VIII. COMPUTER DATA BASE AND PROCESSING

Frank J. Kern

A. Introduction

The computer-generated plots and tables in the appendices are derived from two major data files which were compiled as one of the tasks completed by the project teams. One major file contains all stage-discharge data for all eighteen stations covered in the study and the other file contains dike, revetment and levee data. Since the organization of, and methods used, in processing these major files are quite different, they will be discussed in separate sections.

B. Stage-Discharge Data File

1. Organization

The stage-discharge data file contains numeric stage and discharge data for each of eighteen stations. Data for each station are arranged sequentially by month and year, each year block being separated from the next year by a header card. All stations are arranged in a partitioned, sequential data set with the first eight characters of the station name being used as an identifier. The station names contained in this file are:

> St. Louis, Vicksburg, Red River, Memphis, Keokuk, Arkansas City, Meredosia, Paducah (Flow data only), Little Rock, Simmesport, Alexandria, Helena, Hermann, Metropolis, Thebes, Clarendon, Chester, and Alton.

The data contained in these files are FORTRAN compatible. A complete description of the card format is contained in Appendix 3.1.

2. Debugging

The data cards were punched from written lists. After each station file was completely punched, a listing of the card images was checked by hand against the lists used to punch the data. After corrections were made, the data file was loaded onto disk and a series of debugging programs were run. The first program (called MXMN in Appendix 3.1) was used to check for obvious punch errors, misplaced cards, and order-of-magnitude data errors which result from having entire cards shifted left or right one or more columns. A second program (called BUGPACK in Appendix 3.1) was used to determine if the data were "reasonable", i.e., if successive data points are likely to be on the stage-discharge curves. "Questionable" data points indicated by this program were checked by hand thus enabling some correction of errors which occurred in the original data lists.

3. Processing

All debugging and processing programs for this file are written in FORTRAN IV (G) (IBM 360/50). A complete, documented, listing of each is contained in Appendix 3.1 and will not be discussed in further detail here. All machine-generated plots contained in other appendices were made in a standard 7.5" x 15" format on a Calcomp 565 Plotter using a felt tip pen and then photo reduced for report purposes. With slight modifications, the plotting programs also are fully FORTRAN compatible and can be used in other Calcomp plotter installations.

C. Dike, Levee, and Revetment Files

1. Organization

The dike, levee, and revetment files contain alphanumeric code which is a distillation of information collected on data forms by the project team. Data cards containing this information were punched from coding forms which were in turn prepared by copying data from questionnaire forms which were filled out in the various districts. Since the type of information obtained on dikes, levees, and revetments is slightly different, the main file is organized into dike, levee, and revetment subfiles. Each subfile is further subdivided into blocks of approximately 1000 card images, each block containing data pertaining to a particular district and (variable) span of river miles. The card images in each block are loaded in random order with the information on each card keyed to an identifier number which may be used for sorting purposes.

2. Debugging

After punching, each file block was checked by hand against the coding form using a listing of the card images. No further preliminary debugging was attempted.

3. Processing

Since the generation of the reports pertaining to those data involves primarily character-string manipulation, all programs used in processing have been written in PL-I (IBM 360/50). This language was chosen because of the flexibility it offers in character-string manipulation and also because of local availability. A complete, documented listing of these programs is

contained in Appendix 3.1 along with a detailed description of the card format for each file. It was necessary to develop a standard format for displaying all dikes, all levees, and all revetment data in report form which included all information which could, but did not necessarily, appear in a given file. Therefore, blank entries in the report format are usually interpreted to mean that data are not available from current files.

D. Obtaining Copies of Data Files

The two major files consist of approximately 100 boxes of cards. Copies of these cards, or card images on user supplied tape, are available from University of Missouri-Rolla. There will be some duplication fee, depending on the user's requirements.