U.S. GEOLOGICAL SURVEY NATIONAL COMPUTER TECHNOLOGY MEETING: PROCEEDINGS, PHOENIX, ARIZONA, NOVEMBER 14-18, 1988

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 90-4162

U.S. GEOLOGICAL SURVEY NATIONAL COMPUTER TECHNOLOGY MEETING: PROCEEDINGS, PHOENIX, ARIZONA, NOVEMBER 14-18, 1988

By Barbara H. Balthrop and John E. Terry, editors

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 90-4162

science for a changing world

Nashville, Tennessee

1991

U.S. DEPARTMENT OF THE INTERIOR
MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY
Dallas L. Peck, Director

For additional information write to:

Coordinator, National Computer Technology
Meeting Proceedings 1988
U.S. Geological Survey
300 W. Congress
Federal Building, FB-44
Tucson, Arizona 85701

Copies of this report can be purchased from:

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

FOREWORD

The U.S. Geological Survey National Computer Technology Meetings (NCTM) are sponsored by the Water Resources Division and provide a forum for the presentation of technical papers and the sharing of ideas or experiences related to computer technology. This report serves as a proceedings of the meeting held in November, 1988 at the Crescent Hotel in Phoenix, Arizona. The meeting was attended by more than 200 technical and managerial people representing all Divisions of the U.S. Geological Survey.

Scientists in every Division of the U.S. Geological Survey rely heavily upon state-of-the-art computer technology (both hardware and software). Today the goals of each Division are pursued in an environment where high speed computers, distributed communications, distributed data bases, high technology input/output devices, and very sophisticated simulation tools are used regularly. Therefore, information transfer and the sharing of advances in technology are very important issues that must be addressed regularly.

This report contains complete papers and abstracts of papers that were presented at the 1988 NCTM. The report is divided into topical sections that reflect common areas of interest and application. In each section, papers are presented first followed by abstracts. For these proceedings, the publication of a complete paper or only an abstract was at the discretion of the author, although complete papers were encouraged.

Some papers presented at the 1988 NCTM are not published in these proceedings.

John E. Terry

CONTENTS

CHAPTER A--USGS COMPUTER MANAGEMENT/ADMINISTRATION

CHAPTER B--DIS/DISTRIBUTED ENVIRONMENT

CHAPTER C--GIS APPLICATIONS IN USGS

CHAPTER D--COMPUTER GENERATED GRAPHICS

CHAPTER E--DATA BASES AND AUTOMATED DATA HANDLING

CHAPTER F--ELECTRONIC PUBLISHING/MANUSCRIPT PREPARATION

CONVERSION FACTORS

Multiply	By	To obtain
inch (in.)	25.4	millimeter
inch per day (in/d)	25.4	millimeter per year
foot (ft)	0.3048	meter
foot per day (ft/d)	0.3048	meter per day
foot per mile (ft/mi)	0.1894	meter per kilometer
cubic foot per second (ft^3/s)	0.02832	cubic meter per second
mile (mi)	1.609	kilmometer
square mile (mi^2)	2.590	square kilometer
mile per square mile (mi/mi^2)	0.621	kilometer per square kilometer
gallons per day (gal/d)	0.003785	cubic meter per day
million gallons (Mgal)	3,785	cubic meter
million gallons per day per square mile [$(Mgal/d)mi^2$]	1,460	cubic meter per day per square kilometer

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F = 1.8 \times °C + 32$$

The use of brand, company, or trade names in this report is for indentification purposes only and does not constitute endorsement by the U.S. Geological Survey.

AUTOMATED DATA COLLECTION AND ENTRY TECHNIQUES FOR WATER-USE INFORMATION IN ARKANSAS

By Nancy T. Baker and Terrance W. Holland

U.S. Geological Survey

ABSTRACT

The Arkansas Soil and Water Conservation Commission, in cooperation with the U.S. Geological Survey, annually collects water-use information for more than 25,000 registered wells and surface-water withdrawal sites. To increase the efficiency of collecting and entering large amounts of water-use information, the Survey has developed several automated techniques.

Computer software, designed for use with the Local Applications options of Arkansas' version of the New Site-Specific Water-Use Data System, was developed to automate the printing and distribution of water-use registration forms.

Three methods of using personal computers (located in county Conservation District offices) to improve the efficiency of the data entry process were also developed. One method requires that water-use data entered by each Conservation District office be stored on personal computer diskettes. These data then need to be transferred from the diskette into the New Site-Specific Water-Use Data System data base on the U.S. Geological Survey's minicomputer. The other two methods involve directly accessing the data base by linking personal computers with the use of a modem to the U.S. Geological Survey's computer, which is equipped with a security dial-back system. An advantage of the dial-back system is that costly long distance phone charges can be avoided. It is anticipated that eventually paper copies of the water-use registration form will not be distributed and that all water-use data will be collected and transferred directly to the data base.

INTRODUCTION

A large quantity of potable freshwater is one of Arkansas' most valuable resources. Approximately 5,860 million gallons of freshwater are withdrawn daily from ground- and surface-water sources (Holland, 1987, p. 9). To ensure the protection, development, and management of its water resources, the State has passed legislation requiring registration of diversions of water from streams, lakes, and ponds, and withdrawals of ground water to be registered with the Arkansas Soil and Water Conservation Commission (ASWCC). The ASWCC has been designated as the State agency responsible for water-resources planning, including collection of water-diversion registration information. The U.S. Geological Survey in cooperation with the Arkansas Soil and Water Conservation Commission collects, stores, and manages Arkansas' water-use data for the State. Water-use information for over 25,000 withdrawal sites were reported to the ASWCC during 1986.

The ASWCC entered into its water-use cooperative program with the U.S. Geological Survey in 1985. The State Water Use Data System (SWUDS), available nationwide to U.S. Geological Survey District offices and their cooperators, was the computer program that was initially chosen to manage the data base. Because SWUDS did not meet the specific needs of the ASWCC, the original program (or code) was modified and resulted in a new program called New Site-Specific Water-Use Data System (NEWSWUDS).

Currently (1988), site-specific water-use data are collected by ASWCC by distributing water-diversion registration forms to ground- and surface-water users in the State. Nonagricultural registration forms are mailed directly to the water user. Agricultural forms are distributed through local Conservation District offices in each county. These offices commonly deal closely with the agricultural water users in their district and are able to collect more water-use information than the ASWCC can collect by direct solicitation. In most cases, the water user goes to the Conservation District office and personnel from the office help register their water usage. The Conservation District office sends completed forms back to the ASWCC for entry into NEWSWUDS. Because the large volume of site-specific data reported annually to the ASWCC has become increasingly difficult to collect and process effectively, the U.S. Geological Survey has developed several techniques to automate data collection and entry. This paper describes the techniques that have been developed by the Arkansas District of the U.S. Geological Survey to automate water-use data collection and entry.

AUTOMATED DATA COLLECTION

Applications software was added to the Local Applications system in NEWSWUDS to streamline the data-collection process. The Local Applications system is a feature of NEWSWUDS that allows locally developed applications to be installed in the NEWSWUDS system by using Fortran "envelope" routines and templates that are provided as a part of the system.

The ASWCC Forms Routine is one of the local application software packages installed in NEWSWUDS in Arkansas. The ASWCC Forms Routine prints water-diversion registration forms for ground- and surface-water withdrawal sites (figs. E23 and E24). The routine can print blank registration forms without water-user information or it can preprint pertinent water-user information, such as water-user identification number, name, address, well and pump data, and location of diversion. Preprinting the forms assures that water users who submitted forms the previous year will be sent a form with current information that may be updated as necessary. The amount of time needed to complete the form is reduced considerably, and generally, the only new information that is required is the annual water-use data.

To create the preprinted forms with the ASWCC Forms Routine, the user needs to first retrieve the appropriate information from the NEWSWUDS data base. A "Command Procedure Language" (CPL) program, which operates at the system level, has been set up to retrieve the data from NEWSWUDS. The CPL allows the user to retrieve information quickly and easily, and in the order required by the ASWCC Forms Routine. The CPL also eliminates the typing of a long sequence of commands each time a retrieval is made. The retrieval can be submitted as a batch job; this frees the the system for other uses, and allows

WATER WITHDRAWAL REGISTRATION FORM
AGRICULTURE

1. Name of Landowner ____________________
2. Water User ID. ____________________
3. Use of WATER:(check one)
   ____AG Agriculture (livestock, fish farms) ____IR Irrigation
4. SIC Code 1st ____________. 2nd ____________.
5. ASWCC Permit # ____________ 6. Telephone number ____________
7. Name of Diverter ____________________
8. Hydrologic Unit Code ____________
9. Street Address ____________________
10. City name ____________ 11. ZIP Code ____________
12. FIPS County code ____________ 13. FIPS State Code ____________
14. State _____ 15. Action Code _____ 16. Source type. ____GW ____SW
17. Conservation District Permit #: Topographic Quad # ___ ___ ___ ___
    Operator # ___ ___ ___ WELL # ___ ___
18. Aquifer (GW only) ____________
19. Reclaimed waste (GW only) _____ YES _____ NO
20. Name of lake, ditch, bayou, or tributary (SW only) ____________
21. Was a dam or other obstruction of streamflow utilized (SW only)?
    _____ YES _____ NO
22. Method of Irrigation: ____Flood, ____Furrow, ____Sprinkler, ____Other
23. Diameter of pipe ____________ 24. Depth of well ____________
25. Location of diversion or withdrawal point (Legal description):
    _____ 1/4 of _____ 1/4, Section ____, Township ____, Range ____
26. Location of withdrawal: Latitude ____________ Longitude ____________
27. Method of Diversion:
    ____Stationary pump ____Portable Pump ____Gravity
28. Type of power ____ELC ____LPG ____Diesel ____Other
29. Pump Horsepower ____________
30. Year 19____
31. Annual amount withdrawn(acre ft.) 1987________ 1988________
32. Measuring Method: (circle one)
    __E__ Estimated __M__ Measured __C__ Calculated.
33. Restrictions: (circle one) __Y__ __N__
34. Salinity: (circle one) __F__Fresh __S__ Saline __U__ Unknown
35. Estimate total withdrawal per month.

	1987	1988		1987	1988		1987	1988
JAN			FEB			MAR		
APR			MAY			JUN		
JUL			AUG			SEP		
OCT			NOV			DEC		

36. (a) Kind of crop	(b) Number of irrigated acres	(c) Amount of water applied (ft)	(d) Total acre-ft. (b x c)

____________________ (Signature) ________ (Date) ____________________ (County of Diversion)

Figure E23.--Example of Arkansas Soil and Water Conservation Commission water-withdrawal registration form for agricultural water users.

WATER-USE REGISTRATION FORM

1. Name of Facility ____________________

2. Water-user Id. # ____________________

3. Use of water: (check one)

WS - Public Supply Use	____	PF - Fossil Fuel Power	____
CO - Commercial Use	____	PG - Geothermal Power	____
IN - Industrial Use	____	PN - Nuclear Energy Power	____
MI - Mining	____	PH - Hydroelectric Power	____

4. SIC Code (1)________ (2)________ (3)________

5. Permit# or PWS#________ 6. Telephone Number (___)________

7. Location of Facility: Latitude________ Longitude________

8. Hydrologic Unit Code____________

9. Address ____________________

10. City ______________ 11. ZIP Code ________

********* COMPLETE 12-18 FOR PUBLIC WATER SUPPLY FACILITIES ONLY *******

12. Total Ground & Surface Water withdrawals(MILLION GALS.)________

13. Water Purchased From Other Facilities (Amount in MILLION GALS.):

Facility Name ______________ Amount:________
Facility Name ______________ Amount:________

14. Water Sold To Other Facilities (Amount in MILLION GALS.):

Facility Name ______________ Amount:________
Facility Name ______________ Amount:________

15. Water Used For Facility Maintenance (backflushing, plant operations, losses due to seepage and leaks - MILLION GALS.)________

16. Domestic Population Served ______________

17 Deliveries to Users (Do not include water sold to other facilities):

Water Furnished To:	Total Water Delivered	Number of Connections
Domestic Households	________	________
Commercial	________	________
Industrial	________	________
Mining	________	________
Agriculture	________	________
Irrigation	________	________

18. Current Rate Structure:

________ ________ ________ ________ ________

________________ (Name of Plant Manager) (Date) ________________ (County of Diversion)

FILL OUT A SEPARATE PAGE FOR EACH WELL OR SURFACE WATER DIVERSION POINT

19. Measurement Point ID____________________

20. Local Descriptive ID____________________

*** If CENTROID, amount of water withdrawn (#29-30) represents the sum ***
*** of water withdrawn from all withdrawal points for this facility. ***

21. Action Code______ (If Action Code is DL skip to #29)

22. Source of Water (Ground, Surface, or Transfer) ____________

23. If Surface Water withdrawal:

A. Name of Lake or Stream:____________________

24. If Ground Water withdrawal:

A. Name of Aquifer: ____________________

B. Depth of Well ____________________

C. Name of Driller ____________________

25. Pump Information:

A. Pump Capacity (Horsepower) ______________

B. Type of Power (check one): Electric ____ LPGas ____ Diesel ____ Other ____

C. Type of Pump (check one): Stationary ____ Portable ____ Gravity ____ Other ____

D. Diameter of Well or Intake Pipe: ______________

26. Location of Well or Withdrawal Point (Well or Intake Structure):

Latitude ________ Longitude ________

27. Hydrologic Unit Code ____________________

28. Withdrawal Amounts Obtained From (check one):

Billing Records_____ Pump Capacity + Hours Operation_____
Meter Readings _____ Other (list)______________

**** For questions 29-30, withdrawal values reported in (check one):****

Gallons_____ Thousand Gallons_____ Million Gallons_____

29. Total Water Withdrawn From This Withdrawal Point (nn):________

30. Amount of Water Withdrawn From This Withdrawal Point (nn):

nn	nn	nn
JAN________	FEB________	MAR________
APR________	MAY________	JUN________
JUL________	AUG________	SEP________
OCT________	NOV________	DEC________

Figure E24.--Example of Arkansas Soil and Water Conservation Commission water-withdrawal registration form for nonagriculture water users.

users not familiar with the NEWSWUDS system to retrieve data for the printing of registration forms.

After the appropriate information is retrieved from NEWSWUDS, the ASWCC Forms Routine is selected from the Local Applications menu. The user responds to the queries in the routine and the water-user information retrieved earlier is automatically printed on the water-diversion registration forms. To improve distribution, the forms are retrieved and printed according to (1) the type of water use (agriculture (fig. E23) or nonagriculture (fig. E24)), (2) the source type (ground or surface water), and (3) the county in which the withdrawal site is located.

Another Local Applications feature added to NEWSWUDS is the ASWCC Mail Labels Routine. This routine generates mailing labels from the data retrieved for the ASWCC Forms Routine. Labels are printed in the same order that the forms are printed. In Arkansas, mail labels are generated for nonagricultural water users.

AUTOMATED DATA ENTRY

Although automation of registration form processing and distribution has helped to streamline the data collection process, the ASWCC does not have the personnel necessary to process effectively all the registration forms that are received. Most water-diversion registration forms received by the ASWCC are from the agricultural region in eastern Arkansas. If the ASWCC could divide the task of entering the water-use data collected in the large usage counties of eastern Arkansas among the local Conservation District offices, data entry time could be greatly reduced. Techniques to divide data entry among ASWCC and Conservation District personnel are being tested in a few Arkansas counties. Many of the Conservation District offices are being equipped with personal computers (PC's) and several options using the computers for remote data entry are being investigated by the USGS and the ASWCC.

Personal Computer Diskette Method

The personal computer diskette method requires that each Conservation District office enter water-use data on PC diskettes. Data are entered using Keydisk, which is a software package that formats water-user information for entry into NEWSWUDS. Keydisk is part of NEWSWUDS, however, a PC version of Keydisk is available from the USGS. The data entered into the PC version of Keydisk would be saved on a diskette. The diskettes would then be sent to the ASWCC and loaded into the NEWSWUDS data base. The advantages of this system are: (1) the Conservation District offices do not need to purchase any additional equipment (phone lines or modems), (2) Conservation District personnel could enter the data at their convenience and would not be restricted to a schedule of data entry, which is a possibility for the other methods, and (3) there would not be a slowdown of the U.S. Geological Survey's computer because of increased usage load. The disadvantages include: (1) multiple handling of the diskettes, (2) an extra step is involved in processing the information (the diskettes would be sent to the ASWCC for loading), and (3) there would be several copies (1 for each of the 26 counties in eastern Arkansas) of the Keydisk software to maintain.

Dial-Back Method

The dial-back method requires accessing the NEWSWUDS data base directly by linking the PC's by modem to the U.S. Geological Survey's computer, which is equipped with a security dial-back system. The dial-back system consists of placing a modem at the PC end of the system and a dial-back controller at the U.S. Geological Survey's computer end. To use the system, the person at the PC has to call the dial-back controller from a touch-tone phone, enter an access code, and wait for the prompt to hangup the phone. The access code has to be preprogrammed in the dial-back controller. Upon receiving the code, the controller calls the number in its library for that access code and completes the connection. After the system is connected, a signal is sent to the user to set the terminal baud rate. The computer system in the Arkansas District office of the Geological Survey has baud rate options of 300, 1200, and 2400 baud. Upon setting the baud rate, the user can login to the U.S. Geological Survey's computer system.

There are several advantages to the dial-back system that include: (1) Conservation Districts will not incur costly long distance phone charges, (2) Conservation District personnel would have access to the NEWSWUDS data base and would be able to make their own retrievals, and (3) there would only be one copy of Keydisk to maintain. Eventually, paper copies of the water-diversion registration form might not be needed, and water-use data can be collected and directly transferred to the NEWSWUDS data base.

Disadvantages include the possibility of having to enter data on a restricted schedule. Presently (1988) the computer system in the Arkansas District of the U.S. Geological Survey does not have enough dialup lines to handle 26 additional users accessing the NEWSWUDS software. A schedule may have to be set up to assure that only a few counties are entering data at a given time. Another disadvantage is that many of the counties might have to install an additional phone line and might have to purchase a modem.

Combination Method

The combination method involves using the dial-back system to transfer data from the PC to the U.S. Geological Survey's computer, and using the PC version of Keydisk to enter the data. Data would be entered at the PC without linking to the U.S. Geological Survey's computer. These data would be stored in the PC until all data for a particular county are entered. Conservation District personnel would then transfer all of the data at one time to the U.S. Geological Survey's computer system by using the dial-back controller. Conservation District personnel would still have access to information in the NEWSWUDS data base, and this method would not cause a heavy usage load on the U.S. Geological Survey's computer. Although this method would reduce the multiple handling of PC diskettes, there would still be several copies of the PC version of Keydisk to maintain.

CONCLUSION

The development of techniques to automate data collection and entry should greatly increase the efficiency of data management in Arkansas. By increasing the efficiency of water-use data management, more attention can be given to improving the quality of the data collected. Automation of data collection assures a more thorough distribution and collection of water-use information. Automation of data entry divides the task of entering large quantities of data among Conservation District personnel in each county and reduces the workload of the Arkansas Soil and Water Conservation Commission. Of the three automated methods discussed, no single method will likely be appropriate for every county in eastern Arkansas. The method appropriate for an individual county will probably be dependent on the facilities and needs of the Conservation District office.

REFERENCES

Holland, T.W., 1987, Use of water in Arkansas, 1985: Arkansas Geological Commission Water-Resources Summary 16, 27 p.

AUTOMATION OF THE WATER-USE DATA BASE FOR MINNESOTA

By Lee C. Trotta

U.S. Geological Survey

Water Resources Division

St. Paul, Minnesota

ABSTRACT

Because of the complexity and scope of water use, management of water resources requires classification and automation of these data. Data are collected for 10 different categories of water use for the U.S. Geological Survey's National Water-Use Information Program. Water-use managers in Minnesota needed a system to satisfy complex management needs that would identify users in a specific location, describe the source of water, type of use, and rates and volumes of water used. Summaries of water use within a county or watershed were required for transfer of these data to the U.S. Geological Survey's Aggregated Water-Use Data System. A computer based relational data-base-management system was chosen and developed jointly by the Minnesota Department of Natural Resources and U.S. Geological Survey to automate data-base management of water for the State and national programs. Data aggregation programs, a simulation model based on Standard Industrial Classification Codes to determine the economic value of water, and various utility programs have been developed. Many decisions by Federal, State, and local water-resources managers for mandating water-use restriction, initiating ground-water studies, and developing local comprehensive water plans are streamlined by automation of the water-use data base.

INTRODUCTION

Everyone uses water and water is used in many different ways. Water is withdrawn from sources on the surface or in the ground, used for one or more reasons, and then consumed or discharged back to the ground or surface. The U.S. Geological Survey's National Water-Use Information Program currently (1983) collects data on withdrawal, use, and disposition for 10 different categories of water use. These categories are thermoelectric power generation, hydroelectric power generation, irrigation, livestock, commercial, domestic, industrial, mining, public supply, and sewage treatment.

Users of large amounts of water may affect the quantity and quality of water available to other users. The State of Minnesota monitors water use for its large users. Water-use information is collected by a variety of State agencies to satisfy various legislative mandates. Because of the complexity and scope of water use, management of the water resource requires classification of water-use data and automation of data entry, processing and retrieval.

This report describes the Minnesota Water-Use Data System (MWUDS), computer methods used to automate the system, and programs to use the data base. MWUDS was developed independently from the State Water-Use Data System, distributed by the U.S. Geological Survey (Survey).

Permitting Systems in Minnesota

Permitting systems are used to obtain site-specific water-use information in Minnesota. The State of Minnesota has established one permit system for water appropriations greater than 10,000 gallons per day (or 1 million gallons per year, whichever is less) and another to monitor the quality of water dis-charged. Water appropriations (withdrawals) are reported to the Minnesota Department of Natural Resources (MDNR). Rates of compliance with the reporting provision of permits is about 85 percent among irrigation users and 95 percent among all other users (Young, 1987, p. 6).

The MDNR collects information from permit reports in 6 of the 10 water-use categories defined by the Survey. However, the MDNR has defined about 50 subcategories to allow more detailed analysis (table E-8). The MDNR also collects data on aquaculture (code 72), quality improvement (codes 60, 61 and 71), and preservation (codes 60, 61, and 74). These categories are not part of the current focus of the Survey. Relations among water-use categories specified by both MDNR and the Survey were defined by Marilee Horn (1984, p.4) and are revised and expanded upon in the second part of table E-8. Some Survey-specified categories in this part of the table may contain only limited data (hydroelectric power generation and sewage treatment), or be completely absent (livestock and domestic) because of the nature of MDNR's withdrawal-permit reporting requirements. Withdrawals for three of these Survey-specified categories (hydroelectric power generation, livestock, and domestic) have been and are estimated by the Survey annually since 1984, and on a 5-year schedule before that (Trotta, 1988). A water-quality-based permit system regulated by the Minnesota Pollution Control Agency (MPCA) provides some data on water returned to the environment at sewage treatment plants.

COMPUTER METHODS

Thirty-nine years passed between issuance of the first Minnesota water-use permit and development of a rudimentary data-storage system in 1976. During that time, information needed by State planners was retrieved by leafing through thousands of pages of data in notebooks (Gil Young, Minnesota Department of Natural Resources, written commun., 1987).

In 1982, the Minnesota Department of Natural Resources and the U.S. Geological Survey began a joint project to implement a water-use data-base-management system (Horn, 1986). After a long search for applicable computer-based systems, a computer based relational data-base-management system was chosen (fig. E-25) that allows staff with minimal training or computer experience to input, update, and retrieve data. The Land Management Infor-

Table E-8.--Water-use codes and categories in the Minnesota Water-Use Data System

Code	Category used by the Minnesota Department of Natural Resources	Code	Category used by the Minnesota Department of Natural Resources
3	Well abandoned	43	Mine processing (other than sand and gravel washing)
4	Terminated		
5	Active not under permit	44	Sand and gravel washing
		45	Sewage treatment
10	Waterworks	46	Petroleum (chemical processing)
11	Municipal	47	Metal processing
12	Private waterworks (Trailer courts/ small housing units/and so forth)	48	Non-metallic products (rubber/plastic/glass)
13	Commercial and institutional	50	Temporary
14	Cooperative waterworks	51	Construction (non-dewatering)
15	Fire protection	52	Construction (dewatering)
16	State parks/waysides/highway rest areas	53	Pipeline and tank testing
		54	Landscape watering
20	Power generation	60	Water level maintenance
21	Hydropower	61	Basin (lake) level
22	Steam power cooling (once through)	62	Mine dewatering
		63	Quarry dewatering
23	Steam power cooling (wet tower)	64	Sand/gravel pit dewatering
		65	Tile drainage and pumped sumps
24	Steam power cooling (ponds)		
25	Steam power (other than cooling)	71	Pollution confinement
		72	Hatcheries and fisheries
26	Nuclear power	73	Snow making
		74	Peat fire control
30	Air conditioning		
31	Commercial building	80	Non-crop irrigation
32	Institutions (schools, hospitals)	81	Golf course
		82	Cemetery
33	Heat pumps	83	Landscaping
34	Coolant pumps	84	Sod
35	District heating	85	Nursery
		86	Orchard
40	Industrial processing		
41	Food and livestock (agricultural processing)	90	Major crop irrigation
		96	Wild rice irrigation
42	Paper/Pulp		

Codes	Category used by the U.S. Geological Survey
20, 22-26	Thermoelectric power generation
21	Hydroelectric power generation
54, 80-86, 90, 96	Irrigation
30-35, 73	Commercial
40-42, 46-48, 50-52	Industrial
43, 44, 62-64	Mining
10-16, 53	Public supply
45	Sewage treatment

mation Center (LMIC), part of Minnesota's State Planning Agency, helped design the data-base structure and programs needed. Menu-driven programs provide users access to water-use data for approximately 11,500 withdrawal sites currently (1988) in the system.

Data storage was not the only purpose for Minnesota's water-use program. Water-use managers needed a system that would identify users in a specific location in the State, describe the source of water (such as well, lake, or stream), the type of use, and the rate of use. Summaries of water use within a county, a watershed, or statewide were needed for input into the Survey's Aggregated Water-Use Data System (AWUDS).

One of the first accomplishments of MWUDS automation was the development of data-aggregation programs (Land Management Information Center, 1984; Trotta, 1988), which are useful for calculating county and watershed water-use summaries. After verification, this aggregate information can be automatically transferred to the Survey's Aggregated Water-Use Data System (AWUDS). These data summaries are useful for analyzing regional water use trends.

Each of the water-use categories corresponds to a group of related Standard Industrial Classification (SIC) codes (Snavely, 1986, table 5). SIC codes define and classify the entire field of economic activities (Office of Management and Budget, 1972). These SIC codes are included on tax forms used by businesses and on surveys made by the U.S. Department of Commerce, Bureau of the Census. Individuals or facilities are thus categorized on the basis of the purpose of their water use in relation to their economic activity. Published surveys of use ratios, such as that by Debra Snavely (1986, table 7A), when combined with economic reports, such as those published by the Minnesota Department of Jobs and Training (1986), allow economic-policy-analysis simulation models to estimate water-use amounts for each SIC code (Young, 1987, p. 5-9). These estimates are used when actual data are not available.

Combining information in this way from several data bases depends on a linkage or relation between them. When one data base has an item in common format with another, they are termed "RELATED."

Sometimes stand-alone data bases, such as Well Log Listing System (WELLS) or Permit Information Exchange (PIX), are related to MWUDS see figure E-25). Well Log Listing System contains a brief description (location, aquifer and well-construction data) for approximately 100,000 well logs. Locational search routines are available to verify or transfer aquifer data about a particular user's source of water. Menu-driven systems recently have been developed. Well Log Listing System also has a routine to load data to a geographic information system. Permit Information Exchange contains detailed information from well permits on fees and legal actions for the regional water-data network. The permit-handling software in PIX does the regular name and address maintenance on permit applications. Well Log Listing and PIX are linked with either a common geographic-location code or reference numbers that help identify the same water user in different data bases (Horn, 1986, p. 11). The program

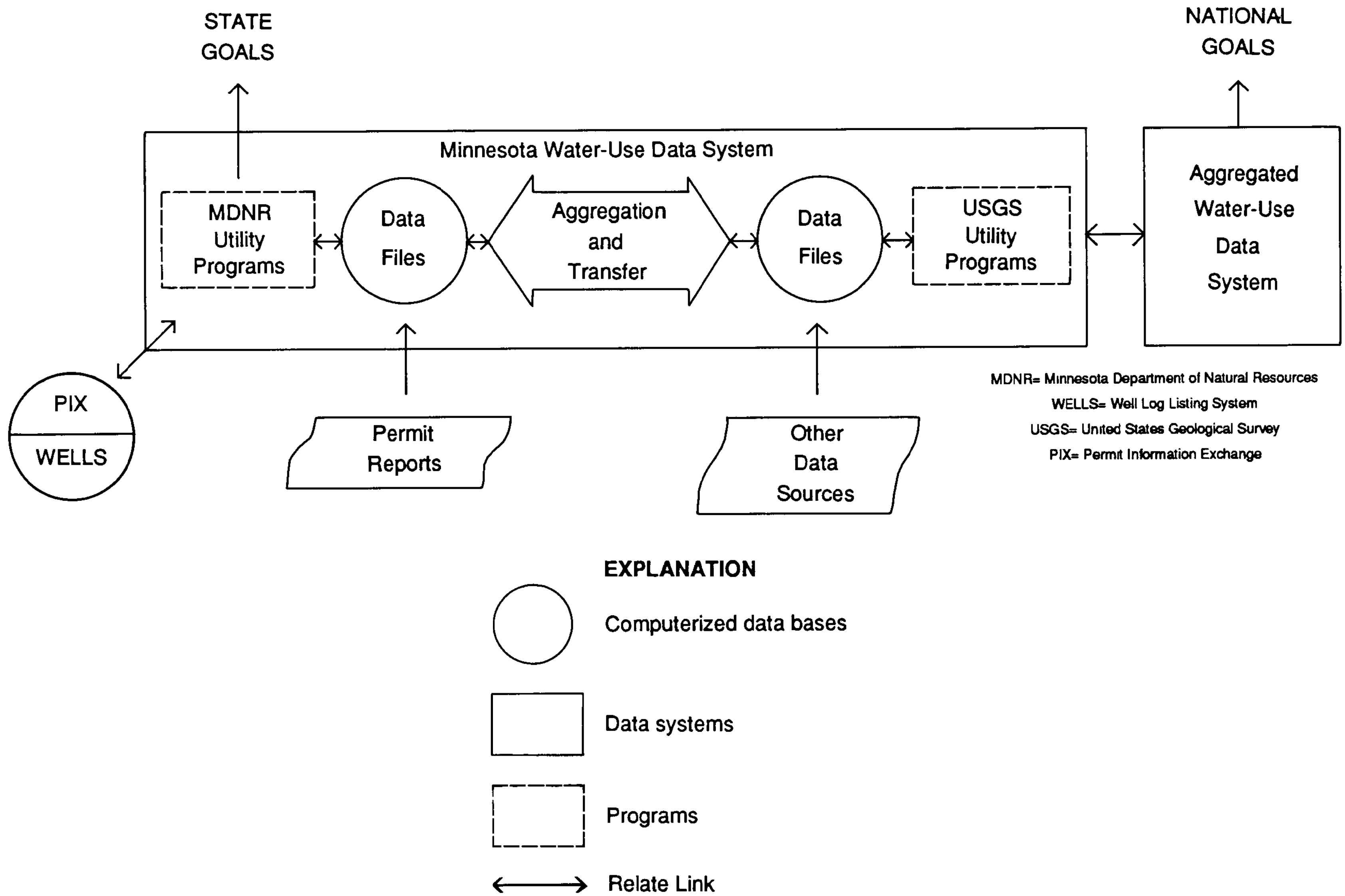


Figure E25.--Computer methods used in the Minnesota Water-Use Data System.

PERMSRCH allows easy access to both historic and active data in a flexible and user-friendly way. A partially menu-driven system called MAILLIST uses information from PIX and MWUDS to produce written communications and forms for reporting pumpage data (Richard Gelbmann, Minnesota Department of Natural Resources, written commun., 1988). Computer linkage to other data bases is described by Horn (1986, p. 18-25).

Utility programs have been developed by MDNR to provide data for many specific applications. Some of the applications are to (1) aid in tracking permits once they have been issued, (2) ensure compliance with regulations or restrictions, and (3) assist the State's education functions with illustrative reports. The MDNR program INDEX, for example, is useful for producing reports that summarize 3 years of site-specific data for any MDNR administrative region (Richard Gelbmann, Minnesota Department of Natural Resources, oral commun., 1988). Search routines in MWUDS can quickly outline, for example, areas where water-level declines may be related to municipal lawn sprinkling or irrigation. The MDNR also has developed input programs for an economic model that determines the economic value of Minnesota water for various water-use scenarios. The economic value of conservation initiatives and reprioritization of use is discussed in a report that summarizes resultsof this model (Adelsman and Bloomgren, 1987, p. 19-22).

U.S. Geological Survey utility programs have been developed that provide information for trend analysis. The AWUDS software, distributed by the National Water-Use Information Program, accepts aggregated data from MWUDS and produces graphical output, pie diagrams, and choropleth maps of the State, depicting county or watershed data. The program QA2, part of the AWUDS, creates a table useful for comparison of water-use data for 1980 (or another reliable base year) and any other data year available (Howard A. Perlman, U.S. Geological Survey, written commun., 1987). Additional programs were developed in the Minnesota District of the U.S. Geological Survey to extract data needed by the national program from specialized data files associated with MWUDS. These programs calculate municipal population by watershed, sewage-treatment facilities by watershed, water withdrawals for livestock by watershed, and diversion of water across watershed boundaries. Predictions of the effect of diversions or changes in pumpage may be made within a short time frame.

SUMMARY

Automation improves access to data and efficiency of data-analysis capabilities. Careful documentation of procedures used in data-search and reporting commonly points out inefficient manual steps correctable by automation. Many management decisions for mandating restrictions, initiating ground-water studies, and developing local comprehensive water plans are streamlined by automation of the water-use data base.

REFERENCES CITED

Adelsman, H.R. and Bloomgren, P.A., 1987, Summary--The economic value of water: Department of Natural Resources, Division of Waters Report to the Legislative Commission on Minnesota Resources, St. Paul, 22 p.

Horn, M.A., 1984, Annual ground-water use in the Twin Cities Metropolitan Area, Minnesota, 1970-79: U.S. Geological Survey Open-File Report 84-577, 130 p.

____ 1986, Development of a water-use data system in Minnesota: U.S. Geological Survey Water-Resources Investigations Report 85-4306, 59 p.

Minnesota Department of Jobs and Training, 1986, Average covered employment and wages by economic region for calendar year 1985 by two digits of industry: St. Paul, 45 p.

Minnesota Land Management Information Center, 1984, State water use data system (SWUDS) system documentation: State Planning Agency, St. Paul, 42 p.

Office of Management and Budget, 1972, Standard industrial classification manual: Executive Office of the President-Office of Management and Budget, Statistical Policy Division, Washington, D.C., 649 p.

Snavely, D.S., 1986, Water-use data-collection programs and regional data base of the Great Lakes-St. Lawrence River Basin States and Provinces: U.S. Geological Survey Open-File Report 86-546, 204 p.

Trotta, L.C. [1988a] Sources of water-use data in Minnesota: U.S. Geological Survey Open-File Report 87-544, 2 p.

____[1988b] Aggregation of Minnesota water-use data and transfer of data to the National Water-Use Data System-Procedures and programs: U.S. Geological Survey Open-File Report 87-40, 16 p.

Young, P. G., 1987, An analysis of water use in Minnesota: Department of Natural Resources, Division of Waters Report to the Legislative Commission on Minnesota Resources, St. Paul, 49 p.

DEVELOPMENT OF A DATA BASE TO ACCOMMODATE MANAGEMENT OF WATER-RESOURCES DATA WITHIN A GEOGRAPHIC INFORMATION SYSTEM (GIS)

--

By Douglas D. Nebert and Joel Frisch

--

ABSTRACT

The U.S. Geological Survey is working with the U.S. Army Corps of Engineers Engineer Topographic Laboratories' Terrain Analysis Center (TAC) to establish a water-resources data base for areas outside of the United States and to integrate the data base with geographic information systems (GIS) technology. TAC wants to automate production of its standard cartographic products and build an extensive attribute data base by use of an integrated spatial data-base design. Such a design would improve the relation between the previously distinct cartographic and tabular data so that production and update of overlays and related tabular data may be conducted more efficiently. Managing the data within a GIS would permit water-resources data-base users to define and generate products in a variety of map scales, formats, and media with minimal lead time, and would provide analysts the opportunity to query the coordinated data base to support spatial and tabular retrievals.

INTRODUCTION

The U.S. Geological Survey is working with the Terrain Analysis Center (TAC) of the U.S. Army Corps of Engineers to design and implement a relational data model for the Department of Defense Water Resources Data Base (WRDB). The automated spatial data base will be used by TAC to generate cartographic and tabular products for use by all military forces. The U.S. Geological Survey's efforts since 1983 have included the preparation of a user-needs assessment, preliminary data-base design using the Oracle data-base management system, hydrologic analyses and data preparation, data-input form coding, retrieval testing, and product generation for test data associated with water resources and facilities in the Middle East.

The data base was originally intended to contain only text attributes for manually-produced map overlays keyed to Department of Defense 1:250,000 scale maps. These overlays portray the location, quantity, and quality of surface-water features (such as canals, streams, and reservoirs), ground-water features (such as wells, springs, and qanats), areas of ground-water development potential, and water- related facilities such as pumping stations, desalination plants, and pipelines. These overlays are used by military planners, engineers, logisticians, and TAC teams to identify available or potential water resources in an area for treatment, storage, and distribution.

As a distinct and separate entity from the cartographic manuscript, the tabular data base in Oracle has a restricted ability to handle spatial queries and makes it difficult to keep both the cartographic manuscript and data base current. Modification to either the map manuscript or the tabular data base requires that each be manually updated in a timely fashion. Having both cartographic and tabular data on one medium would eliminate duplication of data-management responsibilities and reduce confusion in update and analysis. For this reason, a data base containing spatial- and tabular-data elements was determined by the Survey-TAC team to be the most effective solution to resolve coordination problems in production.

A user-requirements study was conducted by U.S. Geological Survey and TAC personnel to evaluate the need for, and the uses of, water- resources data by the Army and other Department of Defense users. As a result of that study, a report by Frisch and Stephens (1985) identified the need for a computerized water-resources data base to provide source data for thematic-overlay production and facilitate interactive retrievals of tabular data to support user queries. The report indicated that consideration be given to using computerized GIS technology to produce overlays and manage tabular data. It also indicated the need for, as the next step, the preparation of a prototype or pilot project to include development of a sample data base and analytical reports, specification of a functional hardware/software environment in which such work could be done routinely, and a test of the products' use to analysts, planners, and other users.

Pilot Study: Appraisal of a Geographic Information System in Overlay Automation

A pilot study was initiated by the U.S. Geological Survey late in 1986 to assess feasibility of TAC to implement an integrated spatial and tabular data base using GIS software and techniques. Three topical overlays (ground water, surface water, and water-resource facilities) to the Al Kuwayt, Kuwait 1:250,000-scale map were prepared using the ARC/INFO software package. High-resolution overlays were produced in an approximation of the official TAC format specifications. A workplan was prepared by the U.S. Geological Survey to define incorporation of the GIS data base into the TAC program over the next several years. Based on an evaluation of the pilot study and investigations by the U.S. Army Corps of Engineers Topographic Developments Laboratory, TAC elected to proceed with GIS technology and acquired a copy of ARC/INFO for use in the preparation of map overlays for hydrologic and terrain analysis.

Implementation Phase

The implementation phase in the WRDB development, conducted by the U.S. Geological Survey includes the following work elements that are scheduled for completion in late 1989:

(1) Install and test GIS software and provide on-site training in the use of the software for the data-base implementation activities;

(2) Modify the existing data structure with respect to the inclusion and management of spatial water-resources features and associated attributes;

(3) Design and implement input, update, and product-generation modules for use with the data base;

(4) Design and implement a tracking system of map-overlay status for the WRDB;

(5) Provide guidance in the conversion of existing spatial and tabular-data manuscripts to the new data-base format;

(6) Conduct a full-system test of the prototype; and

(7) Prepare a user manual to document all aspects of data management in the system.

Elements 1 through 3 have been completed. Only element 2 will be discussed in detail in this report.

Purpose and Scope

The purpose of this report is to document the development to-date of the WRDB data model and to emphasize the utility of spatial (or GIS) approach to data-base management over tabular systems now predominately used. Requirements of TAC will be reviewed, and the spatial data-base design techniques used to meet these requirements are described. An overview of the resulting data model and its operating environment in a worldwide context is presented. Special attention is given to the consideration of spatial features as primary elements of the overall data model to which associated attribute data are linked.

PRIMARY ELEMENTS OF A WATER-RESOURCES DATA BASE

The water-resources data bases used by most agencies share some common characteristics that are inherent to the collection and storage of water-resources data collected at discrete locations. Although many data bases do not explicitly manage their data in a GIS data structure, the organization of the water data commonly includes both the observations and some type of spatial reference for each sampling site, typically stored as point locations.

Traditional Features

Traditional efforts of water-resources data-bases design have treated water features as unique point features, because agencies tend to manage data collected at specific geographic locations. The creation of a single site file that contains non-repeating, basic site information and one or more measurement files that contain repeating records of observations is a common theme in water-resources data bases. Measurement or observation data can be associated to the site file and its characteristics by a common key -- commonly a site identifier -- in both files. This hypothetical file structure is illustrated in figure E26. In this traditional example, the location of the site is stored in the site file, and occurs only once per station. Once linked to the site file, measurement data can be associated with that point in space.

Maintaining water-resources features as unique and separate entities allows standard retrievals to be made based on the attributes in either the site file or the measurement file. Lacking is the capacity to identify measurement sites on a stream network or within an aquifer boundary without explicitly encoding site codes into the site file. Without this important element, the capacity to identify or associate sites that meet common spatial situations is limited to data in the site file. However, identifying all possible common spatial situations is a lengthy and time-consuming task when done manually. In addition, the inclusion of too many elements in the site file can degrade retrieval performance and can require excessive amounts of data storage. The use of GIS techniques to populate the site file -- either statistically or dynamically -- can save substantial time in encoding site characteristics.

Site file

site identifier	station type	latitude	longitude
4201	gage	032345	0171856

Measurement file

site identifier	parameter	units	value	date	time
4201	temperature	degrees Celsius	24.3	830713	1430
4201	pH		7.2	830713	1430
4201	TDS	mg/L	300	830714	0930
4300	temperature	degrees Celsius	9.7	870203	1145

Figure E26 --Relational data diagram showing site and measurement files related by a common key item, "site identifier." Arrows illustrate the one-to-many relation between a site and its measurements. Parameter type is stored in the column "parameter" and units of measurement are stored under "units." TDS = total dissolved solids.

Benefits of Geographic Information Systems to Water-Resources Data Management

The use of a GIS as a spatial data-management system allows maintenance of traditional data files and logical associations between water features. This principle was recognized in the conceptual design of the future U.S. Geological Survey's data base, the National Water Information System (NWIS), version 90.1 (Edwards and others, 1987). Traditional water-resources features can be considered as elements of an infinite number of possible spatial sets; any of these can potentially be used to satisfy a query. Assignment of additional geographic characteristics to spatial entities can be accomplished through the use of GIS concepts of overlay (integration), adjacency, and connectivity with one or more spatial-data themes. For example, by overlaying or integrating measurement points with country boundaries, these points can be automatically considered members of different country sets and thereby be assigned the identity of the country they are in. Rather than manually encoding each feature with the name of the country it lies within, the GIS can be used to assign points with country identities, as needed. There can be as many potential sets as there are polygon or surface data sets by which point data could be classified--elevation, ground-water units, slope, and soil association are a few examples.

Adjacency can be used to associate features with other nearby features, and to provide new opportunities for retrieval classification. Point-in-polygon overlay is useful to assign polygon values to a point, but is limited to the literal position of that point with respect to the areas. Measurement sites can be flagged with the identity of the nearest stream segment or the nearest upstream gage, and the attributes of those related features can be accessible to the user for analysis or retrieval. Another use for the proximity calculation would be to identify all facilities near a given pipeline.

The connectivity of the various conveyances to one another should be clearly made, where known, to enhance the query of features along surface-water or conveyance networks. The Cartesian coordinates of a gaging station indicates little about its position within the stream network. Attributes such as direction (one- or bi-directional), the immediate upstream and downstream segments, and pointers to related features (for example, facilities such as pumping stations or wells) should be maintained where they can be determined. Such a connectivity or network query would be to request all attributes of wells associated with a given pipeline. To satisfy the query, all elements or segments of the pipeline in question would first have to be identified before all wells associated with these segments could be identified. The resulting list of wells then could be used in further analysis or to generate tabular printouts.

Other benefits of integrating cartographic and tabular data in a GIS include the ability to generate map products at a variety of scales, map projections, and with different degrees of generalization on a variety of map media. The management of all cartographic data on-line improves the turn-around time to produce a standard product.

Integrated management of the water-resources and cartographic data facilitates the production of non-standard cartographic and tabular products to satisfy a user query. Water features can be selected on the basis of their attributes in the data base (for example, depth greater than 100 meters, water use classified as agricultural). Data from other spatial data bases also can be used with water data within the GIS for specialized hydrologic analysis.

EVALUATION OF EXISTING DATA ELEMENTS AS SPATIAL FEATURES

An early task in the implementation phase was to characterize the known WRDB features as legitimate spatial features within the ARC/INFO data model. At present, these features include points, lines, and areas, but future versions of the GIS software also probably will allow nodes to have attributes, and will allow all feature types to be kept within one collection of data files (known as a coverage in ARC/INFO).

Data-Management Issues

The next step was to define the many potential associations among the features that may be used in standard queries of the data base. Several management issues needed to be resolved regarding the distinction between points and areas at a given scale, how node attributes could be managed in the current version of the software, and how areas of numerous facilities were to be managed.

Given that the scale of the original cartographic product is 1:250,000, the distinction between points and areas may become blurred. The size of a plotted point symbol on the final overlay could be equal to as much as 0.8 miles on the ground. Where a single coordinate for the feature is known, the feature is stored and generated as a point. Where a feature is determined to have area, multiple coordinates are stored in the data base and generated as a polygon.

Point features on streams, such as gaging stations, that should be managed as attributed nodes are managed at present as a "link" coverage--a digital data set of line and point data. When node attributes are supported by the GIS software at a future date, the data model can be modified and the nodes assume point attributes.

Areas that contain numerous water features too dense to depict as separate features on the overlay are displayed with text annotation on final map products that states "numerous canals" or other feature type. Data are not eliminated from the data base due to crowding at a given scale because most water-feature location information is scale independent. This allows for design of thematic map products at a wide range of map scales. The GIS potentially could suppress the display of some features based on feature density.

Functional Organization of Water Resources Data Base Spatial Features

A functional delineation of the WRDB features is shown in table E9. The majority of the spatial features to be managed by the data base are points, although several line features also exist. By logically grouping features on the basis of their function and feature type, the ability to retrieve information across features is enhanced. This can be considered a type of functional and spatial normalization.

Comparison of potential associations between the features in table 1 indicate that most groups of water features can be associated with any other one, by one characteristic or another. Wells can be associated with pipelines or storage tanks, canals with other canals. A retrieval option desired by TAC is the ability to select all water features related to a water feature of interest.

Table E9.--Department of Defense Water Resources Data Base features organized by functional similarity with spatial feature types

Feature grouping	Feature class	Spatial feature type
Processing facility	Desalination facility	point
	Waste-treatment facility	point
	Other purificationpoint facility	point
Storage facility	Water tank	point
	Water tower	point
	Small reservoir	point
	Cistern	point
	Surface-water body	area
Other facilities	Pumping facility	point
	Miscellaneous source	point
Conveyances	Stream (river)	line
	Canal	line
	Qanat	line
	Pipeline	line
Ground-water features	Well	point
	Well field	area, point
	Spring	point
Other symbols	Dam	line symbol
	Graphic leaders	line symbol
	Surface-water access reach	line symbol

Definition of Data Structure

Once spatial similarities were determined, the existing attributes were reviewed for similarity. Not surprisingly, many similar water features had similar, if not identical, data elements associated with them. Where functional features and attributes were identical, they were stored in the same data files.

Every water-resources feature in the WRDB has a basic set of attributes (much like the site file shown in fig. E26) stored in the identification file. This file contains unique reference information for the feature, such as the feature class (water tower, pipeline, and so forth.), official and common names, and a series of identification numbers. Every feature in the WRDB has a unique number, called REC, that can be used as one of the keys to associate with various data files.

Geographic location, as maintained in the WRDB, is in a file separate from the ID file because of the potential of having more than one reported coordinate and positional accuracy for the given feature. For example, imagery analysis can determine a general location for a feature, whereas subsequent field checking can yield a different location. If more than one location is identified for a feature, the professional judgement of the analyst is used to decide the coordinates used to actually generate the feature. Multiple coordinate references for a feature are managed as attributes of the feature.

Other files exist to store information about the source documents used, update transactions, and quality-assurance testing information. A security level also is stored with each record of each file in the WRDB to assist in producing cartographic and tabular products for audiences with different security classifications.

Diagramming the Normalized Data Structure

A simplified view of data-file organization is shown in figure E27. All files shown contain the common item REC and, therefore, can be directly associated with one another. The ID file is shown at the center to signify the central or unique data elements it contains. This file relates one-to-one (1:1) or one-to-many (1:M) with the feature-attribute tables across the top. Abbreviations and symbology are given in the figure explanation.

Each feature class (for example, water tower, pipeline, or spring) is described by the unique association of a set of related data files. All features have at least two files in common: the identification file (ID) and the remarks file (REMARKS). Additional files are used to carry other attributes in normal form.

Examples of the basic relations required for two different water features are given in figure E28. The top figure, E28A, shows most of the files needed to define a pipeline in the data base. Because a pipeline is a linear feature, its component arc (or arcs) is referenced in an arc attribute table and stored as internal ARC coordinates. Because a pipeline can be composed of several arcs, a 1:M relation is diagrammed. Multiple text remarks are stored in the REMARKS file.

There is a 1:1 correspondence between the ID file and the pipeline characteristics file (PIPELINE), whereas a potential 1:M association exists between ID file and the MEASUREMENT and RELATED FEATURES files. The RELATED FEATURES file is used to build cross references between any two associated WRDB features. This file contains the record number of the current feature and all related features, such as pumping stations or wells, with one record per association.

A well is diagrammed in figure E28B. This well, as a feature not associated with any other feature, has no entry in the RELATED FEATURES file, but is associated with WELL data. As a point feature, it has an entry in the point attribute table (PAT). Standard associations can be made between the ID and the REMARKS files.

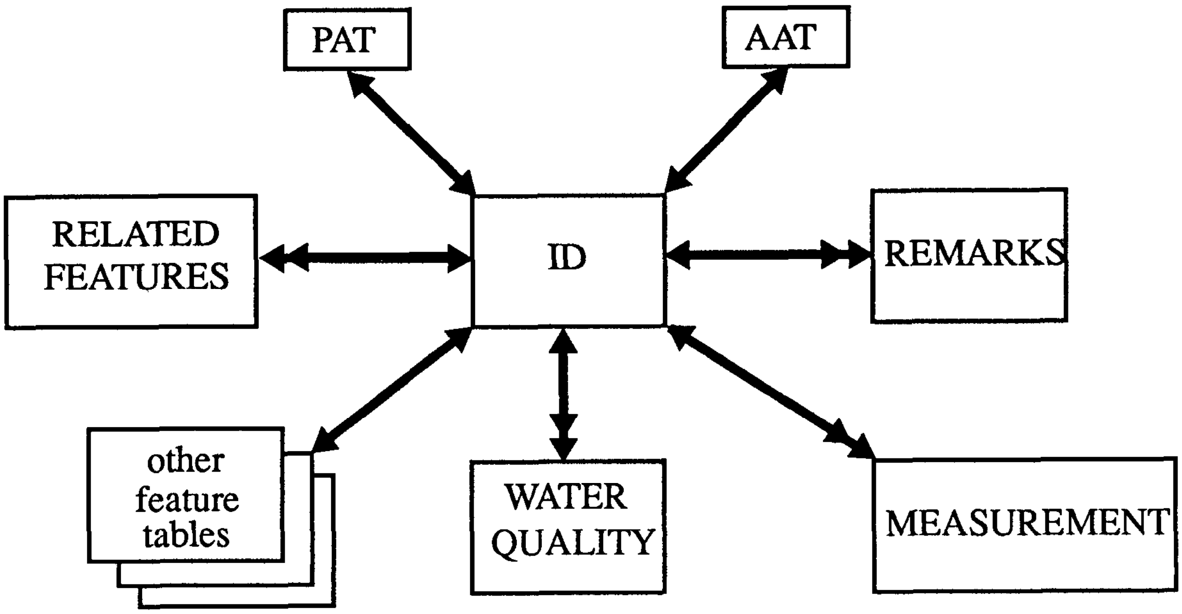
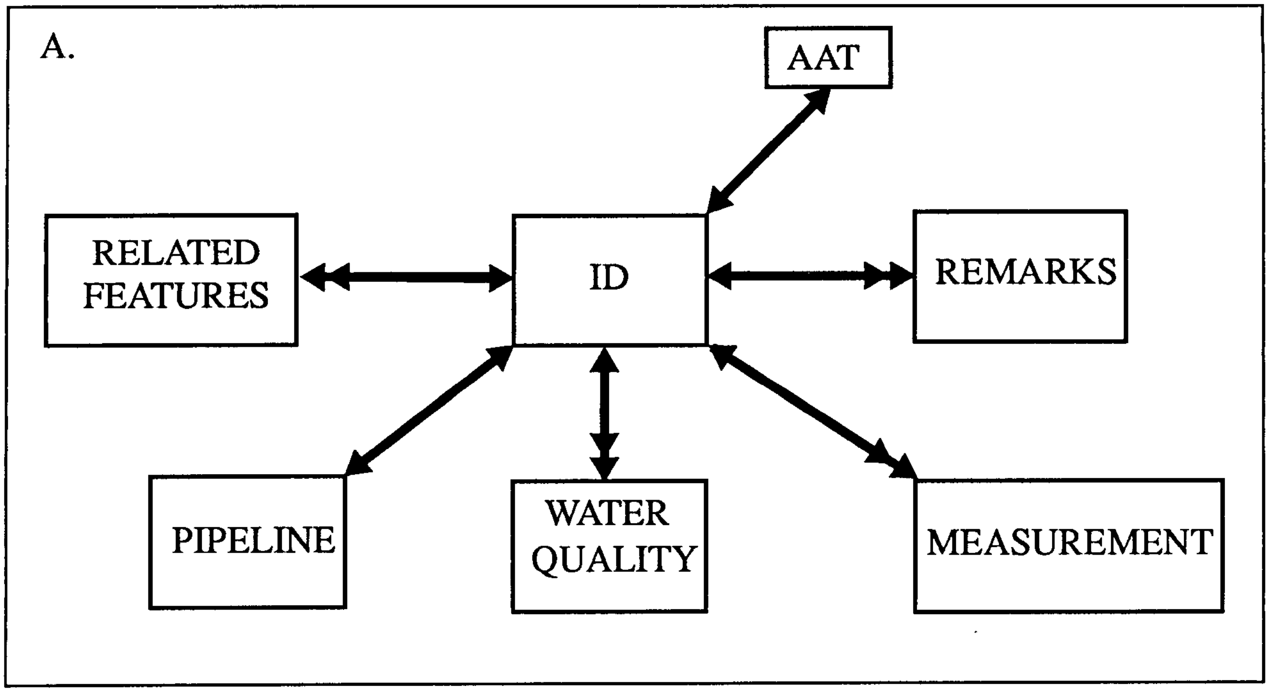
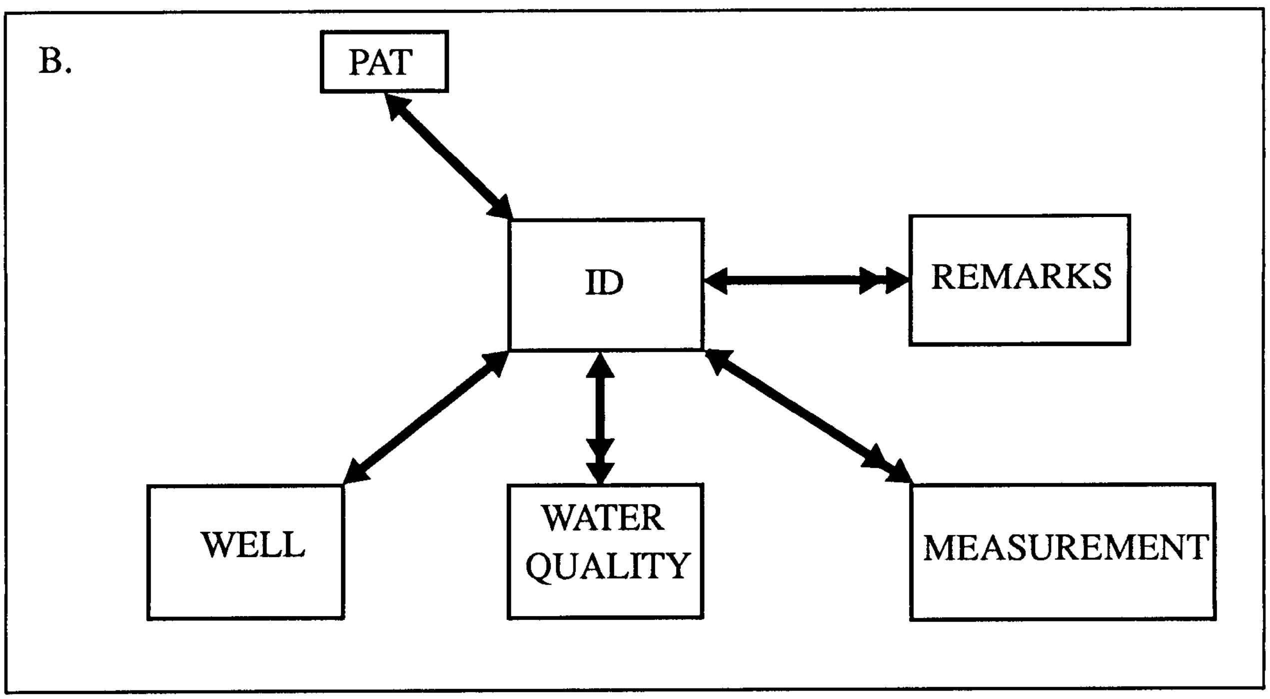
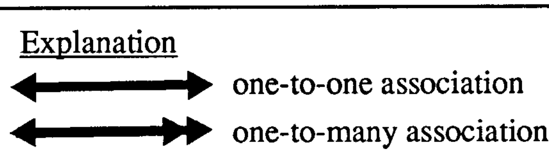


Figure E27 --Generalized diagram of data file organization and dependencies. Each file is represented by a box, arrows signify potential one-to-one and one-to-many relations. PAT = Point or Polygon Attribute Table, AAT = Arc Attribute Table. "Other feature tables" are tables for data unique to a given feature type such as a well. ID is the identity file that stores basic characteristics for all feature types.







AAT is Arc Attribute Table
PAT is Point Attribute Table
ID is primary identification table

Figure E28 --Diagram of data file linkages for two water feature types: Figure E28A is for a pipeline feature; Figure E28B is for a well. Boxes represent data files; arrows represent file association by record.

GLOBAL DATA-BASE CONSIDERATIONS

This report has described the management of water-resources features on a single set of map sheet overlays. The management of a global data set requires the consideration of some additional organizational details. The first of these is the definition of the physical and logical organization of the data base on the computer system at TAC. As all WRDB data are compiled and stored with reference to 1:250,000- or 1:50,000-scale maps, the map boundaries define a logical data organization or map management (tiling) system for the WRDB.

To search the data base, users must know what maps lie within what countries in order to retrieve data in an area of interest. To prepare this cross reference, an unpublished coordinate file of map corners was acquired from Defense Mapping Agency. Coastline and political-boundary data were acquired from the Central Intelligence Agency's World Data Bank II at an approximate scale of 1:3,000,000 (Gorny, 1977). Both data types were converted into ARC/INFO coverages and the two were combined to produce an integrated global coverage that maintains a country code and a quadrangle number for each resulting map/country polygon. From the polygon attribute table, one can identify all maps that fall within each country and the percentage of each map within each country. Using the interactive GIS query functions, one can also identify all maps or countries within a circle, rectangle, or irregular polygon. This important basic capability is useful for browsing the data base. Each quadrangle-based workspace contains one full set of coverages and all associated data files.

There are several advantages to storing the data base in map-based tiles. By storing each map-based tile separately, continuity with the original map manuscript is preserved--including its original map projection. Entire workspaces can be copied to tape or to an analyst's work area without affecting the overall data base. Graphic selection of feature data can be done across tile boundaries, eliminating the need to append adjacent maps before analysis. Finally, the integrity of the standard map-based products is maintained. For analysis that requires several maps to be queried, the contents of adjacent tiles can be joined to produce new products.

Input, update, and plot generation operations are conducted by tile and are monitored via the map overlay status tracking system. The use of the Map Librarian package with the tracking system can facilitate the on-line management of hundreds of tiles planned for automation within the next few years.

CONCLUSIONS

The management of associated water-resources features is a logical task for spatial data-management functions available within state-of- the-art GIS software. The management of the extensive attribute data base can be done adequately within the relational data-base management systems (DBMS) used by GIS packages. As individual data files increase in size, the performance of these DBMS could be enhanced through the use of indexed data structures or through the use of intermediate key files that contain common retrieval items. Given the relatively small size of the tile-based data sets, these techniques have not been implemented.

The use of a coordinated spatial and tabular data base within a GIS processing environment is an ideal means of supporting the data requirements of TAC and its user community. The improvement of timeliness of data products, the versatility in producing them, and the added sophistication in spatial analysis are three areas where the use of GIS techniques should excel over manual methods of production.

As with the U.S. Geological Survey's NWIS 90.1 data-base design, TAC has recognized a need to manage its water-resources data in a single, integrated data model. Unlike the initial versions of NWIS design, TAC's WRDB redesign reflects the need to manage water features as primarily spatial entities, forming the basis of the data model. Because the objectives of the system are to analyze and manage the water-resources data in a specific geographic area, the GIS approach should improve access time, speed the production of map-based and tabular products, and enhance on-line analysis of water-resources data with respect to other digital spatial information (for example, geology, elevation, base thematic data from the Defense Mapping Agency).

SUMMARY

The U.S. Geological Survey and the U.S. Army Corps of Engineers are designing a water-resources data base that integrates extensive tabular data with geographic or mapped data. These data are used with a geographic information system (GIS) to produce standard cartographic products and to provide an environment for query and management of water-resources data.

A user-needs assessment was conducted to evaluate the need for and uses of water-resources data by Department of Defense users. GIS technology is the preferred means of managing the cartographic and tabular data defined in the user-requirements study. A pilot study was conducted on a 1:25,000-scale map of Kuwait to demonstrate the ability of a GIS to manage both spatial and tabular data in a coordinated manner.

Water-resources features defined in the user-requirements study were characterized as spatial features (points, lines, areas), and logical associations between features were made to define the appropriate data structure. Attribute data were normalized with respect to the water-resources features and other attributes to produce an optimal relational data base with most tables containing a common identification data item (REC). This common data item allows virtually any files to be related to support a tabular or graphic query in the GIS.

The GIS approach to water-resources data management will facilitate production of standard maps and tables and interactive analysis of the data with respect to other spacial data; this provides flexible user- views into the integrated data base.

REFERENCES CITED

Edwards, M.D., Putnam, A.L., and Hutchison, N.E., 1987, Conceptual design for the National Water Information System: U.S. Geological Survey Bulletin 1792, 22 p.

Frisch, Joel, and Stephens, Jerry C., 1985, User requirements analysis for the Worldwide Water Resources Data Base: U.S. Army Corps of Engineers, ETL Report ETL-TAC MISC 01-85, 50 p., plus appendixes.

Gorny, Alexander J., 1977, World data bank II, Volume 1 -- North America, and general user guide, Central Intelligence Agency Report CIA/DF-77/001a.

THE USE OF OPTICAL MEDIUM AS A MEANS FOR STORAGE OF IMAGE AND DIGITAL DATA

By Brenda L. Groskinsky and Richard A. Hollway

ABSTRACT

Storage of image and digital data, either for initial collection or after publication, is a requirement of all projects within the U.S. Geological Survey. Optical storage devices can be used as a means for archival storage and also as a means for temporary routine storage of image and digital data. Recent advances in optical storage devices indicate the need for applied research and evaluation of prototypes to assess the viability of optical medium for archival and temporary routine storage of image and digital data.

The major difference between the concepts of archival and routine backup of data occurs in the intent. An archive of data is a method to store historical record for periods greater than 10 years. The concept of backup, however, was devised as a precautionary measure in the event of data loss with information being stored for a period of less than 3 years. There are several physical alternatives for each concept. The objective of this research is to determine the advantages and disadvantages of optical or write-once-read-many drives as a method of image and digital data storage and retrieval when used as both an archival and a backup device.

The assessment of write-once-read-many drives has shown that the 32-bit Unix workstation allows integration of nearly all of the currently (1988) available write-once-read-many drives. At this time there is a limited amount of data base management system software that will store and retrieve both text and image data as well as interface directly to optical devices. Research concerning image scanners as input devices to the data management systems indicates Small Computer System Inteface (SCSI) scanners to be the most acceptable. The functions of a stand-alone optical storage and retrieval system would be isolated from the internal operations of the networked office environment commonly found throughout the U. S. Geological Survey. Recent advances in optical technology imply the adoption of a uniform write-once-read-many formatting standard, but possible interfacing difficulties remain between computing platforms. Because optical disk drive technology is still evolving, the longevity of a write-once-read-many disk could easily extend beyond the life-span of the optical drive.

Optical media have the ideal potential to provide a longer storage life for the archival of data than either paper or microfiche. Optical media also will work well as a means for temporary routine backup storage because of its high volume storage capacity. Therefore, future research of optical storage technology by the U.S. Geological Survey is needed.

CHAPTER F--ELECTRONIC PUBLISHING/MANUSCRIPT PREPARATION

AUTOMATING PROCEDURES FOR ANNUAL WATER DATA REPORT PREPARATION

By Mark L. Farmer and Jim E. Monical

U.S. Geological Survey

ABSTRACT

New procedures for automating many of the processes involved in the annual State water data reports of the U.S. Geological Survey have been developed by the Arkansas District of the U.S. Geological Survey. The automated process is based on the same steps as in manual compilation to establish a sense of continuity in preparation of the report. A directory structure was established to centralize the storage of all data files and software applications involved in the preparation of the data report. All software for the system were designed to be as modular and portable as possible.

Post-processing routines in Fortran 77 and Prime Command Procedure Language were developed to reformat data tables and retrieve from National Water Information System data bases as well as collating tables and text data. Macros have been developed using a full screen text editor for merging and sorting heading and table data.

INTRODUCTION

One of the primary functions of the U.S. Geological Survey is to collect data on the quantity and quality of surface and ground water and publish these data accurately and in a timely manner. Sophisticated methods are continually being developed for data collection and for recording and storing data with the use of high-precision recorders and high-speed computers. Although efficient hydrologic data collection and recording methods are essential, the need for efficient publication of these data also exists. The reporting of the data can be enhanced by automating some of the manual methods currently in use. The automation process incorporates the use of high-speed computers, retrieval and word-processing software, and laser printing technology.

District offices of the Geological Survey currently store most hydrologic data in the data bases of the National Water Information System (NWIS) of the U.S. Geological Survey. These data bases include the Automated Data Processing System (ADAPS), Ground Water Site Inventory (GWSI), Quality of Water (QW), and the New Site Specific Water Use Data System (NEWSWUDS). These data bases are supported through a network of Prime minicomputers, which uses Fortran 77 as the principal programming language for storage and retrieval applications.

Purpose and Scope

The purpose of this paper is to discuss the report preparation procedures developed in the Arkansas District of the Geological Survey, focusing on the preparation of the annual water data report. Also discussed are post-processing routines, written in Fortran 77, Prime's Command Procedure Language (CPL), and EMACS, Prime's full-screen text editor. These routines are used to reformat data tables retrieved from the NWIS data bases as well as collating table and text data for use in the annual water data report.

SOFTWARE DEVELOPMENT OBJECTIVES

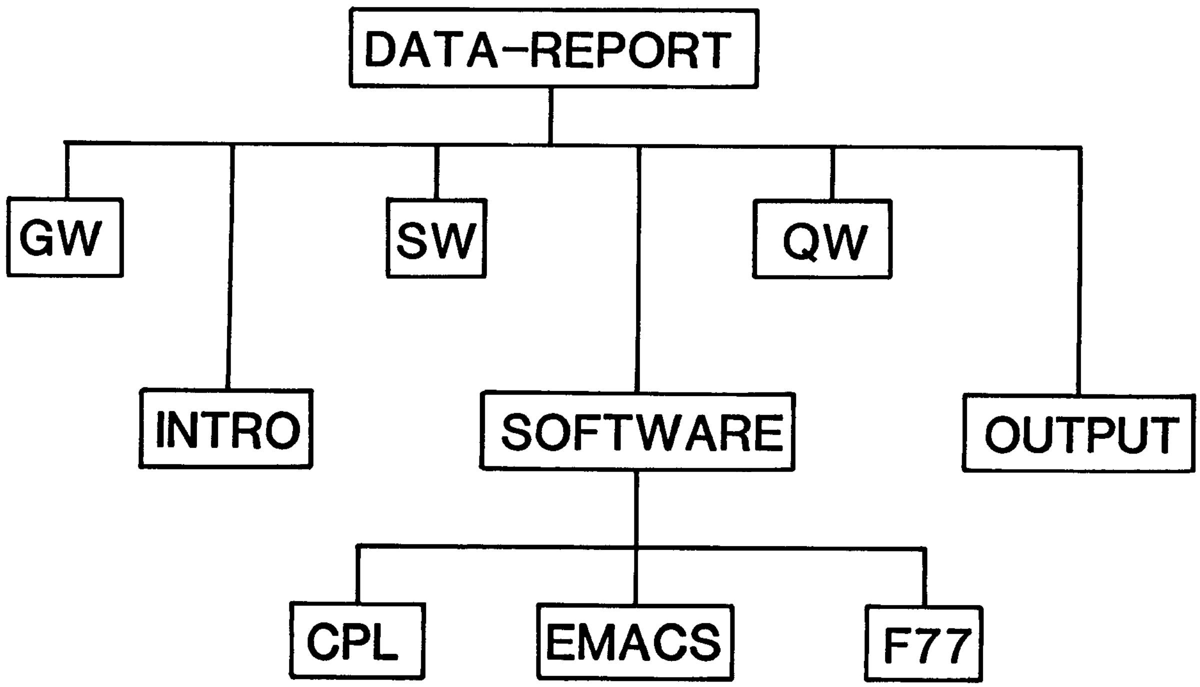
The steps used in the manual compilation methods of report preparation were used as a guide for the development of the automated system. Designing the system in this manner would allow a smooth transition for personnel adapting to the automated system, provide a guideline for software development, and would use experienced personnel who have worked with and developed the manual compilation system.

The first of two objectives, therefore, was to centralize the storage of all data files and software applications involved in the preparation of the data report into a top-level, user-file directory (UFD) on the Prime computer named DATA_REPORT. This UFD would contain six subdirectories that would group data and software files by subject (fig. F1). The SW, QW, and GW subdirectories contain all manuscript and tabular data for the surface-water, quality-of-water, and ground-water disciplines, respectively. The SOFTWARE subdirectory contains all of the application programs for the automated system, whereas the INTRO subdirectory contains all of the introductory text material for the annual report. The OUTPUT subdirectory is a repository for the final report after all text, station description, and table data have been compiled and merged. The centralization of all data files and software into one user-file directory allows control over the access to the data; this prevents accidental corruption or deletion of important files containing data or software.

The second objective was to ensure that all software developed for the automated system would be as modular and portable as possible. All of the software and file structures used in the automated system should be easily portable from one Primos partition to another and be transportable to other Prime systems. The only requirement for transferring software and data files to other systems is that the receiving systems contain the user-file directory DATA_REPORT, so that software references to that directory will successfully access the residing data files.

AUTOMATED WATER DATA REPORT PROCESS

The first step in the automated data report process is to update the introductory text using one of the Prime text editors or a word processing system software package. The textual data includes the introduction, summary, and formulas that are traditionally found at the beginning of the report. These files are placed in the INTRO subdirectory and, once established, require few changes from year to year.



EXPLANATION

GW - Ground water data tables.

SW - Surface water data tables.

QW - Quality of water data tables.

SOFTWARE - Contains all CPL, EMACS, and F77 programs.

OUTPUT - Contains camera ready water data report.

Figure F1.--Directory structure containing data and software used in the automated preparation procedures for the Annual Water Data Report.

Updating the station description text, which includes the downstream order number, station name and location, and period of record, is the second step. This text material complements the corresponding data tables for that particular station in the annual water data report. The subdirectories, SW, GW, and QW contain one large file of the station description text for that particular discipline with the exception of SW, which contains one station description file for partial record stations, and another for complete record stations. All of the station description files are arranged by downstream order number.

After the station description files have been updated, all data tables to be published in the report are retrieved. The ADAPS and QW systems, used to retrieve surface- and ground-water data from the NWIS data bases, are the primary systems used for this purpose. Retrievals are currently handled by authors of the data report on a station by station basis; however, the Arkansas District is currently developing programs to retrieve automatically all the data from both QW and ADAPS. Once the data tables are retrieved, the QW tables are then reformatted by a post-processing routine, which restructures the tables in a more suitable format to be used in the annual water data report. All tables retrieved then are placed in their appropriate discipline subdirectories (SW, QW, and GW).

The merging of all station description data with its corresponding table data then follows. This task is accomplished by first using an EMACS macro to search the manuscript files and isolate an individual station description. The table file is then searched to find the corresponding table data to merge with the station description text. During this process, extraneous lines (blank lines, duplicate data, and so forth) are deleted from the data tables, page headings are automatically generated, and page-division characters are inserted. The completed output files are written to the output subdirectory. The original data files are not disturbed so that if an error is discovered, a correction may be applied and the data can be processed again. The system was designed to be flexible enough to allow large table files to be subdivided into smaller files for processing.

The page numbering process is executed after the introductory material has been compiled and all station description headings are merged with their respective data tables.

The last two processes involve the creation of the table of contents and the index. Although the production of the table of contents is partially automated, the complete automation of these processes is still under development. An EMACS menu program is currently available to assist the preparer with the table of contents, but the index to the report needs to be written to a file using one of the available text editors on the Prime. Once the table of contents and the index file have been created, they are then appended to the finished output file and the file is printed using a laser printer to produce a camera-ready copy to be used for publication.

SOFTWARE

The following is a description of all of the programs used to automate the production of the annual data report. The languages used for the development of these programs are CPL, Prime EMACS Extension Language (PEEL), and Fortran 77.

CPL Programs

The automated procedure is driven by the CPL program, DATARPT.CPL, that displays the following menu options:

(1) Retrieve Data,
(2) Merge Quality Water Manuscripts and Tables,
(3) Merge Surface Water Manuscripts and Tables,
(4) Merge QW and SW files from (2) and (3),
(5) Merge Ground Water Manuscripts and Tables,
(6) Quality of Ground Water,
(7) Number pages in a file, and
(8) Exit and Return to Primos.

The first option in this menu executes the CPL program, RETRIEVALS.CPL, that performs data-table retrievals from either the water-quality data base, QWDATA, or the surface water data base, ADAPS. The other menu selections execute CPL programs that initialize and execute the appropriate EMACS macro for that selection. These CPL programs can be executed as either interactive or non-interactive processes. The programs perform edit-checking, such as testing for the existence of key input files, creating files that contain error messages. The user is allowed the option of selecting a page length or accepting a default. The page length is then passed on to the EMACS macro.

EMACS Programs

EMACS was chosen to perform all of the text processing functions because most of the work involved in the production of the annual data report consists of word processing tasks, such as manipulating blocks of text, and searching for unique character strings. The EMACS programs, called macros, are written in the programming language, PEEL.

The EMACS macros listed below perform the same text processing functions, merging manuscript data with corresponding data tables, for the class of data indicated:

DISCHARG.EM - Surface-water stations,
REGRPT.EM - "Complete" water-quality stations,
PARTRPT.EM - Parital-record water-quality stations, and
GWQW.EM - Ground-water water-quality stations.

Each program uses two input files, the first containing station description data and the second containing data tables for that discipline. Each of the input files are copied into a work area where they may be manipulated without altering the original file.

These programs isolate the station description data for a given site in a work buffer, extracting the site identification. By using the site identification and the station description, the programs construct the headings that will appear at the top of each completed page. The buffer containing all of the data tables then is searched for those tables associated with the current site identification. If those tables are present, they are copied into the work buffer. If no data tables are found for a site, a message is written to an error file and the program continues processing. The manuscript and table data then are merged and written to the output buffer, one page at a time. This process continues until all station description entries in the input file have been merged with their respective data tables.

The EMACS macro, PAGENUMBER.EM, numbers the pages in the output file. The program prompts the user for the beginning page number, left and right margins, and for which side of the page to start placing the page numbers. Given the margins and the number of digits in the page number, the program calculates the correct columnar position to place the page numbers on either side of the page. The program starts at the top of the file, searches for the page delimiter, and inserts the page number in the correct position on the next line. This process is continued until the end of the file is reached.

The macro TOC.EM creates the table of contents. The macro first reads the file containing all of the merged data and builds a separate file that contains the station name followed by the page number of that station. After the file is built, each line of the file is displayed, and the user is prompted to enter the number of spaces to indent that line from the left margin, and to enter the type of data that has been recorded for this station (discharge, chemical, biological, and so forth). The line then is reformatted with the station name indented the correct number of spaces, followed by the abbreviations for the types of data recorded, ending with the page number at the right side of the page.

Fortran 77 Programs

The program QTAB.F77 reformats the data tables output from the QWDATA system. This program deletes all occurrence of "no value" indicators (represented by '--') where the indicators are the only items present for a given time sequence in the data table (tables F1 and F2). The program realigns the columns of the output table for centering purposes, and also includes the 'TIME' column on every table within a given retrieval. QWDATA currently gives the 'TIME' column on only the first page of a station retrieval.

SUMMARY

New procedures for automating the preparation process of the annual state water data report have been developed by the Arkansas District of the U.S. Geological Survey. The automation process uses high-speed computers, retrieval and word processing software, and laser printing technology.

Table F1.--*Data table from QW data retrieval system before processing by QTAB reformatting program*

07032000 - MISSISSIPPI RIVER AT MEMPHIS, TENN.

WATER QUALITY DATA, WATER YEAR OCTOBER 1987 TO SEPTEMBER 1988

DATE	TIME	AGENCY COL-LECTING SAMPLE (CODE NUMBER) (00027)	AGENCY ANA-LYZING SAMPLE (CODE NUMBER) (00028)	STREAM-FLOW, INSTAN-TANEOUS (CFS) (00061)	SPE-CIFIC CON-DUCT-ANCE (US/CM) (00095)	PH (STAND-ARD UNITS) (00400)	TEMPER-ATURE WATER (DEG C) (00010)	TUR-BID-ITY (FTU) (00076)	OXYGEN, DIS-SOLVED (MG/L) (00300)
DEC									
28...	1230	80513	80020	748000	300	7.60	6.5	150	9.6
JAN									
26...	1330	80513	80020	762000	330	7.40	3.5	120	11.2
APR									
19...	1415	80513	80020	584000	419	8.00	15.0	48	9.7
JUN									
27...	1340	80513	80513	--	548	8.08	29.0	--	8.6
27...	1342	80513	80513	--	--	--	--	--	--
27...	1344	80513	80513	--	548	8.14	29.0	--	9.7
27...	1346	80513	80513	--	548	8.11	29.0	--	9.5
27...	1348	80513	80513	--	--	--	--	--	--
27...	1400	80513	80020	123000	548	8.10	30.0	17	8.2
JUL									
18...	1200	80513	80020	141000	492	7.80	30.0	37	7.5
AUG									
16...	1020	80513	80020	130000	538	8.30	31.0	18	8.1

07032000 - MISSISSIPPI RIVER AT MEMPHIS, TENN.

WATER QUALITY DATA, WATER YEAR OCTOBER 1987 TO SEPTEMBER 1988

DATE	SILICA, DIS-SOLVED (MG/L AS SIO2) (00955)	SOLIDS, RESIDUE AT 180 DEG. C DIS-SOLVED (MG/L) (70300)	SOLIDS, SUM OF CONSTI-TUENTS, DIS-SOLVED (MG/L) (70301)	SOLIDS, DIS-SOLVED (TONS PER AC-FT) (70303)	NITRO-GEN, NITRATE DIS-SOLVED (MG/L AS N) (00618)	NITRO-GEN, NITRITE DIS-SOLVED (MG/L AS N) (00613)	NITRO-GEN, NO2+NO3 DIS-SOLVED (MG/L AS N) (00631)	NITRO-GEN, AMMONIA TOTAL (MG/L AS N) (00610)	NITRO-GEN, AMMONIA DIS-SOLVED (MG/L AS N) (00608)	NITRO-GEN, ORGANIC TOTAL (MG/L AS N) (00605)
DEC										
28...	5.7	182	174	0.25	--	<0.010	1.20	0.140	0.150	0.36
JAN										
26...	5.9	195	199	0.27	1.68	0.020	1.70	0.130	0.110	0.67
APR										
19...	5.5	236	231	0.32	2.17	0.030	2.20	0.090	0.050	0.41
JUN										
27...	--	--	--	--	--	--	--	--	--	--
27...	--	--	--	--	--	--	--	--	--	--
27...	--	--	--	--	--	--	--	--	--	--
27...	--	--	--	--	--	--	--	--	--	--
27...	--	--	--	--	--	--	--	--	--	--
27...	2.9	303	310	0.41	--	<0.010	0.360	0.040	0.090	0.76
JUL										
18...	1.4	288	294	0.39	0.410	0.010	0.420	<0.010	<0.010	--
AUG										
16...	0.88	309	305	0.42	--	<0.010	0.300	0.030	0.030	0.67

Table F2.--*Data table from QW data retrieval system after processing by QTAB reformatting program*

07032000 - MISSISSIPPI RIVER AT MEMPHIS, TENN.

WATER QUALITY DATA, WATER YEAR OCTOBER 1987 TO SEPTEMBER 1988

DATE	TIME	AGENCY COL- LECTING SAMPLE (CODE NUMBER) (00027)	AGENCY ANA- LYZING SAMPLE (CODE NUMBER) (00028)	STREAM- FLOW, INSTAN- TANEOUS (CFS) (00061)	SPE- CIFIC CON- DUCT- ANCE (US/CM) (00095)	PH (STAND- ARD UNITS) (00400)	TEMPER- ATURE WATER (DEG C) (00010)	TUR- BID- ITY (FTU) (00076)	OXYGEN, DIS- SOLVED (MG/L) (00300)	SILICA, DIS- SOLVED (MG/L AS SIO2) (00955)
DEC										
28.	1230	80513	80020	748000	300	7.60	6.5	150	9.6	5.7
JAN										
26.	1330	80513	80020	762000	330	7.40	3.5	120	11.2	5.9
APR										
19.	1415	80513	80020	584000	419	8.00	15.0	48	9.7	5.5
JUN										
27.	1340	80513	80513	--	548	8.08	29.0	--	8.6	--
27.	1342	80513	80513	--	--	--	--	--	--	--
27.	1344	80513	80513	--	548	8.14	29.0	--	9.7	--
27.	1346	80513	80513	--	548	8.11	29.0	--	9.5	--
27.	1348	80513	80513	--	--	--	--	--	--	--
27.	1400	80513	80020	123000	548	8.10	30.0	17	8.2	2.9
JUL										
18.	1200	80513	80020	141000	492	7.80	30.0	37	7.5	1.4
AUG										
16.	1020	80513	80020	130000	538	8.30	31.0	18	8.1	0.88

07032000 - MISSISSIPPI RIVER AT MEMPHIS, TENN.

WATER QUALITY DATA, WATER YEAR OCTOBER 1987 TO SEPTEMBER 1988

DATE	TIME	SOLIDS, RESIDUE AT 180 DEG. C DIS- SOLVED (MG/L) (70300)	SOLIDS, SUM OF CONSTI- TUENTS, DIS- SOLVED (MG/L) (70301)	SOLIDS, DIS- SOLVED (TONS PER AC-FT) (70303)	NITRO- GEN, NITRATE DIS- SOLVED (MG/L AS N) (00618)	NITRO- GEN, NITRITE DIS- SOLVED (MG/L AS N) (00613)	NITRO- GEN, NO2+NO3 DIS- SOLVED (MG/L AS N) (00631)	NITRO- GEN, AMMONIA TOTAL (MG/L AS N) (00610)	NITRO- GEN, AMMONIA DIS- SOLVED (MG/L AS N) (00608)	NITRO- GEN, ORGANIC TOTAL (MG/L AS N) (00605)
DEC										
28.	1230	182	174	0.25	--	<0.010	1.20	0.140	0.150	0.36
JAN										
26.	1330	195	199	0.27	1.68	0.020	1.70	0.130	0.110	0.67
APR										
19.	1415	236	231	0.32	2.17	0.030	2.20	0.090	0.050	0.41
JUN										
27.	1400	303	310	0.41	--	<0.010	0.360	0.040	0.090	0.76
JUL										
18.	1200	288	294	0.39	0.410	0.010	0.420	<0.010	<0.010	--
AUG										
16.	1020	309	305	0.42	--	<0.010	0.300	0.030	0.030	0.67

The steps used in manual methods of report preparation were used as a guide for development of software in the automated system and also to establish continuity between the old and new report preparation process.

A directory structure was developed to centralize the storage of all data files and software applications. The directory structure contains six subdirectories that group data and software files by subject. The SW, QW, and GW subdirectories contain all text and tabular data from NWIS systems ADAPS, QW, and GWSI, respectively. The other subdirectories are: SOFTWARE, containing all application programs for the system; INTRO, introductory text material; and OUTPUT, the repository for the final report after text station descriptions and table data have been compiled and merged.

Post-processing routines have been developed in Fortran 77 and Prime Command Procedure Language to reformat data tables retrieved from NWIS data bases. These data bases include ADAPS, GWSI, QW, and NEWSWUDS. The post-processing routine will also collate table and station data suitable for use in the annual report. Macros also have been developed, using EMACS, to perform all text processing functions, merging manuscript with data tables, page numbering and creating tables of contents.

SELECTED REFERENCES

Buck, B., 1982, EMACS Reference Guide IDR 5026, Revision 18.3: Prime Computer, Inc.

Lamb, T.E., Porter, J.E., Lambert, B.F., and Plafcan, Maria, 1988, Water resources data for Arkansas--water year 1987: U.S. Geological Survey Water-Data Report AR-87-1, 575 p.

32-bit Workstations: The Trials, Tribulations, and Triumphs of Converting to an Open-System, Tools Environment

By David R. Boldt

ABSTRACT

The U.S. Geological Survey will use 32-bit supermicrocomputers to complement the existing minicomputer network and to provide for the increasing computer needs of the Geological Survey for the next several years. The Geological Survey office in Ithaca, NY, has been among the first to use a 32-bit workstation for project work and report processing. The computer has been installed, operated, and maintained by hydrologists with no previous experience in site administration. Once the new operating system had been learned, the 32-bit workstation proved to be an extremely powerful tool. The flexibility of the operating system combined with the efficient and productive windowing environment, has made the workstation the preferred computing environment for data manipulation and visualization, programming, modeling, and document preparation.

INTRODUCTION

Why Workstation, and Why Change Operating Systems?

Workstations are computers intermediate in power between microcomputers and minicomputers. They follow the microcomputer paradigm in that every user has a dedicated central processing unit (CPU), but workstations are typically linked together by high-speed networks that allow them to use disk resources and other peripherals that exist elsewhere on the network. A workstation typically has a high-resolution screen for display of graphics and support of a windowing environment, which is the most effective interface currently available for allowing a single user to access the multitasking capabilities of an operating system such as Unix. Workstations excel at providing computing capabilities for computer-aided design, scientific modeling, desktop report processing, and geographic information systems. Distributed CPU's ensure that intensive computational tasks will not affect other users, and permit the high level of graphics feedback that is an integral part of the processing environment.

Unix is a hardware-independent operating system based on the C programming language. This basis in C rather than in an assembly language has made Unix relatively easy to transport from one machine to another. The operating system is available on virtually all varieties and sizes of computers, from large mainframes to home microcomputers. As a result of this hardware independence, the Federal government as a whole, and the U.S. Geological Survey, in particular, have been looking to Unix as a way to avoid the constraints imposed by the proprietary operating systems of most computer manufacturers.

This paper relates the impressions of a geologist with a moderate computer programming background who has been involved with the procurement, installation, and use of a 32-bit workstation, and tries to convey the difficulties and benefits of using a workstation. The interactive graphics interface is not easily described, but must be experienced to be truly appreciated. Once mastered, the window-based interface greatly increases the productivity of computer operations.

At present (1988), workstations are of interest mainly to a rather specialized group who feel comfortable in front of a glowing phosphor screen, but if the U.S. Geological Survey follows its present course, workstations and Unix will affect nearly all Geological Survey employees.

The Windowing Environment

To a new user the biggest difference between a 32-bit workstation operating system and mainframe-style operating systems is the user interface. When mainframe computing evolved from batch-only mode to a time-sharing system in which multiple users could simultaneously access the computer, each user had the illusion of exclusive access to the computer. A multitasking, windowing environment takes this concept a step further and provides the illusion of having access to several computers at once. (In a networked environment this can be literally true.). In the most general sense, a window is a virtual terminal, and a single, oversized, high-resolution screen can have several active windows in which a user can manipulate text files, run programs, display graphics, and so forth. Windows can be moved, covered by other windows, uncovered, resized, or closed. Closing a window changes the appearance of that window into a postage-stamp sized icon that functions as a sort of bookmark. An icon takes up much less space on the screen than a window and can be opened at any time to resume work. Manipulation of a window has no affect on programs that may be running within it.

Windows are manipulated by a small, hand-held device called a mouse. The mouse moves over a 9 by 8 inch gridded pad, while a small light-sensor in the base of the mouse senses movement and transmits this to the computer. (Microcomputers commonly use a mechanical mouse that does not require a special pad.) The position of a pointer or small arrow on the screen corresponds to the position of the mouse. Pressing one of the buttons on the mouse. Pressing one of the three buttons on the mouse will invoke a response that depends on what the arrow is pointing to and which button is pressed.

The windowing environment enables the transfer of text from one area to another. For example, text from one window may be copied into another by highlighting it with the mouse, which copies it to a temporary buffer, and pasting it to the selected location by pressing a different button on the mouse. Complex commands can be generated on the command line by copying pieces of previous commands and file names that are visible on the screen. In addition, previous interactions within a window are saved as they scroll out of view and can be retrieved by scrolling backwards.

An example of the windowing environment in action would be a session using an interactive program debugger. The debugger, which when executed, creates a new window that displays several tiles, or subwindows, showing (1) the source code marked with symbols, (2) the last line executed and the position of existing breakpoints, (3) a menu selection of debugger options, (4) a listing of values of variables, and (5) a space for typing debugger commands. The output of a graphics program, for example, can be directed to a graphics window so that, as the program runs within the debugger, the effect of individual graphics calls can be seen on the drawing as it is created. If a program error is found, it can be corrected with an editor in another window. The program can be edited, recompiled, and re-executed by moving from window the window and either typing the appropriate commands or using the mouse. Workstations are generally acknowledged to provide the most productive programming environment available.

The convenience of multiple windows whose size and locations can be easily manipulated soon becomes addictive because it avoids the exasperation caused by being in a program and needing some piece of information that resides elsewhere in the computer system, but can be accessed only be terminating the program. Instead, one can bring up a new window to access the needed information. A window can be expanded to fill the entire screen, providing up to 140 columns by 52 rows of normal-sized text, or reduced as small as desired. Even the type style and size can be changed for a window, so that one could see more data, or take off one's glasses. Each person who has been using the system long enough to become familiar with it develops an individual style of window management; some people work with so many open and partially obscured windows that the screen appears similar to a paper-cluttered desk. The benefits of an environment that "belongs" to the person working within it may be difficult to quantify, but most certainly improve productivity.

Communications

Communication with other computers by terminal emulation through asynchronous communications ports, is relatively straightforward. A Unix terminal emulation utility allows one to define a port configuration as to baud rate, parity, and so forth, and label it with a single word. The terminal emulator supports the transfer of files without error checking, which has been fast and trouble free for filed of less than 300 lines. For longer files, however, we have noted buffer overflow at a transmission rate of 4,800 bits per second or higher and loss of data. To send longer files, or files for which error checking is important, the staff has used the communications package Kermit, which is available for many varieties of computers from large mainframes to home microcomputers.

The two 32-bit workstations in our office are linked by a local are network, ethernet, which enhances the speed and reliability of communication among systems. The U.S. Geological Survey Distributed Informations System II requires local area network capabilities for communications among workstations.

The Complexity of a Workstation Operating System

Every operating system has individual characteristics that result from the philosophy of those who designed it. Unix was developed by programmers and systems developers to facilitate their own work. Thus, Unix has commands to do many things that might otherwise require working with a conventional programming language. Unix has extensive support for manipulating text files, work processing, and computer programming.

Unix has a reputation for being somewhat user-hostile; one reason is that it is extremely terse. Many operating systems contain shortened forms or abbreviations for commands, but Unix contains only abbreviations, and most commands consist of two or three letters. The reason for this brevity is simply that short commands require less typing--no small consideration to a programmer who may use these commands hundreds of times each day. Another factor contributing to the difficulty of using Unix is that some actions such as listing a file may not have direct commands but result as a side-effect of a different command; for example, the "cat" command can be used to list the contents of a file to the screen but is actually the command to join files, shortened from "catenate." A second example is that the command to move a file and to rename a file are the same (mv); one being a move between directories, and the other a "move" within the same directory. Still another cause of unease for new Unix users is the sheer size of the operating system; it contains many capabilities that add power and flexibility that can understandably intimidate the beginner.

Most occasional users on the Prime computer use no more than a half-dozen Prime operating system (Primos) commands. This minimal interaction with the operating system should ease the transition between operating systems for most people. In fact, the windowing environment can be tailored to guide those uncomfortable with computers through any desired process by means of pop-up menus.

Most of the effort in adapting to a 32-bit workstation will be learning the new text editors and applications programs. Those who have an aptitude for programming will likely find the Unix environment extremely flexible and supportive of their efforts. The 32-bit workstation in the Ithaca, NY, office is the preferred platform for writing programs and for the type of computer activity required to support ground-water modeling. The people who will be most adversely affected by the change in operating systems will likely be those of medium computer skills, that is, those who need to do more than run a few existing programs but who have not generally been writing their own programs. The complexity of the operating system has been a problem in the business world as well and has slowed the acceptance of Unix in the marketplace. As a result, workstation vendors have a strong interest in making Unix more accessible.

Capabilities of a Workstation Operating System

One of the most powerful features of Unix is Input/Output (IO) redirection. For example, program output that would normally be written to the screen can be directed to a file, and conversely, a program that would normally read from the terminal can be directed to read from a file. Fortran file unit 5 is reserved for standard input, whereas unit 6 is reserved for standard output. All programs adhering to Unix system conventions read from standard input and write to standard output to facilitate IO redirection. The idea of interactively routing IO is most powerfully demonstrated with "pipelines," whereby the standard output from one program is connected to the standard input of another so that results from one program can be used as data by another without the use of intermediate files. Any number of programs can be linked in this manner, each functioning as a filter on the data produced by the previous program. Below are some examples of how IO redirection is accomplished.

ls	--	list files in current directory
ls > *suds*	--	place listing into a file called "suds"
sort < *data* > *report*	--	sort the file "data" and place the results into a file called "report"
ls | sort -r	--	list files in current directory in reverse alphabetical order

The operators for IO redirection are:

<	--	use data from the file whose name follows the "<"
>	--	send output to the file whose name follows the ">"
|	--	connect the output from the program preceding the "|" to the input of the program following the "|"

To give some idea of the tools available in Unix, a few Unix commands are described below.

cut *cols filename*	list specified columns from a file
grep *text filename*	list all lines of a file containing "text"
head -n *filename*	list the first n lines of a file
paste *cols filename*	merge two files side by side
sort *filename*	sort the contents of a file
spell *filename*	invoke a spelling checker on a file
su *username*	switch to another userid without logging out
tail -n *filename*	list the last n lines of a file
tr *pattern filename*	apply global edit changes to a file
wc *filename*	count letters, words, and lines

This is just a sample of the variety of commands available. In addition, IO redirection can be used with any of these commands, wildcards may be used to specify groups of files, and pattern-matching wildcards are available as well.

As one gains experience with Unix, one's perception of the interaction between programs and files changes subtly. In most operating systems, files can be conceptualized as being stationary receptacles of information, and programs can be imagined to search for these files and perform actions on them. In the Unix environment, a more useful concept is that the programs are stationary and that the files themselves are in motion, carrying information between programs.

Another powerful feature of Unix is "last command editing," where all previously executed commands are stored in a a queue and may be listed, retrieved, edited, and reissued with a minimum of keystrokes. For example, to rename a file a fumble-fingered (or CTR hypnotized) programmer may have typed:

```
mc sf.data stream-flow-measurements
```

Having mistyped the "mv" command, the mistake can be corrected and the command reissued by typing:

```
^mc^mv
```

which results in the execution of the command:

```
mv sf.data stream-flow-measurements
```

One can also retrieve any of the words used in a previous command. For example, $! refers to the last word of the last command. If the following were typed immediately after the previous command:

```
cat $!
```

The command:

```
cat stream-flow-measurements
```

would be executed, and the file would be listed to the screen.

APPLICATION OF 32-BIT WORKSTATIONS

The Ithaca, NY office of the U.S. Geological Survey has used a 32-bit workstation for several project-related tasks during the year in which it has been available. These tasks are described below.

Ground-Water Modeling

A three-dimensional ground-water flow model by McDonald and Harbaugh (1988) has been installed and used to prepare a cross-sectional simulation of an aquifer response to an aquifer test. It is currently being used to prepare a series of cross-sectional models to estimate the hydraulic conductivity of a riverbed and also in the expansion of a two-dimensional ground-water flow model of a contaminated industrial site in western New York to a full three-dimensional ground-water flow model.

Data Base

A special-purpose data-base program has been written in the C programming language that is keyed to the bedrock stratigraphy in the area of Niagara Falls, in New York and Ontario, and includes stratigraphic horizons, elevations, and formation names. The data base is used to maintain information on formation pressures and geochemical data obtained from multilevel piezometer systems (12 to 21 zones per hole) at 17 locations. The intended use of this keyed file system is to produce maps of individual zones between the 17 well locations and to support the development of a three-dimensional ground-water-flow model.

Programming

Several Fortran programs have been ported to the 32-bit workstation from minicomputers. Most required minor alterations to run on the workstation, depending on the degree to which the programs adhered to standard American National Standards Institute (ANSI) Fortran 77. In many ways the software migration between the two systems has served to highlight exactly which features constitute standard Fortran 77.

The productive programming environment makes the development of programs much easier and faster on 32-bit workstations than on conventional minicomputers. A Unix facility named "make" uses information about the interdependency of subroutines so that, after a program is edited, only those subroutines that are directly affected by the change are recompiled. Another Unix utility named "SCCS" keeps track of all modifications made to a program and thus provides an audit trail so that one can retrieve previous versions or see who has edited a program.

Reports Processing

The 32-bit workstation with reports processing software is replacing outmoded, dedicated word-processing machines. The ability to create and edit on screen special characters; such as Greek symbols, italics, superscripts and subscripts; to create mathematical equations; and to mix text and graphics greatly facilitates the preparation of technical documents. The high quality of the financed product, produced by a laser printer, enhances the appearance of memorandums and correspondence.

The 1988, U.S. Geological Survey annual data report for Ithaca, NY, was prepared using the new hardware-software combination for the first time and was completed on schedule. Most of the time spent was on transferring and reformatting files that were stored on the dedicated word-processor.

Graphics

Graphics software is being used to generate graphs from data stored in a custom-built data base program. The graphics package is a C subroutine library that supports two- and three-dimensional graphics, but lacks facility in font handling and labeling. Interactive graphics packages are available for 32-bit workstations and the office staff has strong interest in the capabilities of interactive graphics software for preparation of charts and diagrams for presentations and reports.

SUMMARY

Unix workstations provide both a promise and a challenge. The promise is that each of us will have the capability of a dedicated computer and at our desk with the proper tools and interface to do our work much more efficiently than at present. The challenge is to manage the complexity of the new power and capability so that we retain what we have already gained while benefitting from new technology.

EVALUATION OF THREE ELECTRONIC REPORT PROCESSING SYSTEMS FOR PREPARING HYDROLOGIC REPORTS OF THE U.S. GEOLOGICAL SURVEY, WATER RESOURCES DIVISION

STILTNER, Gloria J., USGS, WRD, 12201 Sunrise Valley Drive, Reston, VA 22092

In 1987, the Water Resources Division of the U.S. Geological Survey undertook three pilot projects to evaluate electronic report processing systems as a means to improve the quality and timeliness of reports pertaining to water-resources investigations.

The three projects selected for study included the use of the following configuration of software and hardware: Ventura Publisher software on an IBM model AT personal computer, PageMaker software on a Macintosh computer, and FrameMaker software on a Sun Microsystems workstation. The following assessment criteria were to be addressed in the pilot studies: The combined use of text, tables, and graphics; analysis of time; ease of learning; compatibility with the existing minicomputer system; and technical limitations. It was considered essential that the camera-ready copy produced be in a format suitable for publication. Visual improvement alone was not a consideration.

This report consolidates and summarizes the findings of the electronic report processing pilot projects. Text and table files originating on the existing minicomputer system were successfully transferred to the electronic report processing systems in American Standard Code for Information Interchange (ASCII) format. Graphics prepared using a proprietary graphics software package were transferred to all the electronic report processing software through the use of Computer Graphic Metafiles. Graphics from other sources were entered into the systems by scanning paper images. Comparative analysis of time needed to process text and tables by the electronic report processing systems and by conventional methods indicated that, although more time is invested in creating the original page composition for an electronically processed report, substantial time is saved in producing subsequent reports because the format can be stored and re-used by electronic means as a template. Because of the more compact page layouts, costs of printing the reports were 15 to 25 percent less than costs of printing the reports prepared by conventional methods.

Because the largest report workload in the offices conducting water-resources investigation is preparation of Water-Resources Investigations Reports, Open-File Reports, and annual State Data Reports, the pilot studies only involved these kinds of products.

EVALUATION OF A DESKTOP REPORTS PROCESSING SYSTEM FOR PRODUCING EARTH-SCIENCE TECHNICAL REPORTS

By Richard A. Hollway and Denise A. Wiltshire

ABSTRACT

An evaluation of effectiveness of desktop reports processing systems for widespread application in the U.S. Geological Survey was conducted from November 1987 through June 1988 in association with two other studies. For this study, the hardware platform of the desktop reports processing system consisted of a 32-bit, multiuser microcomputer manufactured by Sun Microsystems, Inc.[1], (model 3/160). FrameMaker (version 1.11) software, manufactured by Frame Technology, Inc., was selected for the evaluation because it is an integrated text processing and page composition package that operates on a Sun workstation, and it is reasonably priced.

The evaluation process entailed producing a 92-page technical report with 13 tables and 41 figures using integrated desktop reports processing techniques and comparing procedures to those for preparing the same report using manual and conventional word-processing methods. The desktop reports processing system was evaluated by testing procedures of text processing, graphics, page composition, document management, and the user interface. In comparison to presently used word processing software, FrameMaker software provides excellent text processing and good-to-excellent page composition capabilities. Strengths of FrameMaker software include versatile document structuring, report format templates, and use of the SunView windowing interface. FrameMaker was consistently fast and accurate for global changes of text and page composition. The software offers easy to use pull-down and pop-up menus for selecting typefaces, column layout, and headers and footers. Among FrameMaker's best features is the more complete and faster Spelling Checker utility.

The hardware platform enhanced the performance of the software throughout the text processing procedures used during the evaluation . The 32-bit, virtual memory operating system of the Sun workstation offers broad functionality for an integrated word processing and page composition software package such as FrameMaker. Because FrameMaker operates within the SunView windowing environment, multiple documents may be viewed on the workstation screen and are easily copied into FrameMaker windows. In addition, facing pages of a document may also be viewed on the screen simultaneously, thus facilitating page layout. The virtual memory addressing system of the Sun workstation allows nearly limitless document sizes, and the power of the processing unit is sufficient to ensure rapid loads, saves, and global changes of text and page composition in reports exceeding 100 pages.

Comparisons of desktop reports processing methods to conventional document processing methods used by the U.S. Geological Survey indicate additional advantages of a fully integrated software package such as FrameMaker. Complex document and graphics transfers from a variety of computers and manual cut and paste procedures are eliminated using FrameMaker. The use of proportional typefaces, widow and orphan control, and condensed tables using tab construction techniques resulted in reducing the total number of pages of the final version of a document when compared to conventional document processing methods. As a result, a 20 percent decrease in printing costs may be expected using desktop reports processing software. Although additional time is required for preparing structured templates for page composition during the initial typing stage, the ability to store templates allows rapid production of subsequent reports with a similar format. Hence, substantial time can be saved if a publications office produces many reports requiring the same format specifications. Although improving the visual appearance of report formats was not an objective of the study, the style of the final publication was enhanced and printing costs were reduced when compared to conventional publishing methods.

USE OF AN ELECTRONIC PAGE-COMPOSITION SYSTEM TO PREPARE CAMERA-READY COPY OF SCIENTIFIC REPORTS

By L.H. Geiger[1], P.R. Mixson[1], and S.D. Flagg[1]

ABSTRACT

The U.S. Geological Survey publishes each year several thousand reports, abstracts, and articles that require time-consuming collaboration among authors, cartographers, illustrators, typographers, and editors. In 1989, camera-ready copy for approximately 550 of these reports was prepared using various word-processing software packages and desktop printers.

Three pilot projects were conducted to assess the feasibility of electronic page-composition systems to prepare camera-ready copy. Electronic page-composition was expected to result in low-cost and high-quality documents. Each project tested a different configuration of hardware and software. The projects were conducted in the Florida District, Ohio District, and Oregon offices of the Water Resources Division, U.S. Geological Survey.

In the Florida project, the configuration consisted of a personal computer with a full-page monitor, laser printer, word-processing software, and page-composition software. The scope of the Florida project included the preparation of camera-ready copy of several Water-Resources Investigations Reports, Open-File Reports, an annual State data report, text for map reports, and materials for slide presentations.

In assessing the electronic page-composition system in the Florida project, the following improvements were noted over previous systems:

- *Reusable style sheets reduce the amount of time spent on subsequent reports.*
- *Automatic preparation of table of contents, indexes, and page numbering reduces the amount of time spent on these parts of the report.*
- *Proportional fonts reduce the number of pages per report; this can lower printing costs. The greatest reduction of pages occurs in those reports having a high ratio of text to tables or illustrations.*
- *Importing graphics reduces the amount of time spent to integrate graphics with text and tables when preparing camera-ready copy.*

[1]U.S. Geological Survey, Tallahassee, Fla.

EVALUATION OF A USER-FRIENDLY ELECTRONIC REPORT PROCESSING SYSTEM FOR PREPARATION OF SELECTED REPORTS

By Michael Eberle
U.S. Geological Survey, Water Resources Division
Columbus, Ohio

ABSTRACT

An evaluation of a user-friendly electronic report processing (ERP) system showed that the system is easy to install and learn, readily transfers text data files to and from a minicomputer located in the same office, and adequately performs most of the text-processing and page-layout functions needed to prepare camera-ready copy for selected reports of the U.S. Geological Survey, Water Resources Division. The preparation of camera-ready copy of text and tables for two sample reports by use of the ERP system, however, took about four times longer than preparation of the same text and tables by use of a conventional word-processing system.

The evaluation, which was conducted by the Ohio (Columbus) office of the U.S. Geological Survey, was one of three evaluations that addressed basic questions of cost effectiveness and technical feasibility of introducing ERP systems for applications where use of conventional word-processing techniques has been the norm. The ERP system tested in the Ohio office consisted of Macintosh SE computer equipped with a 20-megabyte internal hard disk; an Apple LaserWriter printer; PageMaker page-layout software (version 2.0a; version 3.0 was not available during most of the evaluation); and ancillary terminal-emulation, word-processing, and graphics software. The conventional word-processing system consisted of an IBM-compatible microcomputer (also with a 20-megabyte internal hard disk), a 55-character-per-second strike-on printer, and WordMARC word-processing software.

An important assumption in the evaluation was that the ERP system should readily transfer text files to and from the network of minicomputers already in place. To test for transferability, all text and table files for the sample reports were transferred from the minicomputer colocated in the Ohio office to the ERP system in American Standard Code for Information Interchange format. Selected files also were transferred from the ERP system to the minicomputer.

For both sample reports prepared with the ERP system, text and table processing time was about equally divided between reformatting of American Standard Code for Information Interchange format files for transfer to PageMaker software and on-screen page layout; however, the shorter of the sample reports (which was entirely in a two-column format) took twice as long per page to complete as the longer report (which contained about 50 pages of data tables in a fixed-space font in single columns).

Evaluation of graphics was restricted to how well files in MacPaint format, PICT format, and Tag Image File Format could be transferred to the PageMaker software. Time and cost comparisons of ERP and conventional illustration-preparation techniques were not feasible. In general, graphics generated with the Macintosh-based software in all of the above-mentioned formats were transferred satisfactorily. Attempts to use images saved from the Macintosh screen in MacPaint and PICT formats, however, resulted in graphics characterized by garbled text and insufficient resolution for publication.

INDEX

☆U.S. GOVERNMENT PRINTING OFFICE 1991-550-938