Code for variable 6
VAR7	INTEGER	Code for variable 7
VAR8	INTEGER	Code for variable 8
VAR9	INTEGER	Code for variable 9
VAR10	INTEGER	Code for variable 10
VAR11	INTEGER	Code for variable 11
VAR12	INTEGER	Code for variable 12
VAR13	INTEGER	Code for variable 13
VAR14	INTEGER	Code for variable 14
VAR15	INTEGER	Code for variable 15
VAR16	INTEGER	Code for variable 16
VAR17	INTEGER	Code for variable 17
VAR18	INTEGER	Code for variable 18
VAR19	INTEGER	Code for variable 19
VAR20	INTEGER	Code for variable 20
LATITUDE	TEXT	Latitude, *e.g.*, 10 30.1 S
LONGITUD	TEXT	Longitude, *e.g.*, 110 01.3 W
DATE	TEXT	Date, *e.g.*, 0000 14 JUN 81

TABLE 2. EPIC DATA FORMAT

Line 1	Cruise id, cast number, date, instrument type, header-type flag
Line 2	Latitude, longitude, meteorological information, file creation date
Line 3	Number of data scans, initial, final, incremental depths, number of variables in each data scan, meteorological information
Line 4	Data origin information, meteorological information
Line 5	Variable codes for each variable in a data scan
Line 6	Ending latitude, longitude, time
Line 7	Comments
Line 8	Comments
Line 9 - Line 9+n	n Sequential data scans