Maine Groundwater Data Management Study
Phase II:
Evaluation of Groundwater Data Management Systems

Final Report

Submitted
to
State of Maine
Land and Water Resources Council
Data Management Committee

Prepared
by
Mark A. Jadkowski, Ph.D.
and
Donald D. Tarbet, Ph.D.

January, 1988

JAMES W. SEWALL COMPANY
P.O. Box 433
Old Town, ME 04468 Tel. (207) 827 4456

Executive Summary

This report provides detailed evaluations of computerized groundwater data management systems for the State of Maine. Included in this report are: (1) generic descriptions of various groundwater data management options; (2) specific recommendations for database management systems and geographic information systems; and (3) system implementation guidelines. Cost estimates are also provided.

Generic Groundwater Data Management Options

Three levels of groundwater data management are recommended for further consideration: (1) a groundwater data index / groundwater data coordinator, (2) a database management system (DBMS) , and (3) an integrated geographic information system / database management system (GIS / DBMS). These levels represent three stages of increasing function, complexity, and cost. Each level also allows upward integration to the next. A "groundwater data index / groundwater data coordinator" management scheme can be built up to a full database management system, which can later be integrated with a geographic information system to provide GIS / DBMS capabilities.

This three-level concept permits a multi-phased implementation. Each element of the overall groundwater data management system can become operational over a relatively short period of time with demonstrable utility and benefits. Additionally, a phased implementation allows periodic assessment of actual benefits realized from the system based on usage experience.

Under this scheme, a groundwater data index is developed to serve as a source of current information on availability, location, and format of groundwater data in the State. This index would be a "road map" for individuals requiring access to groundwater data available from Maine government. The index need not store any of the groundwater data. The role of the groundwater data coordinator would be to facilitate access to data and to maintain the index.

Three DBMS configurations should be considered: (1) stand-alone PC workstations; (2) networked PC workstations; and (3) a centralized mainframe system. Each of these configurations allows upward integration to the next. The State can choose to "buy-in" with the stand-alone PC workstations. Later, if justified by high usage or regulatory requirements, these workstations could be networked to one another or integrated with the State's IBM mainframe.

Any GIS adopted for use by the State should be fully integrated with a DBMS. We recommend that the GIS component be added only after one of the DBMS configurations described above has been implemented and is in operation. The capability to do geographic analysis and digital mapping will be of little state-wide utility for groundwater management until such time as groundwater data have been computerized in a DBMS. This two-phased implementation will also allow time for review of the DBMS and reassessment of the need for an in-house GIS (i.e., a DBMS may satisfy a majority of day-to-day needs of most State agencies - less frequent GIS analyses could be performed by an outside contractor). The reader should note, however, that selection of a particular brand of DBMS will restrict future GIS options, since not all GISs can be integrated with a given DBMS.

System Recommendations

Approximately one hundred and seventy-five database management systems and seventy-five geographic information systems were evaluated against an extensive list of criteria that reflect Maine's near-term and long-term groundwater data management needs. Of these, four DBMSs and four integrated GIS / DBMSs are recommended to the State for further consideration.

DBMS

Four commercial DBMS packages are recommended for groundwater data management: dBase III+, Ingres, Info, and Oracle. Each of these systems provides a relational database structure and a programmable query language. All but dBase III+ are available for a broad range of computers including IBM PC-AT compatibles, mini-computers, and the State's IBM mainframe. IBM PC-AT compatible DBMS workstations can be configured as stand-alone units or as nodes on a local area network. The costs of PC DBMS workstations vary as follows:

• PC-AT DBMS software -	$ 700-1300 / workstation
• PC-AT network -	600-1500 / workstation
• Annual software maintenance, support, updates -	400-500 / workstation

IBM mainframe versions of Oracle, Ingres, and Info are considerably more expensive than their PC counterparts, but these support multiple-users and permit a high level of database centralization:

• IBM mainframe DBMS software -	$ 30-155K
• Annual software maintenance, support, updates -	4-18K

EPA's STORET system is a possible alternative to acquisition of a IBM mainframe DBMS package. Users could access the system over telephone lines with PCs and modems or through dedicated computer links. EPA provides each state using the system with a credit to offset system usage costs ($5K in 1987 for Maine). STORET is widely used on the national level by EPA and by a number of states for groundwater data management. The system is very well suited for groundwater quality data, but it is of limited utility for other types of groundwater data. STORET also lacks provisions for future addition of a GIS component.

GIS / DBMS

Four commercial GIS / DBMS options are recommended to the State: (1) Strings / Ingres, a product of GeoBased Systems; (2) KGIS / Oracle, a product of KORK Systems; (3) Arc / Info, a product of Environmental Systems Research Institute; and (4) GeoVision / Oracle, a product of GeoVision Corp.. Each of these systems was selected based on its versatility and processing power, high level of integration with a programmable relational DBMS, compatibility with State computer standards, and vendor's track record.

Costs depend highly on the exact system configuration chosen. A single-user configuration consists of a GIS / DBMS workstation with full digitizing, editing, analysis, and plotting capabilities. Price ranges for single-user systems are as follows:

• Single-user GIS / DBMS software and hardware -	$ 45-70K
• Annual maintenance and updates -	3-9.2K

A centralized multi-user configuration consists of a mini-computer GIS / DBMS with full digitizing, editing, analysis, and plotting capabilities together with seven remote graphics workstations for editing and geographic information analysis only. Price ranges for multi-user systems are as follows:

• Multi-user GIS / DBMS software and hardware -	$ 200-480K
• Annual maintenance and updates -	12.6-33K

A number of private firms offer geographic information analysis and digital mapping services on a contractual basis. The services offered by these firms provide an alternative to acquisition of a GIS. The rates charged by some firms are comparable with the costs associated with operating an in-house GIS ($40-45 per hour). Additionally, some consultants maintain on-line geographic databases with information relevant to the State groundwater data analysis needs (e.g., political boundaries, topography, land cover). These can be incorporated in a project at costs substantially below those originally required to convert the data to a digital format (costs are spread out among several clients).

System Implementation

System implementation procedures should provide opportunities for periodic evaluation and gradual development of capabilities. Our recommendations involve five steps: (1) implementation of a groundwater data index, (2) addition of stand-alone DBMS workstations, (3) workstation networking, (4) addition of a centralized mainframe DBMS, and (5) integration with GIS capabilities. Networking and a mainframe DBMS are optional.

Data coding formats established by EPA for groundwater information should be adopted by the State for its DBMSs. These data formats have been operationally tested by both state and federal agencies, and their use will result in a high level of compatibility with EPA and USGS.

Measures should be taken to insure that the state-wide groundwater data management system is used properly and used to its fullest advantage. These include system backups, documentation, user logbook, concurrency control, integrity control, data definitions and coding, system audits, map base standards, GIS standards, security control, and accessibility provisions. Consideration should also be given to compatibility and integration with other databases (e.g., forest resources, socio-economics). System standards should be reviewed by users and periodically modified to reflect changing conditions.

Long-term staffing requirements may include: (1) a groundwater data coordinator; (2) mainframe DBMS manager; (3) GIS / DBMS manager; and (4) support staff (programmers, data entry personnel, GIS operators). We recommend that experienced consultants be used for system implementation, but in a manner that also provides adequate training opportunities for State personnel.

Costs for groundwater data index development, DBMS installation, development of DBMS file formats and query and analysis programs may total $37-63K. GIS / DBMS installation and programming will be an additional $10-15K. Staff training may cost as much as $22-32K.

The long-term cumulative cost of entering all existing groundwater data in the system will be approximately $200-250K. Of this figure, $180-210K is DBMS related and $40-70K is GIS related. An optional $600-800K will be required for digitizing surface hydrology, transportation, political boundaries, and cultural features from existing and future USGS 1:24,000 scale maps in order to enhance the GIS. This does not include the cost of digitizing topographic contours.

Conclusion

This report does not attempt to suggest which commercial system should ultimately be purchased for groundwater data management. It does, however, define a range of realistic options.

System purchase costs vary with the options described above. These range from $700 for a single PC DBMS software package to $480K for a seven workstation GIS / DBMS. Annual maintenance and software update costs for these options are between $400 and $33K. System implementation costs depend largely on the volume of data to be entered in the system. A conservative upper limit is in the $1-1.2 M range.

Finally, a phased system implementation will take several years, or longer, during which time user needs will change and technology will evolve. Parts of this report will become out-of-date during this time. The reader should be aware of this fact, and should assess the information presented herein accordingly.

Table of Contents

List of Figures

List of Tables

List of Acronyms

CERCLA - Comprehensive Environmental Response, Compensation, and Liability Act
DAFRR - Department of Agriculture, Food, and Rural Resources
DBMS - Database Management System
DEP - Department of Environmental Protection
DOA - Department of Administration
DHS - Department of Human Services
DOC - Department of Conservation
EPA - Environmental Protection Agency
GIS - Geographic Information System
MDOT - Maine Department of Transportation
MeGIS - Maine Geographic Information System
MGS - Maine Geological Survey
PCS - Permit Compliance System
PUC - Public Utilities Commission
RCRA - Resource Conservation and Recovery Act
SPO - State Planning Office
USGS - United States Geological Survey

Background

Eight state and federal agencies are major generators and/or users of groundwater information in Maine: Department of Environmental Protection (DEP), Department of Agriculture, Food and Rural Resources (DAFRR), Department of Human Services (DHS), Department of Conservation/ Maine Geological Survey (MGS), Maine Department of Transportation (MDOT), Public Utilities Commission (PUC), State Planning Office (SPO), and U.S. Geological Survey (USGS). A number of these agencies are proceeding with computerized data management. Most groundwater data management systems (DBMS) are designed to meet the specific needs of individual agencies.

Current groundwater data management systems are used primarily for data storage and retrieval. Needs which are not being efficiently met include (1) comprehensive knowledge of what groundwater information is available in Maine, (2) mutual access to the most up-to-date information available, (3) ability to answer inquiries and satisfy requests for data, (4) trend analysis of regional groundwater quality and quantity, (5) performance monitoring of pollution abatement systems and remedial containments, (6) rapid access to information for emergency response to hazardous materials spills, and (7) automated analysis and map-making from previously collected information. These needs can be largely satisfied through implementation of a coordinated state-wide groundwater data management system.

The Maine Land and Water Resources Council is an inter-agency coordinating body of the State Planning Office composed of the Commissioners and Directors of State agencies that use, manage, and regulate the State's natural resources. Its Standing Committee on Data Management is committed to improving the effectiveness and overall coordination of the State's natural resources data management systems. Accordingly, the Groundwater Inter-agency Coordination Subcommittee of the Groundwater Standing Committee and the Data Management Committee of the Council initiated a two phase study.

Phase 1 Analyze existing groundwater data management systems and data management needs within the agencies that are currently collecting and / or using groundwater information.

Phase 2 Analyze options and develop recommendations for a State Groundwater Data Management System.

The specific objectives of Phase 1 were (1) evaluation of how existing data management systems are currently being used to support State efforts to manage Maine's groundwater, (2) identification of new arrangements needed to implement any further management programs now under consideration by the council, and (3) provision of a sound foundation for this Phase 2 project. James W. Sewall Company completed Phase 1 in January, 1986.

In order to achieve the objectives of Phase 1, accurate up-to-date information on agencies' needs was required. Sewall designed and distributed a questionnaire to determine (1) what types of groundwater data are generated and how it is managed, (2) which agencies are outside users of this data, (3) which agencies are sources of groundwater data, (4) what opportunities exist for improving accessibility to groundwater data among agencies, (5) what resources are available for data management, (6) what needs are met by current groundwater data management systems, (7) what needs are not being met by current data management systems, (8) what data management functions and features would be most desirable, (9) what additional requirements might arise due to anticipated future programs, (10) if there are any contractual bounds in effect that might restrict data management options, and (11) what the costs are of maintaining current data management systems. Most questionnaire respondents were contacted in person.

Sewall's Phase 1 report, available from the State Planning Office, provides specific information on (1) the types of groundwater data in Maine, (2) current forms of groundwater data management, (3) accessibility to groundwater data, (4) State groundwater data management needs, (5) desirable functions and features of a mutually accessible groundwater data management system, and (6) conclusions and recommendations on how to proceed with Phase 2. These results of the Phase 1 study are summarized below.

Currently, twenty-five state and federal groundwater programs generate groundwater information. The types of data generated include (1) ownership information, (2) map locations, (3) water quality information, (4) well construction and installation information, (5) well yields, (6) hydrogeologic descriptions, and (7) map delineations of bedrock fractures and sand and gravel aquifers.

Most information is filed manually. As for large data collections, searches through these files are very cumbersome. Requests for information are difficult and expensive to fulfill. Complicated queries of the type needed to analyze trends in the quality and quantity of the State's groundwater resources are practically impossible.

A number of computerized data management systems are also used to manage groundwater related information. They include FRAMEWORK, WATSTORE, HONEYWELL DM-IV, MEGIS, BURROUGHS DataManager and RBASE 400, LOTUS, dBase III, PCS, STORET, and a number of nameless IBM PC-XT systems. These vary in capability from simple data storage and retrieval to unsophisticated geographic information system functions.

In general, little compatibility exists between groundwater data management systems currently in use. The lack of a common data organization scheme is the principal cause of this and is a major obstacle to the development of an integrated and mutually accessible groundwater data management system.

A number of opportunities currently exist to improve accessibility to groundwater information. These range in complexity from a simple groundwater newsletter and information index to implementation of a geographic information system to sort data based on map location. The costs of implementation also vary a great deal. However, each element will contribute greatly to overall data accessibility.

The specific objectives of Phase 2 include development of (1) technical designs (generic blueprints) for data management configurations which address the needs listed above; (2) recommendations for specific software / hardware systems; (3) system implementation recommendations; (4) quality control recommendations; (5) security and accessibility recommendations: and (6) staffing roles and recommendations for overall system organization; .

The project was organized in four major tasks. These are (1) review past data management reports and recommendations, and the Phase I report on State agency systems in use, and current and projected information needs; (2) formulate a generic blueprint for data management options which addresses the needs identified in our Phase 1 report, (3) identify specific systems suitable for state-wide groundwater data management, and (4) incorporate performance of Tasks 1 through 3 in a report with findings and recommendations to the Data Management Committee.

Generic Groundwater Data Management Options

Many levels of functionality and complexity can be selected for a State groundwater data management system. These options vary in capability, cost, staffing, and effort required for implementation. The initial selection of a particular system should be a compromise between realistic budgetary constraints and the State's high priority groundwater database management needs (e.g., improved accessibility, reduced duplication of effort). This initial selection should also allow "upward integration" to higher levels of functionality as future priorities and budgetary constraints evolve.

Three levels of data management are recommended to the State (1) groundwater data index and groundwater data coordinator; (2) a database management system; and (3) an integrated geographic information system / database management system.

A groundwater data index would serve as a source of current information on availability, location, and format of groundwater data in the State, but it would not necessarily store all of the data itself. This index would be a road map for individuals requiring access to groundwater data available from State government. The role of the groundwater database coordinator would be to facilitate access to data and maintain the index.

Groundwater data in Maine exists in both mapped and non-map forms. For example, significant sand and gravel aquifer boundaries exist primarily as mapped data. Water quality information, on the other hand, exists as alpha-numeric data. Most groundwater data in Maine are in the non-map category. Non-map data are best managed with a database management system (DBMS). DBMSs are sophisticated computer programs that can be used to store, edit, sort, organize, analyze, and retrieve alpha-numeric data. Commercial DBMS software packages are available for most computers and with some effort can be configured specifically to meet groundwater data management requirements. A typical DBMS configuration consists of a host computer system, DBMS software, and a printer.

A geographic information system (GIS) is also used for storing, editing, sorting, organizing, analyzing, and retrieving information, but it is specifically designed to deal with mapped data. In addition to software and a host-computer, GISs utilize (1) digitizing tablets for transferring the locations of points, lines, and polygons from maps to the computer, (2) graphics terminals for displaying this information once digitized, and (3) plotters for drawing maps from the computerized data. An integrated geographic information system / database management system (GIS / DBMS) combines both GIS and DBMS features in one package. An integrated GIS / DBMS is most useful and appropriate for managing and analyzing a database consisting of both mapped and non-map information (e.g., groundwater data).

Each level of data management described above represents a different stage of functionality, complexity, and cost. Each level also allows upward integration to the next. For example, a groundwater data index can be upgraded to a full database management system, which can in turn be integrated with a geographic information system at still a later time. Each level also requires that subordinate levels be already implemented. For example, an index and coordinator are necessary prerequisites for any state-wide groundwater DBMS. Likewise, an operational DBMS would greatly enhance the State's ability to take full advantage of an integrated GIS / DBMS capability.

If this three level concept is implemented in a multi-phase manner, then each element of the State's overall groundwater data management system can become operational over a relatively short period of time with demonstrable utility and benefits. Additionally, gradual implementation based on a solid experience record will allow the State to carefully assess actual benefits realized from the system and respond appropriately by controlling acquisition, implementation, and staffing costs. The following sections describe each system option in generic terms. Specific system recommendations are given in the following chapter.

Groundwater Data Index / Groundwater Data Coordinator

All viable groundwater database management options available to the State involve an index of groundwater data managed by the State and a groundwater data coordinator. A prototype computerized index has been developed by James W. Sewall Company. This prototype lists all State programs which generate groundwater data, and provides for each program the following information:

- State agency
- Name, address, telephone number of responsible agency contact
- Location of data files
- Categories of groundwater data kept, parameters recorded
- Period of record
- Chain of command from collection to storage
- Frequency of access
- Data quality control indicator
- Indicator of geographic locatability of wells, map boundaries, etc.

Sewall's prototype index is programmed in dBase III+ on an IBM PC-AT micro computer. This relatively simple program supports computerized queries and report generation and is easily used. A listing of index menus and an example of a report generated from the index are provided in Appendix E. Computerized versions of the index are available from Sewall for a nominal fee.

It is recommended that the index be maintained by the State Groundwater Coordinator, who would be responsible for updating the system on a semi-annual basis and distributing computerized files or hardcopy listings to interested parties. This coordinator would also be charged with facilitating inter-agency and outside access to groundwater data listed in the index.

With some additional information (geographic coverage, water quality parameters recorded, etc.) and at a relatively low cost ($6-8K) , the prototype index can be expanded to an operational system according to additional specifications established by the Groundwater Data Management Subcommittee, State Land and Water Resources Council.

Data Management System (DBMS)

Separate databases are currently being used by some State agencies for groundwater data management. Stand-alone systems provide each agency with total control over its database (data formats, access, security, scheduling equipment, cost, etc.). However, these systems are generally not compatible in terms of software and data file organization. The result of this incompatibility is that different agencies cannot easily share or exchange data files and cannot perform complicated queries and analysis using data from more than one agency at a desired level of efficiency. This greatly limits the state-wide utility of these stand-alone systems. To overcome this limitation, we make the following recommendations:

- *Use only one DBMS software package*. Only one DBMS software package should be used by all State agencies involved with groundwater data management. Multiple installations of a single DBMS will simplify procedures for copying and transferring files between computers and agencies. The use of only one software package will also reduce staff training requirements and reduce duplication of DBMS programming efforts.

- *Adopt a DBMS with a relational query language and programmer interface.* A relational DBMS should be used to promote flexibility. Data within such systems are stored in files consisting of multiple "records" which in turn consist of one or more "fields". For example, a groundwater quality database for sand and salt storage piles might consist of one or more data files. These files are organized into records containing water conductivity data and measured chloride concentration and other data. Each record corresponds to a particular well and observation. At a finer level of detail, conductivity and chloride measurements are stored in assigned storage fields within each data record. Relational database systems store such information in fixed-formats and define relationships among data elements by indexes on storage fields. This relational scheme enables a user to make queries by specifying what data he or she requires; the system always knows how to retrieve it. Through a programmable query language, this non-procedural approach enables an operator to phrase queries very quickly and efficiently.

- *Maintain a standardized set of data elements.* All DBMS groundwater records should posses a minimal set of standardized data elements which can be used to relate groundwater records to each other (e.g., agency number, station number, geographic location, date, etc.). This set of data elements should be be standardized between all agencies and should include all minimal file header elements required by EPA's STORET system (see Appendix F).

- *Index DBMS entries.* DBMS entries should be referenced in the State's groundwater data index according to predetermined specifications (see above).

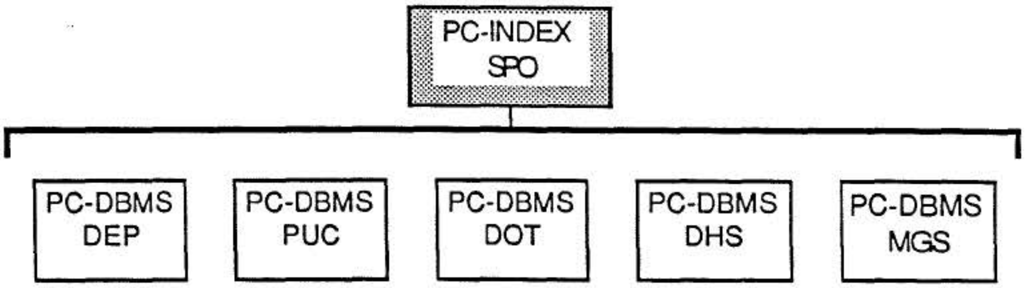
These requirements allow users to access and query multiple groundwater DBMS files that are organized differently. For example, DHS water quality files and MGS well driller's files must be organized differently because of the nature of the data stored. Given a particular task, say "find all wells drilled less than 50 feet deep in MGS files with corresponding chloride concentrations in the DHS files greater than a certain standard", these requirements would work together as follows. First, the groundwater data index is used to determine if the required data exist within these files. Second, having both files compatible with the same DBMS facilities copying and transferring data from both agencies and subsequent access to the actual data. Third, specific data records corresponding to the same well within these files can be identified by site number or geographic location. Finally, a relational file structure and query language allows complicated sorts involving both files in order to identify records corresponding to the same location with data indicating depth greater than 50 feet and excessive chloride concentration.

These requirements are necessary to insure the success of any state-wide DBMS option discussed below. They do not limit the ability of individual agencies to maintain separate databases, and thereby control data formats, accessibility, security, equipment, and cost. These generic recommendations merely impose a structure that permits these separate databases to work together at a desired level of efficiency.

This is not to imply that the implementation of these recommendations is a trivial matter. It is not. At minimum, a new software package has to be purchased and installed, existing computerized data will have to be translated to the new system, data elements may have to be added to some existing records, and staff will require training in use of the DBMS.

The level of efficiency actually realized from a DBMS depends on the configuration adopted for state-wide use. Three generic configurations are described below (1) stand-alone PC workstations; (2) networked PC workstations; and (3) a centralized mainframe system. Each option represents a different level of functionality, complexity, and cost. Each option also allows upward integration to the next. For example, the State can choose to "buy-in" with the stand-alone PC workstations. Later, if justified by high usage or regulatory requirements, these workstations could be networked to one another or integrated with the State's IBM mainframe.

Figure 1. Stand-alone DBMS workstations.



Stand-alone PC workstations

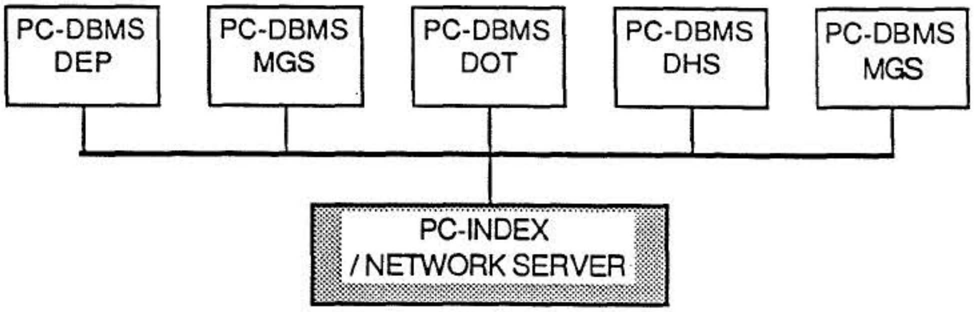
The simplest and least costly configuration consists of independent PC workstations each running the same DBMS software (Figure 1). These workstations would be located with agencies directly involved in groundwater data management. Each workstation would be used to manage only those groundwater files directly attributed to that agency's activities (i.e., DOT's workstation would be used to manage only DOT groundwater data files). All files would be described in the groundwater data index and all groundwater data elements would be related to one another as outlined above.

The main state-wide centralizing feature of this configuration would be the groundwater data index. The groundwater data coordinator would still be responsible for maintaining the index, a function that would involve coordination with the concerned agencies, and would also be responsible for insuring state-wide consistency and quality control. Day-to-day data management would be the responsibility of the individual agency.

With this configuration, permanent groundwater data files could be accessed only through workstations controlled by the responsible agency. However, with approval, these files could be copied onto diskette distributed to other agencies. The groundwater data coordinator would serve as a "project coordinator" for complicated database queries and analyses involving several agencies and multiple groundwater data files.

One recommended option for this stand-alone configuration is to set up one of the workstations with <u>all</u> groundwater DBMS files. Each agency would still maintain its own database on its own workstation, but would provide up-to-date copies of its files for the all-inclusive workstation at regular intervals. This workstation would then serve as a centralized source of groundwater data for any or all inquiries, and it could be used as a platform for staging complicated database queries and analyses involving several agencies and multiple groundwater data files. With currently available and very inexpensive technology, it is also possible to configure this workstation so that remote users could access its data and actually operate the DBMS over the telephone. In order to do so, the remote user would need access to a nearby PC, modem, and

Figure 2. Networked DBMS workstations.



communications software. This mode of operation is analogous to using a mainframe or mini-computer from a remote terminal. The limitation of this scheme is that a PC workstation using the DOS operating system can only support one remote user at a time, during which, it is tie-up and not available to local users. This limitation may not be a problem with proper scheduling.

Networked PC workstations

After stand-alone PC workstations have been in operation for one year, an evaluation of system usage should be performed to assess the need for networking. Networking is a method by which PC workstations can communicate with each other, pass files back and forth, and share peripherals, such as hard disk storage and tape drives (see Figure 2). Some commercial packages allow these workstations to share a single DBMS program and its associated database files. Networks differ from the telephone communication scheme described above in that when one PC accesses another's peripherals or software, the other PC is not tied-up by the first and can be used locally for other tasks.

A number of operationally proven networking packages are commercially available. They generally involve little more than a PC "server", network software (and sometimes hardware), and the cable necessary to span the connected workstations. However, local area networks are currently (1988) limited to distances of approximately 1.3 miles (requires fiberoptic cables). This limitation is insignificant for DEP and MGS, which are located in the same building. Agencies such as DHS, PUC, DOT, and SPO could not become nodes on the network because of distance limitations. However, these agencies could maintain stand-alone PC DBMS workstations which could be linked to the MGS / DEP network server via modems and telephone communications software. This link would facilitate access to groundwater database files for both users of the network and users of stand-alone PC workstations. Two local different local area networks can also communicate with one another via telephone and modem in such a manner that they act a one wide area network. This configuration can be used to overcome the 1.3 mi distance limitation but it tends to be slow.

With a networked configuration each agency would still have total control over its groundwater database. However, with proper approval, other agencies could have direct access to this data if required. Most network packages allow flexible security options that can be used to restrict access to database files. For example, each agency could have two copies of its groundwater database on-line: (1) one for general state-wide access with "read" privileges only , and (2) one for restricted access with full "read-write-edit" privileges.

As with the stand-alone option, the groundwater data coordinator would still be responsible for maintaining the index, a function that would involve coordination with the concerned agencies. The coordinator would also be responsible for insuring state-wide consistency and quality control. The centralizing importance of the index would not be reduced, but the usage of the index would be changed. In this case it would function as an on-line "help" document directly accessible to users of the network, rather than a hardcopy document or diskette updated and circulated at intervals.

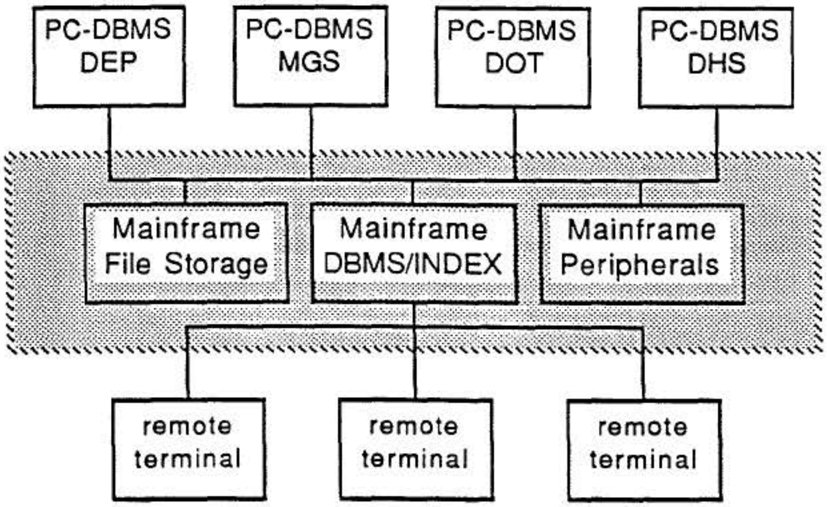
There are several warnings that should be noted about local area networks. First, networking is generally not cost effective for fewer that 10-15 computers. This is particularly true if a large amount of fiberoptic cabling must be installed in order to span nodes separated by long distances. Second, DBMS networks tend to be particularly prone to failures. When a network fails, then all nodes fail (and sometimes visa-versa). Such failures can in some cases result in lost data, and they always result in lost time and effort. However, networking technology is developing rapidly. It is very likely that within the next two to three years the distance limitations described above will no longer be of concern to the State, and that future networks will provide significantly higher levels of performance and reliability than those currently available today.

Centralized mainframe system

Currently, the Bureau of Data Processing is investigating possible acquisition of a programmable relational DBMS for the States IBM mainframe computer. A centralized groundwater data management system would utilize the State's IBM mainframe as an alternative to a local area network of PC workstations. This configuration could utilize the mainframe as a locus for the DBMS or the groundwater database files or both. Groundwater database files residing on the mainframe could be accessed through the several hundred terminals scattered throughout the State or through stand-alone PC workstations. Figure 3 illustrates this concept.

It is recommended that each agency utilize at least one PC workstation that it can use for local management of its permanent groundwater data files. These files would then be up-loaded to the mainframe on a regular schedule from where they could be accessed ("read-only") by other agencies. By using PC workstations in parallel with the mainframe DBMS, each agency would continue to have total control over its permanent data files. Each agency would also be better able to control its costs by performing many database management functions on their PC (e.g., data entry, simple queries).

Figure 3. Centralized mainframe DBMS with PC workstations and remote terminals.



A number of DBMS vendors support very formal and structured networking between PC workstations and a mainframe installation. It is important to note that because these networks utilize the IBM mainframe as a file server, they are not limited by the 1.3 mile distance limitation discussed above.

The mainframe DBMS option will require more planning and cooperation than stand-alone of networked configurations. However, it will greatly facilitate complicated multi-agency queries. The mainframe would also provide maximum "on-line" access to groundwater data and could serve as a link to EPA's STORET and USGS's WATSTORE groundwater data management systems.

As with the options above, the groundwater data coordinator would still be responsible for maintaining the index, insuring state-wide consistency and quality control, but may also be involved in programming, system updating, system backups, and user training. The latter tasks should probably be assigned to a mainframe DBMS coordinator (a Bureau of Data Processing staff member). The groundwater data index would continue to be used as an on-line "help" document for users of the centralized system. However, the index would now be part of the mainframe DBMS and serve as a data dictionary.

Integrated GIS / DBMS

Currently, MeGIS is the State's only geographic information system. This system consists of several software modules which support digitizing, editing, and map making. MeGIS also allows assignment of identifiers, or attributes, to mapped information. Attributes allow the system to distinguish between features, such as lines that are roads and lines that are streams. These attributes can also be cross-indexed to information stored in an outside DBMS. However, MeGIS in not fully integrated with any DBMS. This and its lack of many desirable analytical tools greatly limits its utility as a state-wide tool.

Any GIS adopted for use by the State should be fully integrated with a relational DBMS. A fully integrated GIS / DBMS supports complicated geographic queries and concurrent editing and updating of both map features and their corresponding attributes. We recommend that the GIS component be added only after one of the groundwater DBMS configurations has been implemented and is in operation. The capability to do geographic analysis and digital mapping will be of little state-wide utility until groundwater data have been computerized. This two-phased implementation will also allow time for review of the DBMS and reassessment of the need for a in-house GIS (i.e., a DBMS may satisfy a majority of the needs of most State agencies - less frequent GIS analyses could be performed by an outside service bureau). Phasing will also allow the State to defer the purchase of a GIS for approximately one year during which time the capabilities of most systems will increase and costs will decrease.

Two GIS / DBMS configurations are recommended for consideration : (1) single-user workstations that stand-alone or are networked; or (2) a multiple-user configuration with remote workstations. In either case, the GIS / DBMS workstations would provide complementary GIS capabilities to any of the DBMS configurations described in the previous section. At the same time, State agencies could continue using their DBMS systems for more routine groundwater data management functions (e.g., data entry, simple reports, etc.).

Single-user GIS / DBMS

Single-user GIS / DBMS workstations could be added to the States groundwater data management system as stand-alone workstations (Figure 4) or could be integrated with a DBMS network (Figure 5). Each GIS workstation would be comprised of a micro-computer, graphics monitor, digitizer tablet, plotter, and the necessary GIS / DBMS software. Optional peripherals might include a tape drive, modem, and additional software.

We recommend that the State start with just one GIS / DBMS workstation. A single workstation would be most productive if operated by MGS because of that Agency's existing experience with GIS and its relatively high level mapping related activities. Other GIS / DBMS workstations could be added if justified by high usage levels.

In any case, a GIS Coordinator will be required. The responsibilities of the GIS Coordinator will augment those of the Groundwater Database Coordinator, and will consist largely of computer programming, scheduling equipment usage, system backups, and quality assurance of map products. If more than one single-user GIS / DBMS workstations were established, then each of the respective agencies will require their own GIS coordinator.

Figure 4. Single-user GIS / DBMS with stand-alone PC DBMS workstations.

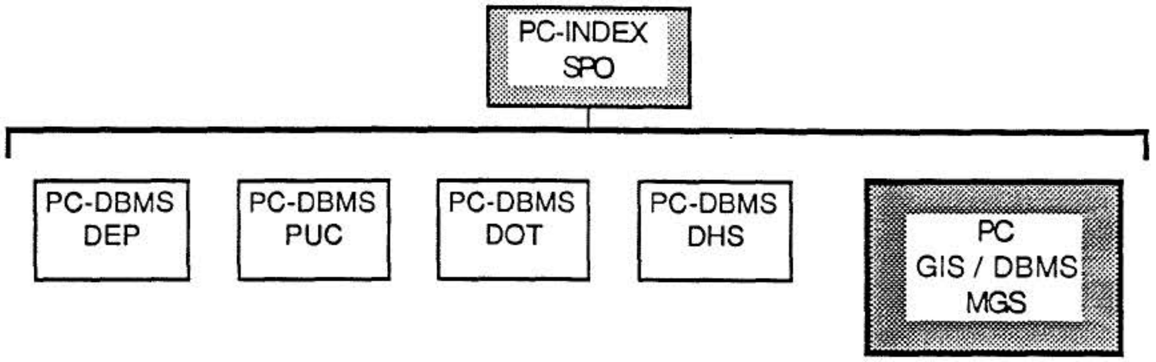
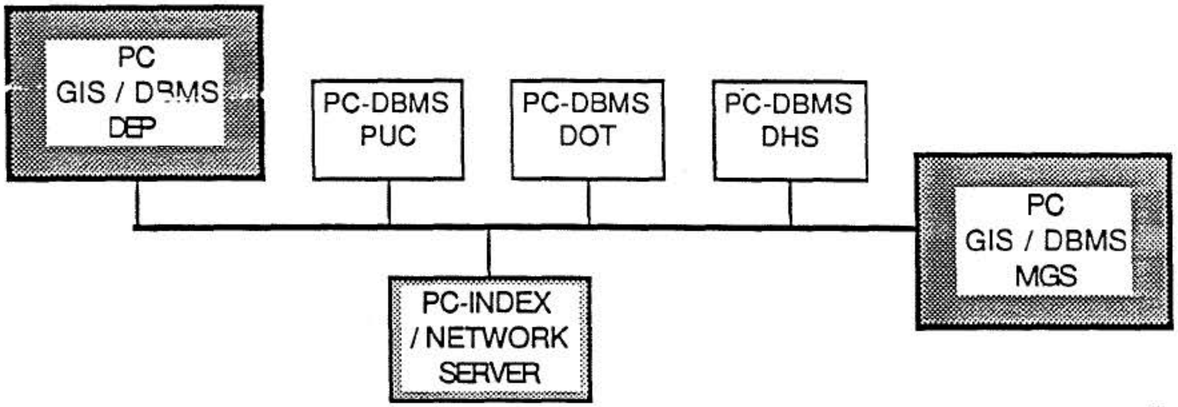


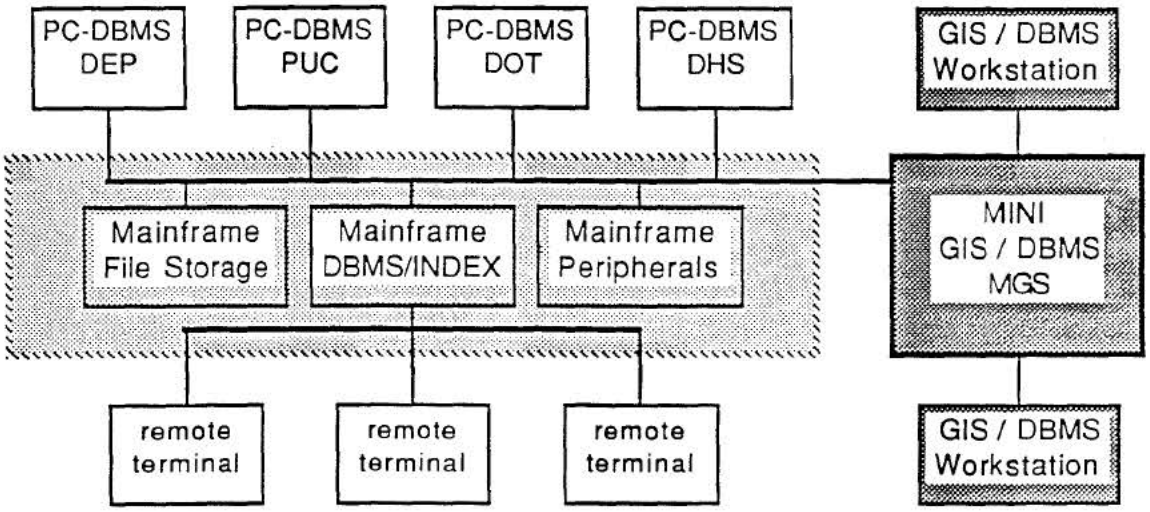
Figure 5. Single-user GIS/ DBMS workstations tied into a PC network. Similar workstation could also be tied into a centralized mainframe DBMS.



Multiple-user GIS / DBMS

A mini-computer based GIS / DBMS could be added to any of the DBMS configurations described in the previous section. A mini-computer based system offers a number of advantages over its single-user counterpart: (1) the system can support multiple workstations; (2) increased computing power; and (3) better utilization of peripherals. However, the buy-in cost and subsequent maintenance costs of a mini-computer system are considerably higher than those of one or even several single-user systems.

Figure 6. Multiple-user GIS / DBMS with remote workstations tied into a centralized mainframe DBMS with PC workstations and remote terminals.



Each of the GIS / DBMS systems recommended in the following section are available in both single and multiple-user configurations. If desired, the State could buy-in with a single-user GIS / DBMS and later upgrade or expand to a mini-computer system capable of supporting more than one user at a time.

Figure 6 illustrates a mini-computer based GIS / DBMS with multiple workstations linked with a centralized mainframe groundwater database management system. This mini-computer configuration could also stand alone or become part of a network. A particularly effective mini-computer configuration consists of (1) a central graphics, digitizing, editing, and plotting workstation; (2) data entry / digitizing workstation with graphics; and (3) several remote graphics workstations for data analysis. We recommend that the central workstation be operated by MGS. Data analysis workstations could be located and operated by other agencies with a need for GIS capabilities.

As with the single-user GIS / DBMS option described above, a GIS Coordinator will be needed for the mini-computer based system. The Coordinator's technical duties would be similar to those described above but would also involve coordinating the use of one system by several agencies. From a staffing efficiency point of view, this may be the State's most desirable option.

System Recommendations

Database Management Systems

A database management system (DBMS) should be used to manage non-graphic / non-map groundwater data. The overall rational for a DBMS and various DBMS configurations were discussed in the previous charter. The following sections provide (1) selection criterion for commercial systems, (2) descriptions of recommended systems together with costs, and (3) an evaluation of EPA's STORET system as an alternative to a State-owned DBMS.

DBMS selection criterion

Approximately one hundred and seventy five database management systems were evaluated (see Appendix A). Five systems are listed below for consideration by the State. Three of these systems (Ingres, Info, Oracle) were selected based on the following criterion:

- Relational structure in which record fields are explicitly defined and can be related to one another in a variety of ways.

- Programmable query language that allows a user to write application programs that perform complicated queries on the related database fields.

- Availability for IBM PC and mini-computer and IBM mainframe to permit upward integration from stand alone to networked to centralized mainframe configurations.

- Fully integrated with a commercial GIS so that geographic information analysis and digital mapping capabilities to be integrated with the DBMS at a later time.

Two systems not meeting the above criterion are also included. dBase III+ is included because it is currently being used operationally by DEP, DOT, and PUC for managing a limited subset of the State's groundwater data. dBase III+ is only available on PCs and is not fully integrated in any commercial GIS / DBMS packages. EPA's STORET system is included because of it is widely used for groundwater data management on the national level, and because its cost is subsidized by EPA. This system is currently used in a limited way by DEP. STORET is not relational in structure and it is not integrated with a commercial GIS. STORET is described under the heading "Outside Vendor Services" because it is operated by EPA and is available for use only on EPA's IBM mainframe.

Recommended DBMS options - descriptions and costs

Oracle, Info, Ingres, dBase III+, and STORET are compared in Table 1 in terms of system type considerations, capabilities, data entry and reporting, data processing capabilities, networking, and general user considerations. Summary

Table 1. DBMS comparison

System	Ingres	Oracle	Info	STORET	dBase III+
Type of system					
Integrated with a reputable GIS	yes	yes	yes	no	no
IBM Mainframe version available	yes	yes	yes	---	no
Honeywell mainframe version avail.	no	yes	yes	---	no
Honeywell mini version available	no	yes	yes	---	no
Mini versions available	yes	yes	yes	---	no
IBM PC version available	yes	yes	yes	---	yes
Supports main-micro-PC network	yes	yes	yes	terminal	no
General capabilities					
Security options	yes	yes	yes	yes	yes
Data dictionary	yes	yes	yes	yes	yes
Large rec & field capacity	yes	yes	yes	fixed	yes
Sort / index multiple fields	yes	yes	yes	yes	yes
Data entry and reporting					
I / O with error checking	good	good	good	good	fair
Import / export ASCII files	yes	yes	yes	yes	yes
Update / edit	good	good	good	good	good
Adequate report generation.	yes	yes	yes	fixed	yes
Processing capabilities					
Internal processing	good	good	good	limited	good
Processing speed and efficiency	good	good	?	poor	good
Query capability	good	good	good	good	good
4th. generation language	good	good	good	no	fair
Networking					
Local area networks	no	yes	no	---	yes
PC - mainframe networks	yes	yes	yes	---	no
User considerations					
Turn-around / access	good	good	good	slow	good
Other groundwater users	national	national	in-state	in-state	in-state
Software updates	yes	yes	yes	yes	yes
Quality documentation	yes	yes	yes	yes	yes
Used for groundwater	no	no	yes	yes	yes

information and costs are provided below for each system. PC DBMS costs cited below do not include computer hardware. IBM PC-AT compatible computers cost $4-8K depending on the vendor and options. Costs for IBM mainframe configurations cited below do not include system charges levied against users by the State's Bureau of Data Processing. Typical IBM mainframe costs are summarized in Table 2.

Table 2. IBM mainframe computer charges

Session connect time	$ 0.00
Dial up access	0.00
Dedicated communications ports	350.00 / month
Electronic mail	0.00
Page printing per page	0.025
Printing per line	0.0012
Disk Storage	3.697 / MB / month
CPU usage per second (depending on operating system)	0.51-0.79

Additionally, it should be recognized that the prices quoted below for commercial systems are subject to a great deal of variation. For example, different prices for the same software configuration were quoted by different sales representatives of the same vendors. There is flexibility when negotiating prices with vendors, and it is likely that costs will be lowered with competitive bidding. For these reasons, the prices given below must be considered as approximations of actual cost (±10%).

DBase III+

DBase was one of the first relational database management systems on the commercial market. Ashton-Tate, the vendor, is a reputable company with an established track record in the DBMS field. Most significant for the user, dBase III + supports a wide range of peripheral software, is extremely well documented, and offers support through the vendor and through users groups.

DBase III+ uses Ashton-Tate's proprietary query language which is easy to learn and easy to use. An experienced programmer can learn to produce substantial application programs relatively quickly. The same language is used in an interactive command mode. Industry reports suggest that dBase III+ will eventually support Structured Query Language (SQL). Compilers, such as Clipper, are currently available to link dBase III+ with high level programming languages as "C" and Fortran thereby producing a great deal of data processing power.

DBase III+ will run on IBM PC-XT and PC-AT computers currently utilized by all State agencies involved in groundwater data management. A single copy of the software costs $695. Multiple copies are listed at $500 per installation. A local area network version is available for $1690 for a group of up to five users. Each additional set of five users would cost an extra $995.

A note of caution should be offered. dBase III+ has not been integrated in any GIS / DBMS packages which meet the State's requirements. A GIS / dBase III+ linkage can be made with most GIS systems through intermediate ASCII files. This process is relatively simple, but operationally inefficient. Additionally, dBase III+ is available for personal computers only. Use of dBase III+ will limit the State's ability to migrate its groundwater database to its IBM mainframe computer in an efficient manner at a future time.

DBase III+ is currently being used by DEP, PUC, DHS, and MGS for managing a very limited subset of the State's groundwater data. The value of this use, particularly in terms of existing user familiarity and satisfaction, should not be underestimated. These considerations may partially offset the negative consequences associated with the PC restriction and the lack of an integrated GIS / dBase III+ package.

We recommend that dBase III+ be adopted for groundwater data management if and only if (1) a centralized mainframe DBMS configuration will not to be considered for future use; and (2) full integration of GIS and DBMS capabilities is not considered important.

Ingres

Ingres, a product of Relational Technology, Inc., has existed as a large mainframe relational database for a number of years. A PC version of the software has appeared on the market recently. Mini-computer versions are also available. Ingres is highly rated and has captured a great deal of attention in the DBMS field. Ingres uses SQL as well as a proprietary visual query language. Modules are available to allow Ingres users to take advantage of the power of fourth-generation programming (4GL) languages such as "C". These modules also allow 4GL programs to have imbedded SQL commands that take advantage of Ingres query capabilities.

The SQL language is quite easy to learn in its basics. It is extremely powerful for database queries and query combinations. Its statistical processing is limited, but easy to implement. For example, a simple linear regression could be programmed in a single command consisting of three lines. For complex processing and advanced statistical interfacing, the 4GL module referred to above or a statistical package such as SPSS would be desirable.

Ingres pricing is extremely variable depending upon configuration. Basic Ingres software for a single PC workstation costs $950. We anticipate that several of these workstations will require an additional programming / applications module at a cost of $500 each. Annual software updates can be purchased for $375 per year. Site support is an additional $2K per year. A discount may be offered for quantity purchases, perhaps 10% for 5 or more packages.

Ingres does not formally support a multi-user local area network (i.e, multiple PC workstations using a single copy of the software). However, PC workstations each running a separate copy of Ingres can be networked to facilitate access to data and sharing of peripherals. The cost of this configuration would be as above, plus the cost of the network (approximately $1.5K per workstation for IBM's Token Ring Network).

Basic Ingres modules for the State IBM mainframe would cost approximately $90K. The additional application / programming language module is an extra $27K. The annual support, maintenance, and update fees total $18K per year. This mainframe version can also be networked to IBM PC workstations running Ingres at a cost of approximately $20K plus the cost of software for each workstation. It is important to note that, because it utilizes the IBM mainframe, this network is not limited by the 1.3 mile distance limitation discussed previously.

In addition to the capabilities described above, Ingres is fully integrated with GeoBased's Strings GIS package. Strings / Ingres provides an excellent solution to the State's geographic information analysis and digital mapping needs.

Oracle

Oracle, a product of Oracle Corp., resembles Ingres in its overall relational functionality and uses SQL as its query language. Oracle also provides an interface to 4GL programming languages such as "C". Technical literature suggests that it differs to some extent in processing procedures making it slower at some tasks and faster at others. As far as its ability to process geographical information goes, those who use ORACLE and are familiar with it recommend it without reservation.

A single PC version of the software costs $1300. ORACLE Corporation recommends local area networking for users with three or more PCs. This network would consist of PC workstations using a single copy of the software which resides on a dedicated PC file server. The price of DBMS software only for this configuration is $2,495 for the server and $695 for each PC node added to the network plus approximately $1.5K per workstation for IBM's Token Ring Network. Maintenance, software updates, and user support are an additional $495 per year per PC. Oracle also offers a corporate plan for $1200 per year plus $120 per year per PC.

As discussed in the previous sections, a local area PC networking scheme would work well for DEP and MGS which are located in the same building. Agencies such as DHS, PUC, DOT, and SPO could not become nodes on the network because of distance limitations. However, these agencies could maintain stand-alone PC Oracle workstations which could be linked to the MGS / DEP server via modems and telephone communications software. This link would facilitate access to groundwater database files by users of the network and users of stand-alone PC workstations.

An IBM mainframe version of Oracle's DBMS software costs $90K. An additional "forms" report generation module costs $22.5K. SQL and PC-mainframe networking capabilities (similar to Ingres) are an additional $18K. A 4GL language module costs $13.5K. Maintenance, software updates, and user support are an additional 12.5% of the purchase price per year.

Oracle's a very well regarded DBMS which has been integrated with two commercial GIS systems - KGIS, a product of KORK Systems of Bangor, and GeoVision, a product of GeoVision Corporation of Ottawa, Canada. Both GIS / DBMS systems are described in later sections of this report.

Info

Info, a product of Henco Inc., has been successfully utilized by USGS and other U.S. Department of Interior agencies for a number of years. It is also currently used by the Connecticut, Massachusetts, and New Hampshire for groundwater data management making it an appealing choice from a compatibility standpoint. Info is functionally similar to Ingres and Oracle in that it is a programmable relational DBMS. It can be used in conjunction with 4GL languages and is available in PC, mini-computer, and IBM mainframe versions. Info does not utilize SQL, but rather it uses Henco's own proprietary query language.

An IBM mainframe version of Info costs approximately $30K. Maintenance, software updates, and user support are an additional 12% per year of the software price ($3.6K).
PC versions of the software are offered either with or without the capability to be networked with the IBM mainframe DBMS. PC software capable of networking with the mainframe costs $995-1200 depending on quantity. PC versions without this capability cost $790-950. Maintenance, software updates, and user support are an additional $400 per year per PC workstation.

Similar to Ingres, Info does not formally support a multi-user local area network (i.e, multiple PC workstations sharing a single copy of the software). However, PC workstations each running a separate copy of Info can be networked to facilitate access to data and sharing of peripherals. The cost of this configuration would be as above, plus the cost of the network (approximately $1500 per workstation for IBM's Token Ring Network). The distance limitations cited above are also applicable to Info. Note that there are no distance limitations on an Info network that utilizes the State's IBM mainframe.

Info is also an integral part of Arc / Info, a product of Environmental Systems Research Institute (ESRI). Arc / Info was selected by USGS as its GIS / DBMS in part because that agency was already using and satisfied with Info. Arc / Info is widely used for groundwater data management.

EPA's DBMS - STORET

STORET (Storage and Retrieval) is one of EPA's computerized database systems for environmental monitoring data relating to the quality of water within the United States. Use of STORET is an alternative to a State-owned DBMS. There are over 9500 unique water quality parameters in STORET's water quality file, and it contains data from more than 270,000 groundwater monitoring stations (160,000 are from USGS's WATSTORE system). The system is widely used on the national level by EPA, and by a number of states, including California, for state-wide groundwater data management.

The system runs on EPA's mainframe computer located at Research Triangle Park, North Carolina. Users access the system using PCs as dumb terminals over telephone lines and modems or through dedicated computer communication links. EPA currently provides Maine with a $5000 per year credit to offset system usage costs. Modems and the software necessary for telephone communications cost less than $1K per workstation. Telephone charges are additional. The cost of establishing a dedicated link to STORET is $10-15K depending on options. Annual linkage fees are additional.

STORET is currently being used by DEP, but in a limited manner for groundwater data. DEP has recently installed a dedicated link to STORET that permits rapid and reliable communications. Other agencies such as DHS and DOT can also utilize STORET for water quality data. Some non-water quality groundwater data can be managed with the system (e.g., sediment, biological, and facility information). However, data formats are fixed by a rigid hierarchical structure. As a result, using this system to manage certain other classes of groundwater information will be very difficult (e.g., test boring logs and well drillers logs).

STORET is very well documented and EPA offers substantial user support. STORET provides many data query and statistical programs for analyzing water quality data. However, these are "fixed" in that they perform very specific functions which may limit their utility for analyzing other classes of groundwater related data. The system does not provide the flexibility of a programmable relational DBMS, but it is well structured and does what it does very well.

STORET provides some geographic analyses and digital mapping options that will satisfy many of the immediate needs of DEP, DOT, and SPO. These options will not, however, satisfy the needs of MGS. STORET has features that will facilitate data transfer to another DBMS or GIS, but it is not integrated in any commercial GIS / DBMS packages. This may limit the State's future groundwater data management options.

STORET 's centralizing features and price makes it a very appealing alternative to current groundwater data management practices. From a technical point of view, however, we recommend that STORET be adopted by the State only as a final alternative to the programmable relational DBMS options described above.

Integrated GIS / DBMS

The need for a mutually accessible groundwater DBMS is widely recognized within the State. The need for enhanced GIS capabilities is also recognized. A GIS / DBMS that integrates these two functions would be considerably more powerful than would either taken separately. Systems that integrate these capabilities also offer an opportunity for a phased implementation (i.e., implement a groundwater DBMS first, then integrate it with GIS capabilities).

This rational was discussed in detail in the preceding chapter. The following sections provide information on (1) GIS / DBMS selection criterion, (2) four recommended commercial systems, (3) approximate costs, and (4) opportunities for utilizing outside vendor services.

GIS / DBMS selection criterion

Approximately seventy-five geographic information systems and digital mapping systems were evaluated (see Appendix A). Of these, most are too specialized to meet the State's needs (e.g., facilities management, land parcel management). Many provide only marginal analysis capabilities. Some have only rudimentary DBMS capabilities. Four systems are recommended to the State based on the following criterion.

- Sufficient versatility and processing power to meet State geographic information analysis and digital mapping needs identified in Sewall's Phase I report.

- Fully integrated with a programmable relational DBMS that can also be purchased and operated independently of the GIS.

- Compatibility with State computer standards (i.e., SNA / SDLC 3270 emulation)

- Software vendor with a strong track record of financial stability, software development, service, and user support.

The commercial GIS / DBMS options recommended to the State are (1) Strings/ Ingres, a product of GeoBased Systems; (2) KGIS / Oracle, a product of KORK Systems; (3) Arc / Info, a product of Environmental Systems Research Institute (ESRI); and (4) GeoVision / Oracle, a product of GeoVision Corp..

Recommended GIS / DBMS options - descriptions and costs

The four recommended GIS / DBMS options are described in Table 3 in terms of data entry procedures, editing options, retrieval and analysis capabilities, map output and display options, and other considerations. These systems should also be compared in terms of the DBMS capabilities listed in Table 1. As can be seen, Strings / Ingres, Arc / Info, and GeoVision / Oracle are very similar in term of capabilities. KGIS / Oracle is new to the commercial market, with some system capabilities still under development. This system is included because its vendor, KORK Systems, is a well established and very highly regarded developer of digital photogrammetric and mapping software. We anticipate that in approximately one year KGIS / Oracle will be equal or superior to the other options described below. Additionally, KORK systems is located in Bangor and can provide a very high level of user support.

Table 3. GIS/ DBMS comparison.

Vendor	GeoBased	KORK	ESRI	GeoVision
GIS Product	**Strings /**	**KGIS /**	**Arc /**	**GeoVision /**
Internal DBMS	Ingres	Oracle	Info	Oracle
Data entry				
manual grid cell	yes	no	yes	yes
digitizer grid cell	yes	no	yes	no
polygon	arc-node	yes	arc-node	arc
input formats (SIF, DLG)	yes	yes	yes	yes
Editing				
auto error detection	yes	yes	yes	yes
find and replace single records	yes	yes	yes	yes
global edits	yes	yes	yes	yes
Retrieval and Analysis				
polygon-raster conversion	yes	no	yes	no
locate data in specific area	yes	yes	yes	yes
edge detection / matching	yes	yes	yes	yes
distance calculation	yes	*	yes	yes
polygon attribute summary	yes	yes	yes	yes
cell or polygon statistics	yes	yes	yes	yes
area statistics	yes	*	yes	yes
nearest neighbor search	yes	*	yes	yes
proximity analysis	yes	*	yes	yes
summarize points in polygon	yes	yes	yes	yes
calculate slope from elevation	no	no	yes[1]	?
calculate length of slope	no	no	yes[1]	yes
define watershed boundaries	yes	no	yes[1]	yes
determine slope aspect	no	no	yes[1]	?
compute acreage	yes	yes	yes	yes
assign weights to categories	yes	yes	yes	yes
change size of grid cell	yes	no	yes	yes
single map layer operations	yes	*	yes	yes
overlay	yes	*	yes	yes
shortest path calculation	yes	*	yes	yes
corridor analysis	yes	*	yes	yes
Map output and display				
grid-cell maps	yes	no	yes	yes
polygon maps	yes	yes	yes	yes
vary map scales	yes	yes	yes	yes
user defined titles	yes	*	yes	yes

? Undetermined

* Being developed -projected 1988 release

** Available with higher level Oracle license

1 Not available for PC Arc / Info

Table 3 continued. GIS/ DBMS comparison.

Vendor GIS Product Internal DBMS	GeoBased **Strings /** Ingres	KORK **KGIS /** Oracle	ESRI **Arc /** Info	GeoVision **GeoVision /** Oracle
Map output and display . . cont.				
user defined symbols	yes	yes	yes	yes
user defined legends	yes	*	yes	yes
perspective views	yes	no	yes	?
optional grid / tic marks	yes	yes	yes	yes
surface relief portrayal	yes	no	yes	?
histograms	yes	yes	yes	**
pie charts	yes	yes	no	**
graphs / scatter plots	yes	yes	no	**
text reports	yes	yes	yes	yes
CRT zoom	yes	yes	yes	yes
define window	yes	yes	yes	yes
trim to fit window	yes	?	yes	yes
mosaic maps	yes	yes	yes	yes
SIF and DLG outputs	yes	yes	yes	yes
Organic considerations				
vendor track record	good	good	good	?
technical / service support	yes	yes	yes	yes
data dictionary	yes	yes	yes	yes
software support and updates	fee	fee	fee	fee
DBMS linkage	good	good	good	acceptable
SNA / SDLC 3270 emulation	yes	yes	yes	yes
tool kit for user applications	yes	yes	yes[1]	yes

? Undetermined
* Being developed -projected 1988 release
** Available with higher level Oracle license

[1] Not available for PC Arc / Info

Each system described below can be installed as one or more independent workstations, or a centralized multi-user configuration with one or more remote workstations. Costs depend highly on the exact system configuration chosen. The discussion below provides expected cost ranges for single user and multiple user configurations. A single-user configuration consists of a GIS / DBMS workstation with full digitizing, editing, analysis, and plotting capabilities. A centralized multi-user configuration consists of a mini-computer GIS / DBMS with full digitizing, editing, analysis, and plotting capabilities together with seven remote graphics workstations for editing and geographic analysis only. The single user configurations described below represent a low cost entry level system. The mini-computer configurations with seven remote workstations represent high cost options at the other end of the spectrum. The exact configuration ultimately chosen by the State will likely be something between these two extremes.

Table 4. Typical hardware costs.

IBM PC compatible -	$ 6-8K
DEC VAXstation 2000 mini-computer-	27K
DEC MicroVAX II mini-computer -	45K
PRIME mini-computer -	80K
Tektronix graphics workstation -	10-15K
SUN graphics workstation -	20-25K
1600 BPI Tape drive	6-8K
Standard printer -	1K
High speed printer -	2-12K
Plotter -	8-12K
Digitizing tablet -	6-10K

Table 4 provides a general guide to costs for specific hardware components. As with the DBMS costs above, it should be recognized that the prices quoted below for specific GIS / DBMS configurations are subject to a great deal of variation. Hardware maintenance costs vary depending on item, number of units, and degree on-site service desired. These are generally 5-10% of the purchase price per year. We strongly recommend maintenance contracts for all major hardware components purchased. There is flexibility when negotiating prices with vendors, and it is likely that costs will be lowered though the competitive bidding process. All system prices given this report are for vendor recommended configurations and must be considered as estimates only (within 10% of actual costs).

KGIS / Oracle

KORK Systems, Bangor, Maine is currently developing a geographic information system, KGIS with DBMS functions handled by Oracle. The first two installations of this system are under contract for delivery at the beginning of 1988 (one at USGS). As it currently exists or will in the near future, KGIS has all the features needed for a first class GIS. James W. Sewall Company has had a long and very satisfactory relationship with KORK Systems, a well established and very highly regarded developer of digital photogrammetric and mapping software. KORK's proximity to Augusta will enhance quick response by the vendor to user needs.

KORK's single and multiple user configurations are respectively described in tables 5 and 6. Both configurations utilized computers from the DEC VAX family. These machines are known for data processing power and reliability. A single-user workstation costs approximately $53K. A multi-user configuration with seven remote Tektronix workstations each with mouse editing and analysis capabilities costs approximately $200K. Installation, training, and annual maintenance and user support fees are additional.

Table 5. KGIS / Oracle single-user configuration

Hardware	Software
1. VAXstation 2000 computer	1. KGIS / Oracle software
2. Plotter	2. System software
3. Digitizer	
4. Graphics display (high resolution)	
5. Printer	
6. 1600 BPI tape drive	

System cost -		$ 52, 500
Other costs -	shipping, installation, training	5,700
	annual hardware maintenance and support	3-5,000
	annual software maintenance and support	2,400
	annual GIS software updates	1,750

Table 6. KGIS / Oracle multi-user configuration

Hardware	Software
1. MicroVAX II computer	1. KGIS / Oracle software
2. Plotter	2. System software
3. Digitizer	
4. Graphics display (high resolution)	
5. Printer	
6. 1600 BPI tape drive	
7. Seven remote Tektronix editing and analysis workstations	
8. Communications hardware	

System cost -		$ 200, 000
Other costs -	shipping, installation, training	6,300
	annual hardware maintenance and support	8-16,000
	annual software maintenance and support	5,100
	annual GIS software updates	3,500

VAXstation 2000 and MircoVAX II computers can be linked to IBM PC compatible Oracle workstations via (1) an application interface which communicates through serial ports on both machines; (2) networking software utilizing special boards on both machines; or (3) serial ports, modems, and communications program such as KERMIT. For a fee, all vendors listed in this section, including KORK, appear to be willing to tailor their systems to enhance linkages between their GIS / DBMS configurations and IBM PC compatible DBMS workstations.

Table 7. GeoVision / Oracle single - user configuration

Hardware	Software
1. VAXstation 2000	1. GIS / Oracle
2. Plotter	2. Advanced mapping software
3. Digitizer	3. System software
4. 15 inch monochrome display	
5. Printer	

System cost -		$ 50,000
Other costs -	shipping, installation, training	1,500
	annual hardware maintenance and support	2-4,000
	annual software maintenance, support and GIS software updates	1,800

GeoVision / Oracle

GeoVision Corp., offers a strongly interactive GIS which is integrated with Oracle. GeoVision has been in business since 1974, with headquarters in Ottawa, Canada and has subsidiaries in Denver, Colorado and Sydney, Australia. GeoVision's clients include city and county governments, utilities, survey and mapping organizations, and consulting engineers.

GeoVision's configurations (tables 7 and 8) are similar to KGIS / Oracle in their use of DEC VAX computers, but their recommended single-user system differs slightly from KGIS / Oracle in its use of monochrome graphics terminals rather than color. The cost of this system is approximately $50K. A color graphics terminal can be added for an extra $20K.

GeoVision offers two versions of its mini-computer based system. The first costs around $200K and supports geographic information analysis from remote workstations but not GIS editing (note that KGIS / Oracle supports both). The second configuration provides full GIS editing and analysis capabilities at remote workstations, but requires VAXstation 2000 computers with software at an additional cost of $35-40K for each station desired. For seven remote workstations, the cost of this second configuration is $450-480K. The cost mini-computer system installation and training will vary between $5-10K depending on the State's specifications ($500 per day plus expenses).

Table 8. GeoVision / Oracle multi-user configuration

Hardware	Software
1. MicroVAX II	1. GeoVision / Oracle
2. Plotter	2. Advanced mapping software
3. Digitizer	3. System software
4. Graphics display	
5. Printer	
6. 1600 BPI tape drive	
7. Seven remote Tektronix analysis workstations (no editing)	
8. Communications hardware	

System cost -		$ 200,000
Other costs -	shipping, installation, training	5-10,000
	annual hardware maintenance and support	8-16,000
	annual software maintenance, support and software updates	5,400

Strings / Ingres

GeoBased Systems, Research Triangle Park, North Carolina, is recognized nationally and internationally as a leader in geographic information systems. GIS software can reside on either a PC AT compatible or MicroVAX II mini-computer. Unlike many companies that built their original software for a large system and then attempted to cut it down to fit on smaller computers, GeoBased develops its systems on personal computers. Hence, GeoBased's stand-alone PC workstations have the same GIS / DBMS capabilities as their more expensive mini-computer configurations.

Single and multiple user workstation configurations are listed in tables 9 and 10. GeoBased's recommended multi-user configuration for the State differs from those of other vendors in its use of multiple PC Strings / Ingres workstations linked to the States IBM mainframe computer. In this configuration, a software package called "Arbitor" is used to connect each PC workstation to the mainframe and through it to all other PC workstations. This linkage allows PCs to share mainframe disk storage in a way that is functionally transparent to each individual user. With this configuration, each Agency could have its desired level of independent GIS / DBMS capabilities on PCs and also the ability to quickly access DBMS and GIS data files created by other agencies (with permission). Under this scheme, the mainframe would play an archiving role for groundwater data (i.e, desired data files would be stored on the mainframe for general access and regularly backed up).

Table 9. Strings / Ingres single-user configuration

Hardware	Software
1. PC with graphics monitor	1. Strings / Ingres software
2. Plotter	2. System software
3. Digitizer	
4. Printer	
5. 1600 BPI tape drive	

System cost -		$ 56,000
Other costs -	shipping, installation, training	5-9,000
	annual hardware maintenance and support	2-5,000
	annual software maintenance, support, and software updates	1,800

Table 10. Strings / Ingress multi-user configuration

Hardware	Software
1. Assumes presence of State IBM mainframe w / tape drives	1. Arbitor (mainframe / PC communication software)
2. Central PC GIS / DBMS workstation including digitizer and plotter	2. Strings / Ingres software
3. Seven remote PC editing and analysis workstations	

System cost -		$ 200,000
Other costs -	shipping, installation, training	5-9,000
	annual hardware maintenance and support	9-17,000
	annual software maintenance, support, and software updates	3,600

Note that GeoBased can also deliver a multi-user MicroVAX configuration with Tektronix and PC remote workstations with prices in the same range as KORK. The advantage of the configuration shown in Table 10 is the opportunity that it offers for incremental phasing-in of independent or linked PC GIS / DBMS workstations using the IBM mainframe without acquisition of a mini-computer.

Table 11. PC Arc / Info single-user configuration

Hardware	Software
1. PC AT or compatible	1. PC Arc / Info software
2. Plotter	2. System software
3. Digitizer	
4. Graphics display (high resolution)	
5. Printer	
6. 1600 BPI tape drive	

System cost -		$ 45,000
Other costs -	shipping, installation, training	see below
	annual hardware maintenance and support	2-4,000
	annual maintenance, support, and GIS software updates	1,000

Arc / Info

Environmental Systems Research Institute (ESRI), Redlands, California, has a long track record of excellence in the GIS field. Arc / Info has been the system of choice for many state and federal governments. Arc / Info has a record of use in the groundwater field, and selection of this system would automatically assure compatibility with most other New England States, EPA, and USGS (although data could also be translated to or from any of the other systems described in this report). The system is reported to be somewhat less interactive than some others, but its wide range of capabilities can be seen from Table 3. The recently released version of the software

Arc / Info is available for a wide range of systems ranging from PCs to mainframe computers. A configuration is available for the State's IBM mainframe, however, it is not recommended because of slowness engendered by the configuration of the IBM. A software remedy to this problem is expected in mid-1988. PC Arc / Info does not possess all of the capabilities available from mini-computer versions of the system. Major differences are related to 3-D topographic data processing and analysis and the sophistication of its user applications development toolbox. A single-user PC Arc / Info system can be configured for approximately $45K plus annual maintenance and support (see Table 11). ESRI does not offer on-site installation and user training for PC Arc / Info, although a video taped tutorial can be purchased for $350.

A multiple-user minicomputer version of the system in listed in Table 12. On first glance, its $330K price tag seems relatively high. But this high cost can be attributed to seven SUN micro computer workstations recommended by ESRI at

Table 12. Arc / Info multi-user configuration

Hardware	Software
1. MicroVAX II or Prime 2455 computer	1. Arc / Info software
2. Plotter	2. System software
3. Digitizer	
4. Graphics display (high resolution)	
5. Printer	
6. 1600 BPI tape drive	
7. Seven remote SUN editing and analysis workstations	
8. Communications hardware	

System cost -		$ 225-330,000
Other costs -	shipping, installation, training	18,000
	annual software and hardware maintenance, support, and software updates	10% of system price

$20-25K each. ESRI can also configure lower cost multi-user systems. SUN workstations can be replaced by Tektronix graphics terminals in order to lower the system costs to approximately $225K. This configuration is comparable to the other mini-computer options described above.

ESRI is currently developing an interface to Oracle. This package will be similar to Arc / Info in its GIS features, but will posses advanced DBMS query capabilities not available from Info (SQL). The interface is due for release in 1988 and will be available as an add-on to Arc / Info. The cost of the interface will be approximately $10K. This does not include the cost of Arc / Info or Oracle. Additional information pertaining to this option will be available from ESRI in mid-1988.

Outside vendor services

A number of private firms offer geographic information analysis and digital mapping services on a contractual basis. Most work closely with clients to define their data analysis needs and desired output formats. These firms accept mapped or non-map data from a client and convert it to their GIS / DBMS system. Some also accept computerized DBMS files (e.g., with groundwater data). A number of these firms operate powerful systems such as Strings / Ingres and Arc / Info with full DBMS, geographic information analysis, and digital mapping capabilities. Several of these also manage geographic databases for clients.

The services offered by these firms are a cost-effective alternative to acquisition of GIS capabilities in addition to a state-wide groundwater DBMS. A 1986 study performed by MGS estimated the hourly cost of operating a MeGIS workstation at approximately $40 for digitizing and editing. Some vendors can provide services at similar rates. Some consultants also have extensive digital geographic databases with information relevant to the State groundwater data analysis needs (e.g., political boundaries, topography, land cover). These can be incorporated in a project at costs substantially below those the State would pay to convert the data to a digital format for its own use (costs are spread out among several clients). The disadvantages of using outside vendor services include potential project scheduling conflicts and cumbersome administrative requirements associated with contracting consultant services on a project by project basis. Open ended contracts for GIS / DBMS support services to multiple State projects may be a partial solution to the latter problem.

System Implementation Recommendations

This section provides general recommendations and considerations for implementation of a state-wide groundwater data management system. It includes (1) a phased system implementation plan, (2) DBMS data format recommendations, (3) quality assurance / quality control recommendations, (4) compatibility and interaction with other data management systems, (5) accessibility and security considerations, (6) long-term staffing requirements, and (7) implementation costs estimates.

Implementation plan

An overall plan for implementing a state-wide groundwater data management system is given in Figure 7. The plan is based on a phased approach that provides opportunities for upward system integration. It consists of five components: (1) implementation of an index listing available data, (2) addition of stand-alone DBMS workstations, (3) PC workstation networking, (4) addition of a centralized mainframe DBMS, and (5) integration with GIS capabilities. Each component is described below. Special attention should be given to monitoring groundwater program requirements and user feedback in order to determine at what point in time each of these elements should be considered for inclusion in the system.

Groundwater data index / groundwater data coordinator

Implementation of the groundwater data index is outlined in Figure 8. The Groundwater Data Coordinator position should be filled using existing staff resources - preferably someone very familiar with the State's existing groundwater programs (e.g., State Groundwater Coordinator). Index implementation starts with development of content and function specifications. We recommend that a qualified DBMS programmer be retained to bring Sewall's prototype index to full operation according to these specifications. Upon completion, the index should be distributed to State agencies involved with groundwater in a hardcopy catalogue form or on diskette. Diskette versions should be accompanied with training materials and usage tutorials.

Stand-alone DBMS workstations

Implementation of stand-alone DBMS workstations is outlined in Figure 9. Detailed DBMS implementation specifications should be developed using information contained in this report and additional user input. DBMS selection should be made with future networking, mainframe, and GIS options in full consideration. Implementation should be phased by groundwater program and prioritized according to program requirements, projected time requirements, data quality, usefulness of data, anticipated costs, and other factors (see Table 13). We recommend that a consultant be contracted to format DBMS files, develop query and analysis programs, translate existing computer files, supervise

Figure 7. Phased implementation plan.

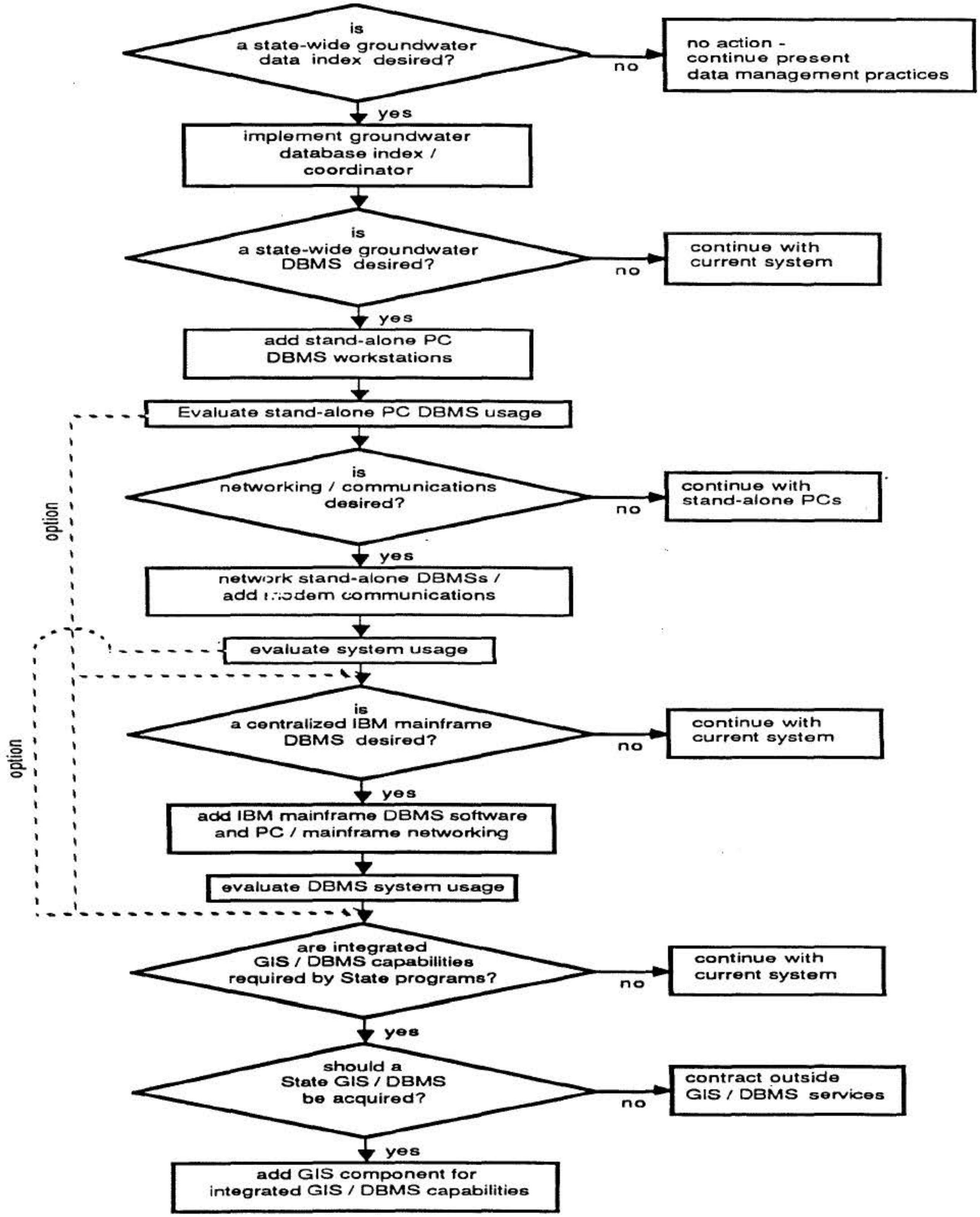


Figure 8. Groundwater data index implementation guidelines

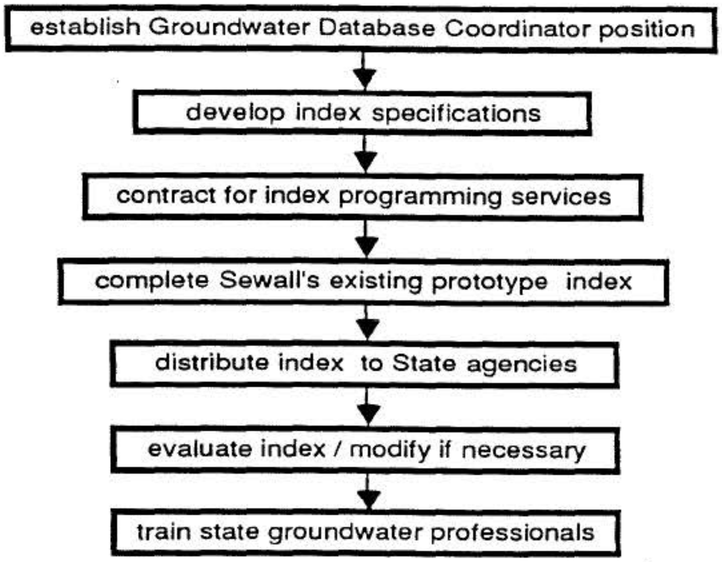


Figure 9. Stand-alone PC DBMS implementation guidelines.

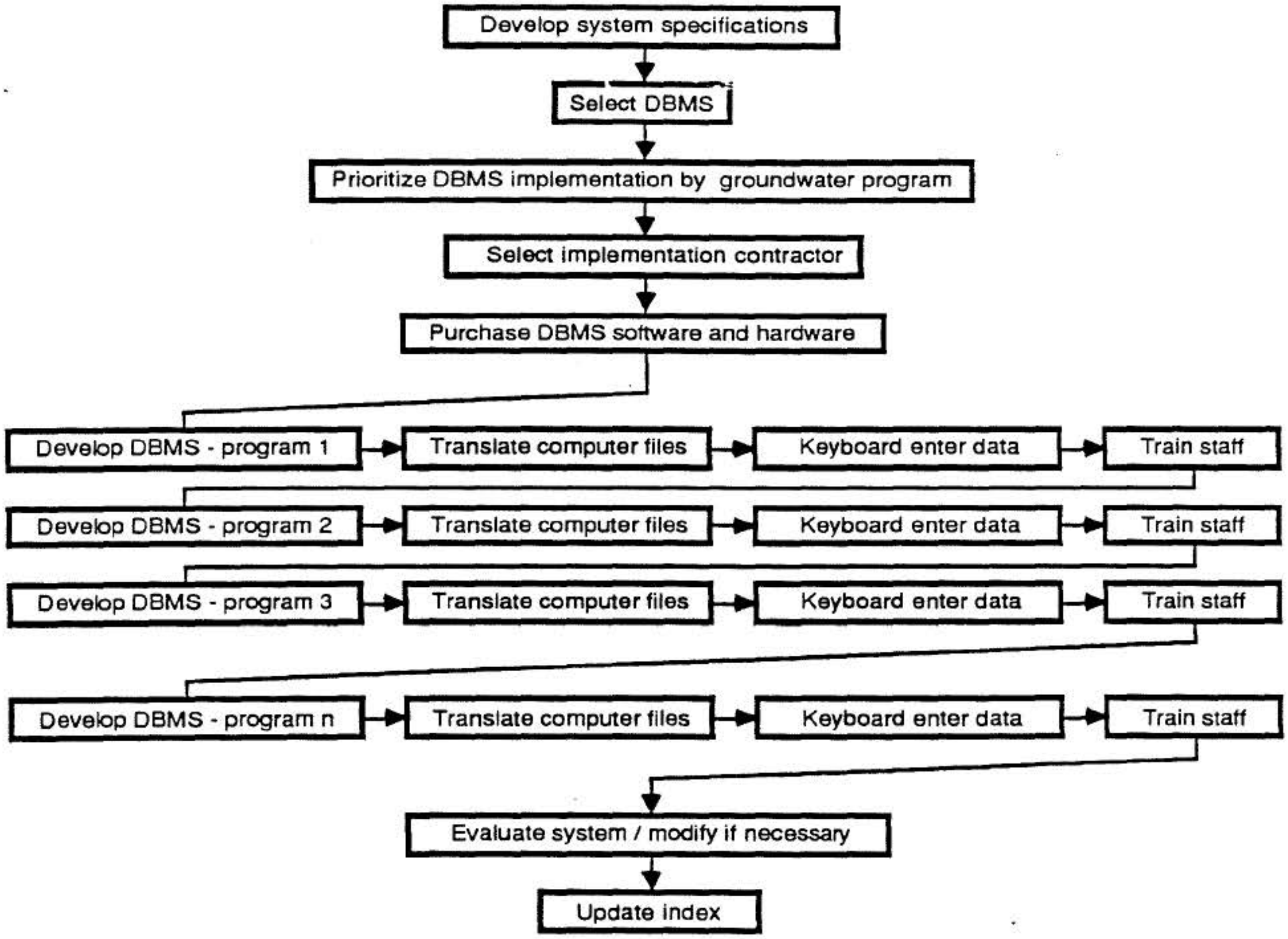


Table 13. DBMS and GIS / DBMS implementation factors .

		GEOGRAPHIC INFORMATION SYSTEMS IMPLEMENTATION					DATABASE MANAGEMENT SYSTEMS IMPLEMENTATION					
AGENCY	PROGRAM	MEGABYTES STORAGE	COSTS	TIME	LOCATABILITY BY COORDINATE	GIS ELEMENT	MEGABYTES STORAGE	COST	TIME	PROCEDURE	QUALITY	ACCESS FREQUENCY
DEP	Sand and Salt Pile Monitoring	-	-	-	GOOD	LOCATION	10	HIGH	LONG	CONVERT/ENTER	HIGH	INFREQUENT
	Sand and Gravel Aquifers	-	-	-	GOOD	SEE MGS	2	LOW	SHORT	CONVERT	HIGH	RARE
	Waste Discharge Licenses	-	-	-	FAIR	LOCATION	2	MEDIUM	SHORT	CONVERT/ENTER	MODERATE	FREQUENT
	Complaint Response	-	-	-	-	-	2	MEDIUM	SHORT	ENTER	HIGH	RARE
	Licensing and Enforcement	-	-	-	GOOD	LOCATION	2	MEDIUM	SHORT	ENTER	HIGH	V FREQUENT
	RCRA Sites	-	-	-	GOOD	LOCATION	2	MEDIUM	SHORT	ENTER	HIGH	INFREQUENT
	Superfund Sites	-	-	-	GOOD	LOCATION	2	MEDIUM	SHORT	ENTER	HIGH	INFREQUENT
	Underground Storage Tanks	-	-	-	-	-	20	MEDIUM	SHORT	CONVERT	MOD-POOR	V FREQUENT
	Oil Spill Sites	-	-	-	TOWN ONLY	LOCATION	4	LOW-MED	SHORT	CONVERT	HIGH	FREQUENT
MGS	Sand and Gravel Aquifer Mapping	60	HIGH	SHORT	GOOD	ENTIRE MAP	-	-	-	-	HIGH	FREQUENT
	Bedrock Aquifer Mapping	9	MEDIUM	SHORT	GOOD	ENTIRE MAP	-	-	-	-	MOD-POOR	FREQUENT
	Regional Lineament (Radioactive Wastes)	10	MEDIUM	SHORT	GOOD	ENTIRE MAP	-	-	-	-	HIGH	FREQUENT
	Well Drillers Information	-	-	-	FAIR	LOCATION	20	HIGH	LONG	CONVERT/ENTER	MODERATE	FREQUENT
DHS	Public Water Supply Development	-	-	-	GOOD	LOCATION	4	MEDIUM	SHORT	CONVERT/ENTER	MODERATE	RARE
	Public Water Supply Monitoring	-	-	-	VARIABLE	LOCATION	6	HIGH	SHORT	CONVERT/ENTER	MODERATE	FREQUENT
	Private Well Analysis	-	-	-	-	-	10+6/yr	HIGH	LONG	CONVERT/ENTER	MODERATE	FREQUENT
MDOT	Sand and Salt Storage Pile Ranking	-	-	-	GOOD	LOCATION	2	MEDIUM	SHORT	CONVERT/ENTER	HIGH	FREQUENT
	Preconstruction Information	-	-	-	GOOD	VARIABLE	4	MEDIUM	SHORT	CONVERT/ENTER	HIGH	INFREQUENT
	Pollution Claims	-	-	-	FAIR	LOCATION	2	MEDIUM	SHORT	CONVERT/ENTER	HIGH	RARE
	Exploratory Borings	-	-	-	FAIR TO GOOD	VARIABLE	4	MED-HI	SHORT	CONVERT/ENTER	HIGH	INFREQUENT
PUC	Water Utilities	-	-	-	TOWN ONLY	LOCATION	2	LOW	SHORT	CONVERT	MODERATE	INFREQUENT

COSTS: Low - Less than $2K
Medium - $2K - $10K
High - Greater than $10K

TIME: Short - Less than 6 Months
Long - Greater than 6 Months

PROCEDURE: Convert - Convert Existing Computer Files
Enter - Keyboard Data Entry

QUALITY: Refers to quality of data that would be included in database

Table 14. System conversion guidelines

DEP Sand and salt pile storage monitoring - Framework currently stores all field and lab data -
Link to DBMS through agency number, map coordinates, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates and station number via ASCII flat files - GIS used for geographic analysis

DEP Underground storage tanks - Honeywell DM-IV currently stores administrative data
Link to DBMS through agency number and station number - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by town centroid map coordinates and station number via ASCII flat files - GIS used for mapping distribution

DEP Oil spill sites - IBM PC-XT Spreadsheet currently stores administrative and laboratory data
Link to DBMS through agency number, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid and station number via ASCII flat files - GIS used for geographic analysis

MGS Sand and gravel aquifer maps - MeGIS currently stores several maps
Link to DBMS not necessary
Link to GIS via DLG translator - GIS used for map development, revision, and analysis

MGS regional lineament mapping - *MeGIS* currently stores several maps
Link to DBMS not necessary
Link to GIS via DLG translator - GIS used for map development, revision, and analysis

MGS Bedrock aquifer mapping - Burroughs data manager currently stores well data
Link to DBMS through agency number, map coordinates, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of center of Delorme grid and station number via ASCII flat files - GIS used for geographic analysis and map development

MGS drillers' well information - Burroughs data manager currently stores well data
Link to DBMS through agency number, map coordinates, station (well) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of center of Delorme grid and station number via ASCII flat files - GIS used for geographic analysis and map development

DHS Public water supply development - dBase III currently stores field and lab data
Link to DBMS through agency number, map coordinates, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates and station number via ASCII flat files - GIS used for

DHS Public water supply monitoring - dBase III currently stores administrative and lab data
Link to DBMS through agency number, map coordinates, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates and station number via ASCII flat files - GIS used for analysis

---

Table 14 continued. System conversion guidelines

DHS Private well analysis - dBase III currently stores administrative data
Link to DBMS through agency number, map coordinates, station number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS not possible due to lack of map information

MDOT Sand and salt pile ranking - IBM PC LOTUS currently stores administrative, lab, and ranking data
Link to DBMS through agency number, station (pile) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid (or map location from DEP) and station number via ASCII flat files - GIS used for geographic analysis

MDOT Preconstruction information - dBase III used for administrative and laboratory data
Link to DBMS through agency number, station (job, well) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid (or map location if available) and station number via ASCII flat files - GIS used for geographic analysis

MDOT Pollution claims - dBase III starting to be used for administrative and laboratory data
Link to DBMS through agency number, station (job, well) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid (or map location if available) and station number via ASCII flat files - GIS used for geographic analysis

MDOT Exploratory borings - dBase III being considered for exploratory boring data
Link to DBMS through agency number, station (job, boring) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid (or map location if available) and station number via ASCII flat files - GIS used for geographic analysis

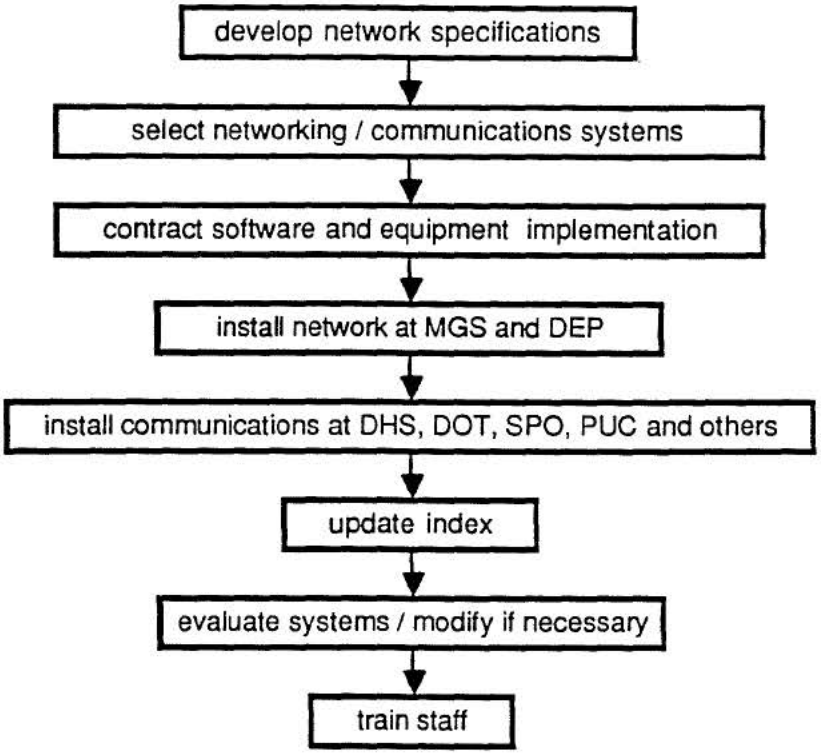
PUC Water utilities - dBase III used to store administrative and water usage information
Link to DBMS through agency number, station (utility, well) number and date - transfer data via ASCII flat files - DBMS used to manage all data
Link to GIS by map coordinates of town centroid (or map location if available) and station number via ASCII flat files - GIS used for geographic analysis

DBMS / GIS linkage with USGS's WATSTORE and ARC / INFO systems
Linkage to WATSTORE should not be considered
Link with INFO through agency number, station number, map location (point data), and date via ASCII flat file
Link with ARC via DLG file format

DBMS / GIS linkage with EPA's STORET and PCS
Link with both via STORET agency code, station code, date and parameter codes via ASCII flat file and modem -
Recommend that State data systems adhere STORET station header standards and overall organization for compatibility

---

Figure 10. PC network implementation guidelines.



keyboard data entry, and train users. Table 14 provides detailed guidelines for converting data already existing in computer compatible formats.

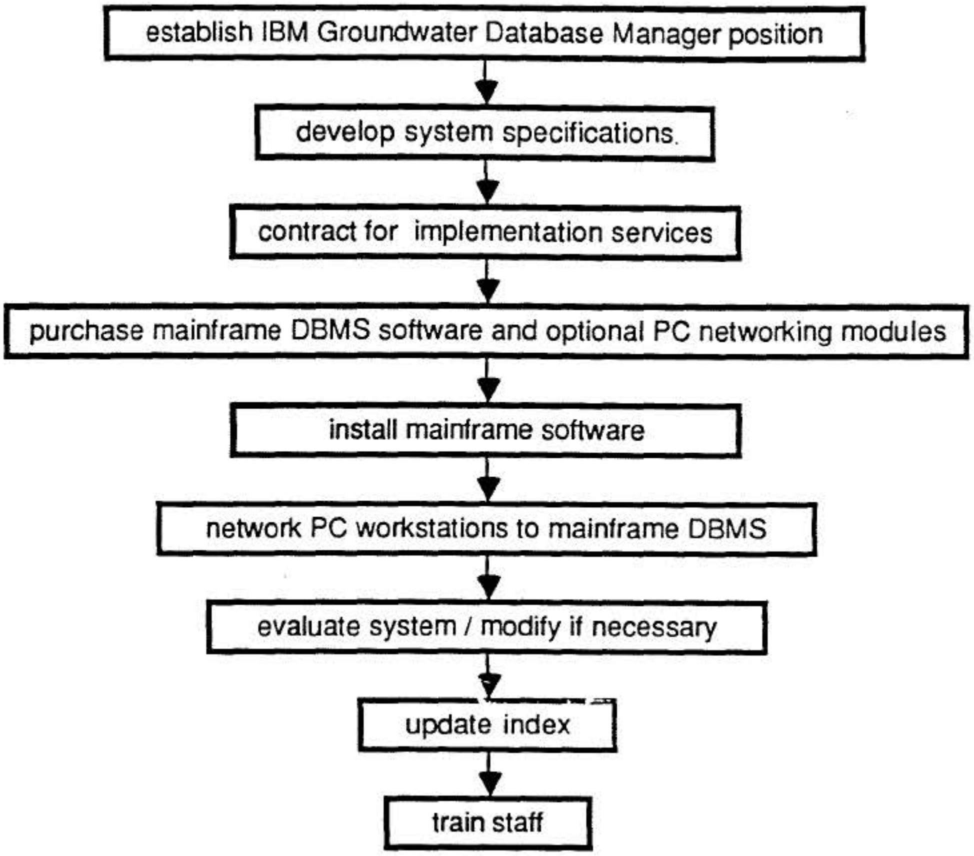
PC workstation network

Figure 10 provides guidelines for networking PC based groundwater DBMS workstations. Network specifications should identify number and placement of PC workstations. These specifications should also provide detailed information on the function of each network node in terms of where software and groundwater data files will reside. Peripherals such as tape drives and disk storage should also be considered. Again, we recommend that implementation be contracted to a qualified consultant. This contractor would be responsible for linking MGS and DEP workstations with a local area network and other agencies via telephone modems and communications software. The contractor would also be charged with training staff in use of the network and communications system.

IBM mainframe DBMS

Figure 11 provides guidelines for IBM mainframe DBMS implementation. These include establishing a part-time mainframe Groundwater Database Manager who would be charged with developing specifications and supervising implementation. Specifications should include detailed information on networking PC workstations and the mainframe DBMS package. Under this plan the Database Manager would be charged with installation of mainframe DBMS

Figure 11. IBM mainframe DBMS implementation guidelines.

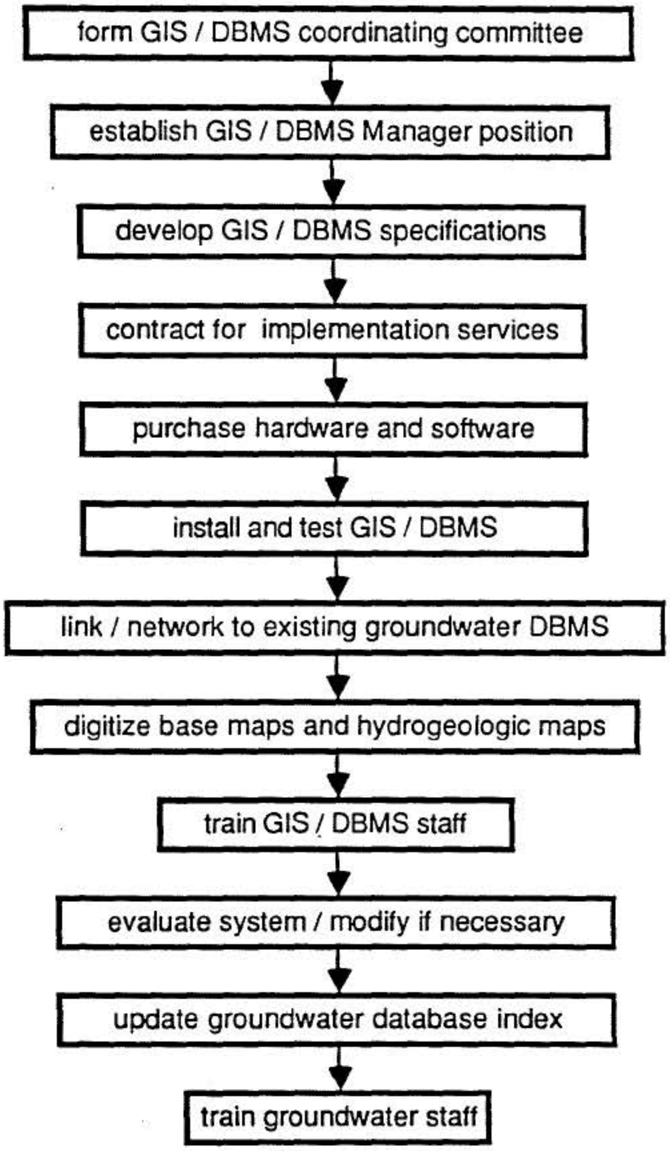


software. However, the State may choose to retain a contractor to install software and hardware for the PC-mainframe network. This should be done under the supervision of the Database Manager. The Database Manager will also be responsible for system user training.

Integrated GIS / DBMS

Guidelines for integrating GIS with an operational groundwater DBMS are given in Figure 12. The operation of a GIS / DBMS will require an inter-agency Coordinating Committee and a GIS / DBMS Manager. The Committee, through the Manager, will be charged with developing specifications for GIS / DBMS software and hardware and integration with operational groundwater DBMSs. A consultant should be retained for system installation, integration, testing, and digitizing base maps and hydrogeologic maps. Extensive training for staff directly involved with GIS / DBMS operation should be provided by the system vendor. Less formal training for other groundwater professionals can be provided either by the system vendor or GIS / DBMS Manager.

Figure 12. GIS / DBMS implementation guidelines.



Recommended DBMS data formats

We recommend that the State adopt EPA's STORET data coding formats for groundwater related data. We do not recommend that STORET's groundwater data file organization be adopted for State use. STORET uses a rigid hierarchical / networked file organization structure which is not well suited to the State's needs. Maine's groundwater database files should be organized to take maximum advantage of relational DBMS capabilities according to State specifications and the professional judgement of the implementation contractor.

STORET groundwater data format descriptions are provided in Appendix F. These data formats have been operationally tested by both state and federal agencies , and their utilization will result in a high level of State compatibility with EPA and USGS. Three broad categories of descriptors are used to document groundwater information available for a particular location. These categories are:

- Station descriptors
- Sample descriptors
- Analytical findings

There are several elements under each category which will enable the DBMS user to describe groundwater data thoroughly. Additional data elements can be added as desired. Elements making up each category of descriptors are described below.

Station descriptors

Factors which are descriptive of the sampling location and which would not change over time are called "station descriptors". There are three types of station descriptors needed to support groundwater database management. They are as follows:

Facility descriptors - descriptors of the operation being monitored, such as type of waste management area (e.g., landfill), facility or site location (e.g., lat / lon, zip code), and type of activity (e.g., private residence, disposer of hazardous waste).

Physical setting descriptors - descriptors of the setting in which the site or facility is located and from which samples were taken, such as aquifer name or geologic formation name.

Well descriptors - descriptors of those characteristics of a well or site which may be an important factor in data analysis and which would not be expected to change over time, such as type of well, well depth, and casing material.

Sample descriptors

Factors that describe a groundwater quality sample at the time it was taken and which are expected to change with each sampling event are called "sample descriptors". Three types of sample descriptors needed by groundwater data managers to support their groundwater monitoring data are:

Sampling purpose descriptors - descriptors of why and by whom a sample was taken.

Sampling condition descriptors - descriptors of the conditions during the sampling event, such as the depth to the top of the water table or water temperature.

Sampling / analysis descriptors - descriptors to document how a sample was taken and/or analyzed, such as how the sample was drawn and whether or not it was replicated.

Analytical Findings

The findings that were determined from each sample at a station are called "analytical findings"/ (e.g., the concentration or arsenic in the sample). Analytical findings will be stored in parametric data fields by using "parameter codes". Descriptions of these codes can be found in Appendix F (STORET Data Storage Requirements). The reader is referred to EPA's STORET User's Handbook (February 1982) for more detailed information.

STORET data formats provide only general provisions for test boring results and well logs. More detailed formats for these types of data can added to the groundwater DBMS according to State specifications.

Quality control recommendations

To insure that the computer database is used properly and used to its fullest advantage, some standards should be noted. These should be reviewed by system users and periodically modified to reflect changing conditions.

Backup

Magnetic disk or tape backups should be performed daily or at a minimum weekly. Hardcopies of each file should be printed periodically. These serve as a textual backup in the event of major database failure. Also, individuals who are not trained to use the computer can refer to these hardcopies.

Program Documentation

Rigid standards of computer program documentation must be observed so that programmers servicing database routines developed specifically for the State can alter programming code, etc., without protracted detective efforts. Critical elements of such documentation include a title block, disclaimer, modification history, general description of program and its function, compile and link instructions, logical unit assignments, description of each subroutine and function used, array descriptions, variable declarations, and common block declarations. This information should be included as part of each computer program.

User Logbook

A system user logbook is an indispensable tool for keeping track of file contents, file locations, and general information about system usage. Upon completion of a given task or procedure, the system user logs his or her actions and results for future reference. The contents of such a logbook becomes particularly important in an environment of high personnel turn-over.

Concurrency control

Concurrency means that all data is current as of the same date. Entries in one data set should be current with entries in another data set. Differences between data sets will produce errors if reports or general queries are made. Concurrency is difficult to control and verify in an operational environment where data are added or edited frequently. Measures to control concurrency are largely procedural in nature and will depend on the final DBMS configuration selected by the State.

Integrity control

Integrity involves the accuracy, correctness, and validity of the data. Data must be reviewed prior to entry on the computer to insure that coding conventions are satisfied and that the numeric ranges, numeric values and spelling are correct. Some error checks can also be programmed into the database management system.

Data dictionary, definitions and coding

A dictionary of required codes and coding conventions should be prepared so that all coding is consistent. The preferred codes should be fully defined so that each attribute can be definitively coded and accurately applied. Strictly enforced definitions should help to eliminate unnecessary redundancy of attribute data.

Audit

There should be a procedure for auditing to document who has used the system and when because potentially sensitive information is included in the database. This can be accomplished with a system logbook as described above, or with the internal auditing functions of some computer operating systems or DBMS packages.

Map base standards

The map base includes a coordinate system and any features which, by their spatial disposition, accurately represent the true position of those features. A fraction of map scale should be quoted as the horizontal accuracy for each map. The coordinate system should be any of three common systems: latitude / longitude; Universal Transverse Mercator; Maine State Plane.

The nominal scale, photographic reference dates, method of construction, map accuracy fraction, date of construction and commissioning agency should be listed for all map bases. As the map bases are transferred digitally or digitized, a plot of the map base should be compared (at the nominal scale of the original) with the original map document.

Thematic map layer organization should be planned so that all currently proposed map layers are coded with unique land symbols. A 100' contour line should be consistently coded from one map to another. A bridge abutment or a well location should likewise be consistently coded and symbolized. This coding must be recorded on a master list for possible update and inclusion of new feature codes and map layers. Where a line may represent several functions (stream, property line, town line, county line), a provision must be made to recognize duplicate purposes.

Additional GIS standards

If a decision is made to computerize graphic / map data, a commitment is automatically made to maintain that data and to "stay current" with software and hardware upgrades. Technology is changing very rapidly and will continue to change. Upgrading will occasionally mean that you can get steered into a cul-de-sac. If data are converted to a system which does not have market strength, system support will diminish, and a search must be initiated for a new system which can utilize that digital data.

Compatibility and interaction with other data management systems

Compatibility between a State Groundwater Data Management System and EPA's STORET system will be insured if the DBMS data format recommendations made earlier are followed. These recommendations are designed to provide meaningful connections between the two databases. This will also insure some level of compatibility with USGS, since WATSTORE is also compatible with STORET. Selection of a GIS / DBMS that supports DLG and SIF file interchange formats will insure compatibility with Arc / Info, which is used extensively by USGS and most other New England states. All GIS / DBMS options recommended in this report support DLG and SIF. Similar measures will ensure compatibility with other databases maintained by the State.

It should also be noted that the hardware and software acquired to support a groundwater data management system may also be used to manage other categories of mapped and non-map information (e.g., forest resources, socio-economic, etc.). This would require a higher level of system coordination and more intensive quality control measures than previously discussed. An assessment of realistic multiple-usage possibilities was beyond the scope of this study. However, these possibilities would only strengthen arguments made to justify the expense of a groundwater data management system.

Lastly, we recognize that at some point in the future a requirement for integrating different State databases will almost certainly arise. In general terms, integration requires the following: (1) access to computers, (2) interaction of operating systems that allow communication of data, (3) interactions between data management systems, and (4) meaningful connections between databases. Items 1 and 2 are insured by the State's computer compatibility standards. Item 3 is the subject of much research and will depend on future developments made by software vendors. Item 4, meaningful connections between databases, depends on data formats and quality control measures implemented by the State. Again, future database integration is beyond the scope of this study, but its requirements should be considered when implementing a groundwater data management system.

Accessibility and security recommendations

Any DBMS configuration adopted by the State should provide accessibility to groundwater data for both State-agency and non-State personnel. Accessibility to the data can be either on-line or off-line depending on the needs and status of the user. General access to groundwater data should be off-line. Personnel unfamiliar with the State's groundwater database or not trained in DBMS operation should start by consulting the groundwater data coordinator and / or groundwater data index. Both would point to the location of hardcopy data, if available, or to the names of designated agency resource personnel who could assist with on-line searches for computerized data.

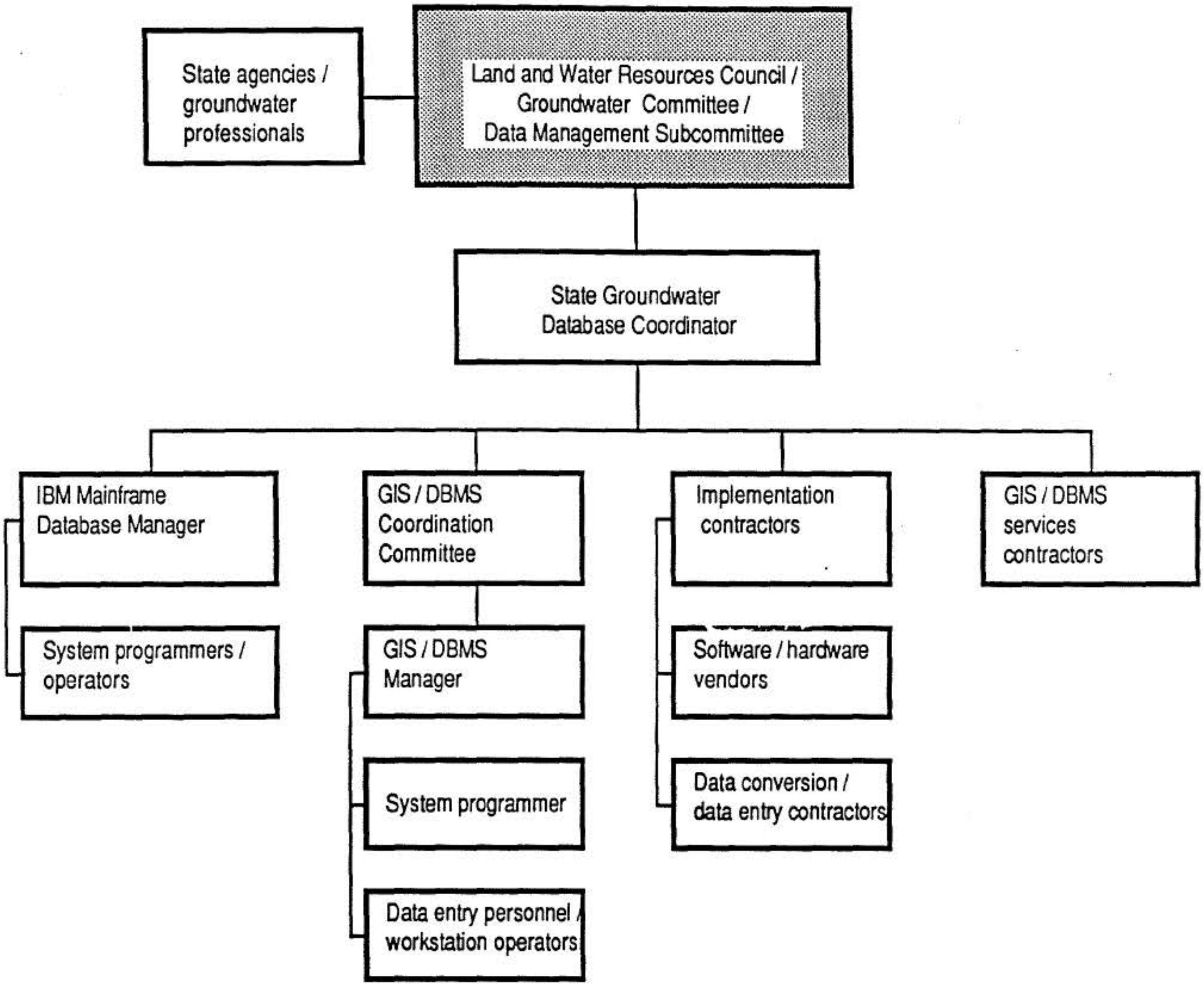
State-agency staff should have on-line access to their own databases as described in the previous chapters. Each of the DBMS configurations recommended in this report also allows inter-agency on-line access, which should be encouraged when appropriate given personnel trained in DBMS operation. The need for additional terminals and workstations should be monitored through periodic review of use patterns.

In special cases, non-State data users may also be allowed on-line access to the groundwater DBMS either through a State agency's workstation or via personal computer, telephone modem, and communications software. Each of the DBMS configurations recommended here allow the latter. However, we recommend that on-line access by non-State data users be avoided when possible and limited to adequately trained personnel when not avoidable.

For reasons of confidentiality, access to some portions of the database must be controlled. Security measures can be defined within the DBMS, or they can be created within a computer's operating systems. The relational DBMSs recommended in this report can be programmed to provide certain security functions. It is suggested that a user matrix be established so that individuals enter unique passwords to gain access to the database. These passwords can be differential in that certain users will have universal access and others will be restricted to "need to know" files. This capability is generally available on most commercial databases.

Overall security will also be enhanced by the following measures: (1) original paper records and micro fiche containing sensitive data and all original electronic media should be kept in secure fireproof areas; (2) on-line data should be time-stamped to document all edits and updates; (3) on-line data should be backed-up at least once a week, with backup-ups stored in a secure fireproof vault

Figure 13. Staffing requirements and tentative organization.



Long-term staffing requirements

Figure 13 provides a tentative organizational chart showing staffing / contractor requirements directly attributed to groundwater database implementation and long-term management. Staff roles have been discussed in context throughout this report. It is not possible to project the exact number of man-year-equivalents (MYE) devoted to establishing and maintaining a groundwater data management system, due to uncertainty in what data management options will be selected by the State. If all options (index, PC DBMSs, network, mainframe DBMS, GIS / DBMS) were adopted, a conservative annual estimate for maintenance alone might be 1.5 MYE.

Table 15. System installation, programming, and training costs estimates.

Option	Installation / programming	Training
Groundwater data index	$ 6-8K	$ 1K
Stand-alone PC DBMSs	15-25K	5-10K
PC network	5K	1K
IBM mainframe DBMS	10-20K	5K
Integrated GIS / DBMS	10-15K	10-15K

Implementation costs

Installation, system programing, and training costs estimates are provided in Table 15. Costs for development of a groundwater data index, DBMS installation, development of DBMS file formats and query and analysis programs for PCs and the IBM mainframe, and PC network implementation will total $37-63K. GIS / DBMS installation will be an additional $10-15K. Staff training sessions will total an additional $22-32K.

Table 13 provides cost estimates for keyboard data entry, computer file conversion, and digitization of existing groundwater data. These costs are listed by groundwater program together with estimated time requirements. The cumulative cost of entering all groundwater data is approximately $200-250K. Of this figure, $180-210K is DBMS related and $40-70K is GIS related. An optional $600-800K will be required for digitizing surface hydrology, transportation, political boundaries, and cultural features from USGS 1:24,000 quad maps to enhance the GIS. This does not include the cost of digitizing topographic contours.

Appendix A

Listing of Evaluated DBMSs and GISs

VENDOR	GIS
AERONCA Electronics	AE-GIS,AE-MAP,VGS-300+
Autometric, Inc.	AUTOGIS
Bliss Associates	BLISSMAP
C. Benson, Lakehead University	Polymapper
Comarc Systems	CIMS,COMPIS,GDMS
ComGrafix	MapGrafix
Criterion, Inc	LANDTRAX
Dipix Systems Ltd.	ARIES GIS
ERDAS, Inc.	ERDAS,ERDAS-PC
ESL, Inc.	GEOMIPS
Environmental Research Institute of Michigan	ERDC-GIS 200
Environmental Systems Research Institute	ARC/INFO;GRID;GRID/TOPO
Forest Data Corporation	TerraPak
GIMMS, Inc	GIMMS
GeoBased Systems	STRINGS
GeoGraphics	GIS overlay for RIPS
GeoVision Corp	GIS,AMS,RAMS/VX
Geogroup	GEO-BASEMAP
HEXXIS, Inc	HEXXIS
Harvard University	ODYSSEY
Hennepin County Bureau of Public Service	UltiMap
Holguin Corp.	Holguin GIS
Integraph Corp.	GPPU,DTM,GDU
Interactive Systems Corp.	AGIS/GRAMS
International Imaging Systems	System 600 GIS
Jet Propulsion Laboratory	IBIS,LUMIS
Lincoln Institute of Land Policy	SOLIR
Minnesota Planning Information Center	EPPL
Morgan Fairfield, Inc.	MICROMAP II
National Park Service GIS Field Unit	SAGIS
PAMAP Graphics, Ltd	PAMAP Systems
Peerless Ingineering Service	PC MAP & PC MAP II
Pennsylvania State University	TOMIS
Planning Data Systems	MULTIMAP
Resources Planning Associates	MAPOVL,RIVBAS,LOWLIB,GDBM,ENDECODE
Riley Datashare International	Micromap
Sammamish Data Systems	DIDS
SecaGraphics	M.A.G.I.C.
South Dakota State University	AREAS
Spatial Information Systems	pMAP
St. Regis Paper Co.	PRO/GIS
Stephen Graham	Schools Enrollment Projection Syst.
Strategic Locations Planning	ATLAS AMP
Synercom	EMIS,SGIS
System Development Corp.	SDCIPS-GIS
TYDAC Technologies, Inc.	Spatial Analysis System (SPANS)
Terra-Mar	T-base,T-mapper,MicroImage
Terrasoft	Digital Terrain Model
Terrasoft	Geographical Information System
The Sidwell Company	SIGNET
Towson State University	MICRO/GIS
US Army Corps of Engineers	Geographic Resources Anal. System
Uni-Graphic Systems, Inc.	Geographics
University of California	SIPS
University of Massachusetts	COMLUP/ILPP
University of Minnesota	County Soil Surv. Inf. System
Utah State University	Micro Sieve
Wild Heerbrugg	System 9

VENDOR	GIS
W. E. Gates and Associates	ADAPT
Yale University School of Forestry	MAP

DBMS	VENDOR
1-2-3, Release 2	Lotus Development Corp.
10 Base	Fox Research, Inc.
ADR Datacom	Applied Data Research, Inc.
Accent R	National Information Systems, Inc.
Adabas	Software AG of North America, Inc
Add System	Westmorland Software International, Inc.
Aladin	Advanced Data Institute, Inc.
Allbase	Hewlett-Packard Co.
Alpha/three	Alpha Software Corp.
Apollo	Schmidt Enterprises
BL8000,BL700,BL300	Britton Lee, Inc.
BOSS	American Planning Corp
BRS/Search	BRS Information Technologies
Base/OE	Information Structures, Inc.
Btrieve	SoftCraft, Inc.
C-Scribe	Unisyn, Inc.
CA-Universe	Computer Assoc. International
CIP	Concentric Data Systems, Inc
Clarion	Barrington Systems, Inc.
ClearCut	Menlosoft/Business Day & Software
Condor #	Condor Computer Corp
Cornerstone	Infocom, Inc.
Courtney Database	Courtney Business Systems, Inc
D The Data Language	Caltex Software, Inc.
DB Aid for DBRC	Financial Technologies International,Inc
DB-FABS	Computer Control Systems, Inc.
DB-General	Bradmark Computer Systems,Inc.
DB2	IBM
DBASE II	Ashton-Tate
DBASE III+	Ashton-Tate
DBS/Experience	Postley Software, Inc
DBXL	Wordtech Systems
DDQuery	Venet-Uiyiams,Inc
DG/DBMS	Data General Corp.
DG/SQL	Data General Corp.
DM	Information Dimensions, Inc
DNA-4	Exact Systems & Programming Corp.
Data Management Software	SAS Institute, Inc.
DataEase	Software Solutions, Inc.
DataPlus-86	Universal Software Co.
Dataflex	Data Access Corp
Datastore:pro	Software Connections
DayOne	Day One, Inc.
Dcare	Microforms Trans-Lingual
Electric Desk	Alpha Software Corp.
Empress 32	Rhodnius, Inc.
Enable	The Software Group
Enrich	Migent Software, Inc.
Executive Card Manager	Hewlett-Packard Personal Software

DBMS	VENDOR
Fastfile	Datamate Company
File Express	Expressware
FilePlan	Chang Laboratories, Inc.
FilePro 16 Plus	The Small Computer Company
Filebase	EWDP Software Inc.
Files & Folders	Starcom Computer Corp.
Filing Assistant	IBM Information Services
First Base	Universal Data Research Inc.
Flexifiler	Sound Decisions
Focus	Information Builders, Inc.
Form Manager	BIT Software, Inc
Formula IV	Dynamic Microporcessor Associates, Inc.
Foxbase Plus	Fox Software Inc.
Framework II	Ashton-Tate
Fulcrum Ful/Text	Fulcrum Technologies, Inc
GDX	General Data Systems, Ltd.
GOLDATAbase	Goldata Computer Services, Inc.
Goldengate	Cullinet PC Software, Inc.
HP SQL	Hewlett Packard Co.
IDMS/R	Cullinet Software, Inc.
IDMS/SQL	Cullinet Software, Inc
IDS/2	Honeywell Bull, Inc.
IIS/Destiny	Intelligent Information Systems, Inc.
IMPRS	Ruf Corp.
IMS/VS-DB	IBM
IPDBMS	International Parallel Machines
Info-DB Plus	Henco Software,Inc.
InfoStar Plus	MicroPro International Corp
Informix 4GL	Informix Software, Inc
Informix SQL	Informix Software, Inc.
Informix-SQL	Relational Database Systems, Inc
Infos II	Data General Corp.
Infoscope	Microstuf Inc
Ingres	Relational Technology, Inc.
Integrated 7	Mosaic Software
Interbase	Interbase Software Corp.
Interel	Honeywell Bull, Inc.
Interrogate	Applications Software, Inc
Kaleidoscope	DBI Software Products
KeepIT	Martin Marietta Data Systems
KnowledgeMan	Micro Data Base Systems, Inc
MAG/base 2	Rocky Mountain Software Systems
MAI Origin	MAI Basic Four, Inc
Mainstay	Mainstay Software Co.
Manager	Manager Software, Inc.
Mbase/9	Century Analysis, Inc.
Metafile	Metafile Information Systems, Inc.
Model 204	Computer Corp. of America
Multrics Relational Data Store	Honeywell Bull, Inc.

DBMS	VENDOR
NPL/R	Database Applications, Inc.
Nomad 2	Must Software International
Nonstop SQL	Tandem Computers, Inc.
Nutshell	Leading Edge Software Products Inc.
OWL A-B-C	Owl Software Corp.
Omnibase	Signal Technology, Inc
Omnifile	SSR Corp
Oracle	Oracle Corp.
Oracle 5.0	Harris Corp.
PAL	Gentry, Inc.
PC-DBMS	Kware
PC-File 'N Report	Jaspir International Inc.
PC-File III	ButtonWare Inc.
PC-File/R	Button Ware Inc.
PC/Focus	Information Builders, Inc.
PFS:File and PFS:Report	Software Publishing Corp.
Pace	Wang Laboratories, Inc.
Palantir Filer	Palantir Software
Paradox	Ansa Software
PeachText 5000	Peachtree Software Inc
Personal Decision Series	IBM Information Services
Personal Pearl	Pearlsoft
Please	Hayes Microcomputer Products Inc
Poise DMS-Plus	Campus America, Inc.
Pounce	Chattahoochee Computer Products Inc
Power-base	CompuWare Corp.
Powerhouse	Cognos
PractiBase	PractiCorp International, Inc
Prime Information	Prime Computer, Inc.
Private Files	Sofistry Inc
Probase	Probase Group Inc.
Prodas	Conceptual Software Inc.
Q & A	Symantec
Q-Pro 4	Quick-N-Easi Products, Inc.
QDMS-R	Quodata Corp.
QINT/SQL	Qint Data Base Systems Corp.
Query III	Hoyle and Hoyle Software, Inc
R:base Series 5000	Microrim, Inc.
R:base Seriews 4000	Microrim, Inc.
RDB, VAX/VMS	Digital Equipment Corp.
RDM1100	Amp
RTFILE	Contel Business Networks, Inc
Ramis Information Systems	On-Line Software International, Inc.
Rank and File	RAF Software Inc
Reflex, The Analyst	Borland International, Inc.
Relate/DB	CRI, Inc.
Reliance Plus	Concurrent Computer Corp.
Revelation	COSMOS, Inc.
SIX	ASAP, Inc.
Saturn-Base	Saturn Systems, Inc.

DBMS	VENDOR
Savvy PC	The Savvy Corp.
Seed DBMS, Proseed	Seed Software Corp.
Session	Public Office Corp.
Simple	Software Merchants Unlimited
Smartstar	Signal Technology, Inc
Sun Ingres	Sun Microsystems, Inc.
Super*List Manager	Pinnacle Software Systems Inc
Superbase	Progressive Peripherals & Software, Inc.
Supra	Cincom Systems, Inc.
Sybase System	Sybase, Inc.
Symphony	Lotus Development Corp.
System 1022	Compuserve Data Technologies
Systgem 1032	Compuserve Data Technologies
T.I.M. IV	Innovative Software
TAS	Business Tools, Inc.
TSM	Dynabase Ltd.
Team-Up	Unlimited Processing Inc.
The Data Reporter	Softwest Programming
The DataFiler	MBS Software
The Officesmith	Officesmiths, Inc.
The Sensible Solution	O'Hanlon Computer Systems
The Smart Data Manager	Innovative Software, Inc.
Turbo Image	Hewlett-Packard Co.
UNI-FILE	Univair Systems Inc.
Ultimate Operating System	The Ultimate Corp.
Ultra	Cincom Systems, Inc.
Unify	Harris Corp.
Unify	Unify Corp.
Universal Data Management System	Unisys Corp.
User-11	Userware International, Inc.
Userbase	Userware International, Inc.
VAX DBMS	Digital Equipment Corp.
VersaForm XL	Applied Software Technology
XDB	Software Systems Technology, Inc.
ZIM	Zanthe Information/Unipress Software

Appendix B

Recommended System Vendors

VENDOR: Ashton-Tate
PRODUCT: **dBase III+**
Address: 2101 Hamilton Ave.
Torrance CA 90502-1319
Phone: (213) 329-8000

VENDOR: Henco Software, Inc.
PRODUCT: **Info**
Address: 100 Fifth Ave
Waltham, MA 02014
Phone: (617) 890-8670

VENDOR: Oracle Corp.
PRODUCT: **Oracle**
Address: 20 Davis Dr.
Belmont, CA 94002
Phone: (800) 345-3267
Contact: Suzanne Hogan

VENDOR: Relational Technology
PRODUCT: **Ingres**
Address: 1080 Marina Village Parkway
Alameda, CA 94501-9891
Phone: (800) 446-4737
Contact: Lori Dryfus

VENDOR: Environmental Systems Research Institute (ESRI)
PRODUCT: **Arc / Info**
Address: 380 New York St.
Redlands, CA 92373
Phone: (714) 793-2853
Contact: Frank Holsmuller

VENDOR: KORK Systems
PRODUCT: **KGIS**
Address: 6 State St.
Bangor, ME 04401
Phone: (207) 945-6353
Contact: Terrance Keating

VENDOR: GeoBased Systems
PRODUCT: **Strings**
Address: P.O. Box 13545
Research Triangle Park, NC 27706
Phone: (919) 361-5717
Contact: Thomas Everly

VENDOR: GeoVision Corp.
PRODUCT: **GeoVision / Oracle**

Address: 1600 Carling Ave, Suite 350
Ottawa, Ontario K12 8R7
Canada
Phone: (613) 722-9518
Contact: Robert Muse

Appendix C

Key State Personnel and Agencies

ADMINISTRATION, DEPARTMENT OF (DOA)

Bureau of Data Processing
State House Station 61, Augusta, ME 04333
Telephone (207) 289-3631
Valton Wood, Systems and Program Manager
Bernard Beaulieu, Project Leader

AGRICULTURE, FOOD & RURAL RESOURCES, DEPT. OF (DAFRR)

State House Station 28, Augusta, ME 04333
Telephone (207) 289-3871
Esther Lacognata, Dir., Bur. Ag. and Rural Resources - 3511
Frank Ricker, Dir., Soil and Water Conservation Com. - 2666

CONSERVATION, DEPARTMENT OF (DOC)

Maine Geological Survey (MGS)
State House Station 22, Augusta, ME 04333
Telephone (207) 289-2801
Walter Anderson, State Geologist
John Williams, Hydrogeologist
Maine Geographic Information System (MeGIS)
State House Station 22, Augusta, ME 04333
Telephone (207) 289-2794
James Rea, Natural Resources Analyst Programmer

ENVIRONMENTAL PROTECTION, DEPARTMENT OF (DEP)

Bureau of Water Quality Control
State House Station 17, Augusta, ME 04333
Telephone (207) 289-3355
Peter Garrett, Senior Geologist - 3901
William Aldrich, Geologist - 3901
Norman Marcotte, ESS IV. - 3355
James Tibbetts, ESS II. - 3355
Gardner Hunt, Director, Division of Lab and Field Studies - 7688

Bureau of Land Quality Control
State House Station 17, Augusta, ME 04333
Telephone (207) 289-2111
Mark Hyland, Director, Division of Technical Services
Florence Hoar, Senior Geologist
Dave Dominie, Director, Division of Licensing and Review

Bureau of Oil and Hazardous Materials Control
State House Station 17, Augusta, ME 04333
Telephone (207) 289-2651
George Seel, Dir., Div. Remedial Planning and Tech. Services
David Sait, Dir., Div. Field Studies
Scott Whittier, ESS IV.
Michael Barden, ESS II.
Cheryl Fontaine, Geologist

Division of Computer Services
State House Station 17, Augusta, ME 04333
Telephone (207) 289-7892
Ron Dolan, Systems Group Manager

HUMAN SERVICES, DEPARTMENT OF (DHS)

157 Capital Street, Augusta, ME 04333
Telephone (207) 289-3826
Donald Hoxie, Director, Division of Health Engineering
Jeff Jenks, Drinking Water Program
Kenneth Meyer, Drinking Water Program
Charles Rossoll, Drinking Water Program

PUBLIC UTILITIES COMMISSION (PUC)

State House Station 18, Augusta, ME 04333
Telephone (207) 289-3831
Raymond Hammond, Senior Utility Engineer

STATE PLANNING OFFICE (SPO)

State House Station 38, Augusta, ME 04333
Telephone (207) 289-3261
Paul Dutram, State Groundwater Coordinator

TRANSPORTATION, MAINE DEPARTMENT OF (MDOT)

State House Station 16, Augusta, ME 04333
Telephone (207) 289-3321
Chris Olson, Supervisor, Well Claims
Melvin Morgan, Geotechnical and Materials Engineer
(Bangor Office, Tel. (207) 941-4545)

UNITED STATES GEOLOGICAL SURVEY (USGS)

Water Resources Division
26 Ganneston, Augusta, ME 04330
Telephone (207) 622-8208
Derrill Cowing, Chief, Maine Office
Thomas Maloney, Supervisory Hydrologist

Appendix D

Groundwater Data Management Bibliography

Blake-Coleman, W. and N. Dee. 1986. Ground-Water Data Management with STORET. EPA / 600 / M-86 / 007. EPA, Office of Ground-Water Protection, Washington, DC 20460.

Dee, N. and C. Miller. 1987. Ground-Water Data Requirements Analysis for the Environmental Protection Agency. GSA contract GS-00K-85AFD-2777, Task W6800-035. EPA, Office of Ground-Water Protection, Washington, DC 20460.

Forkey, B.A., M.S. Gold and M. Dein Bradley. -----. Use of Arc / Info in the Development and Implementation of a Ground Water Protection Program for the State of Rhode Island. Rhode Island Department of Environmental Management, Division of Water Resources, 83 Park Street, Providence, RI 02903.

Higgins, B.R. 1976. The Contribution of a Water Information System for Environmental Planning in Minnesota. PB-254 626. U.S. National Technical Service, U.S. Department of Commerce. 70 p.

Hix, G.L. 1983. A Data Base Management System Approach to Ground Water Data Management. Proceedings of the NWWA Western Regional Conference on Ground Water Management. October 23-26. San Diego, CA. pp 351-355.

Hoffman, J.L. 1984. A Computer Data Base for Overview of Ground Water Pollution Investigations. Ground Water Monitoring Review. Vol 6. No 1. pp 76-79.

Jadkowski, M.A. 1987. Phase 1 Study - State Groundwater Data Management System. Maine State Planning Office, State House Station 38, Augusta, ME 04333. 75 p.

Jadkowski, M.A. and P.W. Dutram. 1987. State Groundwater Data Management. Proceedings of NWWA Focus Conference on Eastern Regional Groundwater Issues. July 14-16. Burlington, VT. pp 65-79.

Kerr, H.A. 1986. Basic Groundwater Data in Alberta (Canada) - From Collection to Use. Third Canadian Hydrogeological Conference. Saskatchewan Research Council. April 20-23. Saskatoon, Saskatchewan. pp 176-183.

Link, L.E. 1982. Water Database Requirements. Proceedings of the Ground-Water Detection Workshop. Department of the Army. 4A762719AT40. January 12-14. Vicksburg, MS. pp 21-22.

Possin, B.N. 1985. The Use of Integrated Software Packages to Track Groundwater Monitoring Program Data for Wisconsin Sanitary Landfills. Eighth Annual Madison Waste Conference on Municipal and Industrial Waste. September 18-19. Department of Engineering Professional Development , University of Wisconson-Madison. pp 187-193.

Appendix E

Prototype Groundwater Data Index Demonstration Model

This routine allows the user to search state groundwater program files by department or by data type. A hardcopy catalog of programs can also be created here. To use the routine, just answer questions and follow instructions. To continue right now, type "CONT"; to quit, type "EXIT"

```
                 GROUNDWATER SEARCH MAIN MENU

                TYPE:           TO:
                _____           ___
                DATA            Search file by data type
                DEPT            Search file by department
                PRNT            Print catalog
                EXIT            Quit
```

```
                      DATA SEARCH MENU

                TYPE:            TO:
                -----            ---
                CONT             Continue Data Search
                DEPT             Do a Department Search
                MENU             Return to Main Menu
                EXIT             Quit
```

```
                   DEPARTMENT SEARCH MENU

                TYPE:            TO:
                -----            ---
                CONT             Continue Department Search
                DATA             Do a Data Search
                MENU             Return to Main Menu
                EXIT             Quit
```

SUMMARY AND TABLE OF CONTENTS

PROG.#	DEPT.	AGENCY	DIV.	PROGRAM	REL. PROG.	DATA KEPT
26	DEP	BLQC		Site Location of Development		CI,HG,ML,O,WQ
22	DEP	BLQC	Tech	Solid Waste Landfills		CI,HG,ML,O,WQ
20	DEP	BOHMC	LI/EN	RCRA Sites		ML,O,WQ
21	DEP	BOHMC	LI/EN	Uncontrolled Sites		ML,O,WQ
14	DEP	BOHMC	LI/EN	Underground Storage Tanks		O
15	DEP	BOHMC	R.Ser	Oil Spills		O,WQ
17	DEP	BWQC		Sand & Gravel Aquifer Monitoring	6	ML,O,WQ
16	DEP	BWQC		Sand & Salt Pile Monitoring	11	ML,O,WQ
18	DEP	BWQC	Eval.	Complaint Response		O,WQ
19	DEP	BWQC	LI/EN	Waste Discharge Licenses		CI,O,WQ
3	DHS		D WAT	Private Well Analysis		O,WQ
1	DHS		D WAT	Public Water Supply Development		ML,O,WL,WQ,WY
2	DHS		D WAT	Public Water Supply Monitoring		ML,O,WQ
7	DOC	MGS		Bedrock Aquifer Mapping	25	AA,BF,HG,ML,WL,WQ,WY
8	DOC	MGS		Radioactive Waste Sites		AA,BF,ML,WY
6	DOC	MGS		Sand & Gravel Aquifer Mapping	17,25	AA,HG,ML,WQ,WY
9	DOC	MGS		Well Drillers Information	24	CI,HG,ML,O,WL,WY
10	MDOT			Exploratory Borings		HG,WL
13	MDOT			Pollution Claims		O,WQ
12	MDOT			Preconstruction Information		CI,O,WQ
11	MDOT			Sand & Salt Storage Facility Rank	16	CI,O,WQ
5	PUC			Miscellaneous Programs		AA,HG,O,WQ,WY
4	PUC			Water Utilities		O,WY
23	USGS			Basic data collection network		HG,ML,WL,WQ
25	USGS			Hydrology studies program	6,20	CI,ML,O,WL,WQ,WY
24	USGS			Well Information	9	CI,ML,O,WL,WY

DHS = Department of Human Services
PUC = Public Utilities Commission
DOC = Department of Conservation
MDOT = Maine Department of Transportation
DEP = Department of Environmental Protection
USGS = United States Geological Survey

AA = Aquifer Areas
BF = Bedrock Fractures
CI = Construction Information
HG = Hydrogeology
ML = Map Location
O = Ownership
WL = Water Level
WQ = Water Quality
WY = Well Yield

The Public Water Supply Development program under DHS collects water quality information for drinking water sources developed for > 25 people or for > 15 outlets. The period of record is approximately 1930 to present. The total information collected is; water quality parameters, ownership, map location, station construction, well yield, and water level. Safe levels of several organics and several inorganics must be met before the source is approved.

Data are collected, usually by a consultant hired by the utility in question, upon establishment of the water supply. Additional information is sometimes required by DHS upon review. Stations can be located to within 10 feet on drawing of the facility. Currently records exist on 500 facilities with 200 more possible within five years.

Data are considered by DHS to be fairly accurate. They are at least sufficiently accurate for the evaluation of the potential water supply done by DHS.

Records are stored manually in archives. It is expected that eventually the records will be entered in a dBASE file with software provided by EPA in Wyoming. The State of Maine has obtained federal funding to upgrade its drinking water records in DHS.

Currently, the access to the manual files is poor, although there is little call for them after the initial evaluation. The data are used for annual department activity reporting.

Other changes are expected due to modification of the Safe Drinking Water Act. Instead of the currently required minimum report of 30 water quality parameters, over 100 will be required increasing data storage needs for the program. It is also likely that more subsurface sources will be used, thereby increasing the amount of groundwater-related data.

After samples are collected by the consultant or by the utility itself, they are sent to the Public Health Lab or to another certified (by the Public Health Lab) lab. The data are reviewed by the person in charge of wellhead protection in consultation with other engineers. Sometimes more data is requested.

PUBLIC WATER SUPPLY MONITORING

The Public Water Supply Monitoring program under DHS maintains records for some 2000 or more public water utilities in the State of Maine. Of these, about 150 use groundwater. This means that there are about 6000 directly relevant records, considering the period of record for present purposes to be 20 years. A sample is taken from each station an average of twice a year with considerable variation in frequency. The purpose of the sampling is to monitor the quality of public drinking water.

Water quality parameters, ownership, and map location are recorded for each station at each facility. The primary interest is in inorganic and organic materials as parameters under water quality.

The locatability of the stations on the map is highly variable, although an effort is made to insure accuracy of location.

Ninety percent of the samples are collected by utility personnel, and 10% by DHS personnel. The samples are then analyzed by an EPA certified lab. The report is then sent to DHS where it is examined and filed manually. If a water quality problem is identified, further testing might be ordered.

The data are stored manually by DHS with the exception that administrative data is stored on an IBM PC/XT. The reports are microfiched and kept for 11 years in the DHS office, after which time they are sent to the state archives for storage. Retrieval is not terribly difficult until the data go to the state archives.

Current access, though not difficult, is somewhat culmbersome for the reporting that could be done using the data. This data is targeted for inclusion in IBM PC dBASE files using software provided by Wyoming EPA. The State of Maine has obtained federal funding for this effort.

There is a great deal of extra agency demand for this data. The Bureau of Health, DEP, and MDOT make the most use of it. EPA uses it for administrative review. There is little call for the data by the private sector. Few extra-departmental sources of data are seen, as the mandate is only to test water being delivered to the public. It is possible that some information maintained by DEP and MGS might be of use.

Plans are underway to require direct computer reporting by labs. The Public Health Lab is automating its own procedures and is taking a lead role in this effort.

Contact: Kenneth Meyer
Department of Human Services
157 Capital st
Augusta, ME 04333
289-3826

PRIVATE WELL ANALYSIS

The Private Well Analysis program tests private wells at the request of a client, usually the owner. A sample is taken by the requesting party and analyzed by the Public Health Lab. The report is sent to the client and also retained as a manual file by DHS. Administrative data, ie, client name, owner's name, town, type of source, type of test requested, serial number for test, and overall pass/fail are kept on an IBM PC/XT.

As many as 7-8000 wells may be analyzed each year. The period of record is about 20 years. Data are generally not accessed once the orignal report has been sent to the owner. The locatability of stations is extremely poor. The station might not even be in the town listed in the files because the requesting party has an address in a town different from that in which the well is located.

The Public Health Lab is automating its procedures. Under the proposed system the program would serve simply an archiving function. The conversion of DHS data to dBASE files on the PC will also affect this program. Also, stricter drinking water standards will effect the progam.

Contact: Kenneth Meyer
Department of Human Services
157 Capital St.
Augusta, ME 04333
289-3826

WATER UTILITIES PROGRAM

The Water Utilities Program stores ownership and well yield information for 150 facilities(public and private water utilities). These data are obtained from the utilities themselves in the form of annual groundwater usage reports. Locatability of stations is poor being immediately locatable only to town within the state of Maine.

The period of record is from the 1930s to present, but data from only the last two years or so are readily available. Records from this time period are maintained in a dBASE III file on a IBM PC/XT by the Computer Technical Analyst. These data are conveniently accessable.

PUC is uncertain of how accurate the data are because all are reported by the utilities themselves and PUC hasn't the manpower to do extensive checking. When the reports arrive in the PUC office, they are entered into the dBASE III file by contract.

Within the agency, financial data are used more than are groundwater data for statistical and financial reports that are done on an as needed basis two to three times per year. The information is little used by outside agencies. The primary users are USGS, SPO, and Water Associations.

In the future, a possible link with USGS's data base is seen. Also expected is a slow, but steady improvement in yield information. Possible drinking water legislation could potentially change the data requirements of this program.

Contact: Raymond Hammond
Public Utilities Commission
State House sta 18
Augusta, ME 04333
289-3831

MISCELLANEOUS PROGRAMS

The primary source of miscellaneous data comes from favorability studies contracted with USGS. If a utility develops source or quality problems, a special study may be undertaken. Results of the study are sent in the form of a letter report and an aquifer map to the utility in question.

A copy of the data is kept in the PUC office in a manual file. This data set is rarely accessed after its initial use in reporting to the utility. The primary data stored are aquifer descriptions. Hydrogeologic data, ownership, water quality, and well yield are also available in most cases.

Contact: Raymond Hammond
Public Utilities Commission
State House Sta. 18
Augusta, ME 04333
289-3831

SAND AND GRAVEL AQUIFER MAPPING 6

The mapping of sand and gravel aquifers is being accomplished under MGS of DOC. The period of record of this program is 1981 to present with expected extension to 1989 or 1990. The maps produced under this program (29 produced so far) contain water quality parameters, well yields, hydrogeological information, aquifer areas, and map locations of stations sampled. Geographic locatability of the stations is extremely good.

Currently the maps are stored manually, mostly in the MGS office. Some, however, are kept by DEP's Bureau of Water Quality and by the US Geological Survey Augusta Field office. MGS also maintains an index to available maps. Access is to raw field data is difficult. In addition to departmental use, consultants, USGS, MDOT, and others make use of the maps. Access to the information stored is approximately twice a month.

The maps produced by MGS form an annual report on mapping activities which is for sale. This serves mandated reporting needs as well.

A priority here is for the obtaining of a GIS so that maps can be updated conveniently.

All data used are collected by MGS or by USGS in cooperation. Seismic lines are analysed by computer. Water quality parameters are analyzed by the DEP lab.

Contact: John Williams
Maine Geological Survey
Department of Conservation
State House Sta 22
Augusta, ME 04333
289-2801

BEDROCK AQUIFER MAPPING

The Bedrock Aquifer Mapping program under MGS of DOC was begun as a pilot program in Presque Isle in 1986, but has not been refunded as of 9/87. The program was initially intended to produce aquifer boundry maps, but bedrock aquifers are so difficult to identify that the objective was changed to one of mapping bedrock characteristics. The main goal remains to show areas favorable for groundwater development in the absence of sand and gravel aquifers.

Most aspects of the mapping is accurate to the limits of the existing topographic maps (7 1/2 or 15 '). It uses the information collected from the well drillers information program (9), which is only accurate to a scale of 1:100,000. Final maps will be produced at the 1:100,000 scale. The data are currently stored on a Burroughs computers in Burroughs Database. This data is hard to access, and it is likely that a change will be made to IBM and another database system.

The information stored consists of hydrogeologic information, map location, well yields, water levels, linear features determined by aerial photographic methods, and geophysical profiles. If the program is refunded, water quality parameters will also be collected.

Contact: John Williams
Maine Geological Survey
Department of Conservation
State House Sta 22
Augusta, ME 04333
289-2801

RADIOACTIVE WASTE SITES

The Radioactive Waste Site investigation was begun because of doubts about the adequacy of rock and groundwater characteristics in the State of Maine to support radioactive waste sites. Maps are produced which show well yields, bedrock fractures, and aquifer areas. The period of record is 1986 to present; the project was refunded in 1987.

The maps are produced in part from airphotos. These photos are interpreted by a private consultant and are digitized by MGS personnel using a link through the University of Maine at Orono. They are then maintained in the state GIS on the Burroughs computer.

The maps are included as part of a larger report on the potential effects of any radioactive waste sites situated in Maine. The maps are to be used by DOE, DEP, the Radioactive Waste Commission, and others in analyzing the suitability of potential sites.

Contact: John Williams
Maine Geological Survey
Department of Conservation
State House Sta 22
Augusta, ME 04333
289-2801

WELL DRILLERS INFORMATION

Well Drillers information is under MGS of DOC. Formerly a voluntary program, it was made manditory by the Maine State Legislature in 9/87. For this reason, a vast increase in number of records is expected. Now, the contracter must complete the drillers log for each new well drilled or dug. There are approximately 20,000 well records in existance presently, and it is estimated that there may be as many as 10,000 new records per year under the new legislation. The geographic locatability of the stations as stored is highly variable. It is likely that the accuracy of all data will decrease with all drillers being required to submit data whereas previously only those with sufficient interest to do this voluntarily submitted reports.

The data are stored currently on the Burroughs computer, but it is expected that a change will be made to IBM PCs to improve access. Due to legal requirements, the raw data can be made available only to other state agencies or to municipalities, but security problems have not arisen as yet.

MGS plans to publish the information on 1:100,000 scale base maps through the Bedrock Aquifer Mapping Program.

Little quality control is possible on this data because of the volume. There is some comparison with existing maps, and personnel can get some idea of the reliability of information provided by each driller. Once the reports reach the MGS office, they are entered into the database by a clerk-typist.Locations are also digitized on state GIS from 7 1/2 or 15 min. topo maps. From there they are accessed by this program, other programs in DOC, by DEP, and MDOT, and to assist in locating alternative water supplies where wells have been contaminated.

Contact: John Williams
Maine Geological Survey
Department of Conservation
State House Sta 22
Augusta, ME 04333
289-2801

EXPLORATORY BORINGS

The Exploratory Borings program of Maine Department of Transportation drills test holes to obtain subsurface information at highway and bridge locations. Geographic locatability is excellent on newer plans, not as good on older ones.

Current files contain 2000-4000 project folders with around 200 stations(borings) per project. These records are filed manually by town and project number. The data they contain are hydrogeologic data and water levels.

A report is published on each project and this serves as the usual means of access to the data for interested persons. Sometimes individual files are examined. Within the department these files are accessed when there are special needs and this happens only occasionaly. Outside the agency, the files are occasionally used by USGS, MGS, or consultants.

MGS maps and USGS reports and maps are the main extra departmental sources of information. Program data are of high quality as they are generated by department personnel.

Currently, these files are being computerized on an IBM PC with a tie in to an AT&T computer in Augusta.

Contact: Melvin Morgan
Maine Department of Transportation
PO Box 1208
Bangor, ME 04401
941-4545

SAND AND SALT STORAGE FACILITY RANKING

Sand and Salt Storage Facility Ranking under MDOT was begun in 1986 and is essentially complete. The purpose is to rank the priorities in covering road sand and salt piles. Variables used to do the ranking were well data, aesthetics, development potential, location, and surface water runoff. The data relevant to groundwater are ownership, station construction, and water quality parameters. Geographic locatability is good as is data adequacy as it was collected by MDOT technicians.

Data are stored manually and on IBM PC in LOTUS format. Some data are also stored on the PC in dBASE III. Currently some data are also being computerized in SAS.

Data in manual storage are accessed weekly. This access is inefficient and is expected to improve when all data are computerized. The LOTUS files are also difficult to access.

Reporting is done as time is available. The data have been used to generate equations dealing with priorities of action. Access by agencies and individuals outside MDOT is frequent, most being by DEP with whom this program is a cooperative. The main extra-agency source of data is also DEP.

Contact: Christine Olson
Maine Department of Transportation
State House sta 16
Augusta, ME 04333
289-3321

PRECONSTRUCTION INFORMATION

The goal of the Preconstruction Information program of MDOT with regard to groundwater is the collection of water quality information on any new construction project that may affect drinking water. The period of record is 1962 to present. Samples are collected by MDOT personnel and analyzed by the Public Health lab. The data are water quality paramenters, ownership, and construction information. Geographic locatability of the stations is only moderately good and somewhat complex with several steps having to be gone through.

Current storage is manual, although this data is being computerized as time permits. Recent information is easily accessable, though, even under manual storage. DBASE III is system to be used. Current files contain about 1200 project records. Growth is about two per year.

There has been little extra-agency demand for the data collected under this program.

Contact: Christine Olson
Maine Department of Transportation
State House Sta 16
Augusta, ME 04333
289-3321

POLLUTION CLAIMS

13

The mission of the Pollution Claims program of MDOT is to respond to complaints that highway construction or a sand or salt pile has caused pollution of drinking water supplies. The program was begun in 1969 by legislative mandate.

Ownership and water quality information are kept. These data are stored manually, although dBASE III and SAS files are being entered as time permits. The data are used to rule on claims and are used as supporting documentation in letters to the complainant.

Little other use has been made of the records to this point, either by the agency or by persons outside the agency.

Contact: Christine Olson
Maine Department of Transportation
State House Sta 16
Augusta, ME 04333
289-3321

Ownership information on underground oil storage tanks is maintained by the Division of Licensing and Enforcement within the Bureau of Oil and Hazardous Materials Control of DEP. The data are maintained both manually and by computer. About 100 sites have mandatory groundwater monitoring, but it is not required that reports be submitted to DEP. The mandate of the program is to locate and register the tanks. Tanks are located by address of tank owner. Directions are given to site.

Those facilities required to monitor groundwater maintain a log of water quality tests. DEP periodically inspects to see that the log is actually maintained, but quality of the data is possible limited.

Contact: David Boulter, Director
or Stacy Ladner
Bureau of Oil and Hazardous Materials
Division of Licensing and Enforcement
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2651

OIL SPILL SITES

The Oil Spill Sites program of DEP maintains monitoring information on 3-15 stations per spill site. The period of record is approximately 4 years. Monitoring is done by DHS and is irregular. Location is by address only within town.

The data are stored manually and on the Honeywell DPS 6/95 computer. Recently storage has begun on an IBM PC. Some personnel find the system confusing and inadequate to their needs. Usability of the data could be greatly enhanced by streamlining of the storage/retrieval system, possibly by the use of a more adequate database.

The data maintained include ownership information and water quality parameters.

Contact: David Sait, Director or
Fred Brann
Bureau of Oil and Hazardous Materials
Division of Response Services
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2651

SAND AND SALT PILE MONITORING

The Sand and Salt Pile Monitoring program of DEP is a cooperative effort with MDOT. The monitoring was begun in 1982 and is relatively complete. About 800 sand and salt piles are monitored with samples taken from private wells nearby that have shown problems. These samples are obtained by DEP personnel and are analyzed by the DEP lab.

Water quality parameters and maps (locatable on 7 1/2 or 15 minute topo maps) are stored manually. The ownership and map location data are entered on an IBM PC FRAMEWORK system. This system is not felt to be tremendously user-friendly. For the frequency of access of manual data, a more efficient system would be desirable. Consideration has been given to adding this data to FRAMEWORK eventually.

All these piles are to by located on GIS maps by the end of 1987 along with other threats to groundwater. Current data are accessed by MDOT, attorneys, consultants, and public interest groups.

Contact: Peter Garrett or Dan Locke
Department of Enviornmental Protection
State House Station 17
Augusta, ME 04333
2890-3901

SAND AND GRAVEL AQUIFER MONITORING 17

Sand and Gravel Aquifer Monitoring is a long standing program of DEP. It is done in cooperation with the MGS mapping program. Each of 16 water quality parameters were taken from 12 stations by DEP personnel. They were then analyzed by the DEP lab.

Once this was completed, the parameters along with ownership and map location were entered into the USGS WATSTORE system and also entered into EPA's STORET in North Carolina.

The data are used by USGS, MGS, EPA, and by the private sector.

Contact: Peter Garrett or William Aldrich
Department of Enviornmental Protection
State House Station 17
Augusta, ME 04333
289-3901

COMPLAINT RESPONSE

The Complaint Response program is designed to handle environmental complaints. These complaints are relatively rare, and the data collected depend upon the situation and are therefore highly variable. Ownership and water quality parameters of varying degrees of detail are collected.

All data are filed manually by case. If a basis for legal action exists, information would be passed to the appropriate agency. After the case is settled, little use is made of the data.

During its active phase, all interested parties to the complaint are likely to use the information.

Contact: Peter Garrett
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-3901

WASTE DISCHARGE LICENSING

The Waste Discharge Licensing program is under the Licencing and Enforcement division of the Bureau of Water Quality Control of DEP. Waste discharge facilities require licenses. Of the approximately 4000 facilities in Maine, about 40 require groundwater monitoring by 3-12 stations per facility. Ownership, construction information, and water quality parameters are maintained. Stations are sampled an average of about two times per year. The total current data files are less than one megabyte in extent and are growing slowly.

Data are stored manually in the Licensing and Enforcement office and on computer -- the Permit Complience System (PCS) -- by DEP at EPA in Research Triangle Park, North Carolina. Access to computerized data is through the Licensing office staff, and requests usually take about 24 hours to fulfill. Access by division staff is frequent; access by other agencies or by the public is much less frequent.

Staff uses primarily station water quality parameters to processes licenses; other people are usually interested in lists of facilities with certain characteristics. The primary water quality parameters of current interest are nitrates.

Data are collected by facility owners or by consultants hired by them. Data are then analyzed by private labs before being sent to the division. The only quality control is by annual (or about that) inspections and by inspectors examining the data as they come into the office. The quality of the data so obtained is considered to be improving, but to have still considerable room for more improvement.

Once a report is accepted, it is given to the computer staff where it is entered into the PCS system, either by phone line or through the IBM mini to Boston thence to North Carolina.

Useful data from other agencies is being obtained by staff from the EPA STORET system with increasing frequency. Facilities monitoring groundwater data can be located to the limits of accuracy of a 7 1/2 minute quad sheet.

Only recent data is computerized. Other data is stored manually. The entire period of record is 1980 to present.

Contact: Norman Marcotte
Dept of Environmental Protection
State House Sta 17
Augusta, ME 04333
289-3355

RESOURCE CONSERVATION AND RECOVERY ACT

The Resource Conservation and Recovery Act(RCRA) program is run by the Division of Licensing and Enforcement within DEP's Bureau of Oil and Hazardous Materials. The period of record for the program is 1980 to present. Under this program an average of 12 wells at 12 hazardous waste disposal facilities are sampled quarterly for water quality parameters. The samples are taken by the facility personnel and are analyzed by EPA-approved labs. Approximately 240 parameters are tested for each station. Little review of the data is done before filing.

The data are stored manually in the Division of Licensing and Enforcement and are accessed by EPA and by public interest groups. For many purposes, the manual storage and retrieval system is felt to be inadquate. The situation may get worse because EPA is increasing its emphasis on groundwater.

This program could make use of a variety of existing data such as maps generated by USGS and MGS.

Contact: George Seel, Stacy Ladner, or
David Boulter, Director
Bureau of Oil and Hazardous Materials
Division of Licensing and Enforcement
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2651

UNCONTROLLED SITES

Twelve or more hazardous waste dumps are monitored under the Uncontrolled Hazardous Substance Sites program within the Division of Licensing and Enforcement of DEP's Bureau of Oil and Hazardous Materials. At each of these facilities, three or more stations are monitored with varying frequency. Although the data set is small currently, it is increasing rapidly. Water quality parameters are collected by DEP personnel, analyzed by the DEP or other approved lab, and stored manually.

The resulting folders are accessed weekly or more frequently by agency personnel, banks, developers, consultants, and municipalities. When a criminal investigation is under way, relevant files kept secure.

Maps produced by MGS and USGS are of potential use to this program.

Contact: David Boulter, Director or
Hank Aho
Bureau of Oil and Hazardous Materials
Division of Licensing and Enforcement
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2651

SOLID WASTE LANDFILLS

DEP's Bureau of Land Quality Control is responsible for the review,licensing, and enforcement of regulations at landfills, transfer stations, sludge lagoons, and septage sites. This period of record for this program is 1974 to present. The waste facilities are locatable to the degree of accuracy of 7 1/2 min topo maps. Stations on site are not further locatable.

Samples are collected at Solid Waste Facilities, usually by a consultant hired by the municipality maintaining the site. The samples are analyzed by an EPA-approved lab. Samples are taken quarterly. While the number of parameters analyzed varies, water quality is taken at each site. Also, map location, ownership, construction information, and hydrogeologic data are taken and stored.

Storage is manual in the Technical Services office. The system is not completely adequate for the almost daily frequency of access by agency personnel, DHS personnel, attorneys, public interest groups, municipalities, consultants, and private individuals. Also needed is some sort of system that would allow the ready calculation of statistics and graphing displaying trends. This information is needed almost on a daily basis and is currently cumbersome to obtain.

It is expected that demand for the data and overall amount of the data will show an increase. Extra-agency sources of relevant data would include DHS and MGS.

Contact: Mark Hyland
Bureau of Land Quality Control
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2111

BASIC DATA COLLECTION

The Basic Data Collection Network of the USGS performs long term monitoring of groundwater throughout the nation. The Agency's primary objectives are to gage the effects of climate, terrain, and man-made stresses on groundwater availability. Depth to water and well drawdown are of primary interest. The Maine monitoring network presently consists of about 20 wells. Ultimately, it will be expanded to 40 wells. This program utilizes the Agency's Water Data Storage and Retrieval System (WATSTORE) for information management.

Water levels at most wells are monitored continuously for the entire period of record. However, only five day high values are retained on WATSTORE for long term storage. Measurement frequencies of wells placed to gage the effects of terrain are sometimes reduced to twice monthly.

Groundwater quality monitoring is a secondary objective. However, at least two chemical analyses are performed on each well: one at high and the other at low water level. Results of these also are stored in WATSTORE.

Contact: Derrill Cowing
United States Geological Survey
Water Resources Division
26 Ganneston
Augusta, ME 04330
622-8208

Well Information. Information from three to four thousand stations inventoried during hydrologic studies is being entered into WATSTORE. Included are parameters such as aquifer type, depth of well, depth to water, well type, and plumbing yield. When completed, this database will be a valuable source of groundwater information.

Contact: Derrill Cowing
United States Geological Survey
Water Resources Division
26 Ganneston
Augusta, ME 04330
622-8208

Hydrogeologic Studies Program. USGS, MGS, and DEP cooperate on a variety of groundwater related projects, including the sand and gravel aquifer mapping program. Information collected for these projects include seismic and other geophysical data, test hole logs, and water level and water quality data from observation wells. Maps, data and interpretative reports resulting from projects are published, distributed and filed manually, but numerical and descriptive information is entered into WATSTORE whenever possible.

Contact: Derrill Cowing
United States Geological Survey
Water Resources Division
26 Ganneston
Augusta, ME 04330
622-8208

SITE LOCATION OF DEVELOPMENT

The Bureau of Land Quality Control of DEP is responsible for the review, licensing, and enforcement of the State's land use laws, in particular the Site Location of Development Act. The site location act reevaluates land use activities such as residential subdivisions and large commercial and industrial facilities. The period of record is 1970 to present. The developments are located to the degree of accuracy of 7 1/2 min topo maps. Stations on site are not further locatable.

Samples are collected at the developments, usually by a consultant hired by the developer or owner. The samples are analyzed by an EPA-approved lab. Samples are taken quarterly, semi-annully, or annually. While the number of parameters analyzed varies, water quality is taken at each site. Also, map location, ownership, construction information, and hydrogeologic data are taken and stored.

Storage is manual in DEP's regional offices and the Technical Services office in Augusta. The system is not completely adequate for the almost daily frequency of access by agency personnel, DHS personnel, attorneys, public interest groups, municipalities, conusltants, and private individuals. Also needed is some sort of system that would allow the ready calculation of statistics and graphing displaying trends. This information is needed almost on a dialy basis and is currently cumbersome to obtain.

It is expected that demand for the data asnd overall amount of the data will show an increase. Extra-agency sources of relevant data would include DHS and MGS.

Contact: Mark Hyland
Bureau of Land Quality Control
Department of Environmental Protection
State House Station 17
Augusta, ME 04333
289-2111

Appendix F

STORET Data Storage Formats

United States
Environmental Protection
Agency

Office of
Ground-Water Protection (WH-550G)
Washington DC 20460

EPA/600/M-86/007
March 1986

Water

EPA

Ground-Water Data Management With STORET

4.0 GROUND-WATER DATA STORAGE FORMATS

This chapter contains an introduction to the formatting and retrieval strategies used in STORET. An explanation of data needed for storage is presented in the next several sections.

4.1 Data Needs Identified for Ground-Water/STORET Users

A station must be thoroughly and correctly described so that data asociated with it can be stored and later retrieved. It should be emphasized that the more information available about a station the more flexibility there will be in retrieving the data stored with that station.

Three broad categories of descriptors have been identified as needed by Ground-Water/STORET users to completely document information available for a particular station. These categories are:

- Station descriptors
- Sample descriptors
- Analytical findings.

There are several elements under each category which will enable the user to describe the station thoroughly. These elements are perhaps more information than would be needed to store surface water data. Elements making up each category of descriptors are described below. A graphic representation of where these descriptors should be entered into the data record are given in Appendix A.

4.1.1 Station Descriptors

Factors which are descriptive of the sampling location and would not change over time are called "station descriptors". There are three types of station descriptors needed by ground-water data managers to support their ground-water monitoring data. They are as follows:

Facility descriptors --

descriptors of the operation being monitored, such as type of waste management area (e.g., landfill), facility location (not the corporate headquarters), (e.g., zip code) and type of business (e.g., disposer of hazardous waste). Facility descriptors, except ownership, will always be stored in either a station header or a descriptive paragraph.

Physical setting descriptors --

descriptors of the setting in which the facility is located and from which samples were taken, such as aquifer name or geologic formation name. For the most part these descriptors are stored in the parametric data field. Two of these descriptors are stored in the descriptive paragraph and one descriptor in the station header.

Well descriptors --

descriptors of those characteristics of a well which may be an important factor in data analysis and which would not be expected to change over time, such as type of well, well depth, and casing material. All but one of these descriptors will be stored in the parametric data "fixed date" field. Note that the term "fixed date" under the parameter field means elements that will not change with time.

4.1.2 Sample Descriptors

Factors that describe a sample at the time it was taken and which are expected to change with each sampling event are called "sample descriptors". Three types of sample descriptors needed by ground-water data managers to support their ground-water monitoring data are:

Sampling purpose descriptors --

descriptors of why and by whom a sample was taken. These descriptors are stored in the parametric "variable date field."

Sampling condition descriptors --

descriptors of the conditions during the sampling event, such as the depth to the top of the water table or the temperature. These descriptors are stored in the parametric "variable date field."

Sampling/Analysis Descriptors --

descriptors to document how a sample was taken and/or analyzed, such as how the sample was drawn and whether or not it was replicated.

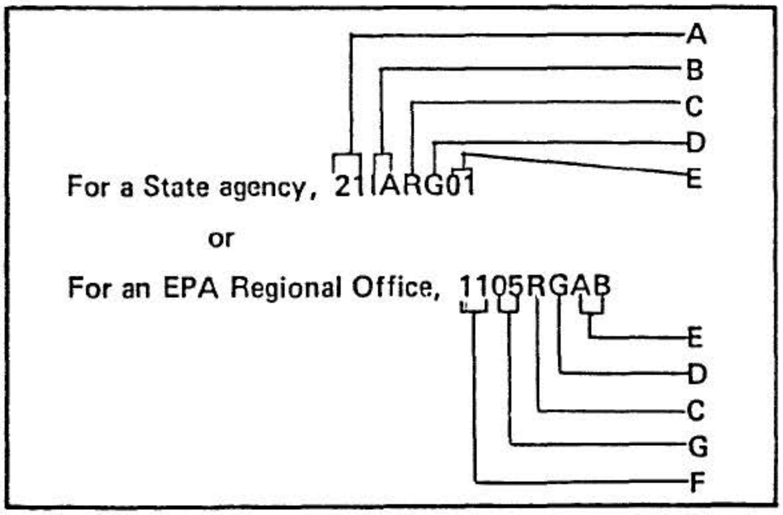
- An alphabetic character to indicate which program the data is collected from as the fifth character. For instance an "R" to indicate that the data is collected under the RCRA program. We recommend the following codes:

R	for RCRA
C	for CERCLA
U	for Underground Injection Control
S	for Public Water Supply
P	for Pesticide data
A	for ambient ground-water monitoring network

- <u>"G" to indicate that the data is ground-water monitoring data</u> as the sixth character

- <u>Two optional "free" characters for the user to code as needed</u> for the seventh and eighth characters.

Examples of STORET agency codes for RCRA ground-water monitoring data are shown below:



Key

A	Indicates a State Agency
B	Federal Information Processing System State Abbreviation
C	Program Code
D	Ground-Water Code
E	Optional free characters
F	Indicates an EPA Regional Office or other Federal Agency
G	EPA Region Represented, in this case Region V

4.1.3 Analytical Findings

The findings that were determined from each sample at a station are called "analytical findings". (e.g., the concentration of arsenic in the sample). Analytical findings will be stored in the STORET parametric data field by using "parameter codes". A printed list of all current STORET parameters and their codes may be obtained by users of the STORET system with the command %Batch with one of the following: PARMALFA, PARNUMER, and PARCAS. Descriptions of these lists are contained in Section 4.6.5 of this report and instructions on storing findings can be found in Chapter WQ-DE of the STORET User's Handbook (February 1982). A list of the parametric codes especially pertinent to RCRA and likely useful to others can be found in Appendices G and H of this manual.

4.2 Data Formatting

There are specific formats that must be used when inputting station and parametric data into STORET. Station header data is always stored and modified with the ?01 format which is a fixed form method. Exhibit 4-1 is the EPA form used for storing station location data in the STORET Water Quality File. Parametric data can be stored with 5 different formats: ?00, ?01, ?02, ?03 and ?04. Whereas any of these formats are acceptable the ?00 format is the most versatile and contains special features to be discussed in a later section that make it the only recommended format for ground-water users.

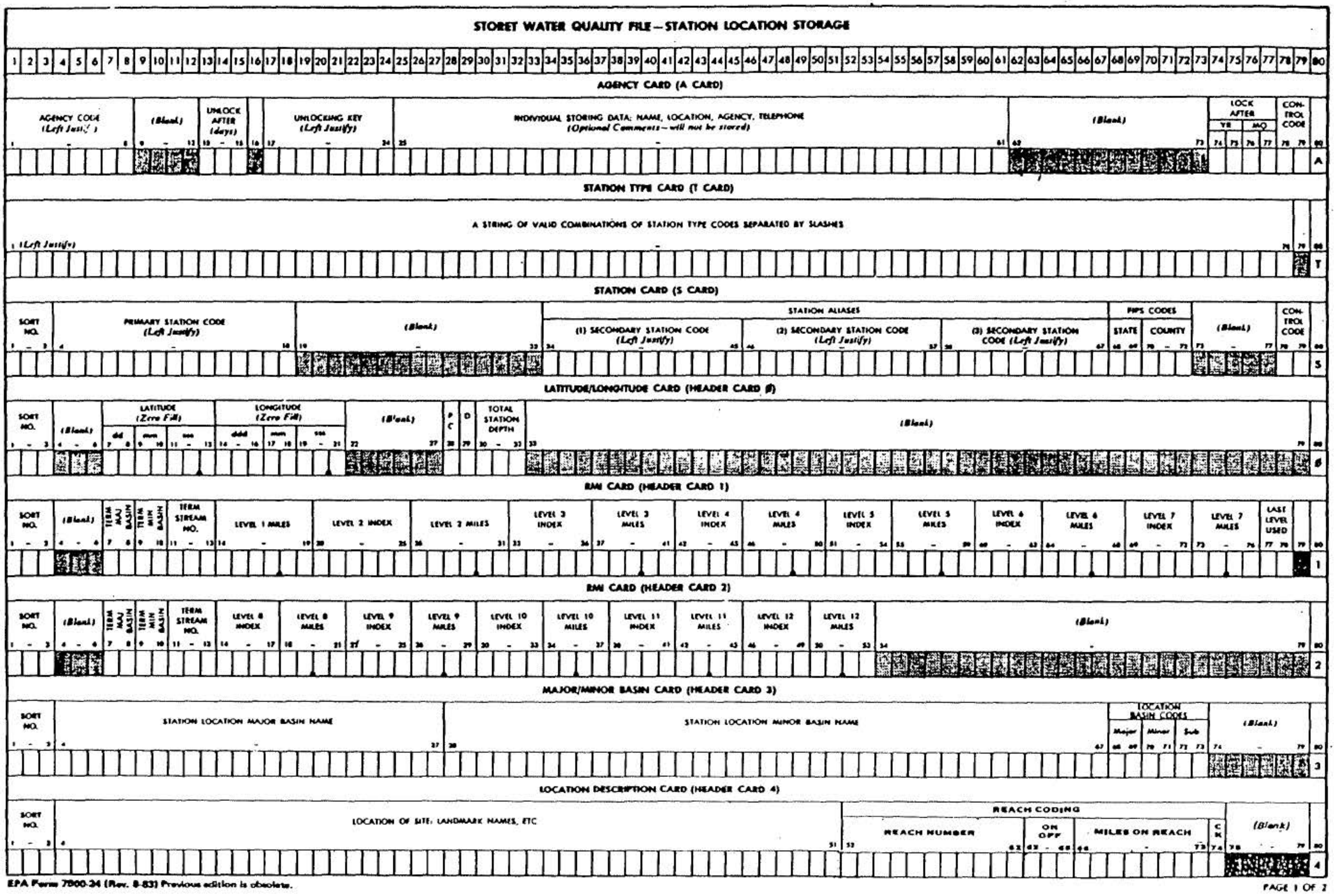
Chapter WQ-DE of the STORET Users' Handbook (February 1982) describes the technical procedures for storing data in STORET. Specifically, it describes the various storage formats, how each one can be used, how to invoke each format and how to enter data. This chapter is meant to be a supplement to chapter WQ-DE of the STORET User's Handbook (Feburary 1982). Rather than describing the technical procedures for entering any data into STORET, this chapter assists ground-water users in understanding what information is needed for data storage and describes what has been determined to be the most useful organization and format for ground-water monitoring data in STORET. Extensive capabilities have been provided for storing station and sampling data. Users should determine which of of these capabilities are appropriate to them.

4.3 Station Header

Before any ground-water monitoring data can be entered in STORET, an identification of each station from which the samples were taken must be "established" in the data base. In other words, a station header must be created for each ground-water monitoring well to which data are attributed.

4-1

Example of a Station Location Storage Form



STORET WATER QUALITY FILE — STATION LOCATION STORAGE

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80

AGENCY CARD (A CARD)

AGENCY CODE (Left Justify)	(Blank)	UNLOCK AFTER (days)		UNLOCKING KEY (Left Justify)	INDIVIDUAL STORING DATA: NAME, LOCATION, AGENCY, TELEPHONE (Optional Comments—will not be stored)	(Blank)	LOCK AFTER YR	LOCK AFTER MO	CONTROL CODE	
1 - 8	9 - 12	13 - 15	16	17 - 24	25 - 61	62 - 73	74 75	76 77	78 79	80
										A

STATION TYPE CARD (T CARD)

A STRING OF VALID COMBINATIONS OF STATION TYPE CODES SEPARATED BY SLASHES		
1 (Left Justify) - 78	79	80
		T

STATION CARD (S CARD)

SORT NO.	PRIMARY STATION CODE (Left Justify)	(Blank)	STATION ALIASES (1) SECONDARY STATION CODE (Left Justify)	STATION ALIASES (2) SECONDARY STATION CODE (Left Justify)	STATION ALIASES (3) SECONDARY STATION CODE (Left Justify)	FIPS CODES STATE	FIPS CODES COUNTY	(Blank)	CONTROL CODE	
1 - 3	4 - 18	19 - 33	34 - 45	46 - 57	58 - 67	68 69	70 - 72	73 - 77	78 79	80
										S

LATITUDE/LONGITUDE CARD (HEADER CARD 0)

SORT NO.	(Blank)	LATITUDE (Zero Fill) dd	LATITUDE mm	LATITUDE sss	LONGITUDE (Zero Fill) ddd	LONGITUDE mm	LONGITUDE sss	(Blank)	P C	D	TOTAL STATION DEPTH	(Blank)	
1 - 3	4 - 6	7 8	9 10	11 - 13	14 - 16	17 18	19 - 21	22 - 27	28	29	30 - 32	33 - 79	80
													0

RMI CARD (HEADER CARD 1)

SORT NO.	(Blank)	TERM MAJ BASIN	TERM MIN BASIN	TERM STREAM NO.	LEVEL 1 MILES	LEVEL 2 INDEX	LEVEL 2 MILES	LEVEL 3 INDEX	LEVEL 3 MILES	LEVEL 4 INDEX	LEVEL 4 MILES	LEVEL 5 INDEX	LEVEL 5 MILES	LEVEL 6 INDEX	LEVEL 6 MILES	LEVEL 7 INDEX	LEVEL 7 MILES	LAST LEVEL USED		
1 - 3	4 - 6	7 8	9 10	11 - 13	14 - 19	20 - 25	26 - 31	32 - 36	37 - 41	42 - 45	46 - 50	51 - 54	55 - 59	60 - 63	64 - 68	69 - 72	73 - 76	77 78	79	80
																				1

RMI CARD (HEADER CARD 2)

SORT NO.	(Blank)	TERM MAJ BASIN	TERM MIN BASIN	TERM STREAM NO.	LEVEL 8 INDEX	LEVEL 8 MILES	LEVEL 9 INDEX	LEVEL 9 MILES	LEVEL 10 INDEX	LEVEL 10 MILES	LEVEL 11 INDEX	LEVEL 11 MILES	LEVEL 12 INDEX	LEVEL 12 MILES	(Blank)	
1 - 3	4 - 6	7 8	9 10	11 - 13	14 - 17	18 - 21	22 - 25	26 - 29	30 - 33	34 - 37	38 - 41	42 - 45	46 - 49	50 - 53	54 - 79	80
																2

MAJOR/MINOR BASIN CARD (HEADER CARD 3)

SORT NO.	STATION LOCATION MAJOR BASIN NAME	STATION LOCATION MINOR BASIN NAME	LOCATION BASIN CODES Major	LOCATION BASIN CODES Minor	LOCATION BASIN CODES Sub	(Blank)	
1 - 3	4 - 37	38 - 67	68 69	70 71	72 73	74 - 79	80
							3

LOCATION DESCRIPTION CARD (HEADER CARD 4)

SORT NO.	LOCATION OF SITE; LANDMARK NAMES, ETC	REACH CODING REACH NUMBER	REACH CODING ON OFF	REACH CODING MILES ON REACH	C K	(Blank)	
1 - 3	4 - 51	52 - 62	63 - 65	66 - 73	74	75 - 79	80
							4

EPA Form 7500-24 (Rev. 8-83) Previous edition is obsolete.

PAGE 1 OF 2

Exhibit 4-1 (continued)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80

DESCRIPTIVE PARAGRAPH CARDS (HEADER CARD 5)

DESCRIPTION (7–78)

SORT NO. 1 - 3	(Blank) 4 - 6	7 - 78	79	80
			0	5
			1	5
			2	5
			3	5
			4	5
			5	5
			6	5
			7	5
			8	5
			9	5
			J	5
			K	5
			L	5
			M	5
			N	5

EPA Form 7500-24 (Rev. 8-83) Previous edition is obsolete.

PAGE 2 OF 2

Data in the station header remains relatively fixed for the life of the station. Although data in the station header may be changed, they are not normally added to or changed once established.

Exhibit 4-1 is an example of the EPA form used to input data for storing station header information in the STORET system.

The information contained on a station header is listed below. The items which are required by the system are starred.

- Agency Code* - identification code for contributor of the data to STORET; this is not the users ID.
- Station Codes*
  - Primary*
  - Secondary
- Station Name*
- Station Location Description*
  - Latitude/longitude coordinates*
  - State and county codes*
  - Major/Minor/Sub Basin Codes*
  - USGS Hydrologic Unit Code*
- Station Type Code

The recommended organization of these items for ground-water data is detailed in the sections that follow.

4.3.1 Agency Code

An Agency Code must be established before any station or sampling information can be entered into STORET.

The agency code in the STORET station header is that element which identifies the "coordinator" of the data. All STORET monitoring station must have an agency code. For ground-water monitoring data, the agency code represents the organization or regulator who is contributing the ground-water data to STORET and not each facility contributing data to the organization or regulator. An organization coordinating data being put into STORET may have one or more agency codes, but each station in STORET may have only one code.

Ground-Water users should store ground-water monitoring data under an agency code that is unique to other agency codes the users' organization may have. This will allow you to segregate and streamline your data base from existing files containing surface water quality data. Users may identify other agency codes that may exist for their organization in several ways:

- By contacting the ADP specialist or the EPA liason in that organization
- By contacting the ADP Coordinator and/or STORET coordinator in the EPA Regional Office
- By contacting STORET User Assistance
- By using a STORET command procedure called "%contacts".

There are four recommended components to an agency code for ground-water monitoring data. In general, the eight-digit STORET agency code for ground-water monitoring data should be composed of:

- A STORET-understood indication in the first two characters of who that agency code represents, which include:
    - "11" for an EPA Regional office or other Federal agency
    - "21" for a State agency.
- A two-letter abbreviation or number representing the agency as the third and fourth characters, which may be either:
    - The Federal Information Processing System two-letter abbreviation for the State (the standard abbreviation, which can be found in Appendix C of the STORET Users' Handbook)
    - "HQ" for EPA Headquarters program offices

      The two-digit number of the EPA Region, such as "01" for Region I or "08" for Region VIII.

4.3.2 Station Codes

Station codes are codes which identify the specific station within the monitoring network from which the data were taken. For ground-water monitoring data, each well is considered a "station". Each station must have at least one, but may have up to four station codes:

- One primary station code, which is required
- Up to three secondary station codes, or "aliases", which may be assigned to a well, but are not required. These are used for the purpose of cross referencing stations for cooperative monitoring and to facilitate the retrieval process.

Station codes may be numeric, alphabetic, or a combination of both. The two important requirements for station codes are that the primary code may not exceed 15 characters in length. A user may create up to 3 secondary station codes. The 1st and 2nd secondary station codes may be up to 12 characters in length, and the 3rd secondary station code may be up to 10 characters in length. Each code, whether it is primary or secondary, must be unique from all other station codes already stored under a particular Agency code.

4.3.2.1 Primary Station Codes

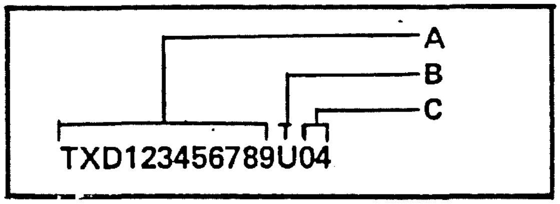
The primary station code for ground-water monitoring wells should be developed in one of two ways. If the monitoring well is located at a facility with a Dun and Bradstreet (D&B) that D & B number should make up the first 12 characters of the code. D & B numbers already exist for most regulated facilities, and should not be "made up" arbitrarily by the user. D & B numbers for facilities may be identified through EPA's FINDS Data Base. Users should contact the group that maintains FINDS in the EPA Regional Office to identify whether or not a D & B number already exists for their facility. If not, the FINDS group will have one created.

If the well is not at a facility with a Dun and Bradstreet number, for example a state ambient ground-water monitoring network, the primary station code may start with the letters "GDW" which will indicate that the data is from a ground-water monitoring station and not registered in the FINDS data base. The prefered method of developing primary station numbers at any facility is through the FINDS data base managers since FINDS serves as a cross-reference between EPA data bases. The "GDW" prefix is an exception to this convention and use should be restricted to circumstances where a D & B number is not appropriate.

The RCRA program has specified that the primary station code for groundwater monitoring should include the following three elements:

- Facility Dun & Bradstreet number of facility
- Relative position of the well to other wells for that waste management area or site (i.e., up- or downgradient)
- Unique number of the well for that waste management area.

For example, a primary station code for a RCRA ground-water monitoring well may look like:



Key

A	Dun and Bradstreet Number
B	Relative position of the well
C	Well number

The first 12 characters of this primary station code make up the Dun & Bradstreet (D & B) number of the facility at which the station is located. The D & B number of each facility is unique, and is used as a standard for many EPA (and other) data bases.

The character that follows the D & B number in the STORET primary station code (the thirteenth of the 15 characters) for wells is an alphabetic character that describes the relative position of the well to the waste management area. The thirteenth character should be a value of "D" for downgradient or a value of "U" for upgradient.

If another code is needed (to represent a different type of station designation), users should contact OSW and inform them of this need. OSW will oversee creation and documentation of the new code.

The last two of the fifteen primary station code characters are for assigning a number to each monitoring station that is unique for that waste management area. Stations may be numbered sequentially starting from the first upgradient station, and increasing by one or more in a clockwise direction. Sequential numbering is particularly meaningful if additions or deletions of stations are not expected. However, as changes are likely to happen during the regulated life of the facility, just assigning 2-digit station numbers that are not sequential (but are unique for each station is acceptable. Ninety-nine stations may be stored in this fashion. Should the user need to number more than 99 stations letters of the alphabet should be used (for example, A0, A1, A2, etc., up to A9, and the B0, B1, etc).

4.3.2.2 Secondary Station Codes

Assignment of the secondary station codes for a ground-water monitoring station is left up to the user. Users may enter any numbering scheme they wish for the station "alias". Users may be interested in having the secondary station code be any of the following:

- The permit number for any permits which have been issued to the facility
- An identifier for the district or region in which the facility is located
- An internal numbering scheme.

Users may keep in mind that data retrievals may be based on any of the station codes, not just the primary, station code. For that reason, each station code, regardless of whether it is primary or secondary, must be unique.

4.3.3 Station Naming Conventions

The "station location" or "station name" is a required data element in the STORET station header. No data selection may be made based on the information appearing in the station name, nor does the station name appear in every output report. However, it is one of the few opportunities for users to enter a brief textual description of a station.

Station names may be up to 48 characters in length. Users may have to abbreviate some words in the station name in order to enter all the needed information within the 48-character length limit.

The RCRA program has developed a specific convention they would like used by users naming RCRA monitoring stations. This format can serve as a model for other programs.

The STORET station name for RCRA stations should have the following items of information:

- Facility name
- Type of waste management area
- An indication that the data is collected from the station under RCRA
- The position and internal number of the station.

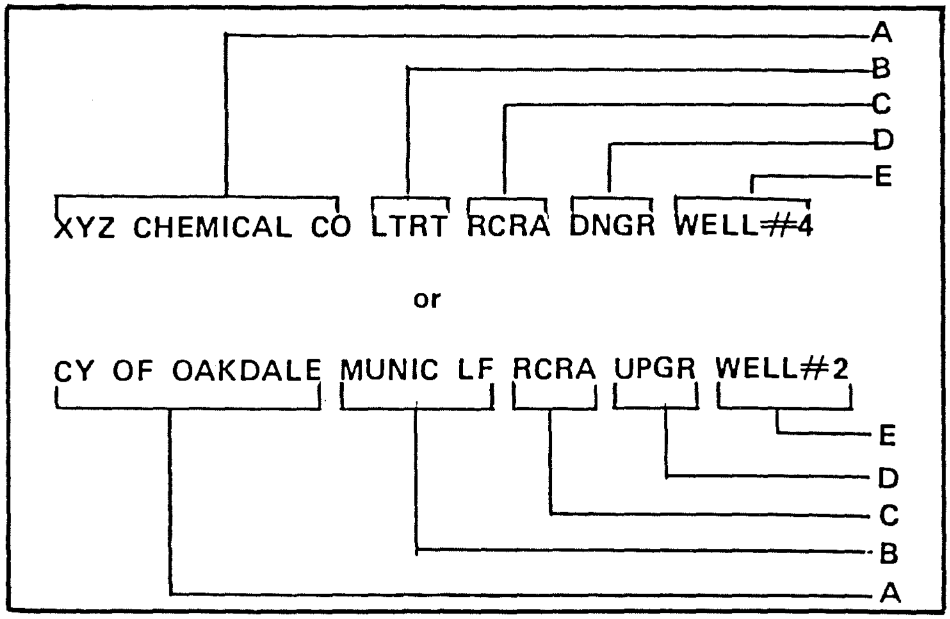
The facility name used for the station name in STORET should resemble the primary facility name in the FINDS data base as closely as possible. Users may find out the exact name of the facility as it has been entered in FINDS by calling the FINDS group in the EPA Regional Office.

The next component of the station name should be the type of waste management area being monitored. For example, the type of waste management area might be a landfill or a surface impoundment.

The third component of the station name should simply be "RCRA". This is entered to distinguish the station as one located in a controlled hazardous waste facility.

The last component of the station name should be an identification of the position and number of the station amongst the RCRA stations for this waste management area. This component should have the same information as that in the last three digits of the primary station code, but presented more explicitly, such as "UPGR WELL#1" instead of "U01".

Some examples of complete station names for RCRA stations include:



Key

A	Facility Name
B	Type of waste management area
C	Program under which data collected
D	Position of station (well)
E	Internal station (well) number

4.3.4 Station Location Identification

There are several data elements in the STORET station header that identify the location of the station. These include:

- The latitude/longitude coordinates of the station (plus a precision code)
- The hydrologic location of the station, expressed as major/ minor/sub-basin codes and names
- The state and county in which the station is located
- The USGS hydrologic unit in which the station is located.

An explanation of how to enter these locational data elements for ground water monitoring stations is given in the following sections.

4.3.4.1 Latitude/Longitude and Precision Codes

Latitude/longitude coordinates are required data elements for all STORET stations (they are two separate station header data elements). The latitude in STORET is formatted as "DDMMSS.S" and the longitude as "DDDMMSS.S", where D equals degrees, M equals minutes, and S equals seconds. Note that expression of degrees greater than one hundred is acceptable for longitude.

It is important to ensure that the latitude/longitude is for the station at a specific facility and not for the corporate office of the company. In the past there has been some errors in the entering of this very important code.

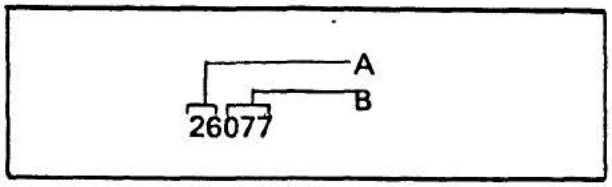
Both latitude and longitude may be entered to the tenths of seconds. This corresponds to approximately 10 feet. In order to distinguish each monitoring station in STORET as clearly as possible, the latitude/longitude coordinates of all ground-water monitoring stations in STORET should be precise to the tenths of seconds. The precision code, which must be entered with the latitude/longitude coordinates reflect the user's knowledge of how precise the coordinates are. An explanation of how to enter latitude/longitude coordinates and the accompanying precision code may be found in Chapter WQ-DE of the STORET User's Handbook (February 1982).

4.3.4.2 Major/Minor/Sub-Basin Codes and Names

Ground-water Users must also identify the hydrologic basin in which the station is located. Users should identify the major, minor, and sub-basins in which the facility or study area at which the station is located. These are defined in Appendix C of the STORET Users' Handbook, (February 1982). Identification of the aquifer which is being monitored is done in the parametric data field. This is described in Section 4.6.

4.3.4.3 FIPS State/County Codes

To identify the geopolitical location of the station, users must use the Federal Information Processing System (FIPS) code for state and county. FIPS codes are five-digit codes, the first two representing the state and the last three representing the county. Users should enter the FIPS codes and state/county names as they are shown in Appendix C of the STORET User's Handbook (February 1982). For example, the total FIPS code for a site in Kalamazoo County, Michigan would be:



Key

A State code for the State of Michigan is 26

B County Code for Kalamazoo in Michigan is 077

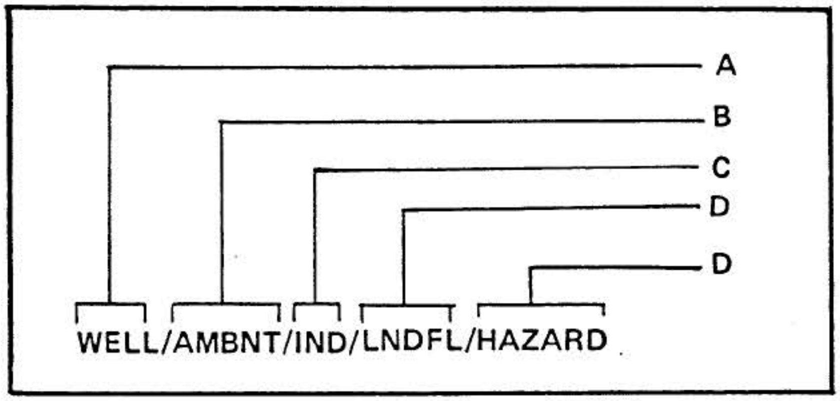
4.3.5 Station Type Codes

Station type codes are those station header data elements that describe the type and purpose of the monitoring station. STORET currently has several categories of station type codes. In STORET terminology each category is called a level. Different levels are required depending on the type of data you are entering (Ground-Water, Surface Water POTW, sediment, or biological)

Existing and planned station type codes of particular importance to Ground-Water/STORET users are highlighted in Exhibit 4-1. Many of the station type codes that are recommended for the station headers of ground-water monitoring wells are not of the "required" type for STORET. However, in order to completely distinguish ground-water monitoring stations from others in STORET, they should be used.

Users are required to specify one level and one level two code for each station entered into STORET. The station codes in levels three through five are optional and may be used in any combination to further describe the sampling site. For example, with ground-water data level 4 codes are not pertinent. However, using two level 5 codes may make station identification easier.

Users must string together station type codes relevant to their station. For example, a station monitoring a hazardous waste landfill located in an industrial facility might have a station type code of:



Key

A	Level 1 Code
B	Level 2 Code
C	Level 3 Code
D	Level 5 Code

The complete list of valid STORET station type codes may be retrieved in the on-line data set called "STORET.HELP.STATION.TYPE".

4.3.6 Station Depth

The station depth field in the STORET station header is used, for surface water stations, to store the total depth (i.e., from surface to bottom) of the point where the sample was taken. Ground-Water/STORET users may store the aquifer thickness, at the point where the well is located, in the station depth field.

4.3.7 Hydrologic Unit Code

The hydrologic unit code of the STORET station header is an eight-digit code representing the USGS hydrologic unit in which the station is located. The components of the codes represent hydrologic region, sub-region, accounting unit, and cataloging unit. This coding scheme represents different basin designations than the major/minor/sub-basin fields and is required by STORET. Exhibit 4-2 depicts the scheme of USGS hydrologic unit designations. Clearer, larger-scale maps titled Accounting Units of the National Water Data Network of Hydrologic Units may be obtained from the Geological Survey.

EXHIBIT 4-2

Recommended STORET Station-Type Codes for Ground-Water Monitoring Stations

Level	Requirement	Code	Definition
1	Required, 1 only	WELL	Station samples from a well.
		SPRING	Station samples from a spring. A natural flow of ground water from the earth which feeds into a stream or other body of water.
2	Required, 1 only	AMBNT	Indicates monitoring of ambient conditions in the environment as opposed to within a facility or in the effluent.
		NONAMB	Indicates monitoring at or within a man-made facility or in the effluent.
3	Optional, 1 only	MUN	Municipal water supply and/or sewage treatment.
		IND	Industrial facility.
		CMBMI	Combined "MUN" and "IND."
		AGRI	Agricultural. Includes raw crops, feedlots, grazing, and silviculture.
		DOMEST	Domestic (residential) domicile or facility. Includes water supplies and on-lot septic systems for private dwellings.
		ABANDN	The station from which samples are gathered is abandoned.
		DISPOS	Waste disposal facility.

EXHIBIT 4-2 (continued)

Recommended STORET Station-Type Codes for Ground-Water Monitoring Stations

Level	Requirement	Code	Definition
4	None recommended for ground-water monitoring stations (wells)		
5	Optional, as many as are appropriate should be entered	HAZARD	Site of hazardous or toxic wastes.
		LNDFL	Landfill.
		IMPDMT	Impoundment. Includes waste pits, treatment lagoons, settling ponds, and evaporation ponds.
		LNDTRT	Land treatment area.
		UPGR	Upgradient of a well or spring.
		DOWNGR	Downgradient of well or spring.
		RCRA	RCRA monitoring site.
		CERCLA	"Superfund" monitoring site.
		INJECT	Site where liquid waste has been injected under ground as a means of disposal.
		SUPPLY	Water supply storage or treatment facility.
		HRZTL	Horizontal well.
		TUNNEL	An underground corridor.
		GALERY	An artificial underground structure implanted to collect ground water.

4.3.8 Unlocking Key and Dates

Many users are concerned about data security. The STORET System contains an element called an "unlocking key" which helps prevent data from either being changed or retrieved by unauthorized persons. The unlocking key is required to store or change any data for an agency and is selected when the agency code is established. No one but a user with the unlocking key may ever store or change any data in STORET. Retrievals can not be made of data that has been flagged as locked by the owner without the unlocking key.

The unlocking key is used in conjuction with a "lock-after" date if users wish no one to look at the data who isn't authorized. The lock-after date is expressed as "YY/MM" (year/month). If a user wants no one to look at the data who is not authorized, a lock-after date of 50/01 (or some other date in the past) should be entered with their unlocking key. This means that data sampled after January 1950, will not be retrieved without including the unlocking key in the retrieval request. If the user wants no one to be able to change the data, but does not mind if others access the data in retrievals, a lock-after date of "99/99" should be entered. As this is an invalid date, 99/99 is used to signal that the data is unlocked for retrieval.

4.4 Descriptive Paragraph

The descriptive paragraph for a STORET station is that field in which a user may store any information in any format about the station (well) they wish. The descriptive paragraph may be up to 1080 characters and composed of 15 lines that are each no more than 72 characters in length. Its use is optional. No automated data selection may be made based on the information in it. Therefore, only data on which no retrievals are likely to be based should be stored in the descriptive paragraph.

The following paragraph contains the information that the RCRA program has indicated RCRA STORET users might want to include in a descriptive paragraph and serves as an example of the type of information other programs might want included in their users descriptive paragraphs. Any or all of the following pieces of information may be included:

- Total number of upgradient and downgradient monitoring stations at the facility, and relation of this station to the others, for example:

> UPGRADIENT WELL #3 OF 20 UPGRADIENT AND
> 60 DOWNGRADIENT RCRA MONITORING WELLS

- Type of business conducted at the facility
- Permit and/or enforcement status of the facility
- History of the facility, including any recent spills or other permit violations
- RCRA-regulating agency (e.g., EPA or State)
- Other permits given to the facility, including Federal and other
- Land use of the surrounding area
- Potential targets or nearby sensitive ecosystems
- Proximity and position relative to drinking water sources
- Contacts from whom further information about the facility may be obtained
- Other non-RCRA monitoring points which may be located at or near the facility.

Complete instuctions on how to create a descriptive paragraph for STORET stations may be found in Chapter WQ-DE of the STORET User's Handbook (February 1982).

4.5 Parametric Data Field

4.5.1 Formatting Information

It should be noted that the storage format to be used for information stored in the parametric data field is not the same as for the station and descriptive paragraph fields, which was ?01. Although there are several formats which can be used with the parametric data field, only the ?00 format is recommended for users inputting ground-water data even though users may have used ?01, ?02, ?03 and ?04 to input surface water quality parametric data. The reasons for this are explained below.

Data stored in the Water Quality File of STORET may be obtained from various types of sampling methodologies. The various types of samples from which data is collected are: grab samples; unqualified composite samples; qualified composite samples; multiple samples with a system multipurpose key (SMK); and samples with a user multipurpose key (UMK). For definitions of these various sample types, see Chapter WQ-DE in the STORET User Handbook (February 1982). Only the ?00 and ?04 storage formats enable to user to input information from all of the sample types. This is particularly important for ground-water users because only the ?00 format will enable you to easily store information on replicate samples used in many ground-water programs. Additionally the ?00 format has the flexibility to store the data utilizing either the parameter code followed by a value approach or a matrix approach which requires fewer keystrokes. It is recommended that the ?00 procedure be used to store all the parametric ground-water data. Complete documentation of the ?00 format, is contained in "STORET.HELP.SEMINAR.DOC.STORAGE"

The STORET command procedure %EASYSTOR allows users to interactively create storage transactions for the entry of either station location information or parametric data into the data base. For station location data storage, this procedure is particularly useful and efficient when a relatively few stations (say, up to a dozen or so) are to be entered into the system. The procedure prompts for each item of station location data, and consequently, users need not be concerned about column placement and other station storage card format requirements. For parametric data storage, data may be entered using any of the five (?00/01/02/03 or 04) storate procedures. Again, this procedure is most efficient (in terms of user time and computer costs) when there is a relatively small amount of parametric data to be added to the system.

4.5.2. Inputting Data

There are a number of types of ground-water monitoring information that should be stored in STORET's parametric data field. These include:

- Station Descriptors:
  - Physical setting descriptors
  - Well descriptors

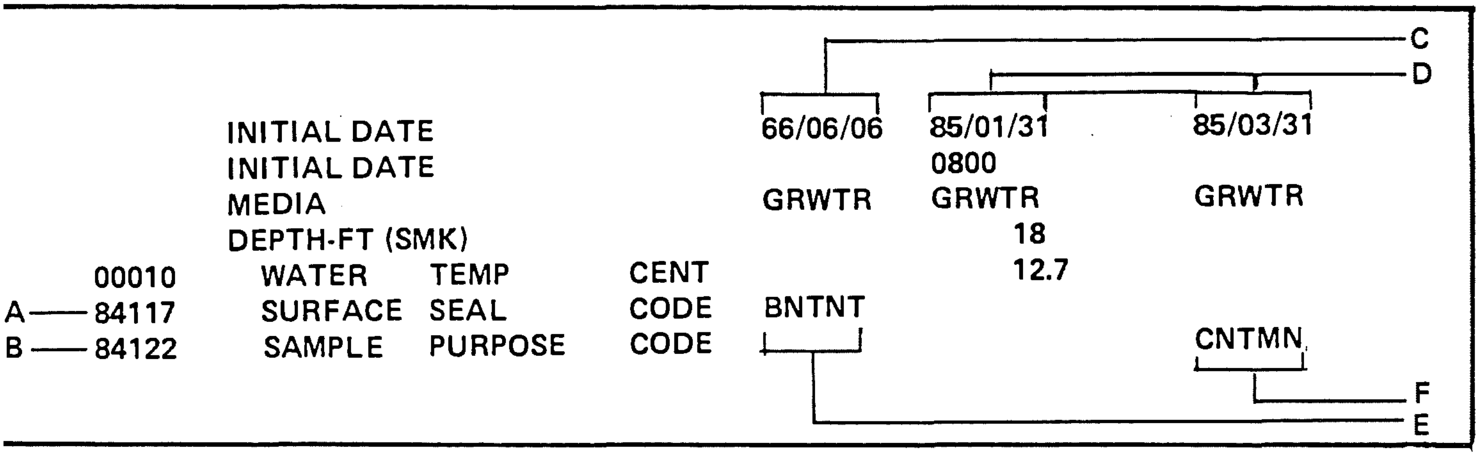
- Sample Descriptors:
  - Sample purpose descriptors
  - Sample conditions descriptors
  - Sampling and analytical methods
- Analytical findings for sampled parameters
- Special data-point qualifiers (remarks).

4.5.3 Station Descriptors

Station descriptors were defined and listed in Section 4.11 and Appendix C of this manual. These are data that describe the facility, geohydrology, and wells from which samples are taken, and are not expected to change over time. As they are not expected to change, it is best that Ground-Water/STORET users store them in the parametric data field with a date specially used for unchanging data. In other words, all unchanging data stored in the parametric data field should be stored as if they were collected on "66/06/06". Dynamic data that will change from sampling event to sampling event will be stored with the date the sample was collected.

Station descriptors include facility descriptors, physical setting descriptors, and well descriptors which were discussed in section 4.1.1 of this Chapter. All facility descriptors, except well ownership, should be stored in the Station Header or descriptive paragraph field. Most physical setting descriptors and well descriptors will be stored in the parametric data field under the "fixed" date ("66/06/06)". Refer to Appendix A to see which physical setting descriptors and well descriptors must be stored in the Station (well) Header, Descriptive Paragraph and Parametric Variable Date Fields. Unchanging date, data stored in the parametric data field might look like:

This example is from a section of a STORET "ALLPARM" retrieval output.



```
               INITIAL DATE                        66/06/06   85/01/31        85/03/31
               INITIAL DATE                                   0800
               MEDIA                               GRWTR      GRWTR           GRWTR
               DEPTH-FT (SMK)                                      18
      00010      WATER     TEMP       CENT                         12.7
A —— 84117       SURFACE   SEAL       CODE         BNTNT
B —— 84122       SAMPLE    PURPOSE    CODE                                    CNTMN
```

Key

A Parametric station descriptor

B Parametric sample descriptor

C Fixed date

D Variable date

E Codes representing Bentenite for the well surface seal

F Code indicating that the well was sampled to determine ground-water contamination

Many of the station descriptors that will be stored in the parametric data field will be stored under parameter codes whose values are coded (e.g., drilling method and drilling fluid from the example). This means that the values given to each data element will be a code that represents something. To identify what the code represents, the user must look up its definition in a table. The site descriptor parameter codes for ground-water data, their coded values and their definitions may be found in Appendix B of this report. For certain physical setting descriptors, namely "Geologic age or formation name" and "Aquifer name", the coded values are presented in Appendix H.

4.5.4 Sample Descriptors

Sample descriptors were defined in Section 4.1.2 of this manual. These are data that describe the sample purpose, sampling conditions, and sampling/ analytical methodology, and are expected to change over time.

Sample purpose descriptors, sampling conditions descriptors, and some sampling/analytical methods descriptors should be stored much the same way as any other "usual" parametric data in STORET, i.e., they should be stored with the analytical findings for each sampling event (special procedures for many of the sampling and analytical methods data are described below). Many of these data elements have coded data values. The parameter codes, value codes, and definitions for sample purpose, sampling conditions descriptors, and testing/analytical methods can be found in Appendix F.

4.5.4.1 Expanded Sample Key

Ground-water sampling and analytical methods for each sampling event will, for the most part, be stored in special STORET key fields which are part of the parametric data field. These fields, known as "media, SMK, UMK", have special ground-water applications which will be explained in the following paragraphs. Additionally they can only be used with the ?00 format.

The ?00 storage format was developed to allow the storage of samples taken from media other than water and to enable storage of additional sample information that is not possible with the other STORET formats. Every sample stored in STORET must be uniquely identified by date, time, and depth. The ?00 format allows you to further identify the sample via an expanded sample key which in addition to date and time also specifies the media, the system multi-purpose key (SMK), and the user multipurpose key (UMK). For samples other than water or water qualified samples (bottom, core etc.), the expanded sample key must be used.

4.5.4.1.1 Media Key

The media key identifies the medium in which sampling was done (e.g. water, sediment, etc). For ground water there are currently two media keys: "GRWTR" and "RCRAGW." "GRWTR" is to be used by all programs except RCRA. The RCRA Program has designated its own media key "RCRAGW" which is to be used for inputting RCRA ground-water monitoring data. Any program that wishes its own media key can develop its own code. This option is currently being investigated at the Agency.

4.5.4.1.2 System Multipurpose Key (SMK)

When doing ground-water monitoring it is common to extract more than one sample per monitoring station. There are several ways to obtain multiple samples and for quality assurance purposes it will be beneficial for STORET/Ground Water users to be able to distinguish the manner in which multiple samples were obtained. The system multipurpose key enables users to make this distinction. A brief review of sampling is included here which will help explain the coding used with the system multipurpose key.

There are three common methods for obtaining mutiple samples from ground-water monitoring stations.

- Several samples may be taken from the same sample point and placed into seperate sample bottles. For the purpose of this manual each individual sample of the total set will carry its own unique number.

- One sample may be taken from the sample point; immediately divided in the field and placed into different sample bottles. Each portion of the original sample now residing in separate sample bottles will be called a "field replicate" in this manual.

- One sample may be taken from a well and not divided into seperate sample bottles until it arrives at the laboratory. Each portion of the original sample now residing in sample bottles will be called a "laboratory replicate" in this manual.

It should be understood that when multiple samples are indicated on a data sheet it may mean that any one of the above methods was used to obtain the multiple sample or that a combination of the above methods was used to obtain the multiple sample. The SMK code will enable a STORET/Ground-Water User to determine whether the ground-water data is a multiple sample, what method(s) was used to obtain the multiple sample, how many multiple samples were taken, and which one of the multiple samples the data you are examining came from.

This information is obtained via the 6 digit SMK code. Each of the first 4 digit positions of the code signifying a specific piece of information. At this time the last 2 digits of the code will appear as zeros because no specific pieces of sampling information have yet been defined for these positions. One SMK will be entered for each sampling event (each set of multiple samples will be considered a sampling event).

The information conveyed in the first 4 digit positions of the SMK is summarized in the table on the following page.

Digit Position	Component Definition
1	Identifies whether the sample is an individual sample or one of a multiple sample set. For example, the sample in question could be: one of a set of samples taken from a station (well) and not further divided, one of a set of multiple samples divided in the field, one of a set of multiple samples divided in the laboratory, or a combination of the above The actual number placed in the first digit position will be a valve ranging from 0-7. The meaning attached to the first digit position number can be determined from the table on page 4-29.
2	Identifies which sample in the set of multiple samples the data you are retrieving comes from. For example, if a sampling event from a single sample point has resulted in 4 undivided samples the data values reported for the first sample would have a SMK code with "1" in the second digit position, the data values reported for the second sample would have a SMK with "2" in the second digit position etc.
3	Identifies which one of the field replicates the data you are retrieving comes from. For example, if one sample was collected at the sample point and divided into several sample bottles for analysis in the field, the data values reported for the first "field replicate" would have a "1" in the third digit position of the SMK code, the second field replicate would have "2" in the third digit position of the SMK code etc.
4	Identifies which one of the lab replicates the data you are retrieving comes from. For example if one sample was collected at the sample point and divided into several sample bottles for analysis in the lab the data values reported for the first "lab replicate" would have a "1" in the fourth digit position of the SMK code, the second "lab replicate" would have a "2" in the fourth digit position of the SMK etc.

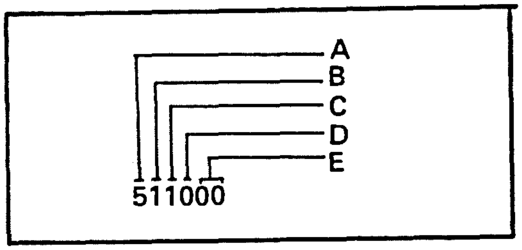
The following table will enable the user to determine the significance of the value appearing in the first digit position of the SMK.

Sample Media Key "SMK" Notation for Ground-Water/STORET User

First digit code	Multiple Sample	Field Replicate	Lab Replicate
0	No	No	No
1	No	No	Yes
2	No	Yes	No
3	Yes	No	No
4	No	Yes	Yes
5	Yes	Yes	No
6	Yes	No	Yes
7	Yes	Yes	Yes

For example, should the first value of a SMK code be 3, the user knows that several separate samples were taken at the sample point. None of these were further divided in either the field or the lab. If the first value of the SMK code were a 5, the user would know that several samples were taken at the station and further that one or all of the samples were subsequently divided in the field so that the data you have retrieved is from a field replicate. Finally, should the first value of the SMK be a 1, the user would know that only one sample was obtained from the sample point but this was divided into several portions in the lab and he/she is retrieving data from one of the lab replicates.

The example of a SMK code illustrated below indicates that the data you are observing is one sample from a set of samples, that this particular sample was the first in the lot. The sample was divided in the field, and that this is the first of the field replicates. There was no division in the lab.



Key

A	5 indicates that this one set of multiple samples which was subsequently divided into field replicates
B	1 indicates that this is the first replicate of the set of samples for this particular station
C	1 indicates that this is the first field replicate
D	0 indicates that the original samples were not divided in the lab.
E	These fields are currently undefined

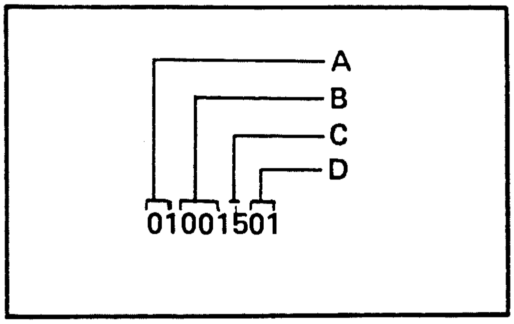
4.5.4.1.3 Users Multipurpose Key (UMK)

The UMK is an 8 digit number which will be used to describe the method in which a ground-water sample was collected and analyzed. There will be one "UMK" per sampling event (each multiple sample may be considered a sampling event). There are 4 components of a UMK code which are defined below.

Digit	Component Definition
1-2	Coded value for sampler type (see Appendix F)
3-5	Identifies up to three different materials making up the sampling equipment, 1-digit each (see Appendix F)
6	Flag for indicating whether the reported values were determined in the lab or the field (see Appendix F)
7-8	Coded value to identify analytical method used to determined reported values (see Appendix F)

The meaning of the Code appearing in each digit position is defined in Appendix F.

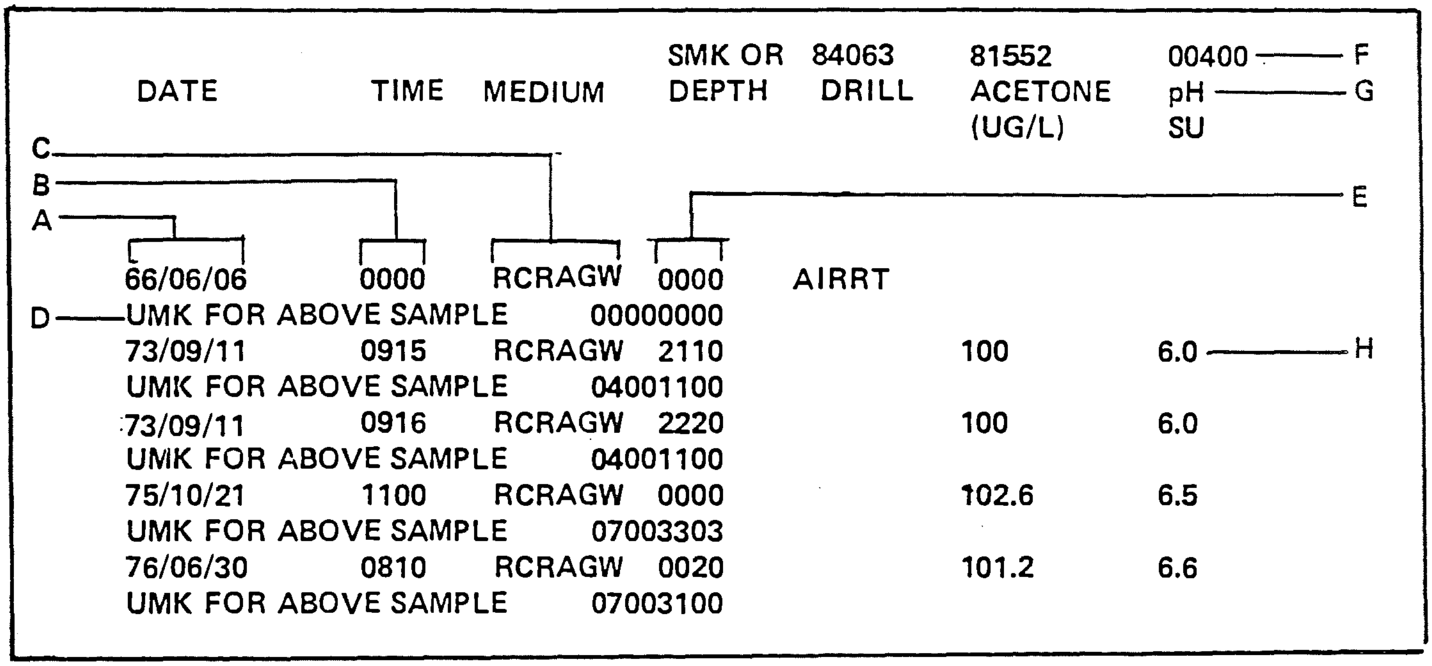
A sample UMK is illustrated below:



Key

A	Sampler type; 01 = Bottom Valve Bailer
B	Sampling equipment material; 001 = stainless steel
C	Lab or field determination flag; 5 = contract, field
D	Analytical method used to determine reported values; 01 = Gas chromatography/mass spectrometry (GC/MS)

An output of data stored in the parametric data field using "media, SMK, UMK" will look like:



```
                                   SMK OR   84063   81552      00400
     DATE          TIME   MEDIUM   DEPTH    DRILL   ACETONE    pH
                                                    (UG/L)     SU

     66/06/06      0000   RCRAGW   0000     AIRRT
     UMK FOR ABOVE SAMPLE     00000000
     73/09/11      0915   RCRAGW   2110             100        6.0
     UMK FOR ABOVE SAMPLE     04001100
     73/09/11      0916   RCRAGW   2220             100        6.0
     UMK FOR ABOVE SAMPLE     04001100
     75/10/21      1100   RCRAGW   0000             102.6      6.5
     UMK FOR ABOVE SAMPLE     07003303
     76/06/30      0810   RCRAGW   0020             101.2      6.6
     UMK FOR ABOVE SAMPLE     07003100
```

Key

A	Date Sample
B	Time Sampled
C	Media Code
D	UMK Code
E	SMK Code or Depth
F	Parameter Code
G	Parameter Abbreviation
H	Parameter Value

Further documentation and training to store data in STORET's "media, SMK, UMK" fields can be obtained from STORET User Assistance. (800-424-9067) codes for UMK fields can be found in Appendix E.

4.5.5 Analytical Findings

STORET "parameter codes", the computer codes used to store data, usually represent a combination of three items:

- Substance
- Medium
- Units

For example, parameter code 01000 represents "arsenic, dissolved fraction of water, ug/l", while a different parameter code, 01002, represents "arsenic, total water, ug/l" because the medium differs. Most ground-water users will be interested in storing their analytical findings under those parameter codes that represent a substance in "whole" or "total" water. Other codes are available should the user find them necessary.

A list of STORET parameter codes may be obtained in any one of the following sequences by using the STORET command %BATCH:

- PARMALFA -- identifies parameter codes in alphabetical order of parameter name
- PARNUMER -- identifies parameters in numerical order of parameter code
- PARCAS -- identifies parameters and parameter codes in order of CAS number

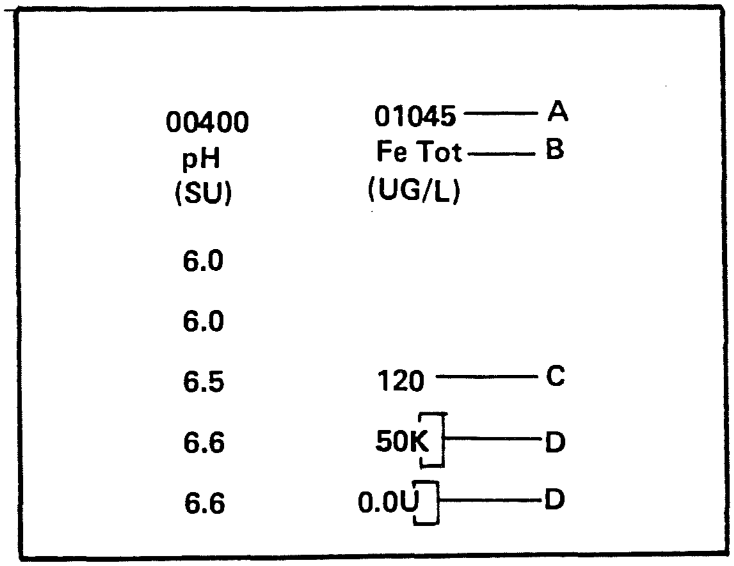
STORET parameter descriptive information and codes may be obtained interactively on the system through the use of the Water Quality Analysis Branch Conversational procedure WQAB PARM and the STORET command %PARABB. The WQAB PARM is quite useful for ground-water data managers and an example is given in Appendix H. If a parameter code for a particular substance does not exist, users may have it created by contacting the STORET Client Services Branch.

Additionally, the parameter codes of interest to RCRA data managers for storing data collected at interim status facilities from fully permitted facilities may be found in Appendices F and G respectively.

4.6 Qualifying "Remarks"

Ground-water data managers may find it necessary to explain something about a particular data point in addition to its numerical value. For example, a data point may be one that is being reported below the detection limit (e.g., "actual value is less than 150 ug/l") or substances may be tested for but not detected.

The STORET System has the capability to store "remarks" with each data value. These remarks flag the data value with such qualifications as "less than", "greater than" and "undetected". Appendix C of this report presents the remark codes that are currently valid in STORET. An example of a parametric data field with data that are remarked could be:



Key

A	Parameter Code
B	Parameter name
C	(Parametric) data value
D	Remark code data value (K = Actual value is known to be less than the value shown, U = material specifically analyzed for but not detected)

Upon data retrieval, users may select any of the following

- All data, regardless of the remark code
- Only data without remarks
- Only data with a specific remark code.

Remark codes may be entered with each data point. They need not always be used. Complete instructions on how to store data with remark codes may be found in Chapter WQ-DE of the STORET Users' Handbook. (February 1982).

It should be noted that some data (Superfund especially) may have remark codes identical to STORET's but with different definitions. These are usually indicated on the lab data forms. For example, remark code "B" in STORET refers to bacterial counts out of range, while Superfund uses "B" to indicate a compound found in a travel or lab blank sample. Another example is that Region IX's office policy is to subsitute "U" for "B" if data value is below the detection level. If the value is above dection level, no data is entered at all. These inconsistancies are mentioned to ensure that users "pre-edit" suspect data so that they maintain consistency with STORET remark codes.

4.7 Quality Assurance/Quality Control

Information on QA/QC for ground-water monitoring such as well construction, sampling methods and laboratory analysis techniques is extremely important because of the numerous factors which may affect the accuracy of the parameter values inputted into STORET. For example, knowledge of the well construction may help the user determine the reliability of the data, and knowing the sampling method used might help the users determine the possibility of sample aeration and a subsequent volitization of organics. Accessibility to this type of information will asist users to determine the usefulness of STORET data for their particular needs.

A parameter QA/QC code named Data Quality (84129) has been added to STORET. As with all parameter codes in STORET, 4 characters of coded values are available for use with the Data Quality parameter code. The presence of this code will enable users to store fairly detailed QA/QC information for each sample.

Quality Assurance/Control is a complex element of a data management system and involves many activities in well location, construction, sampling and laboratory analysis. A decision on how all or some of these activities should be included in the data base has not been made at this time. However, because of the importance of this issue, a preliminary approach has been added to STORET by the Office of Solid Waste. The approach will provide a mechanism for starting to address this topic. This approach is expected to be refined by EPA over the next year.

The 4 digit code contains the following for the specified digit positions.

- The first (left) character will contain a one digit code for the evaluation of well construction. The values in the first digit position will range from 0-2 or be blank. The meaning of each of the possible values is summarized below:

2 -- Well has been EPA/State inspected in the last 5 years and determined to be of high quality

1 -- Well has been properly drilled, constructed of inert materials, properly developed, properly located, and has controls to prevent tampering. Well constructed in accordance to guidance produced by EPA/State.

0 -- Well is known to be inadequate in some manner

blank -- Well information unknown or not stored.

- The second character will contain a 1 digit code for the evaluation of sampling QA/QC. The values of digits in the second position can range from 0-3 or be blank. The meaning for each value is given below:

- 3 -- EPA/State has performed a QA/QC evaluation within the last two years with a positive result
- 2 -- A detailed QA/QC plan with standard procedures and internal checks exists; the objectives of the plan have been verified as being met for at least one year (e.g., RCRA guidance for waste analysis, September 1984)
- 1 -- A detailed QA/QC plan with standard procedures and internal checks exists (e.g., RCRA guidance for waste analysis, September 1984)
- 0 -- No detailed QA/QC plan exists
- blank -- Information unknown or not stored.

- The third position character will contain a 1 digit code for the evaluation of laboratory QA/QC and will have values ranging from 0-3 or be blank. The meaning for these values is identical to position two, described above.

- The fourth position character will contain a 1 digit code for the evaluation of overall QA/QC during the entire sequence of the sampling event. This fourth position character can have values ranging from 0-3 or be blank. The meaning of the values selected for this position is identical to position two above.

Exhibit 4-3

USGS Map-Accounting Units of the National Water Data Network