PROGRAMS FOR GENERATING DATA TABLES FOR THE ANNUAL WATER-RESOURCES DATA REPORT OF THE U.S. GEOLOGICAL SURVEY

By Robert R. Mason and Catherine L. Hill

U.S. GEOLOGICAL SURVEY

Open-File Report 88-173

Raleigh, North Carolina

1988

DEPARTMENT OF THE INTERIOR
DONALD PAUL HODEL, Secretary
U.S. GEOLOGICAL SURVEY
Dallas L. Peck, Director

For additional information
write to:

District Chief
U.S. Geological Survey, WRD
Post Office Box 2857
Raleigh, North Carolina 27602

Copies of this report can be
purchased from:

U.S. Geological Survey
Books and Open-File Reports
Federal Center, Box 25425
Denver, Colorado 80225

CONTENTS

ILLUSTRATIONS

PROGRAMS FOR GENERATING DATA TABLES FOR THE ANNUAL WATER-RESOURCES DATA REPORT OF THE U.S. GEOLOGICAL SURVEY

By Robert R. Mason and Catherine L. Hill

ABSTRACT

The U.S. Geological Survey has developed software that interfaces with the Automated Data Processing System to facilitate and expedite preparation of the annual water-resources data report. This software incorporates a feature that prepares daily values tables and appends them to previously edited files containing station manuscripts. Other features collate the merged files with miscellaneous sections of the report. The report is then printed as page-size, camera-ready copy. All system components reside on a a minicomputer; this provides easy access and use by remote field offices.

Automation of the annual report preparation process results in significant savings of labor and cost. Use of the system for producing the 1986 annual report in the North Carolina District realized a labor savings of over two man-months. A fully implemented system would produce a greater savings and speed release of the report to users.

INTRODUCTION

Historically, annual State data-report preparation in offices of the U.S. Geological Survey was a labor-intensive task. The manuscript headings were updated by word processor and daily discharge data were retrieved from the Geological Survey Headquarters Office in Reston, Virginia. Splicing the manuscript headings and data tables together required time, patience, and a steady hand. The installation of minicomputers in most Geological Survey offices presents an opportunity to improve annual report preparation. Having the minicomputers provides improved communications, a central file system where data can be stored and accessed, and the capability to share responsibilities between offices. Use of an automated annual report preparation system can significantly reduce the number of days required to produce the report. A labor savings of over 2 man-months was realized during preparation of the 1986 annual report. The purpose of this report is

to explain how the North Carolina District of the Geological Survey uses the minicomputer system to facilitate and expedite annual report preparation.

In North Carolina, hydrographers are responsible for processing routine hydrologic data and preparing it for publication. This work includes manuscript editing and preparation of the daily values tables for the annual report. Following supervisor review of the data, the remainder of the annual report process is mostly automated. Programs merge each manuscript with the appropriate daily values table, collate sequential station records, insert miscellaneous report sections, such as the introduction, and, finally, page number and print the camera-ready copy of the entire report. Programs are not yet available for indexing or cross referencing various portions of the report. These are currently being performed manually.

THE AUTOMATED ANNUAL DATA-REPORT PROCESS

The automated annual report system was designed and implemented by personnel in the North Carolina District's Hydrologic Records Section, with assistance from others. The system uses a Prime[1/] minicomputer to prepare the report. It incorporates a highly structured file and directory setup in conjunction with simple modular programs that are compatible with existing software, such as EMACS and WordMarc Composer editors and the Geological Survey's new Automated Data Processing System (ADAPS) and Water-Quality Data System (QWDATA).

The system is unique in that it is designed to be used primarily by the hydrographers who collect, analyze, and prepare the data for the annual report rather than clerical staff who traditionally perform word processing. This change is possible because most of the word processing and document assembly is performed automatically within the system. Unlike most other annual report systems, the North Carolina automated system eliminates manual splicing or electronic transfer of data from one computer to another. These improvements are possible because all report components reside on the District's minicomputer where the entire report is assembled. This system eliminates having to use two computers and is, therefore, easier to access

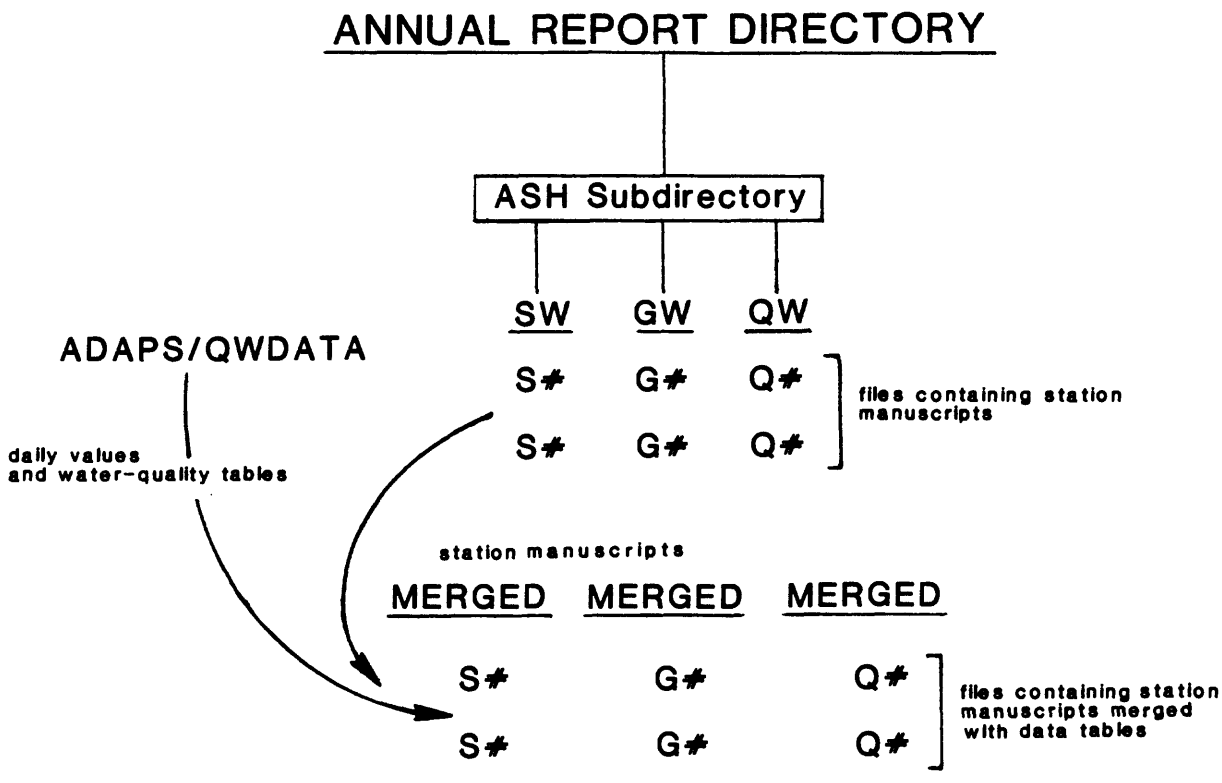
---

[1/] Use of brand/firm/trade names in this report is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.

and use. Similar annual report systems could be installed on other minicomputers.

Annual report activities presently are performed in one master directory. It is a user file directory on a separate disk partition from ADAPS. Overall configuration of the directory is shown in figure 1. The annual report directory contains a series of subdirectories, one for each field office and major project. Figure 1 shows a subdirectory labeled "ASH," the abbreviated name of the Asheville, North Carolina, field office. The subdirectory contains three smaller subdirectories for the three disciplines: SW for surface-water stations, GW for ground-water stations, and QW for water-quality stations. Within each discipline directory are files containing the previous year's station manuscript for each of the stations serviced by the field office. These files contain station data, identified by an appropriate prefix (S for surface water, G for ground water, and Q for water quality), and the station's identification number. A number of other files and directories are also present including files for the annual report introduction, the summary of the water-resources conditions for the year, including graphical displays, and the report index.

Figures 2 and 3 are typical pages from an annual report depicting surface-water and water-quality data, respectively. The textual information occupying the upper half of the page is referred to as the "station manuscript." The lower half of the page is the data table. Each of the manuscripts from the previous year for stations that remain active is retrieved from the appropriate discipline subdirectory and updated manually by hydrographers using the EMACS or WordMarc Composer editor. The corresponding station data table is automatically generated by the annual report system by using ADAPS or QWDATA. Editing of the data table is automatically performed to remove miscellaneous header information, such as "latitude," "longitude," and "Department of the Interior," to transform the data table into publication format. After editing, the data table is electronically appended to a copy of the updated station manuscript, and the merged product is filed in the appropriate discipline directory in a subdirectory called "MERGED" (fig. 1). When all the merging is complete, the MERGED directory will house files for each station containing the updated manuscript and data table merged together for the current year. These files are electronic versions of the



Explanation

ASH - Asheville, North Carolina, field office subdirectory
ADAPS - Automated Data Processing System
QWDATA - Water Quality Data System

SW Surface water subdirectory
GW Ground water subdirectory
QW - Water quality subdirectory
Station number
S# - surface water station file
G# - ground water station file
Q# - waterquality station file

Figure 1.--Annual report directory with example of processing route for a field office.

PEE DEE RIVER BASIN

02123567 DUTCHMANS CREEK NEAR UWHARRIE, NC

LOCATION --Lat 35°22'05", long 80°01'49", Montgomery County, Hydrologic Unit 03040103, near midstream at upstream end of two 6 ft corrugated metal pipe culverts on Secondary Road 1150, 1.0 mi upstream from mouth and 3.0 mi southwest of Uwharrie.

DRAINAGE AREA.--3.44 mi^2.

WATER-DISCHARGE RECORDS

PERIOD OF RECORD.--October 1981 to September 1983, October 1985 to current year.

GAGE.--Water-stage recorder. Elevation of gage is 340 ft above National Geodetic Vertical Datum of 1929, from topographic map.

REMARKS.--Records fair except those for estimated daily discharges, which are poor.

EXTREMES FOR PERIOD OF RECORD.--Maximum discharge, 405 ft^3/s Apr. 15, 1987, gage height, 7.12 ft, from rating curve extended above 85 ft^3/s on basis of computation of peak flow through culvert; minimum, 0.01 ft^3/s July 16, 17, 21, Oct. 6-9, 1986; minimum daily, 0.01 ft^3/s July 15, 16, Oct. 7, 8, 1986.

EXTREMES FOR CURRENT YEAR.--Peak discharges greater than base discharge of 120 ft^3/s and maximum (*):

Date	Time	Discharge (ft^3/s)	Gage height (ft)	Date	Time	Discharge (ft^3/s)	Gage height (ft)
Feb. 28	Unknown	Unknown	Unknown	Apr. 15	2100	*405	*7.12

Minimum discharge, 0.01 ft^3/s Oct. 6-9; minimum daily discharge 0.01 ft^3/s Oct. 7,8.

DISCHARGE, IN CUBIC FEET PER SECOND, WATER YEAR OCTOBER 1986 TO SEPTEMBER 1987
MEAN VALUES

DAY	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP
1	.07	.31	.94	21	1.7	e60	6.1	2.5	2.1	3.3	.35	.95
2	.07	.49	3.9	8.9	1.9	e15	4.7	2.3	1.6	5.2	.36	.33
3	.06	.48	3.3	2.8	1.9	e6.0	4.3	2.2	1.0	2.9	.31	.24
4	.05	.40	1.3	1.8	1.5	e4.5	4.7	2.2	1.1	1.3	.26	.22
5	.04	.38	.94	1.5	1.3	e3.5	3.7	2.2	1.5	1.5	.25	.62
6	.02	.40	.78	1.3	1.3	e3.0	3.3	2.0	.97	.87	.26	1.5
7	.01	.61	.70	1.2	1.5	e2.8	3.0	2.0	.83	.76	4.0	3.7
8	.01	.95	.70	1.2	1.4	e2.8	2.8	1.9	.78	.69	.70	1.4
9	.03	.67	.82	1.1	1.2	e2.7	2.6	1.8	.76	.59	.44	.63
10	.08	.54	1.3	1.2	1.2	e2.6	2.4	1.7	.82	.53	.38	.56
11	.10	.47	2.9	1.2	1.2	e2.5	2.3	1.6	.72	.62	.38	.78
12	.10	1.3	2.3	1.1	1.3	e2.4	2.2	1.7	.90	3.0	.39	2.6
13	.13	1.2	1.6	1.1	1.2	e2.3	2.2	23	.86	3.3	.27	3.1
14	.88	.78	1.1	1.0	1.2	e2.2	2.1	6.3	1.1	1.0	.28	.81
15	.45	1.5	.98	1.1	1.2	e2.0	127	3.4	1.0	.85	.96	.55
16	.24	1.4	.89	1.1	1.5	e1.8	41	2.4	1.0	.71	.56	.45
17	.18	.88	.82	1.2	1.5	e1.7	11	1.9	.86	.58	.38	.43
18	.15	.77	.80	4.8	1.9	e9.0	7.8	1.5	3.5	.59	.31	.40
19	.15	.69	.75	43	3.8	e5.0	6.2	1.9	2.1	.56	.27	.40
20	.15	1.8	.70	8.5	4.0	e3.0	5.1	3.2	1.1	.51	.26	.39
21	.15	1.4	.65	3.9	3.5	e2.5	4.7	1.7	.94	.46	.23	.38
22	.15	.69	.64	39	4.0	e2.3	4.3	1.4	.89	.44	.22	.34
23	.13	.49	.65	13	8.1	e2.0	3.9	1.3	1.0	.39	.23	.31
24	.13	.49	7.6	5.5	3.9	e2.5	4.1	1.1	3.0	.35	.22	.31
25	.16	.83	3.6	7.8	2.8	2.6	4.1	1.0	2.0	.34	.14	.31
26	.70	1.5	2.0	9.8	2.2	2.5	3.5	1.1	1.4	.37	.16	.31
27	.54	1.2	1.5	4.6	e48	2.8	3.2	1.1	1.8	.36	.17	.27
28	.35	.76	1.2	3.3	e80	4.8	3.0	1.1	.90	.31	.15	.26
29	.25	.62	1.1	2.5	---	3.1	2.8	1.0	.73	.28	.21	.26
30	.24	.59	1.1	2.4	---	16	2.7	.92	.63	.28	.52	1.2
31	.24	---	.98	2.0	---	13	---	.87	---	.29	.63	---
TOTAL	5.97	24.59	48.54	199.9	186.2	188.9	280.8	80.29	37.89	33.23	14.25	24.01
MEAN	.19	.82	1.57	6.45	6.65	6.09	9.36	2.59	1.26	1.07	.46	.80
MAX	.88	1.8	7.6	43	80	60	127	23	3.5	5.2	4.0	3.7
MIN	.01	.31	.64	1.0	1.2	1.7	2.1	.87	.63	.28	.14	.22
CFSM	.06	.24	.46	1.87	1.93	1.77	2.72	.75	.37	.31	.13	.23
IN.	.06	.27	.52	2.16	2.01	2.04	3.04	.87	.41	.36	.15	.26

CAL YR 1986 TOTAL 427.12 MEAN 1.39 MAX 13 MIN .01 CFSM .40 IN. 4.62
WTR YR 1987 TOTAL 1124.55 MEAN 3.08 MAX 127 MIN .01 CFSM .90 IN. 12.2

e Estimated

Figure 2.--Typical page from annual report presenting surface-water data.

PEE DEE RIVER BASIN

02123567 DUTCHMAN'S CREEK NEAR UWHARRIE, NC--Continued

WATER-QUALITY RECORDS

PERIOD OF RECORD.--Water years 1982-83, October 1986 to September 1987.

PERIOD OF DAILY RECORD.--
SPECIFIC CONDUCTANCE: November 1981 to September 1983.
WATER TEMPERATURE: November 1981 to September 1983.

INSTRUMENTATION.--Water-quality monitor from November 1981 to September 1983.

REMARKS.--Basin above station is 100 percent forested. Station is part of district's Natural Quality Network for determining the quality of streams relatively unaffected by man.

COOPERATION.--Chemical and biological data shown in last table were provided by the North Carolina Department of Natural Resources and Community Development.

EXTREMES FOR PERIOD OF DAILY RECORD.--
SPECIFIC CONDUCTANCE: Maximum, 66 microsiemens July 22, Aug. 4, 5, 6, 23; minimum, 22 microsiemens Mar. 6.
WATER TEMPERATURE: Maximum, 28.0°C July 21, 22, Aug. 21; minimum, 0.5°C Jan. 19, 20, 21.

WATER QUALITY DATA, WATER YEAR OCTOBER 1986 TO SEPTEMBER 1987

DATE	TIME	STREAM-FLOW, INSTAN-TANEOUS (CFS)	SPE-CIFIC CON-DUCT-ANCE (US/CM)	PH (STAND-ARD UNITS)	TEMPER-ATURE WATER (DEG C)	COLOR (PLAT-INUM-COBALT UNITS)	TUR-BID-ITY (NTU)	BARO-METRIC PRES-SURE (MM OF HG)	OXYGEN, DIS-SOLVED (MG/L)	OXYGEN, DIS-SOLVED (PER-CENT SATUR-ATION)	HARD-NESS (MG/L AS CACO3)
FEB											
27...	1245	26	40	6.60	6.0	--	--	760	13.4	108	--
27...	1635	83	35	6.40	5.0	--	--	760	13.4	105	--
28...	1045	127	30	5.70	7.0	55	17	755	11.9	99	7
28...	1330	178	27	5.85	7.0	60	44	755	12.0	100	7
MAR											
01...	0833	94	33	5.75	7.5	49	9.0	--	--	--	--
JUL											
28. .	1245	0.32	78	7.50	23.5	--	--	750	7.6	--	--
AUG											
07...	1235	6.3	54	7.11	23.5	69	38	752	8.0	96	11

DATE	HARD-NESS NONCARB WH WAT TOT FLD MG/L AS CACO3	CALCIUM DIS-SOLVED (MG/L AS CA)	MAGNE-SIUM, DIS-SOLVED (MG/L AS MG)	SODIUM, DIS-SOLVED (MG/L AS NA)	PERCENT SODIUM	SODIUM AD-SORP-TION RATIO	POTAS-SIUM, DIS-SOLVED (MG/L AS K)	SULFATE DIS-SOLVED (MG/L AS SO4)	CHLO-RIDE, DIS-SOLVED (MG/L AS CL)	FLUO-RIDE, DIS-SOLVED (MG/L AS F)	SILICA, DIS-SOLVED (MG/L AS SIO2)
FEB											
27...	--	--	--	--	--	--	--	--	--	--	--
27...	--	--	--	--	--	--	--	--	--	--	--
28...	3	1.5	0.77	2.1	36	0.4	1.0	12	1.8	<0.10	8.2
28...	3	1.5	0.74	1.6	29	0.3	1.2	13	1.3	<0.10	6.1
MAR											
01...	--	--	--	--	--	--	--	13	1.6	<0.10	--
JUL											
28...	--	--	--	--	--	--	--	--	--	--	--
AUG											
07...	0	2.7	1.1	3.6	37	0.5	1.7	11	2.2	0.10	17

Figure 3.--Typical page from annual report presenting water-quality data.

camera-ready copy that will be sent to the printer. The entire process of producing and editing the data tables and merging them with the manuscript is performed automatically with the minicomputer through the use of command procedure language (CPL) programs.

Printing of the annual data report is done using a laser printer and a special font conversion program developed for the annual report by Patsy Mixson and Michael Hathaway (P.R. Mixson and M. Hathaway, U.S. Geological Survey, oral commun., 1987). The font conversion program, used in conjunction with WordMarc Composer, produces the special scientific symbols used in the annual report, such as the degree and double dagger symbols and superscripted 2's and 3's. The flow chart of the annual report process is shown in figure 4.

The 1985 annual data report, which was produced with the new system on the minicomputer, was sent to the printer only two days earlier than the 1984 annual data report, which was done the "old-fashioned" way--that is, each manuscript was spliced to its respective data table. Extra time was required to design, debug, improve, and implement the automated system. Unforeseen delays in printing the camera-ready copy also were experienced. A slow, letter-quality printer was used, and 12-15 days were spent working with it. The use of the laser printer and font conversion program reduced the time to prepare camera-ready copy to about 8 hours for the entire 500-page report. The program also reduces the printout to page size (8 1/2 x 11 inches) so that no further reduction is required by the printer. The page-size camera-ready copy for the printer eliminates the possibility of reduction errors and should lower printing costs by simplifying the process. According to Tom Nepi, assistant manager for the regional Government Printing Office, providing page-size camera-ready copy for the printer will increase competition among printers and greatly reduce the bid price (T. Nepi, Government Printing Office, oral commun., 1987).

Perhaps the biggest initial problem in 1985 involved merging the water-quality data with the respective manuscript and producing publication format. The 1985 water-quality data were retrieved in batch mode from the Amdahl main-frame computer in Reston. Because the merging and editing process of the resulting printouts could not be automated, it was necessary to

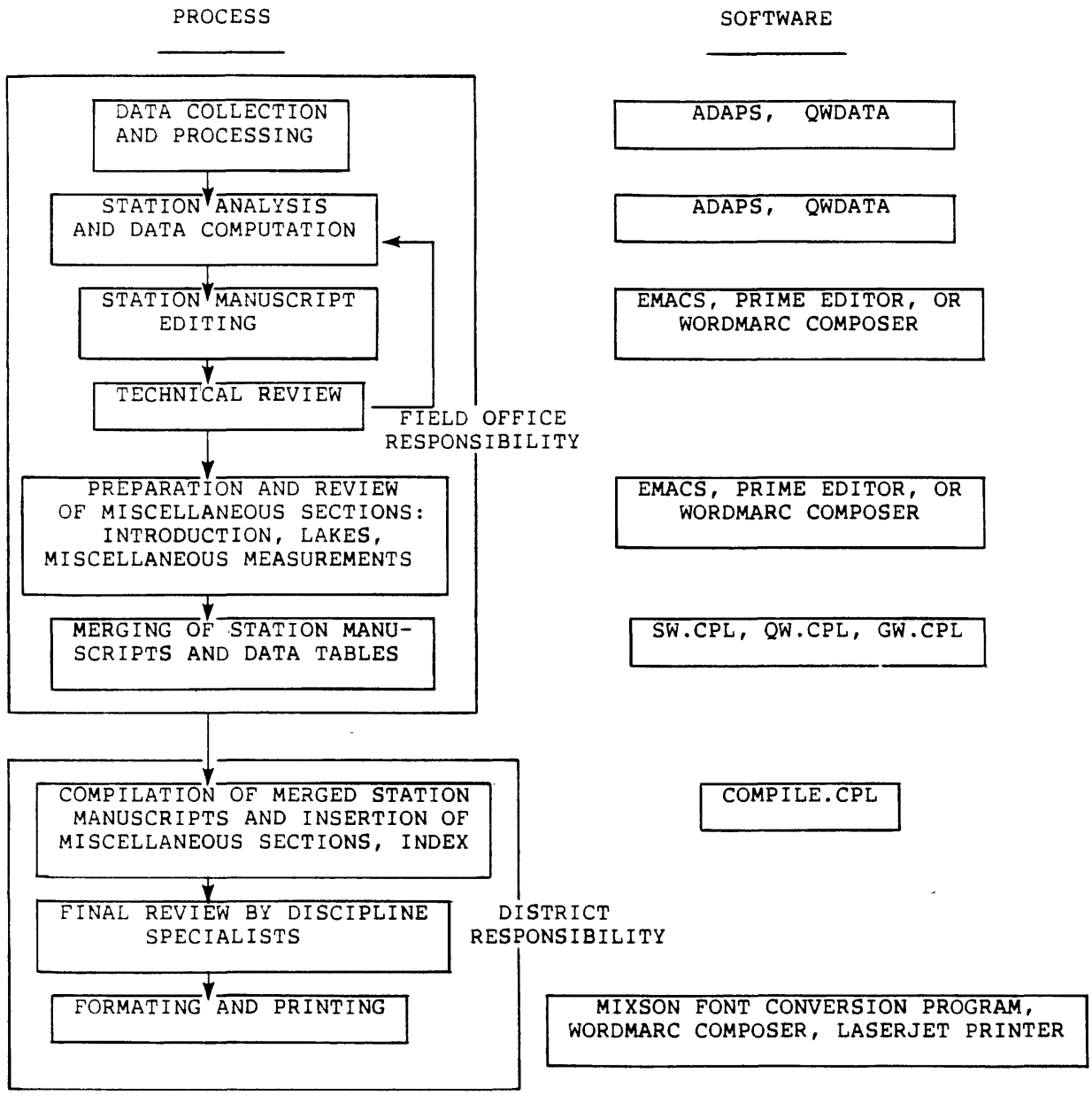


Figure 4.--Flow chart of annual report process.

use EMACS or WordMarc to obtain publication format. The editing process was slow, laborious, and, at times, difficult to track. Fortunately, the availability of the QWDATA system in 1986 eliminated this problem.

Other major problems involved pagination and indexing. Editors used in preparing the 1985 annual data report did not have the ability to place page numbers on alternating sides of consecutive pages. That problem was circumvented by having the machine print the page number on both sides of each page and then painting out the undesired number. The WordMarc Composer software package, which has the capability of alternating page numbers, was used to prepare the 1986 annual data report.

Indexing was done manually for the 1985 and 1986 reports. Consideration was given to automating the process, but the short amount of time available to prepare an automation procedure prevented it. Attempts to cross reference the index created a problem, because there is no simple way to uniquely identify each station by city or town by using the index format now required. In many cases, the same town or stream name appears in several station names. The problem could probably be solved by a change in publication requirements. The use of single-entry listings for all stations rather than multiple-entry listings when there is more than one station on a stream might simplify the cross indexing so it could be automated.

Computer Programs

To effectively use the minicomputer to generate the annual report, it was necessary to write a number of programs to generate data tables and merge them with station manuscripts (fig. 4). Almost all of the programs were written as CPL programs. CPL is the minicomputer's command procedure language (Landy, 1982). CPL programs interact with existing computer systems, such as ADAPS and EMACS, by supplying input instructions for those systems and by passing output from one system to another. The CPLs do what the user would do to activate and execute a job. They were very easy to write and debug and did not require complicated input and output instructions. A discussion of the major programs, including descriptions of their formats and functions follow.

The SW.CPL Program

This program generates a surface-water daily value table from ADAPS, edits it into publication format, and merges the table with the appropriate station manuscript. The user must specify the downstream order number for the surface-water station and the field office subdirectory under which the station file is located. The command, "SW," is a global command on the minicomputer, hence it may be used from any level in the computer. The form of the command sequence is "**SW number field,**"
where:

"**SW**" is the name of the command;
"**number**" is the station number of the station of interest; and
"**field**" is the name of the field office subdirectory in which the station manuscript resides.

A typical command sequence might be "**SW 03161000 ASH.**"

The SW command sequence starts a program which creates and edits an ADAPS batch control file, inserting the values of number and field. ADAPS then uses the batch control file to set printout format options and parameter statistics and generates the daily-values table. Once the daily values table has been constructed, ADAPS uses the variables number and field to define portions of path and file names for the output. When the ADAPS system has generated the daily-values table and deposited it into the appropriate file, the SW.CPL program evokes the EMACS editor to remove unnecessary header information, blank lines and spaces, and inserts the appropriate manuscript at the top of the data table. The entire process is displayed on the user's terminal and requires about 20 seconds per station for execution on the Prime 9955 minicomputer.

The QW.CPL Program

This program operates in a manner similar to that of the SW.CPL program in that it generates water-quality daily-values tables using ADAPS. In addition, the program also generates the water-quality sample analysis tables from the QWDATA system. These tables and the appropriate water-

quality manuscript are then merged and edited into publication format by use of the EMACS editor. The program takes about 2 minutes per station to run. The user must specify the station number and the field office subdirectory. The command is global and must be in the form "**QW number field**,"

where:

"**QW**" is the name of the command;
"**number**" is the station number; and
"**field**" is the field office subdirectory.

A typical command sequence might be "**QW 03453500 ASH.**"

The GW.CPL Program

This program operates similarly to the SW.CPL program and is used to generate a data table from ADAPS for ground-water stations. The user specifies the station number and the field office directory where the station manuscript is located. The command is global. The command sequence is of the form "**GW number field**,"
where:

"**GW**" is the name of the command;
"**number**" is the local station number; and
"**field**" is the field office subdirectory.

A typical command sequence might be "**GW NC.144 ASH.**"

The COMPILE.CPL Program

To facilitate printing and allow pagination, the merged surface-water and water-quality files for each station are appended in downstream order to a single, large file to which other sections of the annual report, such as the introduction, are added. Most of this work is performed by the COMPILE.CPL program. When complete, the large file contains the report introduction, all USGS and cooperator data, lake contents, and miscellaneous streamflow measurements. The COMPILE command is a global command.

A number of other programs have been written to perform various other editing functions such as inserting missing headers on sequential data tables and centering titles and text for camera-ready copy. Most of these programs simply invoke the Prime editors. Some manual editing is still done directly through EMACS or WordMarc, but this is minimized by using consistent formats and automating where possible.

Planned Modifications to Improve System

The North Carolina District plans to use a computer program for the 1987 annual data report that will invoke the ADAPS system to pull the extremes for the year and the average discharge for the period of record and insert those values in the proper places in the manuscript heading. Plans are to modify the COMPILE.CPL program to insert WordMarc Composer and alternate characters used by the font conversion program for printing special scientific symbols, such as the degree sign. Algorithms are being developed to handle indexing and cross referencing of the report. With these modifications to the system, a significant reduction in the time required for preparing each successive annual data report is anticipated.

SUMMARY AND CONCLUSIONS

Use of the automated report preparation system described in this report will significantly reduce the number of days required to produce the U.S. Geological Survey's annual water-resources data report. At this point, an annual labor savings of 2 man-months has been estimated, but an increase in this is anticipated as the system is improved and refined. The programs used by the North Carolina District of the U.S. Geological Survey provide a method of linking various subsystems of ADAPS to existing word processing systems commonly used by the Geological Survey (EMACS and WordMarc). The programs could be incorporated into the existing ADAPS software. The required hardware, a Prime computer and a laser printer, are relatively standard equipment in most Geological Survey offices, which could easily adopt the system for general use.

The Geological Survey's goal is to automate the process of compiling the report to the greatest extent possible. Hopefully, this report will generate interest and discussion in automating the annual data-report process.

REFERENCE CITED

Landy, A., 1982, CPL User's Guide COC 4302-190, Revision 19: PRIME Computer, Inc.