EPA
600/2-89-040

United States
Environmental Protection
Agency

Robert S Kerr Environmental
Research Laboratory
Ada, OK 74820

EPA/600/2-89/040
July 1989

Research and Development

EPA

The Establishment of a Groundwater Research Data Center for Validation of Subsurface Flow and Transport Models

ENVIRONMENTAL PROTECTION AGENCY
DALLAS, TEXAS
LIBRARY

EPA/600/2-89/040
July 1989

THE ESTABLISHMENT OF A GROUNDWATER RESEARCH DATA CENTER FOR VALIDATION OF SUBSURFACE FLOW AND TRANSPORT MODELS

by

Paul K. M. van der Heijde, Wilbert I. M. Elderhorst,
Rachel A. Miller, and Manjit F. Trehan

International Ground Water Modeling Center
Holcomb Research Institute
Butler University
Indianapolis, Indiana

CR-813191

Project Officer

Joe R. Williams
Extramural Activities and Assistance Division
Robert S. Kerr Environmental Research Laboratory
Ada, Oklahoma 74820

U.S. ENVIRONMENTAL PROTECTION AGENCY
ROBERT S. KERR ENVIRONMENTAL RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
ADA, OKLAHOMA 74820

DISCLAIMER

The information in this document has been funded in part by the United States Environmental Protection Agency under CR-813191 to the Holcomb Research Institute, Butler University, Indianapolis, Indiana. It has been subjected to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. It does not necessarily reflect the views of the Agency and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

FOREWORD

EPA is charged by Congress to protect the Nation's land, air and water systems. Under a mandate of national environmental laws focused on air and water quality, solid waste management and the control of toxic substances, pesticides, noise and radiation, the Agency strives to formulate and implement actions which lead to a compatible balance between human activities and the ability of natural systems to support and nurture life.

The Robert S. Kerr Environmental Research Laboratory is the Agency's center of expertise for investigation of the soil and subsurface environment. Personnel at the Laboratory are responsible for management of research programs to: (a) determine the fate, transport, and transformation rates of pollutants in the soil, the unsaturated zone and the saturated zones of the subsurface environment; (b) define the processes to be used in characterizing the soil and subsurface environment as a receptor of pollutants; (c) develop techniques for predicting the effect of pollutants on ground water, soil and indigenous organisms; and (d) define and demonstrate the applicability and limitations of using natural processes, indigenous to the soil and subsurface environment, for the protection of this resource.

This report describes the activities performed at the Holcomb Research Institute to establish the International Ground Water Modeling Center's (IGWMC) Groundwater Research Data Center. The Data Center provides scientists with information regarding existing datasets and assists in accessing selected datasets. This secondary use of data reduces considerably the efforts and costs of data acquisition for model validation and other purposes. The datasets available from the Center or described in the Center's referral database provide a useful basis to test newly developed theories on subsurface flow and contaminant transport processes. Moreover, such datasets should prove useful for evaluation of monitoring equipment, monitoring strategies, sampling techniques of groundwater, and for evaluation of methods and techniques of parameter estimation. This report describes the design of centralized institutional arrangements for collecting, conditioning, documenting, storing, and distributing datasets resulting from field research on soil and groundwater pollution, and the establishment of a referral service for those datasets not managed by the clearinghouse.

During the course of the project a workshop was held to evaluate the Data Center objectives, the role of data in model validation, proposed procedures, quality assurance, and database design. The two-day workshop involved presentations, group discussions, and written comment on provided questionnaires. The workshop results are reported in Appendix A.

Clinton W. Hall

Clinton W. Hall
Director
Robert S. Kerr Environmental
Research Laboratory

ABSTRACT

The International Ground Water Modeling Center (IGWMC) has established a Groundwater Research Data Center which provides information on research datasets resulting from publicly funded field experiments regarding soil and groundwater pollution and related laboratory bench studies, and which distributes selected public domain datasets for testing and validation of models for flow and contaminant transport in the saturated and unsaturated zones of the underground.

To fulfill its advisory role, the Data Center analyzes information and documentation resulting from field and laboratory experiments and evaluates the appropriate datasets for their suitability in model testing and validation. (The Center has identified validation as the major secondary use of such data.) To assure consistency in the analysis and description of these datasets, and to provide an efficient way to search, retrieve, and report information on these datasets, the Data Center has developed a computerized data directory, SATURN, programmed independently from any proprietary software.

As secondary users of such soil water and groundwater data are highly interested in information relevant to the assessment of data quality, a primary concern of the Center is the evaluation and documentation of the level of quality assurance applied during data acquisition, data handling, and data storage.

In addition to providing referral services, the Data Center distributes, on an "as-is" basis, selected, high-quality datasets described in the data directory. The datasets of concern represent different hydrological, geological, and geographic-climatic settings, pollutant compositions, and degrees of contamination.

Because the quality of research data often is of great importance to the end-user, the Data Center has adopted internal QA/QC procedures and related institutional organization tailored to IGWMC's existing, highly successful QA/QC program in model information and software distribution.

The Center's detailed knowledge of the characteristics of a large number of subsurface flow and transport models and research datasets allow it to serve in an advisory role for both data collectors and modelers.

CONTENTS

Disclaimer Notice ii

Foreword iii

Abstract iv

List of Figures viii

List of Tables ix

Acknowledgment x

1. INTRODUCTION 1
   - Project Approach 4
   - Groundwater Research Data 5
   - Secondary Use of Research Data 6
   - Sharing Research Data 7

2. EXISTING DATASETS 9
   - Literature Survey 9
   - Preliminary Database for Information on Research Datasets 9
   - Evaluation of Preliminary Information on Research Datasets 11
   - Specific Large-scale Field Research Sites 16
   - Selection of Datasets for Incorporation into the Data Center 19

3. EXISTING DATA CENTERS 23
   - Data Centers and Data Clearinghouses 23
     - Previous Studies 23
     - Selected Natural Resources and Environmental Data Centers 26
     - NAWDEX 27
     - STORET 27
     - WATSTORE 29
     - Other Water Data Centers 29
     - Project Data Management 30
   - Data Management 30
     - User Access and Data Retrieval 31
     - Data Maintenance 31
     - Quality Assurance 32
     - Documentation 32
     - Standardization 34
     - Geographical Information Systems 34
     - Linking Data Management Systems 36
     - Data Center Maintenance 36
   - Conclusions 37

4. QUALITY ASSURANCE/QUALITY CONTROL FOR RESEARCH DATASETS 39
   - Introduction 39
     - QA/QC 39
     - The Role of Standards and Guidelines 40
     - Assessment of Data Quality for Secondary Use 41
   - Soil Water and Groundwater Chemistry 42
     - Introduction 42

Monitoring System Design and Sampling Programs 43
Effects of Well Construction on Sampling 44
Sampling Soil Water and Groundwater 45
Field Analyses 47
Sample Handling and Documentation 48
Laboratory Analysis 49
Data Quality Indicators 49
Laboratory QA/QC 50
Reporting Analytical Results 51
Discussion 51
Flow Information and Soil and Aquifer Parameters 52
Introduction 52
Water Level Measurement 53
Direct Measurement of Groundwater Flow 54
Vadose Zone Measurements 54
Geomechanical Measurements of Subsurface Materials 54
Determination of Aquifer Hydraulic Characteristics 55
Tracer Tests 56
Sampling Borehole Cuttings, Soils, and Cores 57
Geophysical Surveys 57
Laboratory Bench Studies 58
Data Transfer and Storage 59
Conclusion 59

5. STRUCTURE OF THE REFERRAL DATABASE AND REFERRAL FACILITY DESIGN 60
Introduction 60
Design Criteria 61
Completeness of Data 64
Balance of Information 64
Efficient Searches 65
Selected Search Strategy Criteria 66
Efficient Storage 68
Efficient Memory Usage 68
Other Design Considerations 68
Database Design 70
Structure of the Database 70
Implementation Details 73
User Interface and Database Management Programs 79
The Menu System 79

6. DETAILED ANALYSIS OF THE BORDEN DATASET 90
Introduction 90
General Project Summary and Objectives 90
Aquifer Description 92
Groundwater Flow 92
Monitoring System 92
Groundwater Quality 92
Analysis 96
Quality Assurance 96
Movement of tne Plume 96
Retardation of the Plume 97
Sorption of Organic Solutes 97
Dataset Contents 97
Screening Results 97
Discussion 98

7. DATA CENTER PROCEDURES 99
Introduction 99
Database and Dataset Management 99
Information Management 101
Information Acquisition and Processing 101
Identification of Potential Datasets 101
Information Dissemination 102
Dataset Management 102
Selection, Acquisition, and Evaluation Procedures 102
Potential Problems in Acquiring Datasets 103
Dataset Preparation and Distribution 104
Services Offered by the Data Center 105

8. INTERNAL QUALITY ASSURANCE/QUALITY CONTROL 107
Introduction 107
QA Organization 107
QA Tracking 108
Referral Database 108
Data Entry 108
Information Retrieval 110
Dataset Distribution 110
Acquisition 110
Evaluation 110
Distribution 110
QA Filing 118

9. CONCLUSIONS AND RECOMMENDATIONS 119

10. REFERENCES 121

APPENDIXES

A. Workshop Report Summary 135
B. Dataset Survey 158
C. SATURN Entry Form 200

LIST OF FIGURES

Figure 1. Structure of the IGWMC referral database systems and user interactions 62
Figure 2. Hierarchical structure of the SATURN database 71
Figure 3. SATURN database program: information levels and contents. 72
Figure 4. SATURN database program: use of tables and pointers 74
Figure 5. SATURN database program: binary tree structure 75
Figure 6. SATURN database program: file structure 76
Figure 7. SATURN database program: structure of the user interface. 80
Figure 8a. Main (horizontal) menu 81
Figure 8b. (Vertical) add option menu 81
Figure 8c. Change option menu 83
Figure 8d. Inquire (or search) option menu and display of existing dataset 83
Figure 8e. Report sub-menu of inquire option 87
Figure 8f. Display device menu of report sub-menu 87
Figure 8g. Report option of main menu requires site information 88
Figure 8h. Report option menu 89
Figure 8i. Display device menu of report option 89
Figure 9. The Borden landfill site 91
Figure 10. Cross-sectional view of the extent of the chloride plume at the Borden site 93
Figure 11. Water table maps for the tracer experimental site and vicinity 94
Figure 12. Location of multilevel samplers and injection wells as of January 1986 95
Figure 13. Data Center QA organization 108

LIST OF TABLES

Table 1. General characterization of research studies 12
Table 2. Organizations involved in field research 13
Table 3. Funding of site investigations 13
Table 4. Key pollutants at research sites 14
Table 5. Research sites per state/continent 15
Table 6. Lithology of the research sites 16
Table 7. Preliminary information on field sites and datasets 21
Table 8. ESIS information categories for databases 25
Table 9. ESIS information categories for data systems 26
Table 10. Components of the proposed voluntary standard for abstracts of machine-readable files (OIRA 1983) 35
Table 11. Selected search strategy criteria 67
Table 12. Descriptors included in summary print option 88
Table 13. SATURN annotation processing form 109
Table 14. SATURN search processing form 111
Table 15. IGWMC dataset tracking form 113
Table 16. IGWMC dataset evaluation form 114
Table 17. IGWMC dataset distribution form 117

ACKNOWLEDGMENT

The authors are grateful to Margaret A. Butorac and Karen Ochsenrider for project assistance; to Ginger Williams and Mary Willis for word processing; to James N. Rogers for manuscript editing and production; and to Colleen Baker for graphics.

SECTION 1

INTRODUCTION

Widespread concern about the protection and rehabilitation of groundwater resources is being expressed by both the public and the scientific community. The federal government, through the U.S. Environmental Protection Agency (EPA) and other agencies, is responding to this concern in several ways over and above its National Ground Water Protection Strategy (EPA 1984a). Although the Strategy emphasizes protection, it also stresses mitigation of groundwater contamination, or rehabilitation, at a moderate number of Superfund sites. For policy development, regulation, and enforcement, and for researching the physical, chemical, and biological principles underlying these and other areas of EPA responsibility, the Agency and the scientific community require broad access to the most complete and reliable data on groundwater systems.

Mathematical formulation of acquired scientific understanding forms the basis of the predictive simulation capabilities essential to every regulatory program concerned with groundwater contamination. At the federal level these programs are authorized under the Resource Conservation and Recovery Act (RCRA), the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA or "Superfund"), the Underground Injection Control (UIC) regulations of the Safe Drinking Water Act, and the Clean Water Act (CWA). Other programs focusing on groundwater operate under the authority of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Toxic Substances Control Act (TSCA).

The ability to predict accurately the transport and fate of potential contaminants is critical to the success of most groundwater regulations (EPA 1985). For protecting the integrity of an aquifer or engineered facility, monitoring of groundwater quality often is an ineffective alternative to predictive modeling. Thus, development and adoption of methods for predicting pollutant transport and fate in the saturated and unsaturated zones of the subsurface are key elements of the EPA's groundwater research strategy (EPA/NCGWR 1982) and in the research programs of other agencies. The development and accuracy of such predictive capabilities cannot take place without an equally significant effort in subsurface characterization.

With the growing availability and use of subsurface flow and transport models, concerns regarding their validity and accuracy have increased. Model testing, or more specifically model validation, provides model users, decision makers, policy makers, and legal authorities with information on a model's performance characteristics—information needed to judge the usefulness of the model results for their problem assessments. An extensive discussion of model validation principles, definitions, and procedures can be found for vadose zone models in Hern and Melancon (1986) and for groundwater models in general in van der Heijde et al (1988).

As is discussed in van der Heijde et al. (1988) model, validation is basically part of the scientific discovery process. Determining the validity of a model is in fact accepting the validity of a theory relating quantitatively cause and effect in a observed, natural system. A comprehensive approach to model validation consists of assessing through examination and measurement the correctness of the model concepts, mathematical formulations, and the computer code representing these mathematical expressions. Practically, the objective of model validation is to determine how well a model's theoretical foundation and computer implementation describe actual system behavior by comparing the results of model calculations with numerical data independently derived from specially designed experiments or from detailed observations of a natural system specifically selected for this purpose. In general, the parameters required for the model calculations are estimates based on direct field observations or derived indirectly through subsequent parameter analysis. In both cases the natural or "real" system is observed by sampling the real system inputs (system stresses) and outputs (system responses) resulting in measured input and output values. As these measured data are samples of the real system, they are prone to sampling and measurement errors. Thus, model validity established by comparing calculated values with independently measured values is always subjective.

Ultimately, the "success" of a model validation attempt depends on how accurate and complete the independent system measurements are (which cannot be objectively determined), how small the difference is between the results of the model calculations and the measured values, and the validity criteria accepted by the model evaluators.

To consider a particular computer model valid under site-specific conditions, several tests covering a wide variety of site conditions must be performed under highly controlled circumstances and using quality-assured data. A large number of tests might be involved in the comprehensive validation of a model. The experience gained at the International Ground Water Modeling Center shows that such extensive validation studies are often lacking (van der Heijde et al. 1988).

This finding is confirmed by international model verification and validation studies such as INTRACOIN (International Nuclide Transport Code Intercomparison Study, completed in 1984) and HYDROCOIN (Hydrologic Code Intercomparison Study), organized by the Swedish Nuclear Power Inspectorate (INTRACOIN 1986, HYDROCOIN 1987, Nicholson et al. 1987).

Level 2 of the three-level HYDROCOIN studies was aimed at the validation of mathematical models describing the physical processes involved in groundwater hydrology by comparison of calculations with observations and experimental measurements for five distinct cases. These cases covered heat transfer involving thermal convection and conduction, variable density fluid flow based on a thermal convection experiment, groundwater flow in fractured gneiss, three-dimensional regional flow in low permeable rock, and soil water redistribution near the ground surface (Hydrocoin 1987). It was noted that comprehensive databases sufficient for validation of complex groundwater flow models were not available and that there is a need for experiments specifically designed and planned for validation purposes. The main problems with the available experiments identified by HYDROCOIN were due either to incomplete parameter definition (both spatial and temporal) or the lack of independent datasets useful for both model calibration and for evaluation of the predictions (Nicholson et al 1987).

In recent years an increasing number of datasets derived from specially designed field experiments have become available or are being collected as part of current research on field subsurface systems and

and soil water and groundwater characteristics. Although such data are a necessary and valuable resource for verifying theoretical concepts and for validation of models and modeling approaches, to be used optimally they require complete documentation of collection and analytical procedures, site characteristics, and an assessment of possible sources of error (van der Heijde et al. 1988).

In a report, the Groundwater Review Committee of EPA's Science Advisory Board concluded that regardless of the type of model chosen, increased emphasis should be given to field testing and field validation of each model (EPA 1985a). Data generated in association with remedial action and monitoring Superfund sites may be used to fulfill model validation requirements. The Review Committee commented that these data should be made available for use by other investigators. The need for extant data in the evaluation of tracer-analysis research can be viewed in the same way. (Tracer technology remains one of the principal approaches to obtaining field values for model parameters [Molz et al. 1985]). The Review Committee also found that, surprisingly, the conclusions of many publicly funded research efforts are based on data not available for peer review. Therefore, the Committee recommended that databases from field research projects be made readily available to other groups. As costs of research and environmental monitoring escalate, the spiraling cost of acquiring new data emphasizes the critical need for mechanisms that facilitate access to reliable existing data.

Groundwater research programs are not restricted to the EPA. Major research on groundwater quality issues is carried out under the auspices of the National Science Foundation and within the Water Resources Division of the U.S. Geological Survey. Other federal programs containing significant groundwater quality research are the Subsurface Transport Program of the U.S. Department of Energy (DOE 1985), and the Agricultural Research Program (ARS) of the U.S. Department of Agriculture. Nongovernmental research in this area is exemplified by the Solid Waste Environmertal Studies (SWES) program of the Electric Power Research Institute (EPRI), Palo Alto, California.

No institution currently exists for rapidly locating and searching soil water and groundwater research databases or for standardizing data integrity and documentation of research datasets. Existing centralized database facilities for groundwater resource management do not provide the detail and quality of data required to successfully complete research on contaminant transport and fate. In many research projects, the lack of rapid access to these data causes delays and money unnecessarily spent, resulting in many incomplete model validation initiatives. The groundwater research strategy prepared by U.S. EPA and the National Center for Ground Water Research (EPA/NCGWR 1982) states that the data accumulated through Agency-funded research will be made available to the Agency and to the user community through information transfer. A central data clearinghouse could acquire and distribute such data in error-free, machine-usable form, efficiently and economically.

A report published by the National Academy Press (Fienberg 1985) drew attention to the benefits, costs, and restrictions involved in sharing research data. As the introduction to this publication states, "An open exchange of scientific information encourages others to engage in re-analysis of data that may detect flaws in the original research or may lead to better conclusions as a result of improved methods of analysis." The report continues: "the free flow of raw data can stimulate further research, particularly across disciplines, improve methods of data collection and teaching, and foster greater scientific understanding and progress."

The need for better documentation and information on and access to existing groundwater research data can be met by the establishment of a centralized data facility.

Such a central data center or clearinghouse can provide valuable information on completed or ongoing research projects which have a major data acquisition component. Furthermore, such a facility could acquire and distribute research data in error-free, machine-usable form, efficiently and economically.

In addressing this need, the Holcomb Research Institute of Butler University, with support from the U.S. Environmental Protection Agency, has established the Ground Water Research Data Center within the framework of the International Ground Water Modeling Center (IGWMC). The new Data Center provides information and referral services regarding datasets resulting from publicly funded field and laboratory research on soil and groundwater pollution. In addition, the Data Center has established procedures for selecting, evaluating, documenting, and redistributing such datasets. Creation of the Data Center is expected to lead to additional protocols for error checking, documentation, accessing, and transferring this kind of research data, and for acknowledging the rights that researchers have vested in their data.

PROJECT APPROACH

The project consisted of two phases: (1) determination of the scope and design of the Data Center, and (2) development of facilities and implementation of operational procedures and organizational framework.

The first phase consisted of five elements: analysis of data needs and potential users; survey and analysis of existing datasets; assessment of quality assurance (QA) requirements; determination of computer and other facilities for an operational data center; and operational design of the Data Center.

The analysis of soil and groundwater research data needs and the identification of potential users of high-quality, well-documented datasets provided guidance, justification, and motivation for the development of the Data Center. To determine the required level-of-effort and to obtain baseline information for the design of the Data Center facilities, the availability and status of a number of groundwater datasets resulting from publicly funded research have been evaluated. Current practices in collecting, handling, storing, documenting and distributing these datasets have been studied.

Other data centers utilizing high-quality environmental research and monitoring datasets have been contacted to benefit from their experience in such areas as dataset acquisition, data handling, and quality assurance procedures. Specifically, issues related to the invested rights of researchers involved in the data collection have been discussed.

Quality assurance (QA), an essential task for a central data distribution facility, must be incorporated on two levels: (1) the quality of the datasets of interest needs to be determined and documented; and (2) adequate quality assurance procedures need to be established for the operation of the Data Center in such areas as dataset evaluation, referral, management, and transfer.

To determine the level of detail required for the Data Center in the evaluation of the quality of prospective datasets, an inventory have been made of standards (existing and under development) and current accepted practices as documented in the open literature and technical guidance of regulatory agencies.

Based on the findings in phase 1, the institutional structure for the Data Center has been determined and the database framework created. Two types of database have been developed: (1) a directory-type or referral database containing descriptive information on datasets available from the Data Center or from other sources; and (2) a database containing the datasets selected for distribution by the Data Center. Information resulting from the dataset survey in phase 1 has been incorporated in the referral database.

Arrangements have been made to protect dataset integrity in their transfer from their generators to the Data Center and from the Data Center to secondary users. Furthermore, quality assurance procedures have been implemented for data handling, storing, archiving, and backup. Different levels of implementation are distinguished, dependent on the quality and extent of the datasets, the level of documentation, and the importance of the data. Technical support for format and transfer medium, and to a limited extent for the analysis of the data, will be provided; the level of support will depend on the implementation level selected. Policies have been developed regarding such issues as proprietary rights, conditional use, potential liabilities, and other legal and ethical issues.

As a part of the International Ground Water Modeling Center, the Data Center's activities will be subject to annual review by the IGWMC Policy Board and the International Technical Advisory Committee (ITAC).

GROUNDWATER RESEARCH DATA

Data on groundwater quality and quantity are characterized in both the spatial and temporal domains. Two major types of data are distinguished: site-specific data, and generic, site-independent data. It should be noted that in this report the term groundwater is used for water in both the saturated and unsaturated zones of the aquatic subsurface.

Certain kinds of site-specific data are constant for the time period under consideration, but may vary from location to location. Other site-specific data might show a significant time-dependent behavior. Collection of such data is generally aimed at identifying regional patterns during a certain time period or at studying the time variability at specific locations (Steele 1985). These objectives of site-specific data collection may change during the operation of the data collection network, due to changes in management needs, technology, and institutional arrangements. Subsequently, the design and operation (when and where to sample or measure, and which variable to measure) may be altered. Such variability certainly applies to research data networks, which are often project-oriented and of relatively short duration. Data management, thus confronted with a widely varying set of data characteristics, must be flexible enough to handle them efficiently.

Because water in the underground often moves quite slowly, abiotic or biotic transformations may represent significant attenuation processes in the transport and fate of pollutants. The presence of such processes results in a significant increase in data requirements for the predictive analysis of water quality (as opposed to the data requirements for water quantity problems). Much of this additional data is generic and can be established off-site in controlled laboratory or field experiments in combination with relevant site characteristics. Such generic, site-independent data on specific chemicals (the second type of data mentioned above) are increasingly available from research on the basic processes that govern contaminant transport and fate, and are crucial for successful application of computer-based prediction techniques in specific hydrogeologic environments.

At the beginning of many research projects requiring data acquisition, the establishment of efficient data management practices is often more difficult than anticipated. Traditionally, researchers have had almost total control over the form and documentation of their data; even contractual requirements for data in machine-processible form have had little effect on the ultimate availability and utility of most data. In addition, control by funding agencies over procedures and quality of data collection, storing, and distribution to a large number of institutions, requires extensive organizational arrangements and additional personnel. This is especially true when securing the collected data for distribution after the initial research has been completed and the original research staff is no longer available, or when no funding is available for continuing data management at each individual site.

Datasets for use in transport and fate modeling studies require a high level of detail concerning soil and aquifer properties, density of data points, contaminant behavior, and qualitative data descriptors. Specific data requirements for subsurface models include the need to define precisely the units of measure of each input value; for example, point versus averaged values (Hern and Melancon 1986).

Whenever a model study is performed, the quality of the data used is an issue of concern. Data quality is often critical in model validation due to the sensitivity of most models to changes in certain parameters. Although a given field investigations may result in a large amount of data, the usefulness of the study site for model validation is determined to a large extent by the quality of the data, as reported in the data documentation. However, often the data documentation is lacking in detail, especially with respect to data quality.

SECONDARY USE OF RESEARCH DATA

Current scientific research confirms that many researchers take advantage of existing high quality field studies to test their models and support their research findings. For example, research data sets from the Borden Air Force Base Landfill site in Canada have been used for evaluating monitoring equipment, developing sampling strategies, researching chemical and biological subsurface activity, verifying theoretical concepts and models, evaluating parameter estimation methods, and testing remedial action technologies (Cherry 1983).

The need for soil and groundwater model validation has been stressed by governmental agencies, consultants, and individual researchers. The Science Advisory Board, in its review of the EPA Groundwater Research Program (EPA 1985a), stated that "it is important for EPA to screen computer models and test them for accuracy . . . increased emphasis should be given to field testing and field validation of the models." Many researchers have found that available data are not sufficient to accurately formulate mathematical statements of real-world physical problems (Ross et al. 1982), while Moran and Mezgar (1982) identified the adequacy of datasets as the single most important factor in site-performance model verification and validation. Yet in many cases the modeler must fill in data gaps with estimated, interpolated, and extrapolated values. Moran and Mezgar (1982) concluded that emphasis needs to be placed on the development of a long-term database, one specifically designed for model validation.

Chu et al. (1987) were unable to use field data in their evaluation of data requirements for groundwater contaminant transport modeling due to the complexity of natural aquifers and the absence of extensive field data representing the true conditions of the aquifer. Such concerns were also expressed by the EPA Science Advisory Board, which recommend that "the needed increase of research on the basic processes that govern the transport and fate of contaminants in groundwater should go together with the establishment of data bases for field application" (EPA 1985a).

Although some extensive datasets are available for validation purposes, their present form may cause additional problems. Gelhar (1986), in a review of available field datasets, found that very few sites exist where data have been collected in sufficient detail to be useful for the verification of a new stochastic theory regarding dispersion in aquifers. The site selected, the Borden Air Force Base in Canada, could provide data sufficient only to evaluate part of the proposed theoretical framework. A significant part of this particular study concerned the analysis of the contents and structure of the selected dataset, determining which of the data to use, and reformatting the selected data in preparation of the verification computations.

A recent EPA groundwater protection data-requirements study (EPA 1987a) stressed the importance of improved access to existing soil water and groundwater data and of lowering the transaction costs associated with obtaining and using such data. The report indicates that knowledge about and access to the large volume of groundwater data being generated from federal programs and state initiatives is limited, because the data are managed by many organizations and are stored in many different locations, files, and formats. In addition, relatively few of these soil water and groundwater datasets are computerized, and a central cataloguing facility is lacking. Although the study's conclusions are concerned with all groundwater data useful in the protection of groundwater resources, they apply equally well to research data.

SHARING RESEARCH DATA

Availability and accessibility of environmental research data are discussed in a wide variety of environmental literature (Armentano and Loucks 1979, EPA 1985a, Olson and Millemann 1985). Reviews of data availability indicate that many researchers give little thought to the use of their data other than for immediate research purposes (Armentano and Loucks 1979). The appraisal by researchers of the importance of data accessibility is reflected in their approach to data management. Many consider it an administrative chore to be handled separately from the research, usually at the end of the study (Armentano and Loucks 1979). Other investigators show a keen awareness of the importance of data management both for their own use and the use of others.

Sharing data from detailed groundwater monitoring studies and laboratory bench studies is a subject of concern both economically and with respect to the advancement of scientific research. Due to the ever-increasing cost of field studies and the extensive sampling periods required for transport and fate studies, it has become essential to share groundwater data so that unnecessary duplication can be avoided. Sharing data not only produces cost benefits; it "reinforces open, scientific inquiry; permits verification, refutation, or refinement of original research results; stimulates improvements in measurement and data collection methods; allows more efficient use of resources spent on data collection; encourages interdisciplinary use of data; and strongly discourages the uncommon, but nevertheless serious, problem of fraudulent research" (NRC 1985).

As participants at a CODATA Workshop pointed out (Hopke and Massart 1986), "One of the most critical impediments to using existing data sets is the lack of knowing that they are there." Even when sources of data are known, data access is often complicated, delays occur in obtaining available data, and documentation is rarely available or is very limited (Armentano and Loucks 1979, Hopke and Massart 1986).

A comprehensive referral center as represented by the IGWMC Groundwater Research Data Center, focusing on selected datasets for groundwater model validation and testing, will help to avoid situations where datasets of value to many potential users go unrecognized and therefore unused.

SECTION 2

EXISTING DATASETS

Many groundwater research projects result in field and laboratory data. These studies are conducted by a variety of organizations and for many different reasons. In addition, groundwater management studies (e.g., site characterization in preparation of remedial action) often produce large sets of data. Much of the research data is discussed and referred to in the open scientific literature. However, most of the time only a part of the original data is presented, often in interpreted form, and field notes, laboratory work sheets, and quality control information are absent.

This section presents the results of a preliminary survey of existing groundwater research datasets performed to provide the IGWMC Groundwater Research Data Center with basic information for the development of a referral database and of procedures for information acquisition and distribution. It discusses objectives, type, and extent of the studies from which the data resulted, provides a summary site characterization, lists the kind of data collected, discusses form of and accessibility to the data, its documentation level, the history of primary and secondary use of the datasets, and lists the organizations that funded and collected the data.

LITERATURE SURVEY

In planning the Groundwater Research Data Center, a literature survey was conducted to identify relevant research projects and resulting datasets. The search, which was initiated using IGWMC's computerized literature databases JUPITER (hydrology, geology, water quality, and water management) and MOON (groundwater models), augmented with information from other sources, resulted in bibliographic information on sites possibly useful for further analysis.

Data obtained from this literature survey were entered into a database using the REFLEX (Borland/Analytica Inc.) database system. Selection of sites for the REFLEX database was based on their potential for use in the validation of transport and fate models.

PRELIMINARY DATABASE FOR INFORMATION ON DATASETS

As the project developed and as the level of detail required to describe the research datasets increased, a modified database system was developed using R:BASE V (Microrim, Inc.) database software.

This database contains four files:

(1) general site/dataset information
- site name
- project description
- location
- type of contaminants or tracers
- source(s) of funding
- research organizations involved
- project classification
- number of model applications

(2) bibliographic reference information
- author(s)
- year published
- title
- source/publisher

(3) geographical information
- state
- county
- city

(4) aquifer/soil information
- material
- geologic name/soil type
- geologic time period
- confined/unconfined

A project classification (see file 1 above) was set up to identify "prime" datasets of interest for model testing and validation. The RESEARCH1 classification includes those site studies primarily initiated for research purposes and which have a high degree of quality control and documentation. The RESEARCH2 classification includes those studies that may have originated as groundwater contamination investigations but which evolved into data collection to study researchable questions. Studies initiated for research are also included in this class if they were not as detailed as sites in the RESEARCH1 class or if not enough was known about them to justify their classification as RESEARCH1 sites. RESEARCH3 projects originated as groundwater contamination investigations and their data have been used in groundwater modeling studies and for research studies, without significant additional data gathering. The RESEARCH4 class denotes those studies implemented as site evaluations for waste disposal or for mining operations, generally lacking the level of detail and/or quality control needed for research. The SUPERFUND class includes sites listed on the National Priorities List for which significant characterization have taken place.

EVALUATION OF PRELIMINARY INFORMATION ON RESEARCH DATASETS

As of December 31, 1988, 159 sites or studies have been identified for their potential for model validation and entered in the R:BASE database. Appendix B contains a listing of the sites included in this database. These studies have not been screened for the level of detail present in the datasets or for their general adequacy for validation of groundwater flow and transport models. The characteristics of these 159 studies are summarized in Tables 1-6. Many of these research sites have certain features in common, such as relatively simple geohydrology, location on federal land (usually military installations), or a known history of contamination. These tables are discussed briefly in the following section.

In general, the common purpose of these data collection projects is to better understand physical, chemical, and biological subsurface processes, thus providing a basis for efficient protection of the subsurface environment and for effective remediation of contaminated sites. In addition, the studies identified have distinct objectives related to the framework in which they have been conducted. Similarly, many of the studies at Superfund sites have as their objectives conceptualization of the hydrogeologic system, analysis of the type and extent of the contamination, and evaluation of alternative remedial actions. The purpose of field studies for the siting of municipal and industrial waste disposal facilities is to provide the data for planning and design. However, many research studies are characterized by highly specific goals such as the evaluation of sorption characteristics for a particular combination of aquifer material and chemical composition, or the identification of spatial variability of hydrodynamic parameters.

Table 1 gives keyword characterizations of the various research sites analyzed. It contains information on the hydrogeologic conditions, kind of pollutants, source of contaminants, processes studied, and special interests.

TABLE 1. GENERAL CHARACTERIZATION OF RESEARCH STUDIES

General characterization	Number of Sites
fractured rock	9
unsaturated zone	8
soil	11
groundwater pollution	20
radioactive waste	17
fly ash	11
hydrocarbons	10
landfill/waste disposal	14
tracer test	11
flow/advection	5
biodegradation	5
geochemistry	8
physical processes	5
monitoring	4
drilling	3
geotechnics	2
(strip)mining	5
miscellaneous	6
not described	5

Tables 2–6 provide additional detail regarding the main site characteristics.

Table 2 list the organizations involved. It clearly shows that most of the activities at the various sites are performed either by universities or research organizations. The U.S. EPA and the USGS are actively involved in many of these projects. Often, EPA joins with others or functions as coordinator of multi-agency projects. The USGS performs many of its own field research studies. Consultants are somewhat less involved in such projects, which might indicate that most of the research is performed by universities and research organizations.

TABLE 2. ORGANIZATIONS INVOLVED IN FIELD RESEARCH

Organization	Projects executed by one organization	Projects executed in cooperation with other organizations
EPA	4	18
DOE	3	2
EPRI	2	2
USGS	17	15
Research Institutes	22	21
Industry	6	11
Consultants	16	14
Universities	25	26
States	1	9
Miscellaneous	0	2
Unknown	5	

Most research projects in the survey are sponsored by the EPA, while DOE and USGS also fund a considerable number of research projects. Fewer projects surveyed are sponsored by industries, although industries are often involved in pollution problems.

TABLE 3. FUNDING OF SITE INVESTIGATIONS

	Completely funded Projects	Partially funded Projects
EPA	34	15
DOE	17	4
USGS	15	3
EPRI	3	8
States	3	4
Industry	2	6
Universities	0	4
Miscellaneous	11	4
Unknown	47	

Of the research sites focused on the study of pollutant behavior, most deal with existing radioactive waste or with potential of future storage of such waste. When present in the subsurface, these relate to nuclear power plants or nuclear material production facilities for national defense. Other sites identified contain fly-ash, oils/hydrocarbons, halogenated organics, miscellaneous organics, pesticides, and insecticides.

TABLE 4. KEY POLLUTANTS AT RESEARCH SITES

Pollutant type	Number of sites
Inorganics	
(heavy) metals	6
fly ash	12
conservative tracers	3
radioactive wastes	28
miscellaneous inorganics	11
Organics	
halogenated organics	9
oils/hydrocarbons	15
pesticides	5
solvents	3
aldicarb (insecticide)	8
miscellaneous organics	19
Organics and Inorganics	
landfill leachate	6
sewage	2
chemical waste	4
VOC	2
miscellaneous	5
Does not apply	21

TABLE 5. RESEARCH SITES PER STATE/CONTINENT

State/ continent	Number of sites	State/ continent	Number of sites
Alabama	1	New Hampshire	1
Alaska	0	New Jersey	3
Arizona	3	New Mexico	4
Arkansas	0	New York	4
California	11	North Carolina	1
Colorado	5	North Dakota	10
Connecticut	0	Ohio	2
Delaware	1	Oklahoma	0
District of Columbia	0	Oregon	2
Florida	5	Pennsylvania	2
Georgia	3	Rhode Island	1
Hawaii	0	South Carolina	3
Idaho	3	South Dakota	0
Illinois	5	Tennessee	4
Indiana	3	Texas	4
Iowa	1	Utah	1
Kansas	0	Vermont	0
Kentucky	1	Virginia	0
Louisiana	0	Washington	6
Maine	0	West Virginia	0
Maryland	0	Wisconsin	4
Massachusetts	2	Wyoming	3
Michigan	1		
Minnesota	4	Canada	10
Mississippi	2	Africa	0
Missouri	1	Asia	4
Montana	1	Australia	0
Nebraska	2	Europe	8
Nevada	4	South America	0
New Hampshire	1		
Laboratory studies:	20		

The number of sites per state varies considerably. The state with the most sites identified is California (11). The large number of sites in North Dakota results from extensive pollution from mine wastes.

Most of the sites are characterized by porous media. However, 13 sites are found in regions with fractured rock formations. At another 19 sites, research has been performed in the unsaturated zone.

TABLE 6. LITHOLOGY OF THE RESEARCH SITES

Lithology	Number of sites
sand and gravel	6
sand	6
sand, gravel, silt, and clay	31
shale, limestone, sandstone	7
granite	4
total fractured materials	13
total soils and/or unsaturated zone	19
miscellaneous	10
does not apply/unknown	81

Some typical fractured rock studies identified include one at the Fanay-Augeres uranium mine in France, the Creux De Chippis tracer study in the unsaturated zone at a complex site in Switzerland, and the Underground Research Laboratory (URL) facility of Atomic Energy of Canada Limited, situated in granite rock near Lac du Bonnet, Manitoba.

Many of the studies identified are soil column studies performed in the laboratory, using undisturbed or repacked soil samples taken from specific field sites. These studies were performed for a variety of reasons, such as determining the chemical attenuation in soil of inorganic and organic contaminants, evaluating the effects of cracking on flow and contaminant transport, studying dispersion in unsaturated soils, determining nitrate migration, evaluating the effect of stratified soils on flow and transport of contaminants, and determining the hydraulic parameters in the unsaturated zone.

SPECIFIC LARGE-SCALE FIELD RESEARCH SITES

An increasing number of field research projects take place in a multidisciplinary and multitask framework. At some of these sites different studies may have taken place, conducted by various teams of the same research organization or even independently by various research institutions. These research facilities are sometimes called field experimental stations or sites. An excellent overview of experimental field studies can be found in Anderson (1987).

The Lawrence Livermore National Laboratory is the site of various studies of groundwater contamination with VOCs (volatile organic compounds). The site contains alluvial sediments of highly variable permeability. The Auburn University field study site at Mobile, Alabama, is the location of several single and double-well tracer tests in stratified aquifers. These studies, supported by the U.S. EPA, involved methods to determine vertical velocity distributions.

EPRI's Solid Waste Environmental Studies (SWES) program is also a potential source for research datasets. The program was set up to study the release of solutes from utility waste disposal sites and their subsequent transformation and transport in the subsurface environment. The Macro Dispersion Experiment (MADE) is a tracer study with seven different tracers. The study site is located at the Columbus Air Force Base, Mississippi, and is in a sand, clay, and gravel aquifer.

Currently, a field study underway at Moffett Air Force Base, California, focuses on in situ remediation of an aquifer system polluted with halogenated aliphatic compounds. The study, conducted by a Stanford University team, includes testing of biostimulation techniques and Bromide and TCE transport experiments. The geohydrology of this site consists of interlayered coarse and fine sediments in a shallow, confined aquifer, an alluvial aquifer, and a buried stream channel.

The database contained 23 sites concerned with radioactive or nuclear waste contamination, mostly focused on the effects of subsurface nuclear waste storage (both high- and low-level nuclear waste) such as deep percolation of radionuclides and the geochemistry of the waste and the geologic environment. Research has been conducted for the selection of possible nuclear waste repositories. The preliminary database for this study contains six references to studies performed for the Palo Duro Basin, Texas, one for Richton Dome site, and thirteen for Yucca Mountain, Nevada. The geohydrology at these sites generally involves thick, unsaturated geologic sequences and deep brine aquifers. Another site extensively studied for its potential as a nuclear waste storage facility is at Hanford, Washington. The focus of the studies is an aquifer averaging 70 m thick and composed of glaciofluvial sand and gravel, and silts. The basis of this aquifer system is formed by basalt.

The Idaho National Engineering Laboratory (INEL) is conducting studies on the presence and behavior of low-level radioactive wastes in fractured basalt aquifers. Other studies at INEL include investigations of on-site hazardous waste disposal.

Creosote wastes are a major concern of the regulatory agencies. The U.S. EPA supports studies at a site in Conroe, Texas, involving delineation of resulting groundwater contamination, presence of microbial degradation, and the importance of adsorption. Pensacola, Florida, is the site of many studies of creosote contamination, sponsored by the USGS. The site is characterized by 80 m of surficial deposits consisting of nonhomogeneous sand and gravel deposited as fluvial and deltaic sediments. Studies here include determining the interactions of native clay minerals with organic contaminants, the presence of dissolved inorganic species, improving understanding of contaminant distributions in aquifers, and determining the rates of anaerobic degradation of phenolic compounds. Ongoing research includes work on a wide range of hydrologic, geochemical, and microbial processes affecting the distribution, movement, and fate of organic contaminants in aquifers.

The majority of transport models assume that soil water and groundwater move through a relatively simple porous media. Most research into subsurface processes has been conducted at sites consisting of a relatively simple hydrogeology and soils, such as homogeneous unconsolidated sand aquifers, which are atypical for most contaminated aquifers. Current research indicates that even these "homogeneous" aquifers have complexities such as macroscopic layering, which must be considered for the simulation of transport and fate of contaminants. Such considerations also apply to soils, where the existence of macropores is a significant phenomenon. Hence, a need exists to share data collected from sites of greater complexity such as for consolidated fractured rock aquifers or a complex sequence of layers.

Some hazardous waste sites are especially suitable for the study of particular contamination problems. For example, the Federal Pioneer site is located in unconsolidated materials and is the site of a PCB spill; the contamination is multiphase. Research projects here were supported by the government of Saskatchewan and the National Research Council of Canada. Much data was collected in the unsaturated zone.

The Borden landfill site, located on the Canadian Forces Base at Borden, Ontario, was actively used from 1940 to 1976. The aquifer is unconfined and composed of sand of glaciofluvial origin. The contaminant plume extends for approximately 750 m longitudinally and to the lower confining layer vertically. The landfill has been monitored extensively and the data have been used in several referenced modeling studies. The Borden tracer experiment is probably the first attempt at defining the three-dimensional aspects of a contaminant event using conservative tracers as well as more complex organic compounds that exhibit sorption effects. Other studies at the site include a microcosm simulating the behavior of VOCs.

The Cape Cod, Massachusetts, area has many contamination sites, due to its high potential for groundwater contamination caused by the high permeability of the unconfined sand aquifer, and to its dense population. A field site at the Otis Air Force Base at Cape Cod, Massachusetts, was chosen by the USGS Toxic-Waste Ground-Water Contamination Program for studying contaminant transport and attenuation in aquifers. It was selected because of its 45-year documented history, its relatively simple hydrogeologic setting, and its similarity to many contaminant sites nationwide. The aquifer consists of sand, gravel, silt, and clay deposited as a glacial outwash plain. A sewage leachate plume, 20 to 25 m thick, 750 to 1,100 m wide, and 3,500 m long, has received a great deal of attention. The contaminants have been transported with groundwater while being altered by chemical reactions. A complication is the definition of the source behavior, as contaminants have entered the aquifer at different rates at different times. Studies are being conducted of sorption by phosphorus, transport of bacteria, rates of microbial activity, volatile organic compound mobility, and degradation of organics, among others. A natural-gradient tracer experiment was initiated to measure dispersion and to determine geochemical controls on nonconservative transport in a heterogeneous aquifer.

Another USGS field study is being conducted at Bemidji, Minnesota, the site of a crude oil spill. The aquifer consists of glacial outwash materials, is approximately 20 m thick and is underlain by low-permeability till. Volatile constituents are migrating through the unsaturated zone by diffusion, and studies are attempting a comprehensive understanding of the physical, chemical, and biological processes that will be needed to develop predictive models of contaminant mobilization, transport, and fate.

Traverse City, Michigan, is the site of an aviation gasoline spill. The U.S. EPA continues to conduct research here concerning the biodegradation of organic pollutants, in situ biotransformation, remedial action effects, and sorption processes. Detailed studies are attempting to describe the rates of organic transformations and characteristics of microorganisms able to degrade the pollutants. Other EPA-sponsored research involves the mobility of metals, especially with respect to coal-fired power plant wastes, including four high-level contaminant sites for the study of metal transport processes.

A final example of the kind of field and laboratory experiments of interest for one of the Data Center's primary uses, model validation, can be found in the HYDROCOIN study described in section 1. The five level-2 validation cases selected by the HYDROCOIN group were based on information obtained in laboratory and field experiments conducted in different countries (HYDROCOIN 1986). The first case is a heater experiment in a borehole cluster in a quarry in Cornwall, UK, where temperature gradients were determined and linked to convective heat transport. The second case is an analogue for variable density fluid flow based on a laboratory experiment with thermal convection. Case 3 involves transient pumping tests and steady-state piezometric data in a monozonitic gneiss block at the Chalk River Site of the Atomic Energy of Canada Limited. The block of 150 x 150 x 50 meters contains five major structural discontinuities. Hydraulic and piezometric data as well as core logging and fracture orientation statistics have been collected in 17 boreholes. The fourth case is the Piceance basin in Colorado, a regional scale flow system of about 50 by 100km. The purpose of this case is to compare computed piezometric heads with heads obtained from kriging the heads measured in boreholes. Finally, case 5 focuses on an irrigation experiment in Central Valley, California. Plots of soil were ponded until steady-state saturated conditions were reached in the soil beneath the plots. The water supply was then terminated and the change in moisture content in the soil underneath the plots was followed in time. This information was augmented by laboratory soil characteristic curves determined as part of the study.

SELECTION OF DATASETS FOR INCORPORATION INTO THE DATA CENTER

From the analysis of the information gathered in this preliminary database, some criteria can be derived for the selection of datasets suitable for incorporation in the final referral database and for characterization of these datasets. Many suggestions resulted from the workshop discussions reported in Appendix A.

One of the workshop conclusions was that field research datasets should have the highest priority and be the focus of the Groundwater Data Center; laboratory bench studies connected to these field research sites should be next in priority. Independent laboratory data should have a low priority, although effective model validation may involve laboratory-scale experiments; such data should basically be accepted on an "as-is" basis with no effort made to complete them. Whatever the kind of dataset acquired by the Center, its quality will always be a major selection criteria.

As a result of extensive analysis and discussion, the Center considers the following criteria for dataset selection:

- significance to current soil and groundwater pollution or model validation problems
- accessibility of information regarding the dataset
- quality of the data
- completeness of data
- availability in automated format or suitability for electronic filing
- permission for distribution/publicly funded data collection
- availability of documentation of the data collection and dataset
- availability of reports on peer review of data
- timeliness of dataset

Selection based on these criteria requires a standardized approach to information acquisition and processing. The descriptors used in the final database (see Section 5) are based on the preliminary analysis and additional experience gained in using this database, as illustrated in Table 7.

TABLE 7. PRELIMINARY INFORMATION ON FIELD SITES AND DATASETS

Contaminant Sources

- type of source
- dates of activity

Studies Conducted

- funding
- contacts
- dates of activity
- list of personnel

Type of Data

- parameters measured
- datatype (raw, summarized, derived)
- variables and units used
- temporal and spatial resolution

Form of Data

- dataset name
- history and maintenance of dataset
- technical description of files, etc.
- examples

Amount of Data

Area of Study, Geographic Coverage

Original Data and Documentation Available—Reports, Articles, etc.

- authors/researchers involved

Database Software Used

- name
- source
- example
- associated capabilities

Maintenance of Data

- Automated format, file structure, file size
- Dates
- Organizations/contacts
- Storage

Availability of Data

- is data available for secondary use?

Accessibility of Data

- automated format
- remote access
- transfer of files
- special requests
- publications
- fee information
- other

(continued)

QA/QC Information

documentation
program standards

Anticipated Updates or Ongoing Studies

Cost Information of Datasets and Reports

SECTION 3

EXISTING DATA CENTERS

This section discusses some of the issues and concerns related to the establishment and operation of data centers focused on environmental data, or more specifically, soil water and groundwater data. The experience gained in studying existing extant data centers, combined with past IGWMC experience, is crucial to the successful development of the IGWMC Groundwater Research Data Center.

DATA CENTERS AND DATA CLEARINGHOUSES

The process of gathering and organizing environmental information for dissemination to a community of users takes on several different forms and labels (not always well defined). Data and information centers or clearinghouses may provide a wide or limited range of services.

Typical data center services include accessing databases and information directories (e.g., a referral service), providing data abstracts, and facilitating data exchange. In addition, data centers may provide access to or even operate integrated database clusters, assist in data integration, perform data evaluation or analysis, and develop project-specific databases or offer advice on such developments. Often, data centers are part of larger organizations such as governmental agencies and research institutions. Ideally, integrated database systems contain data selected from several sources, formatted and documented to conform to existing standards in order to facilitate regional, integrated analyses (Merrill 1985). A data center may be extensively involved in data processing, e.g., converting datasets into machine-readable formats, or in data interpretation and data quality assessments.

Some confusion reigns as to what constitutes a data center or clearinghouse. For example, Olson and Millemann (1985) define a clearinghouse as a data referral center containing data descriptions but not actual data. However, in a review of data centers Olson (1984) states that some clearinghouses provide direct access to data while others do not.

In this report we use the terminology of the Earth Sciences Information System of the U.S. Geological Survey (USGS 1982). Here, a clearinghouse is defined as being a directory only; a repository operating as a directory with data; a data center as a directory containing data and providing evaluation services; and an integrated database system as a directory containing data along with evaluation and analytical services.

Previous Studies

In the past, a number of studies have provided an overview of existing data distribution facilities and have assessed the usefulness of data centers. Most of these studies have been based on information collected in literature reviews, from questionnaires, and from interviews; sometimes, such studies included the "trial" use of the data center concept. The extent and completeness of the studies varies from small-scope intra- or interagency studies to statewide and nationwide surveys. The following section discusses the findings of some of these studies.

In 1982 the Committee on Data for Science and Technology (CODATA) of the International Council of Scientific Unions (ICSU) conducted a large international survey of data sources. Over 650 data centers and referral centers in 94 countries were identified. The study found widespread international concern over access to numerical data and stressed the importance of clearinghouses in providing an efficient mechanism to locate and access data resources (CODATA 1982).

In another study, Armentano and Loucks (1979) found an overlapping hierarchy of databases throughout federal and state government, academia, and industry; the reliability and associated documentation of these databases varied widely. They concluded that the potential contribution of these databases to the furtherance of research, resource management, and regulatory purposes is perhaps only "touched upon." They noted that investigators gave little thought to the use of data other than for immediate research purposes. Databases were computerized in some cases but were often not documented and published, resulting in limited user-accessibility. According to this study, poor coordination and lack of planning were prominently illustrated by the number of agencies that stored their often uncomputerized data "in-house," and that had no way to enter their data into some type of master storage system. The study by Armentano and Loucks proposed a network of regional or thematic databases. This national network of regional environmental data centers should be designed to meet the national need for locating and cataloging existing data and models and for increasing their accessibility and utilization. Although such a system would help in locating data resources, the user would still be faced with acquiring data, standardizing formats and units of measure, and reconciling differences in spatial and temporal attributes.

In reviewing existing environmental and natural resource databases Olson (1984) focused on the type of data available, funding statutes, and long-term database maintenance. The study was based on information from mail surveys, telephone interviews, and the outcome of the 1983 Integrated Data Users Workshop (Olson and Millemann 1985). Twenty-four federal agency clearinghouses, referral centers, or data centers were identified which maintain machine-readable inventories of environmental and natural resources data files. The study lists a variety of management approaches and data center objectives among these facilities, including clearinghouses or referral centers, compilations of dispersed national datasets to serve specific agency needs, and integrated database systems containing collections of datasets. A major issue appeared to be the long-term maintenance and funding of data systems. As Olson (1984) concluded: "Either the funding agencies will need to explicitly allocate resources for this activity or a more effective mechanism to recover these costs from data users will need to be established."

The U.S. Geological Survey has been active in earth science data collection and dissemination for several years. Many of the resulting data sets have been computerized. To provide rapid access to these databases the USGS has published a number of surveys (e.g., USGS 1979, 1983). This information is also accessible through the Earth Science Information System (ESIS), a comprehensive data management facility designed to support the coordination, integration, and standardization of data within the USGS. ESIS maintains a central and uniform repository of detailed information on the earth science databases and systems and on associated data elements. The on-line retrieval capabilities of this referral system provide for efficient multiple keyword searching in a user-friendly, menu-driven environment with user-activated help screens and customized reporting. The descriptive characterization of databases and data systems used by ESIS is listed in Tables 8 and 9.

TABLE 8. ESIS INFORMATION CATEGORIES FOR DATABASES

Category	Comments
Database name	short descriptive name or title
Acronym	database acronym, short name, or name abbreviation
Database type	spatial, scientific, and technical; bibliographic; or mission support
Data structure	logical structure of database records
Contact	name, address, and telephone of contact person
Spatial data types	e.g., point, line, cell, polygon, grid
Coordinate system	
Time span of data collection	
Status of database	e.g., operational, under development, under modification, closed (not further maintained)
Users	known or intended users
Availability for use/access	
Access method	on-line, batch, indirect, etc.
Output media	printout, magnetic tape, etc.
Storage media	media where database is stored
Size of database	logical records, bytes per record, expected yearly growth
Computer	make and model on which database resides and location of computer
Language	used in software for database management and access
DBMS	database management systems used
Subject coverage	
Geographical coverage	
Sources of data	individual or organizational sources
Documentation	user manual and descriptive references
Keywords or descriptors	
Definition or description	database abstract

TABLE 9. ESIS INFORMATION CATEGORIES FOR DATA SYSTEMS

Category	Comments
System name	common name of system
Acronym	system acronym, short name, or name abbreviation
Application category	type of applications for system data
Contact	name, address, and telephone of contact person
Databases accessed	either utilized or integral part of system
Date implemented	
Last modified	data of last major modifications or enhancements
Expected application life	
Application status	e.g., operational, under development, under modification, closed (not further maintained)
Annual operating costs/budget	
Availability for use/access	
Access methods	on-line, batch, indirect, etc.
Users	known or intended users
Frequency of input	
Frequency of output	
Computer	make and model on which database resides and location of computer
Language	used in software for database management and access
Output media	printout, magnetic tape, etc.
Input media	keyboard, tapes, remote-transmission
Input locations	geographic locations of system input devices
Related systems	systems that are routinely linked or interfaced with this system
Definition or description	database abstract

Selected Natural Resources and Environmental Data Centers

Data centers designed to store and distribute environmental and natural resources data are operated by a wide range of public and private organizations. The following discussion highlights the objectives and services of some of these centers.

NAWDEX

The USGS/National Water Data Exchange (NAWDEX) system was established in 1976 to link data collectors and users for more efficient use of the nation's water data (Edwards 1978). The system contains information on agencies collecting water resources data. As a management and information system NAWDEX provides a wide variety of services including search and referral, facilitating direct linkages to data sources, assisting in data format conversions, performing data quality evaluations, and providing data manipulation software. Although NAWDEX assumes the role of a data bank for data generators unable to respond directly to data requests, the system is not a database. The NAWDEX system includes more than 50 local assistance centers. A computerized index describes more than 350,000 sites for which water data are available and includes descriptions of the geographic location, data collectors, types of data available for each location, the time periods for which data are available, the major variables measured and their measurement frequencies, and the media on which the data are stored. NAWDEX also coordinates direct access to two major water databases: EPA's STORET and the National Water Data Storage and Retrieval System WATSTORE of the USGS (Cardin et al. 1986). NAWDEX operates an extensive user-support system including annual user meetings, news bulletins, and various on-line and printed information sources.

In 1983, the USGS Water Resources Division began to design and develop a National Water Information System (NWIS) (Edwards et al. 1987). The NWIS will replace and integrate the existing USGS water data systems, including NAWDEX an WATSTORE. The NWIS, which is planned as an interactive, distributed data system allowing multiple use of the various facilities, is scheduled to go on-line in 1990. The major objectives in designing the NWIS include:

- a single flexible and expandable system that is easy to use and that improves productivity and use over existing systems.
- provisions for standardization and uniformity of data handling, data storage, and software procedures, thus increasing the integrity of the databases and software systems.
- modular database and software systems, allowing the development of thoroughly tested and efficient software, that are easy to change, enhance, expand, and maintain.

STORET

STORET (STOrage and RETrieval), a central, inter-agency water quality data bank, was developed in the early 1960s by the Federal Water Pollution Control Administration (FWPCA), the predecessor of the U.S. Environmental Protection Agency, and is administered by the EPA through its National Computer Center (NCC) in Research Triangle Park, North Carolina. As of 1985, the original 140-site database now contains data on more than 500,000 sampling points (EPA 1985b). The data in STORET are usually measurements of concentration of a particular substance at a particular site defined in space (3-D) and time. The STORET system supports both scientific and regulatory applications.

Originally, STORET was designed to store information on surface-water quality and sampling procedures (Shirley 1982). This is reflected in the way it manages general site information: quality assurance and groundwater-related information, such as aquifer name, geologic formation and age, date of well construction, well screen interval, etc., must be entered as parametric data. Up to 5,000 parametric values can be attached to any sample. The principal data storage file within STORET is the Water-Quality File (WQF), which contains 1,800 unique water quality parameters. The file is updated weekly with monthly transfers from the USGS-WATSTORE system. STORET provides the user with various options in retrieval report format, station selection, and the selection of a specific group of samples and observations. In 1982, more than 40 states and approximately 50 federal, regional, and interstate agencies actively used STORET, primarily for surface water data. Increasingly, states are utilizing the STORET database for groundwater management. Groundwater monitoring data for compliance with federal (RCRA, UIC) and other regulations comprise most of the groundwater database in STORET (EPA 1985b).

STORET data is collected and entered into the database by a multitude of federal, state, and local government agencies and their contractors. To a large extent, the reliability of the data depends on the level of care employed by these organizations in the processes of sampling, laboratory analysis, and data entry. However, many problems occur resulting from decentralization of QA/QC (Armentano and Loucks 1979). To prevent storing of erroneous values, the system performs automatic data checking for about 200 of the most frequently sampled water quality parameters. Locations are checked within the STORET system by comparing the latitude/longitude entries against state and county latitude/longitude boundaries (Shirley 1982). Although STORET has some provisions for screening the data, this does not prevent data quality problems. STORET is programmed to reject data outside a certain range, but unfortunately, data outside such a range might still be valid, as in case of groundwater contamination. Thus, to cover such events, the limits on the range of acceptable parameters might be set such that for other sites they do not function as a quality control.

The large number of parameters in STORET has proven to be a source of problems and complaints. In selecting data, users who are often not familiar with the details of the data system structure and parameter definitions, easily err in selecting data. They find STORET cumbersome to use and sometimes supplying the wrong kind of data or the data in units other than those requested.

An analysis of U.S. EPA data requirements (EPA 1987a) recommended many initiatives to improve STORET. For example, groundwater data that meet the current data quality standards adopted by EPA, should be stored separately from historic data for which information on QA/QC is lacking or which do not meet current standards. Over time this process will produce a database that meets the "good housekeeping seal of approval" (EPA 1987a). Other improvements suggested include user-friendly access and retrieval capabilities aimed at non-computer-oriented professionals. In addition, the study concluded that STORET should be modified to obtain consistency with EPA's groundwater data quality requirements.

WATSTORE

The National Water Data Storage and Retrieval System (WATSTORE) was established for the management of water data collected through the activities of the U.S. Geological Survey and to provide an improved means of releasing the collected data to the public. It is operated and maintained at the USGS headquarters at Reston, Virginia. WATSTORE contains mainly data from USGS projects and state cooperative studies. The groundwater data in WATSTORE are collected in a separate data file, GWSI (Ground-Water Site-Inventory), with up to 270 data items per site (Mercer and Morgan 1982). Most of the groundwater data in GWSI are raw hydrogeologic data entered directly by the data-measuring agency. Although several field-collected parameters of water-quality data are stored in GWSI, the majority of water quality data reside in STORET.

WATSTORE contains a computerized verification system for providing several kinds of error checks, including syntax, compatibility, and out-of-range checks for physical, chemical, and locational consistency (Edwards 1983). Services provided range from distribution of simple data tables to complex statistical analyses. Text summaries presented in USGS annual reports include location, aquifer, well characteristics, datum, remarks, period of record, and extremes. Water-quality data are transferred monthly from WATSTORE to STORET.

Other Water Data Centers

A typical example of an integrated approach to data management is provided by the state of Texas, where a comprehensive computerized information system has been established: the Texas Natural Resources Information System (TNRIS 1982). This information system provides groundwater managers and their technical advisors with interpreted or processed data and with extensive support for problem analysis. At its inception it was expected that such centralized natural resources information system would improve consistency in data format and quality, and would introduce and enforce data storage and documentation standards (TNRIS 1983). The current system provides for standardized scales between and within various categories of data, thus facilitating analysis of complex interrelationships between the various kinds of geographic data. Through its services and facilities the system promotes increased interdisciplinary data sharing, allows well-targeted data retrieval for specialized applications, and facilitates increased accuracy in data interpretation by nontechnical users. The TNRIS system includes six data/information categories: base data (including aerial photographs and other remotely sensed data, cartographic materials, and geographic subdivision schemes for locating and assessing natural resources), meteorological data (daily/monthly records, obtained primarily from the National Weather Service), biological or ecological data, information on geology and land use (much of this data is not yet in machine-readable format), socio-economic data (census information), and water data (including USGS stream flow data). Files pertaining to groundwater include water level measurements, groundwater quality data, location and characteristics of subsurface injection wells, well drilling logs, and a well sample and core library. The state-wide water level network is comprised of more than 7,000 wells; the water quality monitoring network includes over 5,000 wells. Water level measurements may be annual, biannual, bimonthly, or continuous, and generally make use of automatic recorders.

Most states take a less integrated approach than Texas. For example, the Nebraska Natural Resources Data Bank Information System Data Access Manual (NNRC 1981) is a state-wide facility for storing, processing, and retrieving basic water and soil data. Both groundwater levels and water quality data are available. Data may be retrieved by county or specific stations. Data summary services include number of observations, maximums and minimums, means, standard deviations, variance, variation coefficient, and beginning and ending record dates.

Project Data Management

In some cases, project-oriented databases have reached a high level of sophistication and detail and thus provide an useful example for other data acquisition and data management projects. One such project is the Wisconsin Power Plant Impact Study (WPPIS) (Shacham and McLellan 1979). To make the datasets collected during this study accessible to other scientists, investigators at the University of Wisconsin-Madison prepared a reference volume. The description of each dataset includes name of dataset, names of investigators, brief abstract, experimental method, dates of research, and data form. The project's data management was designed to facilitate continuous data synthesis during the data gathering. Procedures were put in place to prevent duplication of unnecessary data collection, to facilitate the transfer of data and other information from one disciplinary subproject to another, and to organize the data to facilitate their maximum use. Protocols were established to assure that data were as free from interpretation as possible, and that data were in a format that allowed assessing uncertainties in the data independently. In addition, the project data documentation describes data form and format in detail, provides a data sample, and contains a QA report and other supporting documents. Such a detailed documentation makes these datasets highly useful for secondary analysis, synthesis, or theoretical studies.

DATA MANAGEMENT

The availability of data and the form of the datasets resulting from research projects are related to their projected use. Often, however, not much thought is given to data use beyond the project's research objectives. To address this issue the W.K. Kellogg Biological Station (1982) requires in the planning stage of each research project that measures be taken to facilitate later access to the collected data. In the management of the datasets with which it is involved, the Station distinguishes between four major responsibilities: (1) research data analysis by the data-collecting team (primary use); (2) compilation of databases for secondary use; (3) data referral through directories and catalogs (information services that help researchers locate and obtain datasets); and (4) archiving in data banks (established to maintain [at least] those datasets that have no other means of long-term care).

Data management considerations should include identification of potential data (center) users; identification of data that should be included in the database or center; alternatives for finding a suitable method of data presentation; determination of how data is to be documented; determination of how the center is to be described and made known to users; and selection of storage and retrieval methods.

Furthermore, the underlying goals of the data center should be well understood by the data center management; any changes in these goals should be clearly reflected in management considerations. Management issues include user access and data retrieval, data maintenance, quality assurance, documentation, standardization, geographic information systems, and system linkage. Each of these issues is discussed separately on the following pages.

User Access and Data Retrieval

Moffett (1983) stated that masses of data without ready access and analytic capability have very limited universal value. Unfortunately, this is often the state of most scientific research data because most databases are being created by agencies, industry, or research organizations for a specific purpose, with little effort to benefit users beyond the immediate need.

Armentano and Loucks (1979) encountered many problems associated with the access of data, even from data centers with some type of access procedure in place. Case trial runs to acquire selected information from a few existing information centers exposed problems in contacting personnel, obtaining initial access, accessing on-line databases during prime-time hours, specifying output formats, knowing what datasets were available, user manuals, software system changes, hardware system updates, and terminal procedures. Identifying a principal contact within the data management organization often proved to be a frustrating exercise and required a significant amount of time. Delays of up to three weeks were experienced for the receipt of the requested data. Moreover, for many databases, the access mechanisms, contact persons, and available resources change frequently. Armentano and Loucks (1979) concluded that contacts with potential data sources should be formalized and that emphasis should be placed on determining the accessibility of databases, identifying principal and secondary contacts, and providing complete data documentation.

Procedures for accessing on-line data systems range from simple to highly complex, depending on hardware and software environment, data security, and administrative concerns, among other factors. Issues involved include obtaining computer access authorization and passwords, ensuring compatibility of user equipment with the host system, and learning the search, retrieval, and report procedures of the particular data management systems. Commercially offered data systems usually require some contractual agreement up front, with charges based on connect times. Often, access to these commercial systems is enhanced by user-friendly menus, on-line help, and extensive system manuals.

Data Maintenance

Data maintenance is the continuing process of adding and reviewing data, updating system hardware and software, and reviewing and updating quality assurance, data access procedures, and data documentation. Data maintenance procedures should be documented and this information should be made available to data system users.

Data maintenance requires permanent or long-term operation of the data center. If a data center project represents only a one-time effort, it may already be out-of-date, or may rapidly become so. Typically, this is the case if the data, data directory, or user instructions are available only in printed form.

Data entry, modification, and deletion procedures comprise a large part of the computer maintenance system. Procedures in place for data entry might include automatic syntax checks (e.g., versus a subset of the 256 ASCII-defined characters), compatibility checks within the database (e.g., depths to water cannot exceed depth of well), and out-of-range checks (Mercer and Morgan 1982). Some databases have more extensive error checking capabilities including checking new data for outliers and trends (e.g., Mitchell et al. 1985).

Quality Assurance

A major issue in the secondary use of research data is quality assurance/quality control (QA/QC). How does a data center assure the quality of the data offered? In fact, the QA/QC issue can be split up in two related concerns: (1) the quality of the data acquired by the center and (2) the measures taken to preserve the integrity of the data within the data center. Armentano and Loucks (1979) concluded that quality control in database management often appears to be inadequate. They recommended that procedures be established to certify data files and that mechanisms should be found to fund related data center management. The researchers collecting the data and providing them to the center must also be considered in any plan to maintain quality control.

The need for high-quality data varies by program and by decision within programs. QA documentation often includes descriptions of monitoring equipment such as names and technical specifications (detection limit, zero drift, accuracy, and precision). For example, the WPPIS study included a separate QA report for each dataset, thus providing the information necessary for an assessment of the validity and reliability of the measurements (Shacham and McLellan 1979). Typical topics addressed in these QA reports include the objectives of the data acquisition; descriptions of the system studied, and of the methods, equipment, and procedures used; evaluation of the equipment performance; and the names and qualifications of the investigators.

Another example of data center QA/QC is present in the Texas Department of Water Resources' Industrial Waste System (TNRIS 1983). Information submitted for storage is examined through edit checks performed on most data elements to insure that they meet known characteristics and that essential information is present; cross-checks are made to ensure that the various codes used are defined in the data system; and related information is matched for economy of storage and to facilitate subsequent retrieval.

Documentation

The main objective of data documentation is to provide a secondary user the necessary information to decide for whether a particular dataset is of sufficient quality. Whatever data verification techniques are used, the data documentation should make clear to the user what procedures have been used (W.K. Kellogg Biological Station 1982).

Research project data sets are usually documented in abstract form and include such details as experimental methods, physical setting, sampling duration, QA procedures, data storage methods, and data variability. The abstract should provide a brief description of the dataset and include the purpose for collecting the data and the context of the dataset in the research project as a whole. Armentano and Loucks (1979) considered data to be best abstracted as raw data.

Olson (1984) concluded that dataset descriptors developed by various clearinghouses or agencies differ significantly. Most of the descriptions used included a dataset title, abstract, spatial and temporal coverage, and contact person; some contain detailed descriptions of individual variables within each dataset.

The Wisconsin Power Plant Impact Study recommends a twofold approach to the data documentation (Shacham and McLellan 1979): (1) a short description of all datasets in the database for initial screening and selection by secondary users; and (2) a more detailed description for final selection from a short list and as initial guidance for data use. The extended version includes description of data form and format, a data sample, a QA report, and supporting documents. It also contains information on funding agency type, project classification, temporal data, and units of data.

In addition to printed versions, documentation may take the form of help features in the data entry and retrieval system and text files accessible directly on the computer system used. An example is the TNRIS "File Description Report" (TNRIS 1982) which includes the file name, kinds of data in the file, units of measure, geographic coverage, period of record, and capabilities associated with each file (printed reports, plots, and terminal access).

According to Hoffman (1986), dataset documentation should include dataset name, title, files, research locations, investigator, other researchers, contact person, project, source of funding, methods, storage location and medium, data collection time period, voucher material, processing and revision history, and usage history. Furthermore, data files should be documented with respect to file name, constituent variables, key variables, subject, storage location, physical size, file creation methods, update history, and summary statistics.

The W.K. Kellogg Biological Station Workshop (1982) recommended that documentation for data variables should cover such items as variable name, definition, units of measurement, precision of measurement, range or list of values, data type, position in file and/or format, missing data codes, and computational method. For data catalogs and directories, information should be included on dataset name/title, data collection time period, types of parameter or variable, investigators, contact person, relevant bibliographic references, dataset storage location, research location, keywords, and research site characteristics.

Standardization

Armentano and Loucks (1979) found that environmental and natural resource data files are generally widely dispersed and hard to obtain, and that these files reflect a general lack of standards or coordination (units of measure, spatial identifiers, data formats, etc.). According to a recent study of EPA's groundwater data management requirements (EPA 1987a), one of the major problems with the current dissemination of data is the lack of consistency and standardization in data collection, coding, and reporting among EPA Program Offices and Regions, and the states. This lack of standards reduces the value of the data for sharing and limits the integration of data originating from many sources. The study concludes that one of the fundamental building blocks for improving groundwater data management is the adoption of standards and common formats for data collection and storage.

In reporting the characteristics of datasets it is necessary to document the extent of data interpretation used before its storage in the database. In addition, uncertainties in the data should be expressed in a clear manner so they may be assessed independently by secondary users. Observational data should be prepared and documented in such a manner that they can be re-analyzed independently, e.g., in terms of a different hypothesis than that presented by the original investigators.

In the development of a data center, standards or conventions for reporting null data, missing data, numerical values of zero, and values below detection limits must be seriously considered. Although the use of standards will help integrate different data systems, their development and application should be carefully considered. As Armentano and Loucks (1979) pointed out, a data center can take aggressive action to encourage the development and use of data collection, storage, and reporting standards.

Another standardization issue raised in the past is the abstracting of machine-readable files (MRDFs). An interagency committee on data access and use proposed a voluntary standard for use in preparing abstracts to describe MRDFs (OIRA 1983) (see Table 10).

Geographical Information Systems

A geographical information system (GIS) is a computerized data handling and data analysis system designed to accept large volumes of spatial data derived from various sources (Marble and Peuquet 1983). A GIS allows efficient storage, retrieval, manipulation, analysis, and display of these data according to user-defined specifications. A GIS is basically a combination of an efficient, dedicated database management system, standardized data processing, procedures for analyzing relationships between the data elements, and (graphic) display. Such systems differ from other management information systems by their focus on spatial characteristics and relationships and in their ability to facilitate the analysis of a large number of interrelationships among the spatial variables. They allow projections of different data types over each other (overlays), and extensive statistical comparisons.

TABLE 10. COMPONENTS OF THE PROPOSED VOLUNTARY STANDARD FOR ABSTRACTS OF MACHINE-READABLE FILES (OIRA 1983)

1. Bibliographic Citation
   1.1. Title, followed by the general material designation (machine-readable data file)
   1.2. Statement of responsibility (authorship, sponsor, collaborator, etc.)
   1.3. Place of production
   1.4. Name of producer
   1.5. Date of production
   1.6. Place of distribution (if different from place of production)
   1.7. Name of distributor, followed by qualifier (distributor)
   1.8. Date of distribution (if different from date of production)
2. Date of abstract and file number
3. General description and subject matter coverage
4. Descriptors
5. Geographic coverage
6. Time coverage
7. Technical description: file structure and file size
8. Reference materials
9. Related printed reports
10. Related files
11. OMB clearance number
12. Contacts
13. Availability information

The GIS concept is interesting to environmental data centers because of its efficient handling of the highly diverse, spatial field data of interest to environmental research and management. A GIS provides a framework for data analysis and QA, and drives a certain level of standardization in data formatting, reporting, and display.

Over the years various types of GIS have been developed, ranging from simple, limited-capability systems to large multipurpose implementations. Some of these systems have been designed for a specific management objective or area, while others are completely generic. An example of a large, dedicated GIS is the Texas Natural Resources Information System mentioned before (TNRIS 1982, 1983). In the "Base Data Resources" section of TNRIS, digitized water-well locations and county boundaries are present; these, when combined with latitude/longitude grid lines, provide effective information for local managers. The system's GIS capabilities can be used to generate products that would be difficult for the user to prepare (e.g., because of lack of expertise) or that would require considerable time and expense.

ARC/INFO (Environmental Systems Research Institute, Redlands, California) is a generic GIS with a wide variety of user-specified options, available on a large range of computer environments. It is well-suited for data management related to groundwater modeling projects (e.g., ESRI 1986). It contains routines to perform topological overlay functions useful in building a model grid, performing advanced analysis, and allocating spatial variables to the model nodes and cells. Model results can be put back into the GIS and analyzed together with the existing spatial data (e.g., location of pollution sources, surface water features, etc.)

Linking Data Management Systems

In some cases, the linkage of different data centers might be more efficient and feasible than complete standardization or centralization. Such linkage requires extensive coordination between the system elements to be linked, as well as a clear view of the structure and operation of each individual element. The U.S. EPA Data Requirements Analysis (EPA 1987a) made several recommendations with regard to such system linkages around STORET. These will improve remote access to groundwater data (in a kind of "one-stop shopping") and will facilitate extended data exchange, transfer, and sharing among different organizations. The benefits become clear if links are initiated between STORET and other national groundwater data systems, e.g., WATSTORE, or selected state groundwater data systems. The central coordinating agency (e.g., EPA) should be charged with implementation and operation of interfaces for the routine transfer of data between all system elements.

Another example of a linked system is the Chemical Information System (CIS), a privately operated system of 22 databases and software analytical programs available to environmental scientists. The system incorporates software to solve problems inherent in system linkage. These problems are generally that each database may require use of a unique approach to extract information (search and retrieval), and that various levels of data quality evaluation are required, as well as differences in methods of updating and expanding the databases.

Data Center Maintenance

An important aspect of a data center is assuring its continuity. Often, this equates with assuring a certain level of funding for an extended period (e.g., for the center's anticipated lifetime or until a predetermined date when its utility is re-evaluated). According to Olson (1984), inadequate funding might result in:

- reduction of the scope of data collection
- decrease in services
- delay in new data additions or less frequent updating
- delay in incorporating new technologies and system updates
- reduced access or unavailability of data for secondary users (i.e., delays in providing datasets in a timely manner)

- reduced quality control
- loss of experienced staff
- obsolescence of data

However, decrease in funding might be justified or compensated by better coordination with data providers, improvements in operational procedures, and adapting to the continuing advances in computer technology.

In general, three mechanisms are available to fund a data center's maintenance: (1) funding from institutional overhead; (2) allocation of project funds to database management; and (3) separate cost recovery from data users. Often, a combination of cost coverage is sought.

CONCLUSIONS

The range of information or datasets offered by different centers is determined by the particular goals or purpose of the data center, the means by which data are obtained, and the sources of the data or information. The requirements of management, especially for government activities, can be short-term and may be affected by rapid changes in funding or organizational structure. The management and organizational approaches of a center are initially developed to meet the specific objectives. Although these objectives may change over the course of a center's operation, the center's progress toward meeting its original goals must be rapid for a center to be successful.

Many researchers have pointed out the uselessness of data for which the quality is unknown. Unfortunately, the computer seems to have given data an aura of authenticity it does not always deserve. Unless a data center documents the quality of its data and is willing to provide detailed characteristics relating to it, its utility is in doubt. Adhering to rigid QA procedures, and maintaining hardware and software system security, can help to assure secondary users of overall data quality. However, where possible, such procedures should not replace direct involvement of the data generators in the secondary use, or at least their use as information source and guidance.

In order to interest potential users in the service of a data center, these services must be widely publicized. The clarity, style, and scope of its promotion affects the extent of a center's use.

There is a need to share not only data, but also expertise on information management itself, with several possible ways to share this expertise. One possibility is to establish training courses, consulting services, and internships that benefit from the experience and expertise of leaders in scientific information management. Conferences or workshops are also of great value.

Finally, data providers are concerned about data integrity and preserving their vested rights in their data. Armentano and Loucks (1979) suggested that a protocol for third-party use should be adopted, aimed at preserving the original intentions of the investigator as discussed in research reports associated with the database. Such a protocol is aimed at securing the integrity of the dataset (and the investigator) and preventing the out-of-context use of the data. One such rule might be that subdividing complex tables into more than one dataset would require the approval of the original investigator.

SECTION 4

QUALITY ASSURANCE/QUALITY CONTROL FOR RESEARCH DATASETS

INTRODUCTION

To determine the validity of models and their underlying theory, data are used from field observations and laboratory experiments. The research data in which the IGWMC Data Center is interested might result from research aimed at any of several objectives: determination of pollutant concentrations from various sources in the environment, pollutant transport and fate in the subsurface, response of organisms to pollutants, exposure levels, effects of pollutants on human health and ecosystems, the environment, analysis of risks and benefits, and evaluation of economic impact. These data must be scientifically valid, defensible, and of known and acceptable accuracy and precision and, where feasible, reproducible (Stanley and Verner 1985). Thus, the quality of data needs to be described in terms such that these requirements can be met.

To be able to evaluate the quality of data, the data acquisition, data processing, and data storage procedures must be documented. Because most researchers tend to emphasize the analysis of their data rather than its collection and validation, important qualifying information about the data is often lacking or hard to obtain. This problem is of special concern since many research projects focusing on observation and quantification of natural processes employ new, often experimental data-acquisition techniques and procedures.

Generally, the quality of data is evaluated by assessing its uncertainty relative to the requirements for its specific use (which for the same data might vary, depending on the projected use). From a modeling perspective, a distinction can be made between model validation, use of models in regulation development, technical design, and legal enforcement (van der Heijde et al. 1988). If the data have consistency and a small uncertainty when compared with those requirements, they are considered to be of adequate quality (Taylor 1985). As quantitative measurements are always estimates of the true value of the parameter of interest, the measurements must be made in such a way that the uncertainty can be posed in terms of probability. At the same time, assessing the quality of data is a rather subjective process that depends on the objectives, experience, and personal bias of the assessor.

QA/QC

To ensure that data collection meets project objectives, it is essential that a systematic, well-defined, and controlled approach be taken to all steps in the data gathering process. The formulation, implementation, and control of such an approach is the objective of quality assurance/quality control (QA/QC). Taylor (1985) defines quality assurance as "those operations and procedures which are undertaken to provide measurement data of stated quality with stated probability of being right." Thus, a primary goal in developing quality assurance in data collection programs is to ensure that such data is of known quality.

Important to the quality of the data collected is the use of appropriate methodology, and adequate calibration and proper usage of the equipment involved. Quality control refers to the overall system of standards, guidelines, procedures, and practices designed to regulate and control the quality of the collected data (Taylor 1985). Proper documentation of all elements of the data acquisition, data processing, and data storage procedures is essential for a good quality control program.

To facilitate screening of datasets for their utility in model validation by secondary users (including the Center's staff), this section provides background information on the various standards, guidelines, practices, methods, and equipment applicable to the various relevant data acquisition and processing procedures. Additional QA/QC issues are also discussed, especially with respect to proper documentation.

A distinction is made between data necessary for model-based calculations (numerical values for model variables and parameters such as concentrations, piezometric head, transmissivities, dispersivities, etc.) and data that are used for conceptualization only (e.g., to determine hydrogeological schematization and type of system boundaries; this includes bore hole and morphologic sample descriptions and data from geophysical surveys). This distinction is made because different levels of quality apply to these two data categories. The level of detail required in peripheral data such as available through remote sensing and from meteorological measurements depends on their use in model validation. Often, accurate data regarding evaporation and recharge from precipitation are essential to successfully complete the testing of a model for a particular field site.

The Role of Standards and Guidelines

Increasingly, environmental data acquisition is subject to a framework of standards, regulatory requirements, and technical guidance. In the past the development of standards for field procedures in groundwater measurements has received little attention. Recently, however, developments in the scientific and regulatory arenas have accelerated the discussions on standardization and quality assurance in this area. This recent emphasis stands in stark contrast to the field of analytical chemistry, where over the years a considerable effort took place to standardize and regulate laboratory methods and practices (Nielson 1987).

The current developments in standardization and establishment of groundwater quality assurance guidelines are illustrated by the activities of the Subcommittee on Groundwater Monitoring (D18.21) of the American Society for Testing and Materials (ASTM) in developing standards for monitoring methods and equipment (Lorenzen et al. 1986, Perket 1986, ASTM 1987a, 1987b, 1987c, 1988, Friedman 1988, Collins and Johnson 1988). The U.S. Environmental Protection Agency (CDM 1987, van Ee and McMillion 1988), the Electric Power Research Institute (Summers and Gherini 1987), and the U.S. Geological Survey (Collins and Johnson 1988, pp. 17-21) also study and document groundwater measurements and QA/QC methodology.

Assessment of Data Quality for Secondary Use

One of the most difficult issues in the secondary use of data is determining if the quality of the data under consideration is sufficient for the intended use. Often, such determinations are hampered by incomplete definition of the user's objectives, and show inconsistencies between the various data types evaluated and between the data from different data sources. Even if the documentation of the data is relatively comprehensive, it often proves to be difficult to make a quantitative assessment of the data quality. In most cases, assessing the usefulness of a dataset based on a quantitative comparison of data uncertainty with the limitations on uncertainty placed by their projected use, is not possible.

Describing the quality of data quantitatively requires the definition of measures for data accuracy and precision. If these measures are not known, or when intuitive, incomplete, or inconsistent approaches have been taken in deriving them, their value may be limited. Thus, even if measures for data accuracy and precision are available, the manner in which the data collecting team derived those measures should be discussed in the documentation. If such quantitative measures are absent, the data user needs to make an assessment based on qualitative information. In this latter case circumstantial documentation (e.g., field conditions, methodology applied, etc.) is required. However, such a qualitative assessment of data quality by the user is often biased and subject to inconsistencies.

As the "true" value of a particular measured data item is never known, its accuracy is established by ensuring use of the most appropriate method and applying that method as efficiently as possible. Examples of a systematic approach in evaluating data quality can be found in Campbell and Mabey (1985), Kaplan et al. (1985), and CDM (1987). Campbell and Mabey 1985) proposed the following four-level evaluation procedure for measured field and laboratory data:

- basis of measurement: identification of method and limitations of the method

- application of method: adherence to the method, controls and calibration, and data conversions

- statistical information: statistical methods and reproducibility

- corroborative information: independent measurements, alternate methods, and related information.

For each of these categories, the authors discussed the measurement process points to consider.

An additional complexity in evaluating data for model validation is caused by the interpretations which often take place between the collection of raw data and its transfer to the secondary user. Evaluating the quality of these processed data includes scrutinizing such data manipulations. Again, proper documentation of the data manipulations is an absolute requirement.

Quality assurance (QA) is a major concern for the successful operation of a central data distribution facility. This concern pertains to both the quality of the data acquired by the Center and the data handling within the Center. To assess the quality of data considered for distribution by the Center, information needs to be analyzed with respect to the kind of data collected and the methods of collection, both from the perspective of adequate choice of techniques and instruments, and measurement execution.

To this purpose this chapter discusses the role of QA in data collection and handling, and briefly reviews the major soil water and groundwater data types and related field measurement techniques. Where possible, discussion of existing standards and guidance has been incorporated. This information served as a basis for the development of the descriptive system used in the SATURN referral database (discussed in section 5) and provides guidance to information sources needed in evaluating the quality of the datasets considered for future incorporation in the Data Center.

SOIL WATER AND GROUNDWATER CHEMISTRY

Introduction

Soil water and groundwater quality at a specific location may depend on factors such as the heterogeneity of the soil and rock, and past and present flow regimes. In designing a monitoring network and sampling strategy, the representativeness of samples for the sampled environment is a critical issue. Standard protocols might remove or control some sources of error and therefore set limits to the uncertainties present; however, standard protocols do not and cannot address fundamental issues of representativeness.

The preferred method of determining the quality of groundwater is in-situ sampling, in conjunction with chemical analysis in a laboratory or on-site. To allow for independent assessment of data quality, the following items should be documented:

- sampling (network and sampling design, procedures, execution, special operating conditions)
- transportation, preservation, and storage
- laboratory or in-situ analysis (including calibration procedures and frequency, use of reference and quality control standards)
- data reduction, validation, and reporting

Documentation of both sample collection and sample analysis should include: (1) statistical descriptions providing an indication of the representativeness of the data and the level of confidence placed on the data; and (2) standard operating procedures to ensure sampling integrity and data compatibility and to reduce sampling and analytical error.

Monitoring System Design and Sampling Programs

A perfectly designed monitoring system would provide the optimum number of samples for a sufficient but not excessive amount of data to characterize the contaminant/tracer plume, and would also provide the necessary degree of confidence in the quality of data to support its intended use (Schweitzer and Black 1985). Locating the wells optimally in a monitoring network is a challenging problem, especially taking into consideration uncertainties in the physical system (Meyer and Brill, Jr. 1988).

The selection of materials and equipment for obtaining the appropriate information must take into account state-of-the-art techniques and equipment as well as the necessary scope or detail for the project. The acceptable level of uncertainty must be considered, and the quantification and documentation of this uncertainty must be established. Decisions regarding analytical requirements for sampling should be made in conjunction with the design of the monitoring system.

The identification of all possible sources of error is essential to the QA/QC of groundwater monitoring data. Many researchers (Nelson and Ward 1981, Keely 1982, Schweitzer and Black 1985, Montgomery and Sanders 1986) have recently attempted to design quantitative groundwater monitoring systems that take into account data uncertainty and the statistical aspects of field sampling. In defining data uncertainty in terms of how well representative observed values reflect true population characteristics, Montgomery and Sanders (1986) propose a method to estimate this uncertainty as a function of both sampling and non-sampling errors.

Nelson and Ward (1981) state that with careful planning an iterative procedure can be followed in developing monitoring strategies. In documenting the design of a monitoring system, the following criteria and considerations should be addressed:

Spatial Soil Water and Groundwater Monitoring

- Number of vertical sampling points in relation to the thickness of the aquifer or soil zone, plume depth, and the physical character of the system

- Number of horizontal sampling points in relation to plume area and soil and geohydrological conditions

- Surface-water-related features, drainage, infiltration, and nearby pumpage

Temporal Soil Water and Groundwater Monitoring

- Frequency of sampling

- Seasonal variations

- Presence and effects of sorption, degradation, and (bio)chemical processes
- Velocity of contaminant plume

Sample Representation of the True Population Characteristics

- Statistical evaluation
- Soil and geohydrological conditions
- Shape and extent of plume

Methods Used for Design and Evaluation of Data Uncertainty

- Division into sampling and non-sampling errors
- Addressing spatial and temporal aspects of system
- Method of quantification

Background Water Quality

- Comparison with concentrations of contaminants or tracers

Effects of Well Construction on Sampling

The selection of well construction materials and drilling techniques is determined by the hydrogeology of the site and by the project objectives. In particular, the drilling technique selected depends on designed well depth and diameter, lithology present, chemical constituents being sampled, the type of monitoring system being installed, time limitations, need for rock samples, potential effects of circulating fluid (if present) on the formation, water quality, mobility of the drilling rig, integrity of the borehole, ability to sample formation waters during drilling, and economic limitations (Barcelona et al. 1985, Driscoll 1986).

Recent studies addressing the effects of well construction on the quality of groundwater samples include Fetter (1983) and Dunbar et al. (1985). Barcelona and Helfrich (1986) and Fetter (1983) have pointed out that casing materials may affect groundwater quality, and thus sampling, especially with respect to certain organic chemicals.

While studying cement grout contamination in stagnant well water, Barcelona and Helfrich (1986) found that drilling fluids may significantly affect the groundwater quality and the functioning of the well, regardless of how often attempts to develop the well are repeated. Additional discussion regarding the selection of a drilling method and well construction materials can be found in Barcelona and Gibb (1988) and Kerfoot (1988).

The documentation of the employed drilling technique should address the following items, among others:

- drilling technique
- drilling fluid: type, quantity
- cleaning/decontamination procedures
- well development technique: volume of aquifer water evacuated, rate, time between well completion and development
- casing materials: type, brand, specifications
- screen material: type, size, specifications
- grouting material and installation: volume, type, procedures
- gravel pack material and installation: volume, size, procedures
- construction procedures
- procedures for special situations: e.g., drilling to avoid mixing of polluted and clean water
- well logging, flow measurements, etc.

Sampling Soil Water and Groundwater

Although QC procedures for groundwater sampling should be based on proven field measurements and sampling procedures, this is often impossible because of the wide variety of hydrogeologic and geochemical conditions, and even though many field sampling methods are currently available, their reliability, precision, and accuracy have not been established (Murarka and McIntosh 1987). There is still a lack of standardization of groundwater sampling procedures following systematic development and controlled evaluation trials, although discussion on standards are presently taking place (e.g., Collins and Johnson 1988). Therefore, it is essential to document these procedures and the field sampling conditions with the incorporation of QA/QC techniques used.

The field sampling procedure must account for possible cross-contamination between different wells. This can occur as a result of incomplete or inadequate equipment decontamination procedures. Bryden et al. (1986) discuss the need to decontaminate purge-and-sampling equipment in the field when sampling a number of wells at a site.

The partial pressures of certain gases (particularly carbon dioxide, oxygen, and volatile organics), which are often significant constituents of groundwater, may change dramatically as the water is exposed to the atmosphere or when vacuum is applied to remove a sample. These changes may be significant in and of themselves (e.g., loss of volatile organics), or they may produce other, equally significant changes in groundwater quality, e.g., pH which in turn affects precipitation of other elements. Groundwater sampling must minimize the loss of any volatile substances, especially if they are the subject of analysis. Specially designed equipment, methods, and procedures to assure this are described by Barcelona et al. (1984), Barker et al. 1987), and Pearsall and Eckhardt (1987).

Barcelona et al. (1985) have suggested that a truly quality-assured sampling protocol should be backed up by laboratory and field experiments that are designed to quantify any random or systematic errors, and that the most direct way to evaluate sampling protocols is to specify and document that class of chemical constituents most prone to sampling error. A research tracer study (Roberts and Mackay 1986) offers an example of documentation of a field experiment on purging requirements and the laboratory testing of a prototype sampling station. The validation of field procedures depends on techniques already documented and accepted and on the individual concerns of the project personnel. Any special efforts to validate field sampling methods in order to meet project goals should also be documented.

Although sampling protocols may vary from well to well due to local hydrogeologic factors and well construction, once they are defined, wells should be sampled consistently since trends in water quality data can be created by changing sampling procedures (Slawson et al. 1982).

For sampling soil water, special measures have to be taken. Vadose zone monitoring can involve either sample material removed from the test location or in situ measurements (Barth and Mason 1984, Everett et al. 1984a). Extraction techniques of pore fluid from soil depends on whether inorganics, organics, or microorganisms are under investigation. Liquid sampling methods depend on whether fluid is taken from the saturated or unsaturated regions of the vadose zone (Everett et al. 1984a). Many of the documentation requirements listed above apply equally to soil water sampling.

Everett (1981) provides an introduction regarding instruments and methods for monitoring in the vadose zone. Wilson (1980) presents a detailed catalog of methods for monitoring or estimating waste water fluxes in the vadose zone, including a description of the method, principles on which it is based, advantages and disadvantages, and pertinent references. Everett et al. (1984a) reviewed over 50 types of vadose zone monitoring devices and methods. Everett et al. (1984b) described the factors limiting the operation of suction-type sampling devices and discuss problems related to the representativeness of samples.

In summary, documentation of groundwater sampling should address:

Purging

- number of well volumes removed, rate removed, time between purging and sampling, consideration of hydraulic properties, well design, and changes in color or cloudiness of water

- continuous parameter measurements to check effectiveness of purging (e.g., to determine when representative parameter stabilize): type, equipment used, specifications, calibration frequency, standards or procedures used

- field/laboratory experiments made to determine purging requirements

- equipment used to purge well

- destination of evacuated water

Sampling System

- equipment and materials (pump or bailer specifications)

- consistency of sampling

- considerations for the sampling of multiphase constituents

- decontamination/cleaning of equipment

- sampling containers

- handling

- labeling

Field Analyses

Field analyses of samples effectively avoids bias in the determination of parameters for constituents that do not store well (e.g., alkalinity and pH) (Barcelona et al. 1985). Optimal field measurements of pH, Eh, and specific conductance should be made in the well or in a closed cell to avoid sample oxidation. Field procedures for measuring unstable constituents are given in USGS (1977) and Barcelona et al. (1985), among others.

An external QA program for USGS field measurements is described by Cordes (1984). It recommends that documentation of field analysis should include a description of equipment and methods used, calibration of instruments, field conditions, reference samples, replicate measurements taken, equipment maintenance, and inter-comparison studies performed. Logs should detail all field measurements.

Sample Handling and Documentation

Sample handling depends on the particular constituents to be analyzed and may include filtration, temperature control (cooling), chemical fixatives, special sample containers, and elimination of headspace. In some cases, no special precautions are required. Decisions concerning sample preservation and containers should be coordinated with the laboratory where the analysis will take place and with the monitoring team.

Details about the filtering methods can be found in Nacht (1983) and Barcelona et al. (1985), while USGS (1977) describes container materials to be used for different groundwater constituents. A purge and trap device for isolating volatiles is described by Semprini et al. (1987).

To assist in interpreting water quality data, the USGS (1977) and ASTM (1987b) give recommendations for sample identification and documentation. This should include method of collection, exact location of well or source, sample number, depth and diameter of well, casing record (including screened intervals and type of screens), water-bearing formation(s), water level, rate of discharge and duration of pumping prior to sampling, water temperature and other field measurements, date and time of collection, appearance, type and quantity of preservative added, and any other relevant observations such as use of water.

Methods of sample identification should be consistent and provide for adequate numbering systems of the sampling points and the samples. Field and laboratory notes should indicate any deviations from prescribed labeling protocols.

The transportation of a sample may create data uncertainty from such factors as agitation, temperature, light, and time until filtering and analyses (Montgomery and Sanders 1986). ASTM (1987b) gives recommendations for sample shipping. Summers and Gherini (1987) outline directions for labeling sampling containers, packing samples for transport, and sample documentation through chain-of-custody forms. Scalf et al. (1981) provide guidelines for maintaining sample records.

Data uncertainty from sample storage can be attributed to chemicals added to preserve the sample, filtration prior to analysis, the type of storage container, the method of storage, and changing conditions of storage due to such effects as temperature, light, and length of time of storage (EPA 1982, Friedman et al. 1986).

Attention to cleaning procedures for sampling containers is crucial as a prerequisite for sample collection. Pre-cleaned sample containers are detergent-washed, and acid, solvent, and deionized water-rinsed under quality controlled conditions.

In summary, documentation of sample handling should address:

Preservation

- filtration in field (filter type used, method and equipment, time lag, filter size, prewashing of filter)

- preservatives added (in field or lab)
- type of storage container
- transport and storage conditions
- "headspace" specifications
- temperature control
- coordination between lab and monitoring team

Sample Identification

- documentation system
- sample number and sample type
- well number, sampling point, time, and date
- location
- personnel

Laboratory Analysis

There has been a phenomenal increase in analytical sensitivity for environmental measurements over the past decade, resulting in very low detection limits for many constituents. However, great disparities are found between the ability of present techniques or methods to detect and measure contaminants at ultra-trace levels and the data actually reported at very low concentration levels. In addition, it is rare to find reported sensitivities, detection limits, or data associated with degree of uncertainty (e.g., in terms of confidence levels or intervals) (NWQL 1986).

Data Quality Indicators--

Data quality indicators such as precision, accuracy, completeness, and detection limit of the method, are determined through laboratory procedures for analyzing duplicate, control, blank, and spiked samples, and through subsequent statistical analysis to quantify the results. Where possible, these data quality descriptors are expressed and reported in terms of standard deviation, variance, and relationships between precision and concentration levels (EPA 1983, Paulsen et al. 1988).

Precision is the degree of agreement of repeated measurements of the same property, expressed in terms of dispersion of test results about the arithmetical mean result obtained by repetitive testing of a homogeneous sample under specified conditions (ASTM 1987c). The precision of a method is expressed quantitatively as the standard deviation computed from the results of a series of controlled determinations. Accuracy refers to the correctness of the data and defines the degree of agreement of the measurement with the true value of the magnitude of the quantity concerned. Bias is the systematic deviation of the average of one set of data from another. Bias is defined by ASTM (1987a) as the persistent positive or negative deviation of the method average value from the assumed or accepted value.

Sensitivity is a measure of instrument response factor as a function of concentration. It is commonly measured as the slope of the calibration curve (NWQL 1986).

A confidence interval is an expression of the range of values defined by upper and lower limits at a statistically defined confidence level (NWQL 1986).

The Limit of Detection (LOD) is defined as the lowest concentration level that can be determined to be statistically different from a blank at a specified confidence level. A distinction can be made between instrument, method, and practical detection limits (NWQL 1986).

The Limit of Quantification (LOQ) is the level above which quantification results may be obtained with a specified degree of confidence. It is calculated using the standard deviation and average level for the blank. This defines the level above which quantification is reliable and also a region between MDL and LOQ, where detection is reliable but quantification is not (NWQL 1986).

The Method Detection Limit (MDL) is defined as the minimum concentration of a substance that can be measured and reported with 99% confidence, for which the true value, corresponding to a single measurement, is above zero (EPA 1983). The MDL is the lowest concentration of analyte in distilled water that a method can detect reliably and is statistically different from the response obtained from a blank carried through the complete method including chemical extraction or pretreatment of the sample (NWQL 1986).

NWQL (1986) describes the Practical Detection Limit (PDL) as the lowest concentration of analyte in a real sample-matrix that a method can detect reliably and is statistically different from the response obtained from a blank carried through the complete method.

Finally, completeness is a measure for valid data obtained from a measurement system expressed as a percentage of the amount of data that should have been collected. Completeness is of particular importance to multi-year intensive monitoring programs.

Laboratory QA/QC--

Analytical methodologies and QA/QC procedures have been standardized and modified over a longer period than groundwater monitoring field procedures. The quantification of data uncertainties is described using measures of accuracy, precision, bias, sensitivity, confidence intervals, detection limits, and completeness. Quality control samples, including replicates, blanks, controls, spiked samples, rinse samples, and reference materials, are used to quantify data uncertainty in the laboratory. Eggenberger (1985) discusses fundamental principles and methods to ensure accurate and reliable laboratory analysis results. The estimation of analytical procedure variability may be attained by duplicate analyses, by multiple analyses of a stable standard, or by using spike recoveries. Control of analytical performance may be defined using the repeatability—the agreement among replicate observations of the test by a single laboratory or analyst expressed in standard deviations and the percent recovery—describing the ability of the analyst and the procedure of the analytical system to recover a known amount of a constituent added to a natural sample (may be obtained by spiking). EPA (1983) and ASTM (1987b) give recommendations for frequency of duplicate samples and control samples.

The American Chemical Society (ACS), the National Bureau of Standards (NBS), the American Public Health Association (APHA), and the Water Pollution Control Federation (WPCF) have long been concerned with the establishment of general laboratory practices and analytical procedures in determining the chemical composition of sampled systems. Furthermore, the American Association of Laboratory Accreditation (AALA) has developed a program to accredit environmental laboratories (Locke 1987). To relieve secondary users of the task of evaluating the capabilities of an analytical laboratory and their compliance with the guidance and standards developed by these organizations, the EPA has embarked on a certification program for laboratories involved in drinking water analyses (Hillman 1985).

Reporting Analytical Results--

Documentation of analytical measurements should provide information sufficient to support all claims made for all the results. Documentation requires all information necessary to: (1) trace the sample from the field to the final results; (2) describe the methodology used; (3) describe the confirmatory evidence; (4) support statements about detectability; (5) describe the QA program and demonstrate adherence to it; (6) support confidence statements for the data (Keith et al. 1983). The identification of the analytical method should be made in sufficient detail to allow its specific duplication.

Measurement results should be expressed so that their meaning is not distorted by the reporting process. Reports should make clear which results, if any, have been corrected for blank and recovery measurements. For example, in all cases for which no analyte was detected, the expression "ND" (not detected) and "L" (less than) should be used to report the result (NWQL 1986).

Discussion

Many available manuals, handbooks, and references describe groundwater quality monitoring and sampling techniques (Barcelona et al. 1985, Campbell and Lehr 1973, Dunlap et al. 1977, Scalf et al. 1981, U.S. EPA 1981, Summers and Gherini 1987, and USGS 1977). More general guidance for selecting the appropriate measuring techniques is found in Nelson (1988), who indexes over 700 air, water, and waste measurement methods.

Field techniques, equipment, and instrumentation associated with groundwater quality monitoring are rarely validated to the point that their associated errors can be quantified. Because standards for groundwater monitoring are still in a state of development, the present discussion and guidance for the proper consideration and documentation of these activities are solely intended to summarize the current status of groundwater monitoring activities.

Efficient, well-documented laboratory management procedures are essential to provide quality-assured environmental data. These procedures include communication between the sample monitoring personnel and the laboratory personnel, participation in inter- and intra-laboratory comparison studies, procedures for reporting analysis results, qualifications and assignments of laboratory personnel, validation of analysis results and analytical methods, and incorporation of automated data storage systems. Efforts are being made to establish national accreditation or certification programs for laboratories performing environmental analyses.

In order that secondary data users may assess the quality and accompanying uncertainties associated with analytical data, proper documentation, including sensitivities, detection limits, precision, accuracy, general laboratory procedures, analytical method identification, and other documentation must be available. The ultimate user of environmental measurements must have access to data quality assessments. All major environmental databases must be capable of accepting, storing, and retrieving these assessments with each measurement (EPA 1983).

FLOW INFORMATION AND SOIL AND AQUIFER PARAMETERS

Introduction

The current situation with respect to the QA/QC of hydrogeological field work is comparable to that for field monitoring of water quality and is significantly less well defined than the QA/QC status of laboratory experiments and chemical analysis. Field conditions vary from place to place, and even at one site different situations may occur. Furthermore, technologies to characterize field conditions vary widely in concept, sophistication, complexity, and reliability. For these reasons regulations, standard procedures, and QA/QC requirements are more difficult to establish than for laboratory experiments where all conditions are controllable.

As is the case in acquiring data on the composition of groundwater, in providing data of interest to the research community, all stages of field work and all related activities should be well documented. For the data to be of value for secondary uses (e.g., model validation) documentation on geohydrological characterization should address such issues as methodology, equipment used, and field conditions encountered.

Regarding the measurements, information should be present on parameters measured, the type of measurements and measurement methodology, field-procedures, applicable standards and adherence to them, the use of reference and replicate measurements, and noise-sources and noise level. Among other areas, information should be present on frequency and total number of measurements, location of measurements (in x,y,z coordinates), date and time of measurements, and statistical evaluation.

The equipment used for both the measurements and data recording should be described in terms of equipment type and specifications, their calibration, maintenance records present, and inter-comparison of equipment. Of high interest is information on the precision and accuracy of the measuring devices.

Furthermore, documentation should address the field bias present, methodology and instrument detection limits, and data confidence limits and resolution. Often, secondary users are interested also in the qualifications of personnel involved.

This general listing of documentation elements is valid for various types of measurement techniques. Specific kinds of measurements may require additional documentation.

Water Level Measurement

Water level measurements may be made to determine the phreatic surface in an area, the piezometric head in a confined or semi-confined aquifer, for analyzing aquifer tests, or to determine drawdowns in a well. In general, these measurements are performed in a well either developed as a production or injection well or as an observation well (Powers 1981, Driscoll 1986). The technique of measurement will generally depend on the well type, aquifer conditions, depth of water table or hydraulic head present, frequency of measurements, and other relevant field and logistic conditions. Water level measurements should be taken with great care because hydraulic gradients, critical for correctly estimating groundwater flow rates and direction, normally are determined from measured level differences between wells, which often are very small (Bryden et al. 1986).

Important considerations in evaluating water level measurements are the density of the observation network and the observation frequency. Statistical analyses, including geostatistics and time-series analyses, are used increasingly in order to optimize groundwater level measurement frequency and network density. Other factors that influence the effectiveness of an observation network are the vertical piezometric head gradients (and thus vertical flow) present in the aquifer, seasonal water level fluctuations, trends in water level changes, unknown screen locations, unknown or excessively long screened intervals, and density differences of the groundwater due to varying salinity. Additional considerations that need to be taken into account include well construction integrity and extraneous geohydrological influences (e.g., pressure changes due to nearby pumping, surface water level variations, surface water control, drainage, and recharge variations).

Important aspects of water level measurements are the selection of measurement sites, performing well integrity checks, types of production well access (if applicable), definition and selection of measuring points and reference points, recording devices, and determination of the optimal measurement frequency (USGS 1977).

Direct Measurement of Groundwater Flow

Various instruments are used to measure groundwater flow direction and velocity inside an existing well or a specially designed well-point (USGS 1977, Kerfoot 1988). Examples of direct flow meter measurements are presented by Kerfoot (1984). Kerfoot and Massard (1985) discuss sensitivity of flow meter measurements for various kinds of wells and differences in construction quality. Melville et al. (1985) have evaluated a heat-pulsing flow meter in a controlled laboratory environment.

Vadose Zone Measurements

Measurements in the vadose zone involve water affected by different forces including gravity water, capillary water (matric), and hygroscopic water (governed by molecular attraction and removed by heating). The measurement of soil water may involve the volume of water contained, the movement of water, and the energy with which water is held in the soil. The soil water content is inversely related to matric potential. The rate of water movement for saturated and unsaturated conditions is determined by the head gradient and the permeability. The volumetric soil water content is given as a ratio (between water-occupied pore volume and total volume). Soil water retention is often described using moisture characteristic curves of water content versus matric potential. Other soil parameters of importance for determining soil water transport include soil salinity and temperature. A recent EPRI publication (EPRI 1987) discusses field testing of different vadose zone measurement systems, including tensiometers, resistance cells (yielded moisture content), and neutron probes. Information on measuring soil-water content, soil-water potential, soil-water retention, soil-water movement can be found in USGS (1977), Everett et al. (1984a), and ASA and SSSA (1986), among others.

The unsaturated hydraulic conductivity may be determined in the laboratory by means of steady flow in soil columns (Nimmo et al. 1987) and in the field by the instantaneous profile method, unit gradient methods, disc permeameter, and the air entry permeameter. Soil inhomogeneities might substantially influence the results. Among others, Olson and Daniel (1981) discuss the measurement of the hydraulic conductivity K in the laboratory and field (in situ) both for fine and coarse grain materials, and discuss sources of error in laboratory tests using saturated soils. Fang and Evans (1988) describe laboratory permeability tests and routine soil analysis tests using ASTM standard techniques. Additional discussions of permeability and hydraulic conductivity in saturated and unsaturated rocks and soils can be found in USGS 1977, Zimmie and Riggs (1981) and Kashef (1986).

Because many soil properties are established in the laboratory from field samples, documentation of soil water measurements should specifically address the size of test specimens compared to scale of the main soil features such as soil stratification, size and spacing of cracks and fissures, and other features of secondary porosity.

Geomechanical Measurements of Subsurface Materials

Geomechanical measurements are mainly made for evaluation of soils and rock properties at construction sites. However, many of these properties are identical to those needed in hydrogeologic evaluations, e.g., porosity, bulk density, particle size, and specific surface area. Other geomechanical properties which might be of concern in model validation are clay mineralogical properties and saturation, compaction, total organic carbon, and specific gravity of solids. Determination of various parameters is described by USGS (1977) and Mills et al. (1985). The ASTM Annual Book of Standards series (e.g., ASTM 1988) presents the latest developments in geomechanical testing procedures; standard procedures are described for soil sampling and field investigation as well as for determining soil texture, plasticity, and density characteristics, soil water content, and hydrological and structural properties.

Determination of Aquifer Hydraulic Characteristics

Aquifer testing is often done by stressing the aquifer by pumping a well or a system of wells, while observing the aquifer response by measuring drawdowns in observation points. Currently, aquifer testing is the most widely used method for determining in situ transmissivity, hydraulic conductivity, hydraulic resistance of a semi-pervious layer, and storativity. Other methods for determining hydraulic aquifer properties in situ include slug or bail tests, and various borehole tests (auger hole tests, packer tests, tracer tests, and permeability tests). Measurements conducted as part of an aquifer test include water level, pumping rate, drawdown in well, and barometric pressure. Each time such a measurement is made, the time since the beginning of the aquifer test should also be recorded.

Of great importance for an optimal analysis of aquifer test results are the proper timing of water level measurements, duration of the tests, and number of measurements. Other considerations affecting the results include the location of observation and pumping wells, the design and construction of the wells, and proper recording of the raw data (Kruseman and De Ridder 1970, USGS 1977, Hamill and Bell 1986).

Different factors may affect aquifer test results, such as barometric pressure changes, natural water level changes (in the vicinity of rivers or other surface water bodies, or resulting from variations in precipitation-induced recharge), and abstractions from nearby wells during the test. Other factors to be considered in evaluating the quality of aquifer test data are the possibility of improper well development (resulting in a large skin factor) and variations in the discharge rates caused by pump problems during the test.

The use of automated computerized systems to conduct aquifer tests and collect the data is rapidly increasing. A major benefit of this trend is the increase in data quality, as observer inaccuracies and bias are eliminated and high consistency during the measurements can be obtained (Way et al. 1984).

Many analysis techniques are used to determine hydraulic parameters from aquifer test results (Ferris et al. 1962, Kruseman and De Ridder 1970, USGS 1977, Mills et al. 1985, Kashef 1986, Walton 1987). Both time-drawdown or drawdown-distance measurements may be used in these analyses. Different equations and solution techniques are available, depending on the geohydrological situation. Analyses of aquifer tests is increasingly done by means of computer programs which calculate the parameters that best fit the drawdown measurements (Clarke 1987, Walton 1987).

Although there are many references for these types of tests and many analytical techniques have been developed to determine aquifer parameters, there are currently no standards for performance quality of these tests. Further complications in setting standards are caused by the wide variety of geohydrological conditions encountered in the field.

Less expensive means of determining aquifer permeability include bailer or slug tests (i.e., the removal from or addition of water to a well, respectively). These tests involve measurement of the static water level, the volume of water added or removed, and the number of bailers taken during a given time period (USGS 1977, Freeze and Cherry 1979, Leap 1984). Analytical techniques are used to analyze slug test data to determine the hydraulic conductivity representing the immediate aquifer area surrounding the test well.

Documentation of aquifer tests should include description of aquifer characteristics, including aquifer type, the position of the well screens with respect to aquifer depth, and geometry of the aquifer layering and boundaries.

Tracer Tests

A tracer, injected as a slug into a groundwater body, can provide an efficient way to accurately describe the movement of groundwater and dissolved chemicals subject to such processes as diffusion, dispersion, dilution, adsorption, decay, and other (bio)chemical interactions with the soil or rock. The tracer must be carefully selected to support the purpose of the experiment and not to reflect hydrogeologic conditions and tracer characteristics. In some groundwater systems, such as those in karst terrain, tracer tests may be the only way to determine direction and travel time. Tracer tests can derive valuable information on transport and fate characteristics of pollutants in the subsurface, especially in heterogeneous and anisotropic aquifer systems. The kind of information that may be derived from tracer tests often concerns the layering of the aquifer, relative permeabilities of layers, longitudinal and transverse horizontal and transverse vertical dispersion coefficients, scale dependence of dispersion coefficients, location of the plume, displacement of the plume, velocity of the plume, and the retardation factor. Detailed information on tracer test procedures can be found in Smart and Laidlaw (1977), Davis et al. (1985), Roberts and MacKay (1986), Gaspar (1987), Leap and Kaplan (1988).

In addition to the issues described at the beginning of the section on groundwater levels and groundwater parameters, documentation of tracer tests should include information regarding:

Injection System: volume injected, rate of injection, injection method, construction details

Tracers: composition, concentration, method of mixing, possible interaction with aquifer material and resident groundwater, and special conditions

SAMPLING BOREHOLE CUTTINGS, SOILS, AND CORES

The objective in sampling aquifer materials, soils, and various geologic zones varies widely from project to project. Objectives include determining intrinsic permeability, hydraulic conductivity, porosity, bulk density, or the mineralogical characteristics of the formation; characterizing grain size distribution; identifying the lithology; and conducting geomechanical measurements.

There are many factors to consider in the sampling of soils, aquifer materials, or other solid subsurface materials (Campbell and Lehr 1983). The purpose of the sampling should be clearly stated and the sampling technique and equipment selected should reflect this purpose. The most likely sources of sampling error are sample location, sample size, collection technique, preservation technique, and the number of QC samples.

USGS (1977) presents the requirements for formation sampling, addressing such issues as the relationship between sampling technique and drilling method selected, lithology, drilling rate, and fluid-loss logs, and describing the identification and preservation of samples. Smith (1982) discusses examination of borehole cuttings. Dunlap et al. (1977) provides a discussion of core taking, handling, and preparation of laboratory analysis for organic pollutants and microorganisms. Wilson (1980) discusses solids sampling in the vadose zone, using a variety of techniques and equipment: hand auger samplers, split spoon sampler-barrel type augers, tube type samplers, and engine-driven augers and drilling equipment. Bruner (1986) discusses QA/QC pertinent considerations. Flatman (1986) reports on statistical considerations for design of soil sampling programs.

GEOPHYSICAL SURVEYS

In essence, a geophysical survey is the interpretation of variations in measured response at the earth's surface or in a borehole resulting from certain (geo)physical forces, either naturally occurring or artificially induced within the earth's crust. Such variations might result from differences in physical characteristics within the earth's crust such as density, elasticity, magnetism, radioactivity, and electrical resistivity of the underlying materials.

Geophysical surveys are being used for many different geohydrological purposes such as determining and mapping vertical and horizontal changes in lithology, shallow or deep stratigraphy, delineating contaminant plumes, determining depth to water table, mapping an aquifer, and detecting the location of a fresh-water/salt water interface. Two groups of geophysical methods can be distinguished: surface methods and borehole methods. Different techniques are chosen depending on the geohydrological setting and the goal of the survey (USGS 1977, Campbell and Lehr 1983, Fritterman 1987, and Olhoeft 1988).

The most commonly used surface techniques include resistivity, seismic reflection, low induction conductivity mapping, transient (or time-domain) electromagnetic sounding, and electromagnetic techniques (Zohdy et al. 1974, USGS 1977, Benson et al. 1982, Smith 1982, Violette 1987, van Ee and McMillion 1988, Stierman and Ruedisili 1988, and Benson et al. 1988). White (1988) discusses the use of surface resistivity techniques to determine hydraulic groundwater parameters.

The most common borehole techniques are spontaneous potential, electric resistivity, natural gamma, neutron, and gamma-gamma. Often, additional caliper, acoustic, and flow logs are made (Keys and MacCary 1971, USGS 1977, Campbell and Lehr 1973, Benson et al. 1982).

LABORATORY BENCH STUDIES

Generally, laboratory bench studies are conducted to improve the understanding of the individual processes occurring in the subsurface and their interaction. In the past, such studies have focused on contaminant transport and fate, including hydrodynamic dispersion, adsorption and desorption, ion exchange, and the mechanisms, directions, and rates of chemical reactions. Other processes studied in laboratory experiments have focused on vapor diffusion, liquid diffusion, and volatilization.

Collection methods include chemical analysis, column experiments, and leaching tests. For example, a large volume of lab results has been compiled on the adsorption of organic compounds onto numerous materials; however, very few results are reported on the equilibrium adsorption properties of subsurface formation rocks. Collins and Crocker (1988) describe an experiment using both dynamic flow and static systems. They, among others, designed experimental protocols for application in evaluating mobility, adsorption, and degradation of hazardous organic chemicals in a simulated deep subsurface environment.

To illustrate the variety of laboratory studies performed, the following are some of the studies performed of interest for model validation. Schiegg and McBride (1987) describe a laboratory setup to study two-dimensional multiphase flow in porous media. Bohn et al. (1979) describe adsorption isotherms including the Langmuir and Freundlich equations. Francis et al. (1988) discuss waste leaching tests. Jenkins and Schumacher (1987) discuss extraction solvents for determination of volatile organics in soil. Olsen et al. (1988) describe a lab bench study on the effects of the permanent composition on pore-fluid movement. Crocker and Marchin (1988) conducted a lab study to simulate the adsorption and degradation of enhanced oil recovery chemicals.

The various issues to be addressed in conducting and documenting laboratory bench studies vary among the different types of studies. When studying pH-dependent reactions, attention should be given to temperature fluctuations, order of reactions, ionic strength, sorption present, permeability variations for different fluids, desired precision, and level of confidence. Documentation for research on volatile organic compounds (VOCs) in soils should include recoveries, sorption, extraction material (fluid), purging times, temperature, and decomposition. When the focus is on soil pH and conductivity, issues of interest include the texture variability, water content, salinity, and possible alterations in the chemical equilibria existing between solid matrix and ions.

Documentation of solubility experiments on radioactive elements should cover hydrolysis, carbonate complexation, and redox reactions. Bench studies regarding the injection of hazardous waste often focus on the compatibility issues between the waste and resident water and rock and on factors affecting fate, such as the effects of pH-Eh relationships for the fluids, brine concentration, clay type and clay amount, presence or absence of iron oxide and organic complexing agents, molecular character of organic materials, presence of an anaerobic or aerobic environment, and the plugging potential of precipitates formed.

DATA TRANSFER AND STORAGE

There are many steps within the flow of data from their acquisition in the field or in the laboratory to their use in decision making, the most important of which are the actual data acquisition, site (or laboratory data processing, site (or initial laboratory) data storage, site-to-office transmission (or transport), office data storage and archiving, data processing, and data distribution. It is essential that at each of these steps, the data are managed efficiently. To maintain data integrity attention should be given to the communication and transmission methods and equipment used, computers systems involved, and database systems and other software used. For each stage in this sequence procedures should be established to minimize errors and to obtain a high level of QA/QC in data handling. An in-depth discussion of proper data handling and storage procedures and QA/QC can be found in Sexton et al. (1987).

The effective sharing of data in an increasingly automated environment requires standards and conventions for the identification and transmission of data. The Standard Hydrologic Exchange Format (SHEF) is a standardized system of encoding hydrological data transmissions for both manual and automated processing (Bissell et al. 1983). Features of SHEF include being readable by both man and machine, supporting a wide variety of parameters and data types, having a flexible time identification, using either SI or English units, and allowing flexible use of spaces and comments within the code text to enhance readability.

A more detailed discussion of the QA/QC aspects of data handling is presented in Section 7.

CONCLUSION

This section has focused on quality assurance aspects of the main variables and parameters needed in the validation of the current generation of groundwater models. Many aspects of measurement methods have been touched upon, primarily to indicate information sources regarding current techniques, procedures, and equipment. Many of the issues discussed apply equally to other data that might be required for specific models. For example, QA/QC in meteorological data acquisition programs has been the topic of a number of publications (EPA 1976, USGS 1977, Finkelstein et al. 1983, Lockhart 1985, EPA 1987b).

SECTION 5

STRUCTURE OF THE REFERRAL DATABASE AND REFERRAL FACILITY DESIGN

INTRODUCTION

This chapter describes the operational design of the groundwater research referral database or computerized data directory SATURN and its hierarchical database structure. Also discussed are the SATURN database objectives, database design criteria, the choice of a relational database structure, and the effect of these elements on the overall design approach.

The SATURN database system utilizes current microcomputer technology and relational database management techniques to identify field research datasets suitable for the validation of groundwater and soil water flow and contaminant transport models.

In general, a database is a structure that contains information about specific topics and that includes mechanisms to store data or retrieve it. A database should also contain options to manipulate the data and to display or print it in a useful form.

A referral-type database does not contain the actual data collected; it contains information characterizing the actual data and information regarding its availability.

According to standard database design procedures, a functional analysis and a data analysis was performed (Parker and Parker 1986). The functional analysis is a "top-down" approach to identify the needs for information to be served by the database. The data analysis is a "bottom-up" process to obtain an understanding of the form of the information (or data) to be incorporated in the database.

To achieve the objectives, three steps are taken (Parker and Parker 1986):

- data collection and normalization; gathering and grouping data into small, logical groups resulting in data tables

- entity-relationship mapping; associating entities that are related to each other

- data modeling; designing a logical structure for the data elements

Based on the analysis of the needs for information on field research datasets, five types of potential use have been identified (see Appendix A):

- model validation

- education/training in the use of models

- comparison of field sites
- example/guidance for collection of field data (QA/QC)
- example for the planning of other types of field studies

The first type of use is the primary objective of the database; the other uses are to be considered as spinoffs.

The content of the database is thus a consequence of its primary objective: identifying datasets that can be used in groundwater models. Because many different types of groundwater models exist (e.g., see van der Heijde et al. 1988), and still other types of models will be developed, the data requirements of the models vary widely (e.g., Mercer et al. 1982).

DESIGN CRITERIA

The first step in the development of the referral database facility is identifying applicable standards and formulating design criteria. Five tasks can be distinguished in the development and the operation of IGWMC referral-type databases (see Figure 1):

- information analysis: analyzing objects to be described in the database and update of existing descriptions
- information need or search request analysis: determining the appropriate search strategy and reporting format, and performing the searches.
- system analysis: developing the database structure and programming criteria to meet the design criteria
- programming; developing, updating and testing the computer programs, using either existing database management systems (dbms) software or an independent programming language
- data entry; adding information to the database and updating its content

In the IGWMC information-processing approach, the information analyst functions as an interface between information providers (e.g., data generators) and the databases. Search request analysts intermediate between users (e.g., modelers in search of suitable datasets for their model validation efforts) and the databases, translating the request into searchable questions.

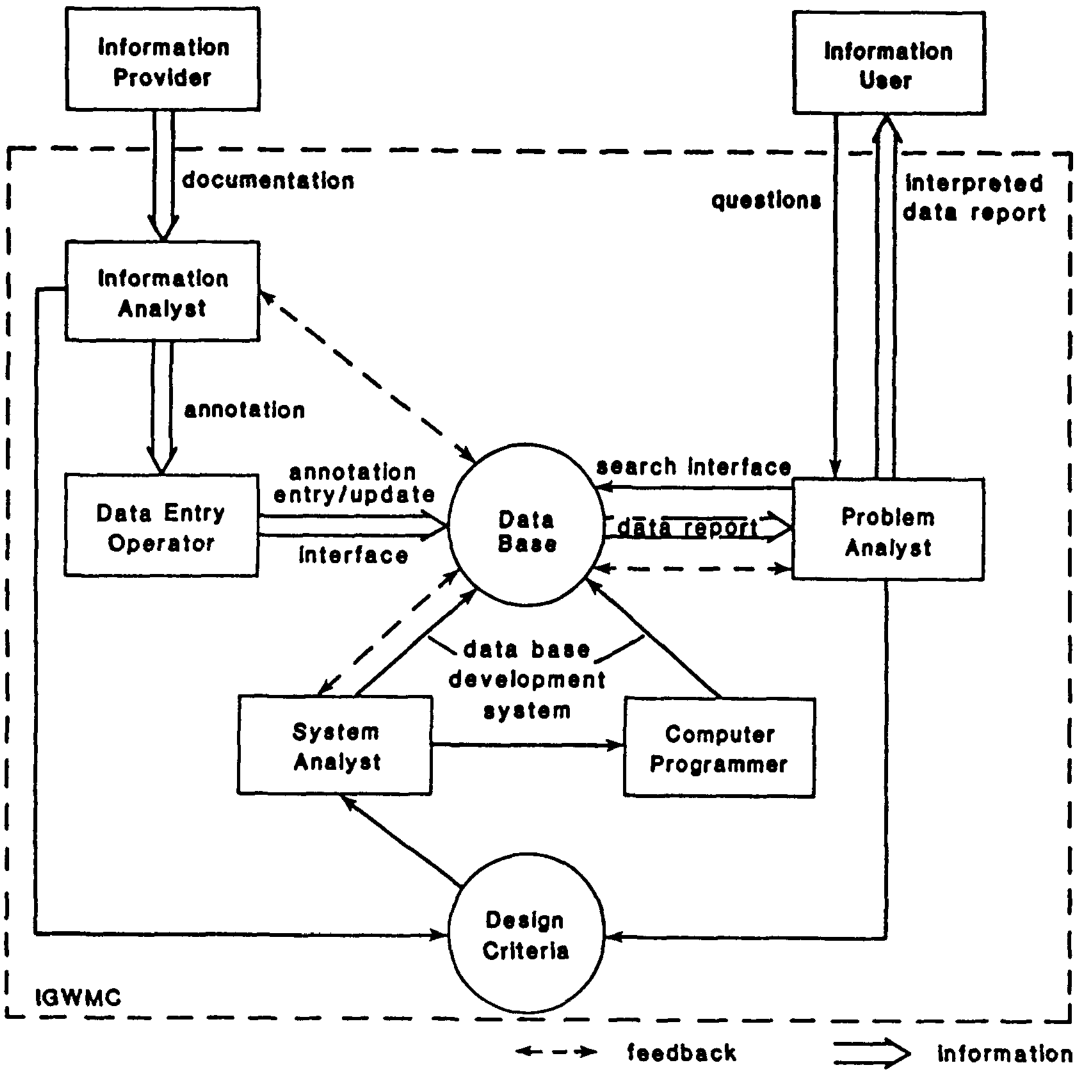

Figure 1. Structure of the IGWMC referral database systems and user interactions

The design philosophy of the referral-type databases is to maximize the performance of these different tasks. More specifically, the databases should include options to facilitate data storage and maintenance, data search and retrieval, and reporting. The databases should be written such that its structure and programs can be modified and updated with minimal effort, even when the original designers and programmers are no longer available for these tasks.

The design criteria are also formulated to assure optimal functioning of the databases for various categories of users. These design criteria will:

- assure completeness of data
- obtain a balance of information stored
- allow intuitive operation
- facilitate optimal user-computer interaction (e.g., effective screen layout, command structure, and command execution)
- permit efficient, useful, and "neat" reporting
- facilitate efficient searches
- facilitate efficient, multi-location updating of database content
- realize efficient internal data storage in terms of computer core memory use and mass storage
- facilitate fast operation (e.g., efficient CPU-mass storage device interaction)
- allow portability within the hardware/software environment in which the database is developed

The role of information transfer in the database design is shown in Figure 1, which depicts both the data stream and feedback features in the setup of the IGWMC referral database systems. Note that the descriptors required for user identification of datasets (searches) may be quite different from those needed in the reporting (see discussion reported in Appendix A).

The Data Center will benefit from involving researchers early in the fine-tuning of its procedures and in promoting of the use of its facilities. Later on, when a referral database is fully operational, user experience and feedback may lead to changes in the descriptor list, database structure, or operational procedures.

Some of the design criteria for the research dataset referral database SATURN are discussed in detail below.

Completeness of Data

The section on identification of potential datasets (Section 2), QA/QC (Section 4), and the analysis of an example dataset (Section 6) provide an extensive analysis of the different kinds of data that might be present in the various research datasets. The discussion on QA/QC provides extensive guidelines for documentation on which the actual choice of the referral database descriptors, though less detailed, is based.

The level of detail in the database should be sufficient to identify datasets of interest to the user. Only descriptors needed for this purpose should be included in the referral database. If the database were too elaborate, a number of problems could arise: (1) few if any personnel present at the Data Center would have an overview of the entire database or would be able to provide sufficient assistance to potential users; (2) manuals would be more difficult to prepare and update; and (3) extracting all the required information from the original dataset and its documentation and entering it in the database would be difficult and very time consuming.

During the 1988 Data Center workshop (Appendix A), where a draft version of the list of database descriptors was discussed with experts from different research organizations, it was noted that the set of descriptors should be selected to cover a reasonably detailed part of the given research areas, but that such a list should not be too elaborate. By avoiding unnecessary complexities and detail and by assuring coverage of a broad area of related research and field investigations, the initial problems in the operation of the database system can be minimized.

Thus, in the initial stage of the project, the database should have a flexible structure rather than an overly extensive or complex set of descriptors. Such flexibility provides an additional advantage: new developments in any of the involved research areas, or changes of interest in specific aspects of research or specific parameters, will be easy to incorporate.

Balance of Information

Although special care should be taken to select a set of database descriptors that is well balanced for the different research areas, the different sections of the database need not contain the same level of detail. On the contrary, the descriptors presently incorporated reflect the experience and views of the project staff and some may prove to be quite subjective. As the main purpose of the database is to identify datasets useful for the validation of particular kind of models, especially transport and fate models, some research areas are covered in more detail than others by more descriptors, and with a higher level of quality assurance/quality control. Again, the flexible, modular database structure adopted facilitates future modifications and enhancements. The final list of database descriptors is presented in Appendix C.

Efficient Searches

The SATURN database contains many descriptors, and though it is possible to use each of these during a search session (using a "query-by-example" option), in practice only a few descriptors will be used frequently. Once it is known which descriptors will be used primarily to identify the datasets of interest, this knowledge can be used to optimize the searches. Search strategies thus identified are implemented in the database system by using so-called pointers, thus accelerating considerably the speed of a search. Search strategies consist of a sequence of hierarchical searches using specific search criteria, based on the most frequently used database descriptors. The various search strategies implemented represent the different search paths users might take to arrive at the information they are seeking. The search strategies selected for implementing in the SATURN database system include the following categories:

1. Site Name or Geography

The user may know through a literature study or from professional and personal contacts that a specific site is well suited for the user's model validation objectives. Such a user might want to select the dataset by site name, city name, or possibly by (approximate) coordinates, and use the referral database to obtain additional information to judge if the dataset indeed fulfills the user's requirements.

2. Hydrogeology/Lithology/Morphology

Physical, chemical, and to a lesser extent biological processes that affect solute transport all depend on the lithology of the porous or fractured medium. For this reason lithology will be a primary search criterion. Related to lithology is the geologic formation. The user may want to select datasets that are collected for a particular formation. However, such a formation might contain different lithologies. Moreover, formation naming often varies from region to region and the user might be unfamiliar with some names. Therefore, such a search will usually be combined with lithology. In addition, the user may be interested in the hydrogeological schematization and related terminology. For soils, a somewhat different approach is taken, based on soil type using a taxonomic description.

3. Type and Source of Pollutant or Tracer

When the transport and fate characteristics of a subsurface system depend strongly on the chemical compounds involved (e.g., type of pollutant or tracer), the search can be performed using the chemical(s) involved as a primary search criterion. Eventually, the user may wish to search for more general compound characteristics such as volatility, radioactivity, sorbtivity, etc..

The characteristics of a contamination source have a significant influence on transport and fate of contaminants. Both the spatial and time-dependent aspects of pollutant transport depend highly on the location, the time-dependent strength, and the composition of the source. Determination of a contaminant source is often a problem in case of an extensive pollution situation, as in a multiple source area, or where the source has been active for several decades.

4. Extent of Database and Datatype

Determining the validity of a subsurface flow or transport model requires an extensive and detailed dataset (van der Heijde et al. 1988). A typical example of field datasets used for this purpose results from three-dimensional transient aquifer tracer tests or from the detailed tracking in time of well-defined contamination plumes. Another example can be found in the one-, two-, or three-dimensional transient tracer experiments in soils. Such experiments and investigations are characterized by the results from the analysis of hundreds or even thousands of water samples. In addition, these experiments require the collection of extensive information on geology, hydrology, and physical and mineralogical properties of the system, and the chemical and microbiological processes present. For these reasons, the user may wish to select a dataset based on such criteria as the number of chemical analyses performed, the number of observation points, or the total monitoring time.

5. Completeness and Quality of the Dataset

In some cases, the user may be interested primarily in well documented, high quality datasets in terms of consistency, completeness, and data quality. A general, single indicator for the QA/QC level of each dataset would identify such datasets.

6. Previous Model Testing and Validation Applications

Some users may want to validate specific parts of their models with data that has been used for other model validation purposes. This enables the modeler to compare his or her modeling results with the results of others, and also an indicates the usefulness of the dataset for such model validation purposes. Even if the model is designed for more complex systems than the dataset represents, partial validation of the model for simplified situations could prove beneficial. Comparison of modeling results for those datasets used for other models might allow the modeler to assess the correctness of numerical approximations, schematizations, assumptions, discretization, boundary conditions, and model parameters (van der Heijde et al. 1988).

Selected Search Strategy Criteria

Based on the above-mentioned considerations, a set of search criteria has been selected from the complete list of dataset descriptors to be included in the adopted search facility. These criteria are listed in Table 11.

TABLE 11. SELECTED SEARCH STRATEGY CRITERIA

Geographic identification
Site name
City/town name nearby
Coordinates (approximately)
Study type
Purpose and scope
Duration
Data type
Pollutant/tracer
Type
Source
Hydrogeology/soil characteristics
Lithology
Soil type
Aquifer/soil material
Extent of database
Types of measurements
Number of measurements/sampling points for main variables in areal plane
Number of measurements/sampling points or zones in vertical direction
Total number of measurements/samples taken in saturated zone
Total number of measurements/samples taken in unsaturated zone
Timespan of measurements/sampling
Frequency of measurements/sampling
Completeness and quality of datasets
QA/QC indicator for data quality
QA/QC indicator for data consistency
QA/QC indicator for data completeness
QA/QC indicator for data documentation
Known model testing applications
Number of reported secondary dataset use projects
Characterization of application

Users should always be able to extract information of interest from the database by using other search strategies based on their own criteria selected from the complete list of database descriptors.

Efficient Storage

In designing the database structure, the amount of disk storage required should be controlled. The SATURN database incorporates several features to assure efficient storage. Detailed information on the datasets is stored in separate tables, independently called in by the program manager. As the amount of detail available may vary considerably for the different datasets, this reduces the disk storage required at any one time.

Another measure taken to reduce disk storage is adoption of a hierarchical structure based on the distinction between sites, studies, and investigations. Different research teams may have conducted various studies at a particular site, each with its own principal investigator or project leader and funding. Also, within a single study (e.g., with a single funding source) more than one type of investigation or task may have taken place (e.g., microbiological characterization, geophysics, aquifer testing). In a hierarchical structure, certain information (e.g., geographic site description) needs to be entered in the system only once, even if more then one study generating its own dataset has taken place at that site.

Efficient Memory Usage

To include a wide range of users, and to make the program more efficient, memory should be used wisely. This has been taken into consideration in the design of the SATURN database. Portions of memory no longer needed are de-allocated. Efficient memory usage has been realized in both code and data. Both, the use of assembly language routines and a design modular structure ensure maximal memory and search speed.

Other Design Considerations

In designing the database, several other considerations have been taken into account:

- Flexibility

The database design ensures that future corrections, modifications, etc. are easy to perform, along with adjustments of the database structure, descriptors, software, manuals, and the reporting formats. This flexibility has been achieved by using a modular design, a relational database structure, and clearly defined relations between the different tables of the database.

- Clarity of Structure

The overall database structure should be clear to all users in order to provide easy handling, easy understanding of the system, and to provide for easy modification.

- Relationship with Other IGWMC Databases

For design efficiency and maintenance reasons the structure of the database for research datasets is to a large extent identical to the IGWMC groundwater model referral system MARS (Model Annotation and Retrieval System). Many subroutines are shared by both systems. The similarity of the two designs eases maintenance of both systems, as the programmer/system analyst has to be familiar with only one type of program structure and a single programming language.

- External Users

To maintain integrity of the database, a version of the database and executable versions of the search and reporting facilities of the database system will be made available on a subscription basis for general distribution. The complete system including source codes will be available for researchers interested in maintaining their own database or in modifying the accompanying software. Measures will be taken to provide regular updates of the database content, as well as support in its use. The modular design of the database structure meets the requirement that only a part of all database system programs will be widely distributed. To assure the integrity of the database system and the quality of the information stored, maintenance of the database content and programs should remain under the control of the (IGWMC) Data Center.

- User-Friendliness

Measures have been taken to make the software as easy to use as possible. For example, "help" is available at any point in the program through the use of the <F1> key. Helpful messages are displayed on the bottom line of the screen throughout the interactive sessions (e.g., data entry session, search, and reporting session). To avoid accidental corruption of data, error checking is automatically performed wherever possible.

Procedures have been designed and programmed to allow for highly flexible updating of the database content while maintaining data integrity and complying with QA/QC requirements. These procedures are designed such that any updating of existing records or adding of new records to the Data Center's master database can be initiated by different operators either at the Data Center or at any other location. These operators work with their own version of the database (not necessarily the latest version). When the update of their satellite database is finished, this database is transferred to the Data Center and checked against the Center's master database. Utility programs are run to identify and isolate new records in the satellite database not yet present in the master database, display their content for evaluation by the Data Center QA staff, and upon acceptance renumber them (to ensure a unique site number) and add them to the master database. Other utility programs are used to check the satellite database against the master database for changes in the content of existing records. Again, records which have been modified in the satellite database are displayed for evaluation by the Center's QA staff and upon their acceptance are integrated in the master database. SATURN's main program includes routines to automatically update a log file to keep track of date, record number, and type of modifications made to the database content.

DATABASE DESIGN

The SATURN referral database system contains information that describes available research datasets. Every dataset is represented by so-called descriptors. Values for these descriptors are stored in tables which are organized in such a way that for each "type" of user and for each activity an optimum configuration is established.

Structure of the Database

A relational database consists of several interrelated tables. The relation between the different tables is expressed by common columns in two or more tables, the so-called keys (Date 1986). Such relational databases have proven to be the most effective and efficient type of databases and as a result, lately almost all databases are relational. Although SATURN has been designed as a relational database, it has some nonstandard features to avoid duplication of data fields. Specifically, the use of common data tables or columns has been replaced by direct pointers between data tables in order to avoid searching for the keys in the separate tables.

In SATURN, three levels of information are distinguished and stored in 13 data tables, using the information fields identified in Section 2 and compiled in Table 12. On the first level, general information about each site is stored (see Figure 2). At the second level, general information is stored about each study that has resulted in a separate dataset. Because more than one study might have been conducted at a given site, the table on study information might contain more records than the table on site information (Figure 2). The third level of the hierarchical structure gives information on the different investigations that might have been performed during the various studies. Level three information is stored in a general task information table and nine separate tables according to the tasks identified. In addition to the tables containing these three levels of information, the database contains two external tables with literature references pertinent to a specific site, study, or task, and references regarding the use of the dataset in model validation, respectively. The relationships between the three main tables and the two external tables are realized through the use of pointers (see arrows in Figure 4).

Figure 3 indicates how information is divided over the three main tables. The complete descriptor list for the database is given in Appendix C.

The SATURN database consists of 14 tables, represented by 14 separate files. The general information table contains a record for each site in the database. The file containing this table is constructed as a binary tree (see Figure 5) with the first record of the file the root of the tree. Each record contains two long integers, one pointing to the root of the left sub-tree, and the other to the root of the right sub-tree. These pointers are the positions of the records in the file. Consider the following tree: In each of the cells, the top number represents the key for that record and the bottom two numbers represent the pointers to the two sub-trees. When the tree is placed in a sequential file structure, the root-record is placed as the very first record or record zero of the file. The file structure for the above tree is shown in Figure 6.

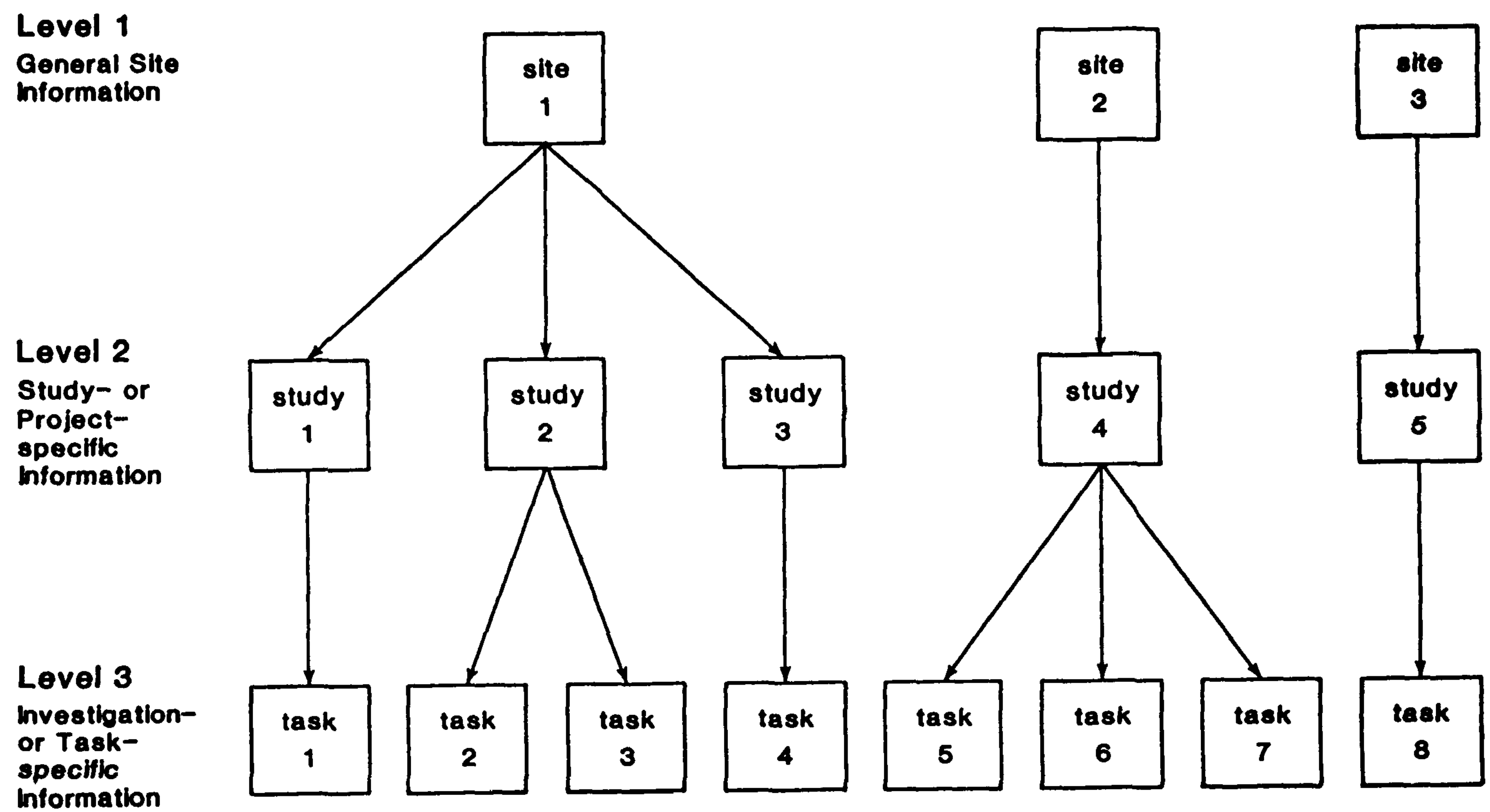

Figure 2. Hierarchical structure of the SATURN database

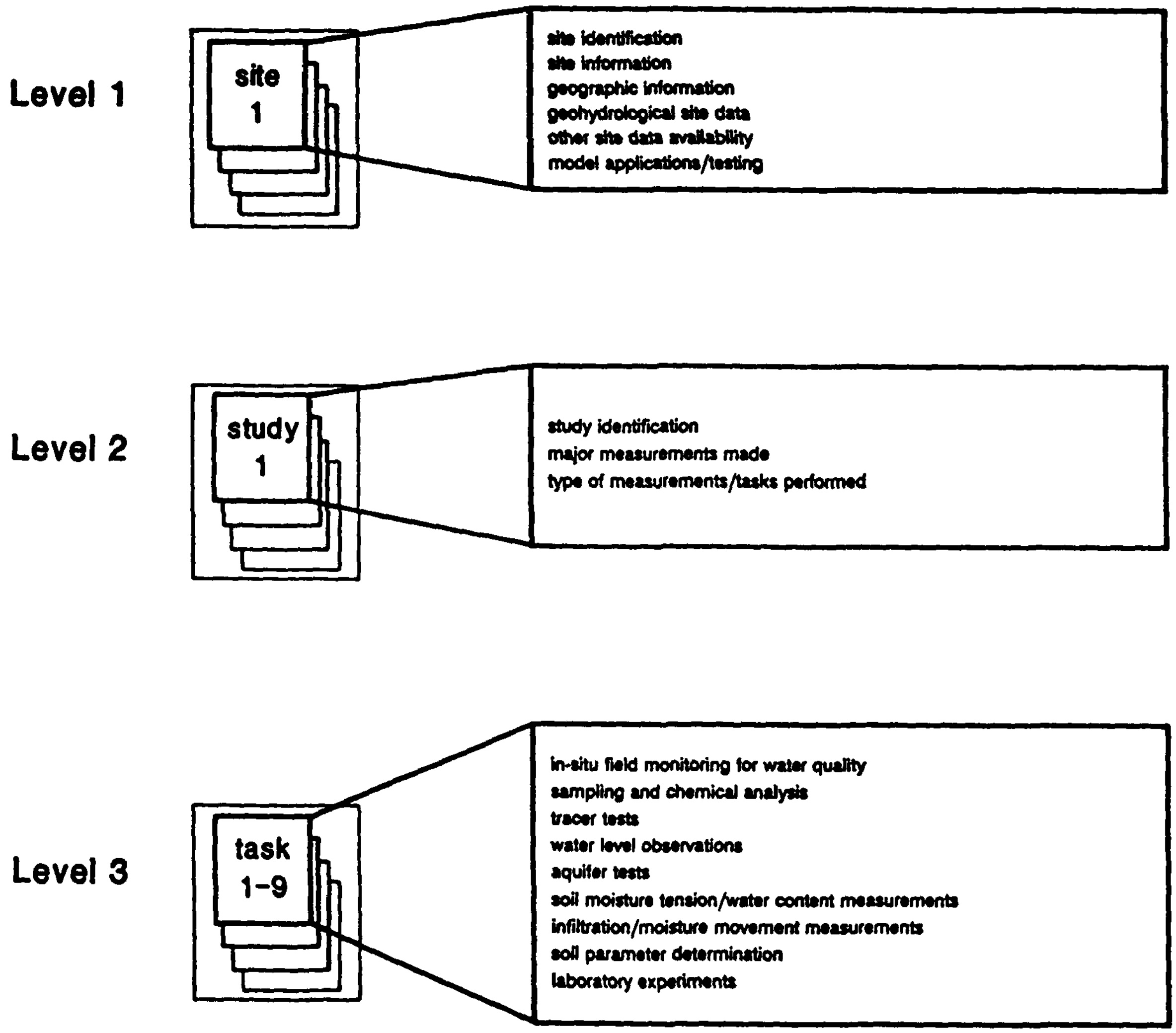

Figure 3. SATURN database program: Information levels and contents.

The other 13 tables are entered sequentially in files. Each new record is added to the end of each. These tables contain the records for each of the following data types: study description, individual investigation descriptions, and relevant publications.

The records are added to these tables in the following manner: A general information record is added for each site. Then, each site may have several study records (or other type of records) associated with it. The general information record contains pointers to its record(s) in the study (sub -) table. The other records are handled in the same manner.

The following list contains the 14 files including the general information task file, and describes their contents.

SITE	SDB -> general information (site information)
STUDY	SDB -> study information
OBS	SDB -> water level observation
MON	SDB -> In-situ monitoring
SAMP	SDB -> sampling and analysis
TRAC	SDB -> Tracer tests
TASK	SDB -> general information on individual tasks
PARM	SDB -> aquifer tests and parameter identification
TENS	SDB -> water content/soil moisture tension measurements
MOV	SDB -> infiltration/moisture movement measurement
MMT	SDB -> measurement of soil parameters
LAB	SDB -> laboratory experiments and measurements
USER	SDB -> user references
SST	SDB -> site/study/task references

Implementation Details

Each of the records relates to a data definition array which describes the fields for its corresponding record. Each element of the array is a record containing the following pieces of information:

'Scr' is the screen number for data entry/update. 'Row' and 'Col' constitute the screen location where the field should be input. 'FType', the field type, can be alpha, numeric, or logical. 'Num' and 'Size' have different meanings for the different field types. For alpha fields, 'Num' is the size of window in which the field will be entered, and 'Size' is the maximum number of characters allowed. For numeric fields, 'Num' is the size of window in which the value will be entered, and 'Size' is the maximum number of digits allowed. For logical fields, 'Num' indicates the bit number and 'Size' is always equal to one.

The alpha fields are represented by arrays of characters. The first two bytes of the array constitute an integer value indicating the number of useful characters in the array. This allows an alpha field to contain

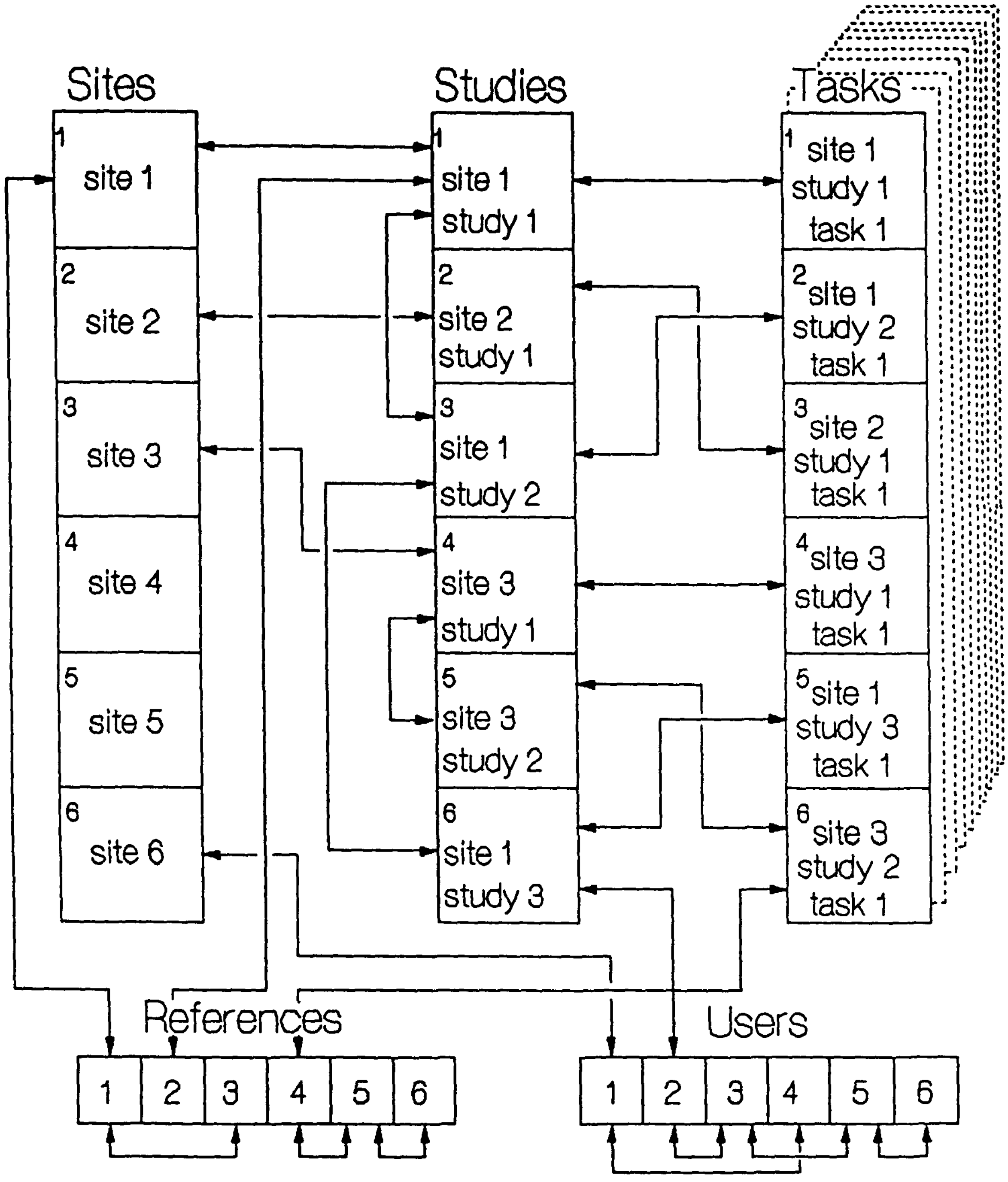

Figure 4. SATURN database program: use of tables and pointers

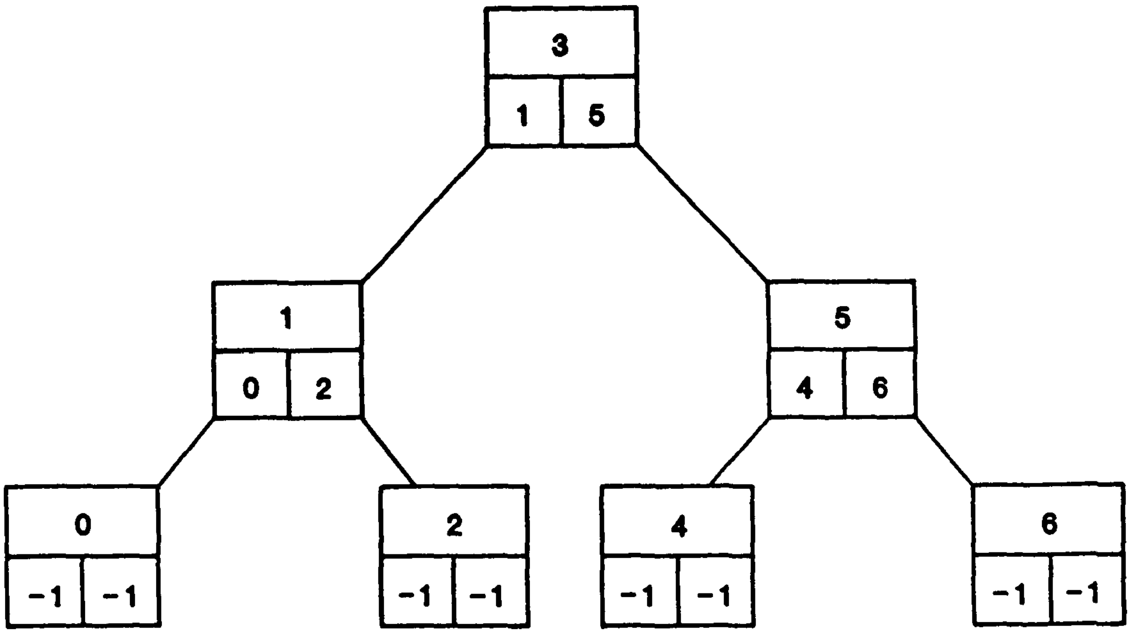

Figure 5. SATURN database program: binary tree structure

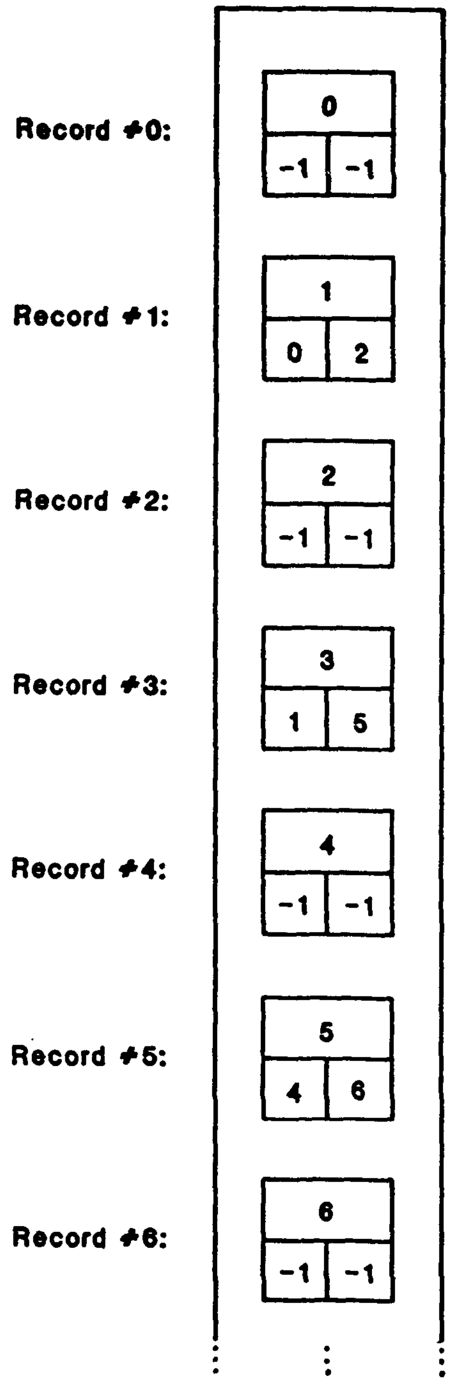

Figure 6. SATURN database program: file structure

up to 65536 characters. The values for alpha fields can be entered using a window smaller than the length of the field. For example, an 80-character field can be input in a 20-character window by scrolling the contents left or right as needed.

The numeric fields are represented by words. A word consists of two bytes, and, the value may range from 0 to 65535.

The following key strokes are allowed for editing alpha and numeric fields during add and change:

<Crsr-Left>	- move left one character
<Crsr-Right>	- move right one character
<Home>	- move to the first character
<End>	- move to the last character
<Ctrl> <Crsr-Left>	- move left one word (alpha only)
<Ctrl> <Crsr-Right>	- move right one word (alpha only)
<Ins>	- toggle insert mode (on/off)
<Del>	- delete character under the cursor
<Backspace>	- delete character to the left of the cursor

The logical fields are represented by bits. Each record contains an array of words. Each word has sixteen bits and therefore can accommodate up to sixteen logical fields. A bit value of one indicates true, and zero indicates false.

The bit number from the data definition record is a number ranging from 0 to 255. Since there are 16 bits per word, the first 16 logical values can be stored in the first element of the array, the next 16 values in the second element, and so on. Therefore, for any given bit number, we can calculate the array element number as the quotient of dividing the number by 16, and the bit number within that word as the remainder. For example, to calculate the bit location of bit number 58, divide 58 by 16.

```
      3
16 / 58
     48
     --
     10
```

Bit 58, therefore, is the 10th bit in the fourth element of the array, marked by "X" in the following example.

```
Bit#:             1111110000000000
                  5432109876543210
                  +----------------+
Element# 0:       |0000000000000000|
                  +----------------+
   --    1:       |0000000000000000|
                  +----------------+
   --    2:       |0000000000000000|
                  +----------------+
   --    3:       |00000X0000000000|
                  +----------------+
   --    4:       |0000000000000000|
                  +----------------+
                  |                |
                  ~                ~
                  ~                ~
                  |                ~
                  +----------------+
```

NOTE: The array elements are numbered starting with zero.

SATURN has been developed mainly with Borland's Turbo Pascal v5.0. The repetitive calculations and time-consuming string handling routines have been developed in assembly language for the 8088/86.

The text fields are represented by character arrays. The first two characters of the array indicate the number of useful characters in the array (i.e., length of the text); this allows text fields to be longer than 256 characters. Some of the routines (for searching, copying, etc.) are written in assembly language. As an example, when searching a database for records that contain a specified word or phrase in one of the text fields, the pattern matching is done by assembly language routines. Other utility routines for string manipulation, such as finding the Nth character of a string, setting all characters of a string to upper-case, etc., are also written in assembly language.

Two functions frequently used, MIN and MAX to find the minimum and maximum of two values, are also developed in assembly language. These functions had to be fast because they are used by the pop-up and pull-down menu routines. Some other routines used to save portions of the screen for displaying menus and restoring the screen are also written in assembly language.

Every time ADD is chosen from the main menu, or Query By Example is chosen from the Inquire menu, several data structures must be initialized. From Pascal, the initialization of large records or arrays can be very slow. For example, to initialize a two-dimensional array, nested loops must be used, as follows:

```
FOR I := 1 TO ROWS DO
      FOR J := 1 TO COLS DO
            Two_Dim_Array[I, J] := 0;
```

An assembly language routine is used instead of such Pascal loops. The routines uses the size of the whole structure in bytes and the starting address in memory to fill the structure with NULL characters.

Every time a record is added or changed in the database, the files containing the modified tables are duplicated. In an attempt to eliminate the delay from this frequent backup, two assembly language routines are used.

User Interface and Database Management Programs

The main menu contains items that allow the user to add information, change information, query the database, and generate reports on the sites and studies in the database. The menus are structured in the following manner (also presented in Figure 7):

SATURN is operated from the keyboard. Commands are executed by bringing the cursor into the field of choice and hitting the <enter> or <return> key (see Figure 8). The system responds either with the execution of the command or with a new menu for further selection.

The Menu System

All menus in SATURN are either horizontal or vertical. The horizontal menus are displayed with all menu items on the same line (Figure 8a), while the items in a vertical menu are displayed on consecutive lines, starting in the same column (e.g., Figure 8b). When a menu is displayed, the first item is highlighted. To highlight a different item, the cursor keys are used: for horizontal menus the <cursor-left> or <cursor-right> keys, and for vertical menus the <cursor-up> or <cursor-down> keys. In either menu type, the <home> key highlights the first item and the <end> key highlights the last item. Alternately, the first character of the menu item to be selected can be used to highlight that item. Once an item is highlighted, it can be selected by pressing the <enter> key. All menus except the main menu and the inquire sub-menu can be removed from the screen by pressing the <escape> key. Doing so returns the user to the previous step. The <F1> key calls up help on the menu currently active.

The main menu items are: ADD, CHANGE, INQUIRE, REPORT, and EXIT. The following section explains the use of these options.

ADD: Adding a record to any of the 14 tables.

A menu of 5 of these database tables is displayed (Figure 8b).

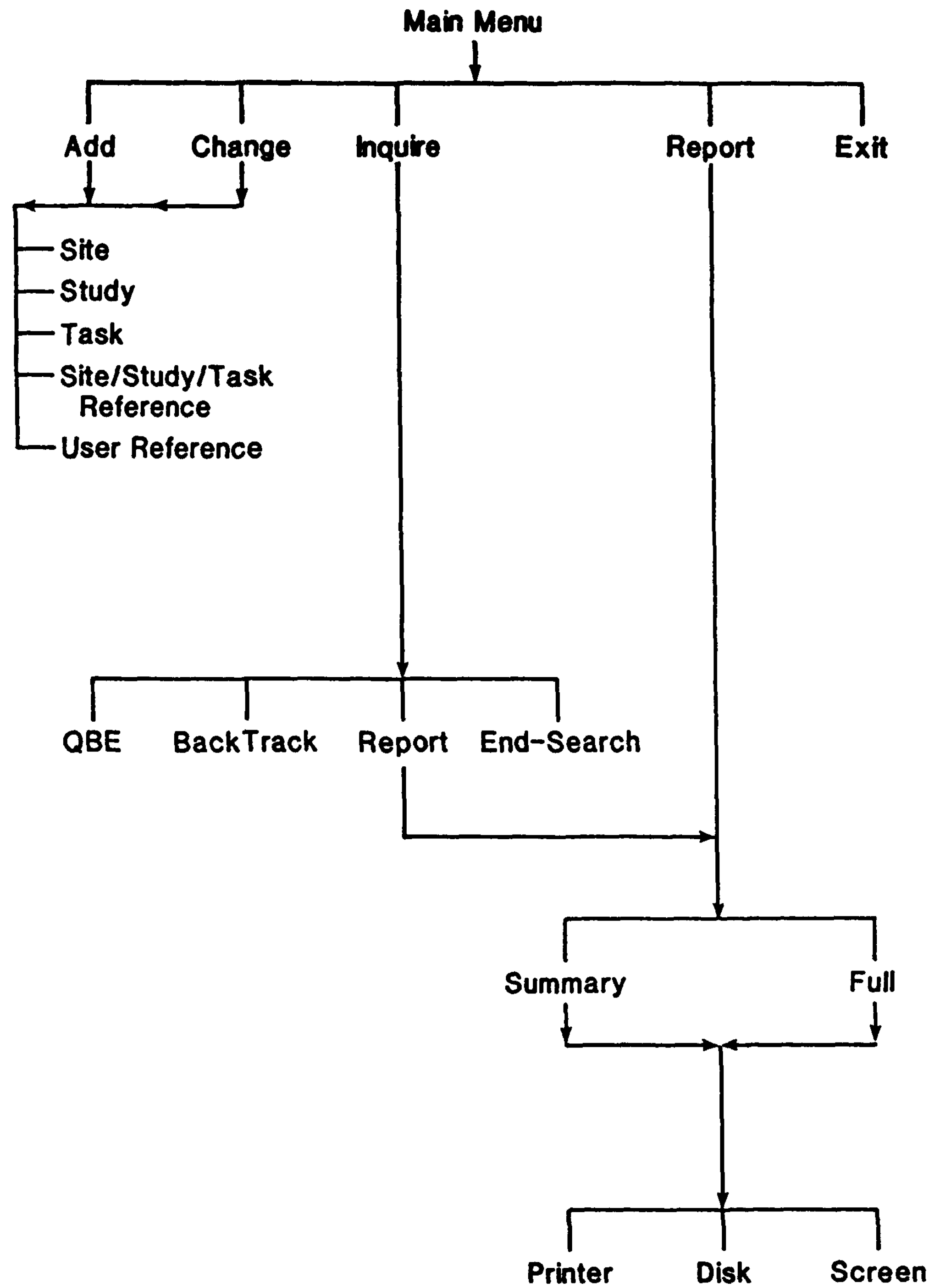

Figure 7. SATURN database program: structure of the user interface.

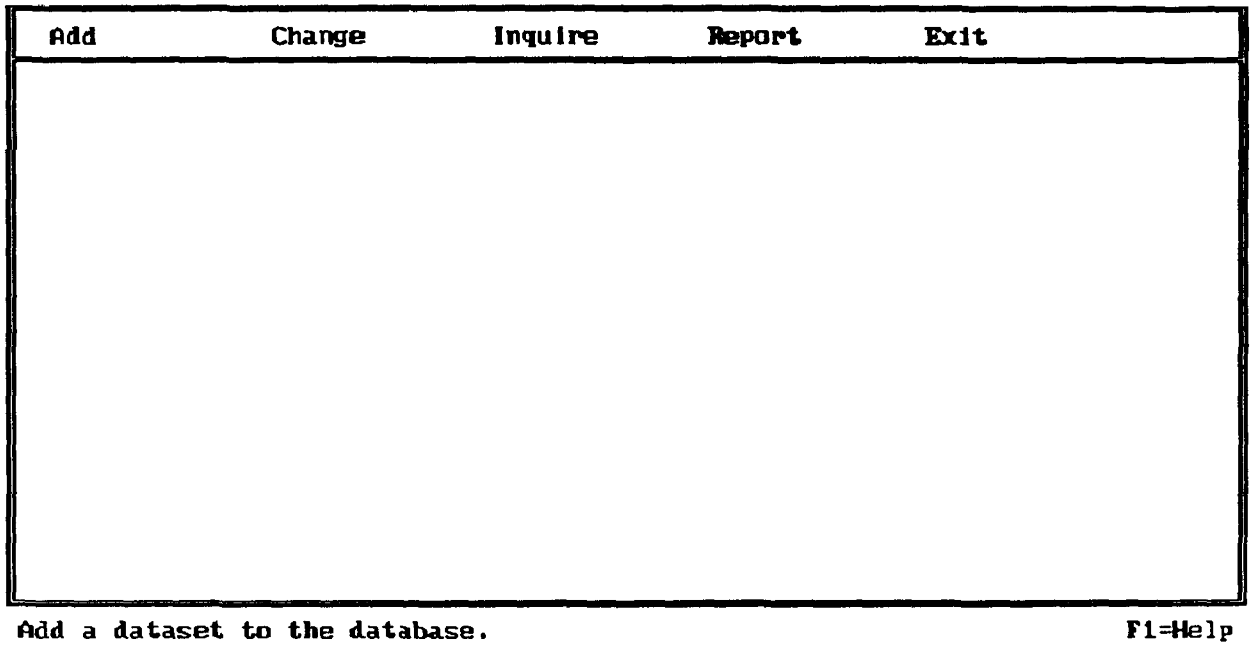

Figure 8a. Main (horizontal) menu

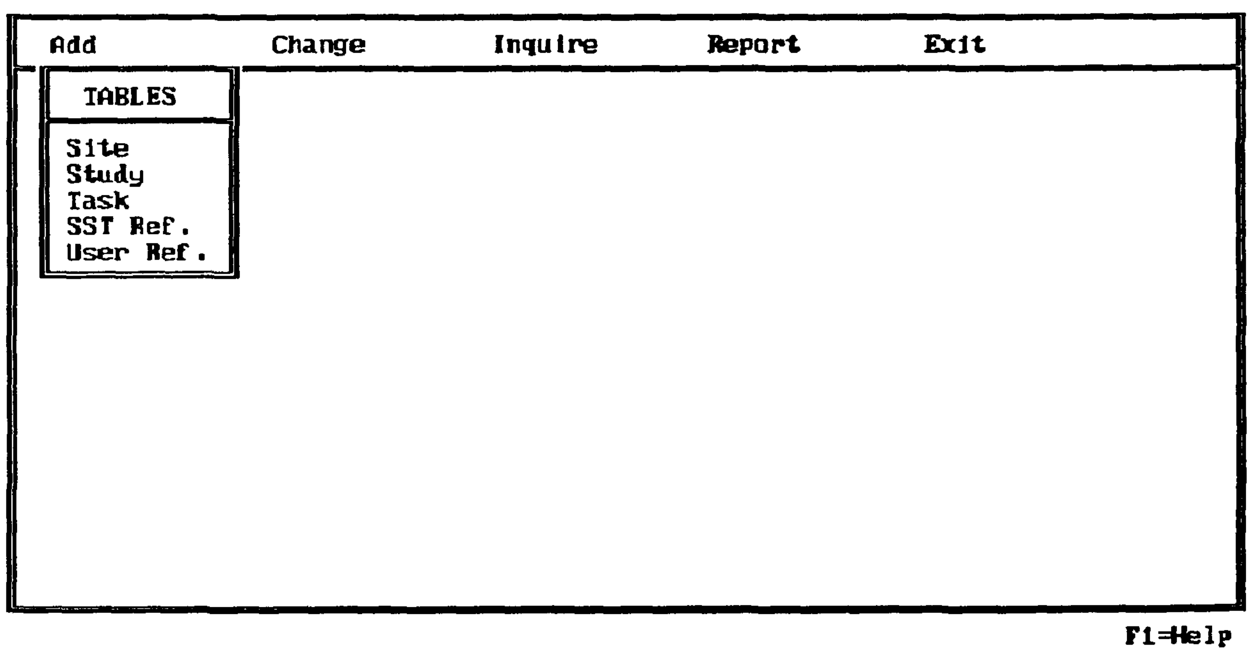

Figure 8b. (Vertical) add option menu

After a table is selected, the first screen of the record layout is displayed (according to Appendix C). Each screen of the layout contains several fields. These fields may be one of three types: alpha, numeric, or logical.

Field editing is allowed for both alpha and numeric fields. Each field is assigned a window smaller or equal in length to the maximum length of the field. While entering information, any of the following keys may be used to correct mistakes:

<F1>	- help!
<crsr-left>	- move left one character
<crsr-right>	- move right one character
<home>	- move to the first character
<end>	- move to the last character
<ctrl><crsr-left>	- move left one word (alpha only)
<ctrl><crsr-right>	- move right one word (alpha only)
<ins>	- toggle insert mode (on/off)
<del>	- delete character under the cursor
<backspace>	- delete character to the left of the cursor

Once the information has been typed, the <enter> key is pressed to make it part of the record.

Any of the following keys can be used to move from one field to the next or from one screen to the next:

<F1>	- help!
<tab>	- move to next field
<shift-tab>	- move to previous field
<page-down>	- display next screen
<page-up>	- display previous screen
<f10>	- save record
<escape>	- do not save record

CHANGE: Changing the information already in the database.

Upon selecting the CHANGE option of the main menu, the same menu of 5 tables is displayed as in the ADD option (Figure 8c). After the user selects a table and the record to be modified, the same layout screens as in ADD are displayed, and the keys defined above are allowed for editing and moving from one field to another and from one screen to another.

INQUIRE: Searching the database for any of the fields in the record and building a search-list.

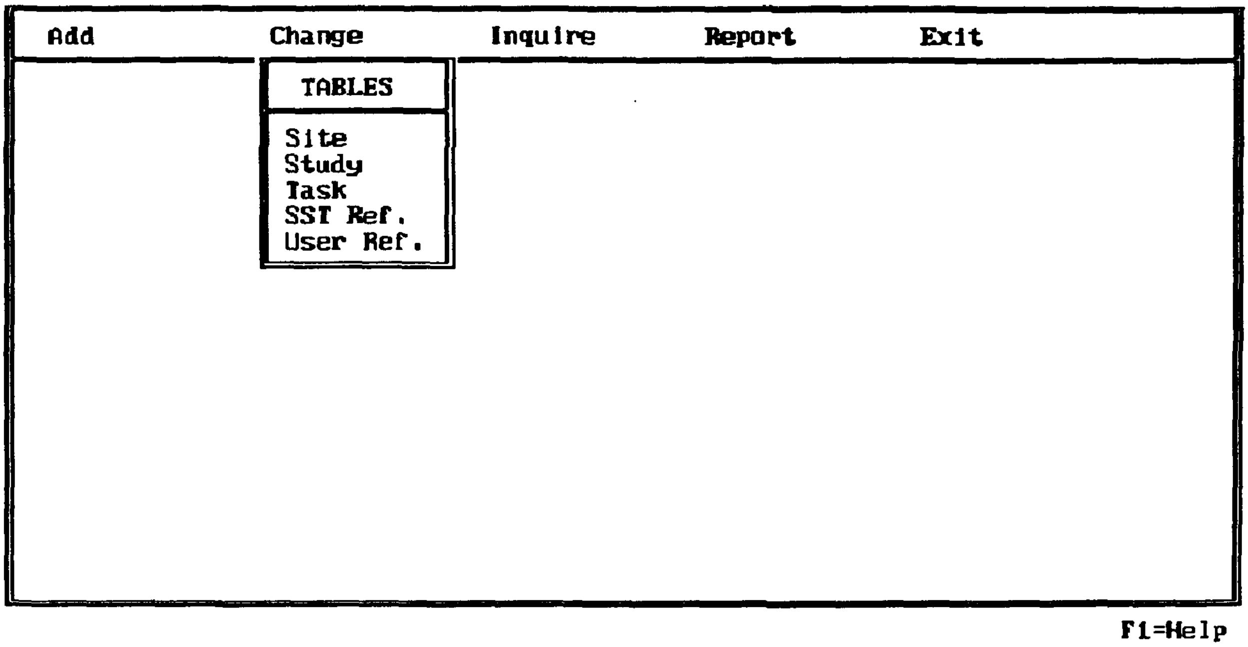

Figure 8c. Change option menu

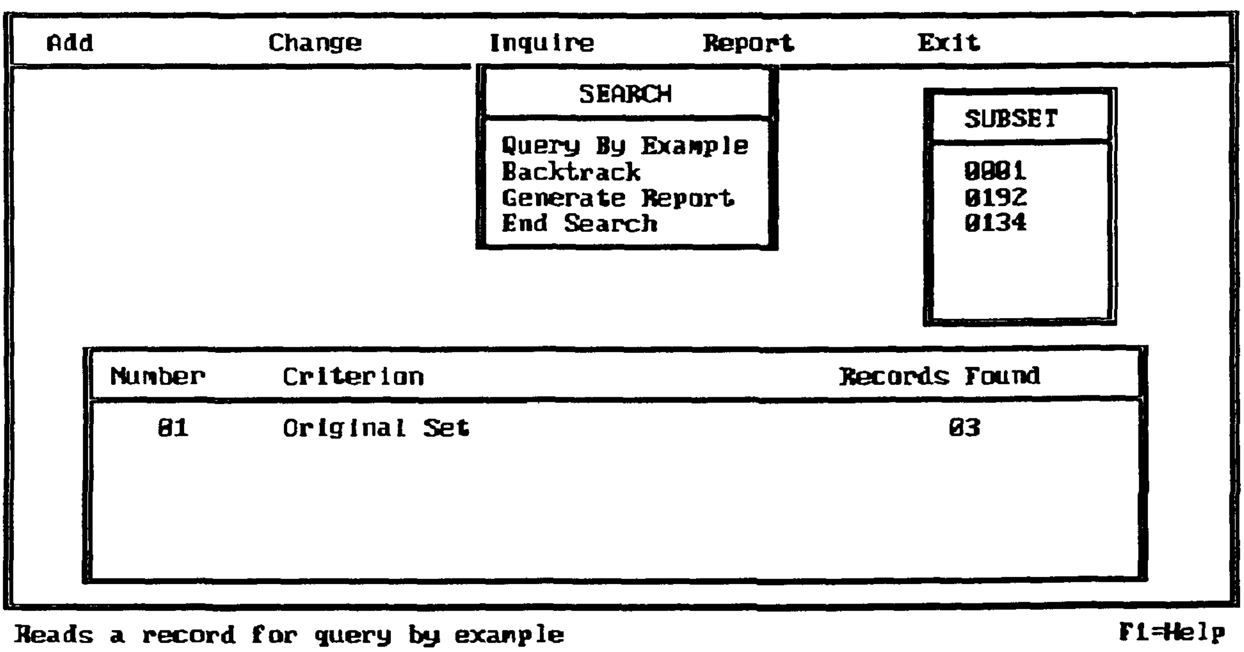

Figure 8d. Inquire (or search) option menu and display of existing dataset

When the user selects INQUIRE, a list of all the keys in the database is built. With each subsequent search, a sub-list is built at the next (sub-)level. This allows the user to backtrack (Figure 8d) one level at a time during the search (e.g., if the last level appears to be empty).

Query-by-example (Figure 8d) is a search method that allows the user to search using any combination of fields with the specified values. When this item is selected from the INQUIRE sub-menu, a record layout, similar to the one in ADD and CHANGE, is displayed (see for details Appendix C). The user can mark the fields to be searched by moving to those fields and pressing the <F3> key. After a field has been marked, the user must specify a value to be searched for.

Alpha fields are searched to see if the specified value is equal to or a sub-string of the field in the record being searched. This allows searching for words and phrases in the alpha fields.

Numeric and logical fields are searched for the specified value.

For example, assume that the database contains five records with keys numbered one through five. The main search list would then be:

```
          +---+---+---+---+---+
Level 0:  | 1 | 2 | 3 | 4 | 5 |
          +---+---+---+---+---+
```

A search on one (or more) of the fields may result in a sub-list containing records two, four, and five. This would produce two levels of search-list in the following manner:

```
          +---+---+---+---+---+
Level 0:  | 1 | 2 | 3 | 4 | 5 |
          +---+---+---+---+---+
Level 1:  | 2 | 4 | 5 |
          +---+---+---+
```

A search at the third level may result in a sub-list with two records, as follows:

```
          +---+---+---+---+---+
Level 0:  | 1 | 2 | 3 | 4 | 5 |
          +---+---+---+---+---+
Level 1:  | 2 | 4 | 5 |
          +---+---+---+
Level 2:  | 2 | 5 |
          +---+---+
```

If a search does not find any records that satisfy the search criteria, the user is forced to backtrack before continuing with the search. Suppose that another search on the example database produces an empty sub-list:

```
               +---+---+---+---+---+
Level 0:       | 1 | 2 | 3 | 4 | 5 |
               +---+---+---+---+---+
Level 1:       | 2 | 4 | 5 |
               +---+---+---+
Level 2:       | 2 | 5 |
               +---+---+
Level 3:       |
               +
```

Backtracking removes the last level from the search-list, resulting in the same search-list as in the previous step:

```
               +---+---+---+---+---+
Level 0:       | 1 | 2 | 3 | 4 | 5 |
               +---+---+---+---+---+
Level 1:       | 2 | 4 | 5 |
               +---+---+---+
Level 2:       | 2 | 5 |
               +---+---+
```

The 'Generate Report' item on the INQUIRE sub-menu (Figure 8d) generates a report for one or all of the records in the last level. A window next to the sub-menu shows a list of the keys for records in the last level. When 'Generate Report' is selected, an arrow is placed next to the first key in the list. The user might press the <F10> key to generate a report for all of the records in the list. To select one record, the arrow should be placed next to its key by using the <cursor-up> and <cursor-down> keys (Figure 8d) and the <enter> key pressed. Once the records to be listed in the report have been selected, there are two more questions—the type of report and its destination. A menu of the type of reports is displayed: 'Summary' or 'Full' (Figure 8e). A list of descriptors used in the summary report is given in Table 12. The last menu contains the three possible destinations of the report--'Printer', 'Disk File', or 'Screen' (Figure 8f). If the user selects 'Disk File', then he/she will be prompted for a file name.

REPORT: Generating a report for a record specified by a predetermined key.

After the user enters the key for the record to be printed (Figure 8g), the same two menus as in 'Generate Report' (above) are displayed, one for the type of report (Figure 8h) and the other for its destination (Figure 8i).

An option exists to obtain for each annotation a hard copy of the database content, using the same format as in the ADD and CHANGE Screens. This option is created specifically for error checking as part of internal IGWMC QA/QC.

EXIT: Closes all database files and returns control back to the calling program (DOS).

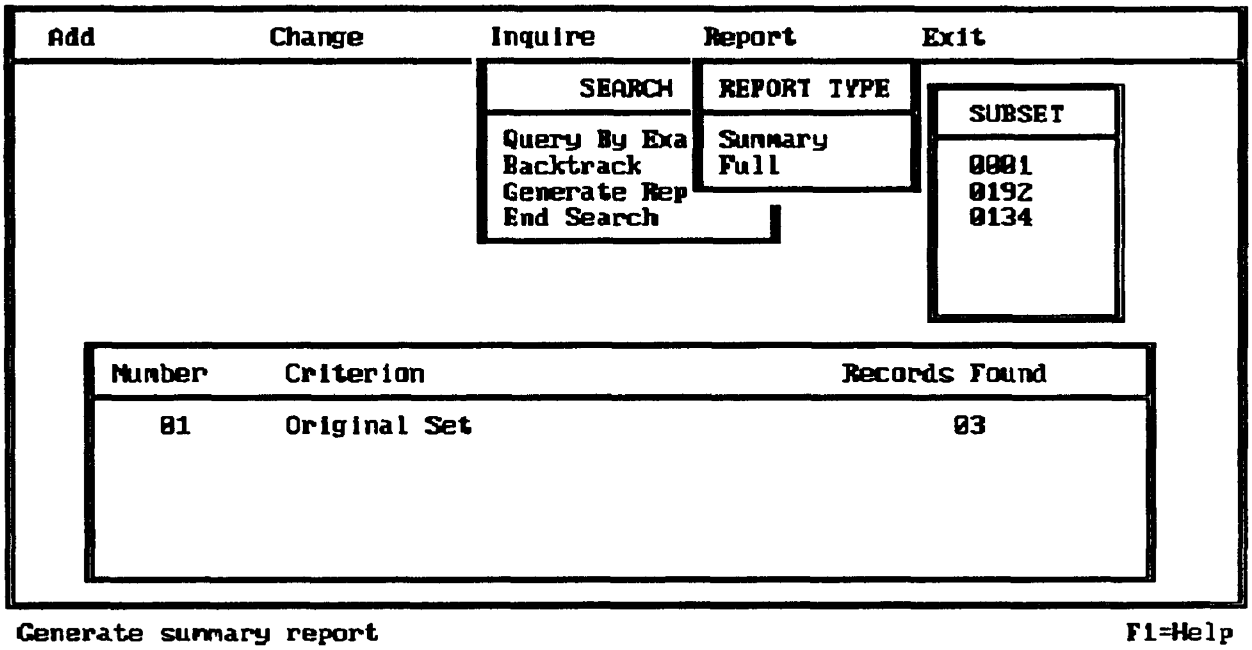

Figure 8e. Report sub-menu of inquire option

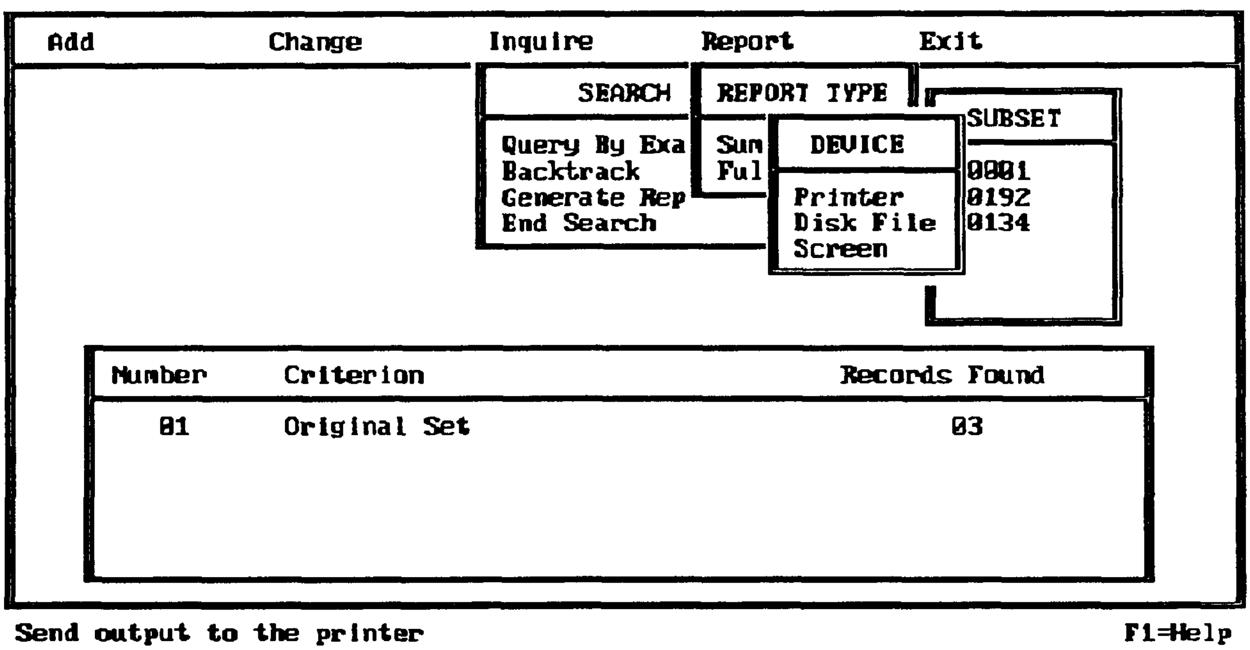

Figure 8f. Display device menu of report sub-menu

TABLE 12. DESCRIPTORS INCLUDED IN SUMMARY PRINT OPTION

Site name and IGWMC key number
Geographic information
Kind of pollutants/source of pollution
Purpose and duration of investigations
Lithology, main geologic formations and soil types
Hydrogeological situation (e.g. schematization, boundaries)
Aquifer and soil characteristics (e.g. porous medium/fractures/macropores)
Parameters measured
Type of data (raw, summarized, derived)
Spatial variability and resolution
Temporal information on dataset (e.g. sampling or monitoring frequency, length and completeness of timeseries)
Technical description of data files (format, amount)
Restrictions on accessibility and secondary use
QA/QC information
Publication (key reference on both the site and the selected study)

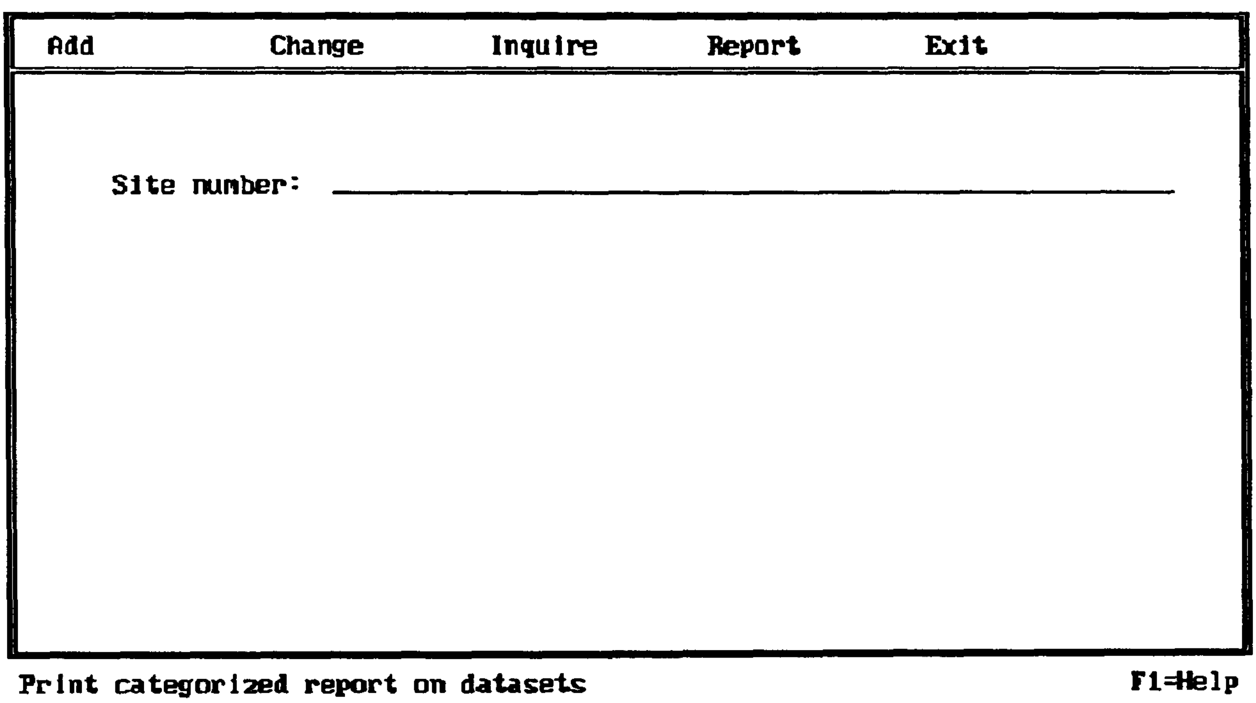

Figure 8g. Report option of main menu requires site information

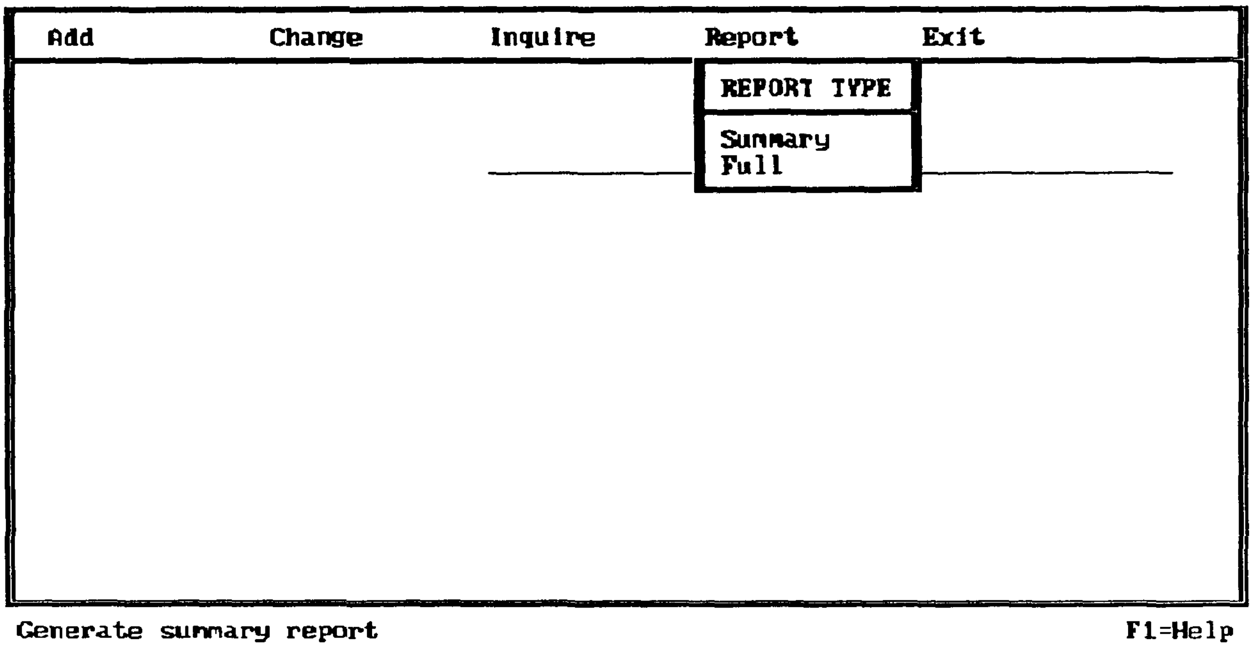

Figure 8h. Report option menu

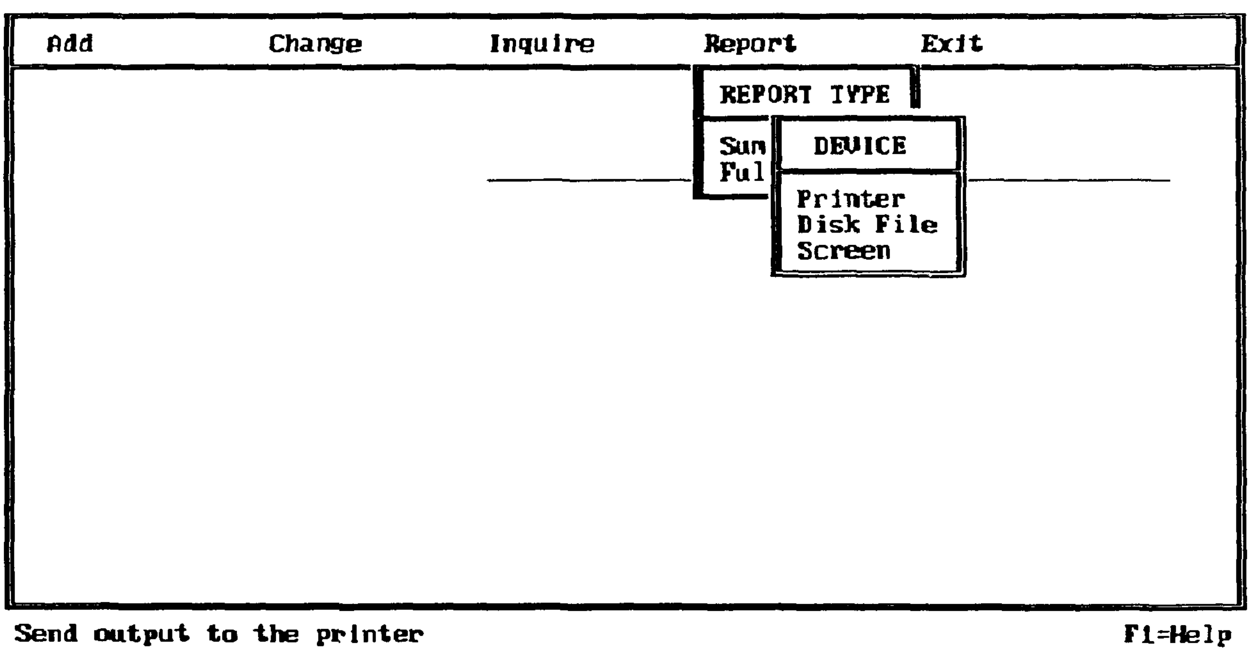

Figure 8i. Display device menu of report option

SECTION 6

DETAILED ANALYSIS OF THE BORDEN DATASET

INTRODUCTION

The Borden site is an abandoned landfill situated on a military base about 80 km northwest of Toronto in the Canadian Province of Ontario. A large zone of shallow contamination in the sandy aquifer underlying the base has provided opportunities for a number of research groups to perform extensive investigations on transport and fate of pollutants in the subsurface. Various studies were aimed at detection and delineation of the extent of certain contaminating chemicals present in the groundwater. Of particular relevance to the present report are the detailed tracer studies of extended duration that have been performed, resulting in various extensive, well-documented datasets (Sudicky et al. 1983, Mackay et al. 1986a,b). Many of the datasets resulting from these studies have been used for testing of groundwater transport and fate models (e.g., Sykes et al. 1982, 1983, Huyakorn et al. 1984, Tompson and Gray 1986, Frind and Hokkanen 1987). One of the natural gradient tracer study datasets from the Borden Site (Mackay et al. 1986a,b) has been selected for analysis in order to develop and test the Data Center's procedures and to evaluate its data storage and retrieval facilities.

GENERAL PROJECT SUMMARY AND OBJECTIVES

A long-term, large-scale field experiment was conducted in the saturated zone of an unconfined sandy aquifer located at the Borden landfill site (Figure 9). The study, a collaboration between the Civil Engineering Department of Stanford University and the School of Earth Sciences of the University of Waterloo, consisted of a natural gradient tracer test, supported in part by the U.S. EPA. The injected tracers included bromide (3870 g), chloride (106.60 g), bromoform (0.384 g), carbon tetrachloride (0.367 g), tetrachloroethylene (0.361 g), 1,2-dichlorobenzene (3.97 g), and hexachloroethane (0.234 g) (Mackay et al. 1986b). The concentration data were collected to further knowledge in the areas of physical, chemical, and microbiological processes controlling transport in the groundwater environment of the site; to test laboratory-scale predictions of field-scale transport; and to develop a database incorporating the effects of chemical interactions, microbiological transformations, and the spatial variability of aquifer parameters (Roberts and Mackay 1986). By design, the organic solutes varied in mobility and potential for biotransformation. Injection took place on August 23, 1982, over a period of 14.75 hrs. The total flow rate equaled 13.5 l/min (9 injection wells), and the total volume of the pulse solution was 11.95 cubic meters. There were 20 synoptic sampling episodes between 8/24/82 and 6/26/85; the number of samples analyzed for each episode ranged from 233 to 1883. A time-series monitoring program included 12 sampling points. Approximately 19,900 samples (solute concentration data) were collected over the 3-yr period. The approximate sampling domain was 100 m (longitudinal), 20 m (transverse horizontal), and 2–4.5 m (vertical). The elapsed travel time was 1038 days, with the maximum travel distance exceeding 110 m.

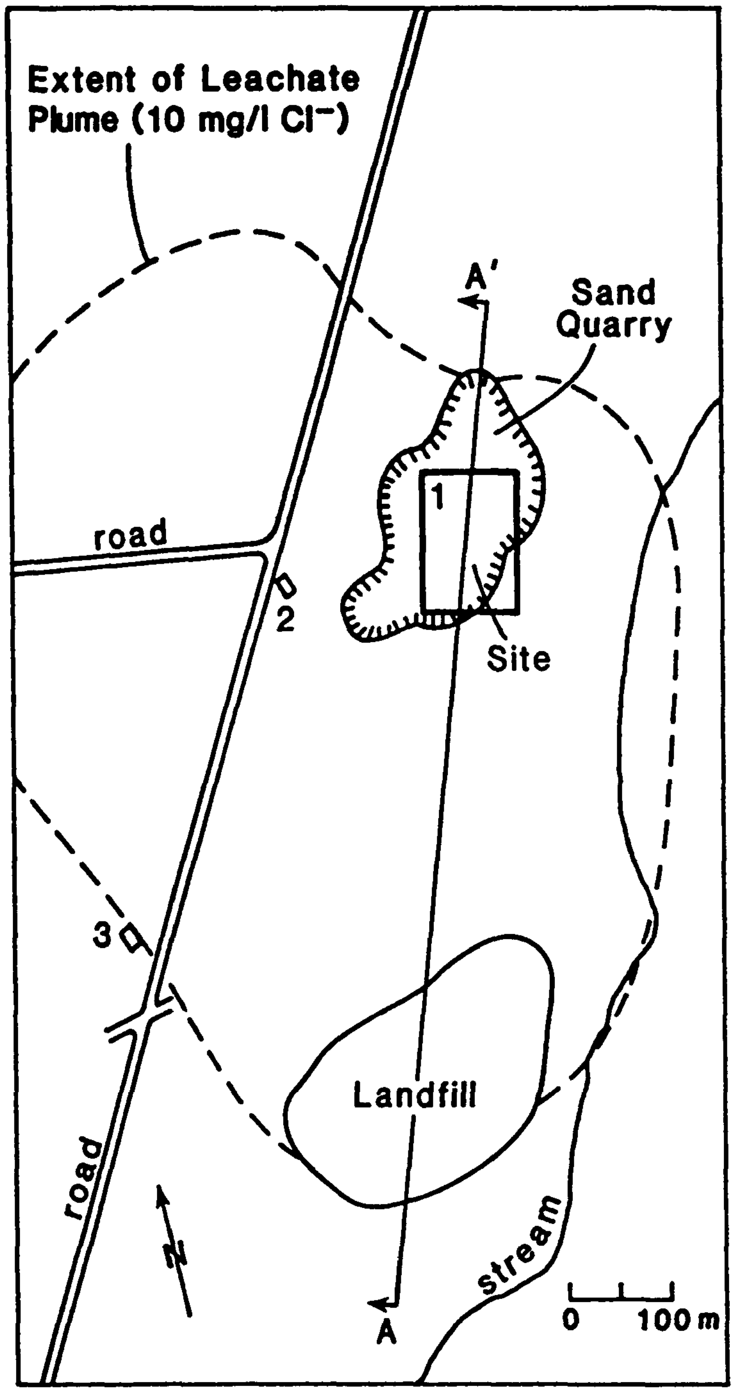

Figure 9. The Borden landfill site and the location of the natural gradient tracer experiment. The rectangle within the sand quarry illustrates the location of the transport experiment and matches the frame of Figure 11. Also shown is the approximate extent of contamination from the landfill in 1979, as delineated by a 10mg/l chloride isopleth. (After Roberts and Mackay 1986)

AQUIFER DESCRIPTION

The total aquifer volume included in the experiment was approximately 3000 cubic meters. The aquifer material can generally be described as clean, well-sorted fine to medium sand that is fairly homogeneous, with some horizontal layering. The phreatic aquifer is underlain by low-permeability clay layers which contain silts and pebbles. As the top of these clay layers slopes, the saturated thickness of the aquifer decreases from ca. 20 m at the injection point to ca. 10 m at the tracer zone, which is situated ca. 500 m downstream of the injection point (see Figure 10). Core samples revealed the properties of the aquifer: (grain size (0.070–0.69 mm with very low clay content), bulk mineralogy (determined by X-ray diffraction), particle density (2.71g/cm^3), bulk density (1.81g/cm^3), porosity (0.33 - calculated), organic carbon (average = 0.02%), specific surface area (average = 0.8 m^2/g), cation-exchange capacity (0.52 +/- 0.09 μeq/100g). Apparent dispersivity (after 11 m advection) = 0.08 m (longitudinal) and 0.03 m (horizontal transverse) with a vertical dispersion coefficient of 10E-10 m^2/sec.

GROUNDWATER FLOW

The average water table (depth below ground surface) is 1.0 m with a yearly fluctuation of approximately 1.0 m. The horizontal hydraulic gradient ranges from 0.0035 to 0.0054 and a vertical hydraulic gradient was not detected. During the period of the study the flow direction varied approximately 10 degrees (see Figure 11). The hydraulic conductivity was determined from 26 slug tests, 2 falling-head permeameter tests, and grain-size distribution analyses. The overall geometric mean hydraulic conductivity was calculated to be 9.75 x 10E-5 m/s with a standard deviation SD = 0.62.

MONITORING SYSTEM

The jetted casing method was used to install the nine injection wells (3.8 cm ID PVC) and the bundle piezometers (1.3 cm ID PVC centerstocks with 12–18 individual sampling tubes of 3.2mm Teflon or polyethylene). The monitoring network consists of the bundle piezometers/multilevel samplers (assembled in the laboratory) in a 3-D sampling array of 5000 individual sampling points (Figure 12). Two portable sampling manifolds, each containing 14 separate sampling stations, allowed all the vertical points at a given multilevel sampler to be sampled simultaneously.

GROUNDWATER QUALITY

The background water quality characteristics were summarized in these ranges: alkalinity (as $CaCO_3$) = 100–250 mg/l, TDS = 380–500 mg/l, DOC < 0.7 mg/l, DO = 0.85 mg/l, and pH = 7.3–7.9.

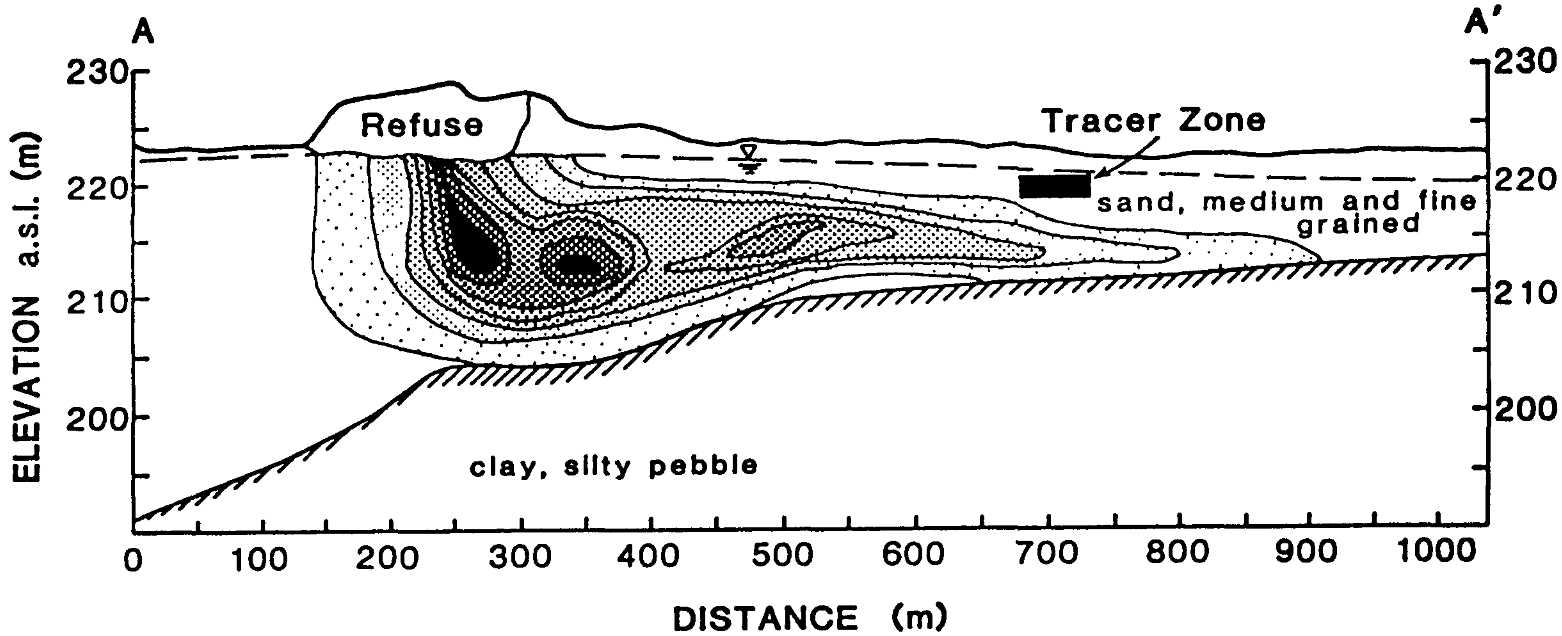

Figure 10. Cross-sectional view of the extent of the chloride plume at the Borden site (after Frind and Hokkanen 1987)

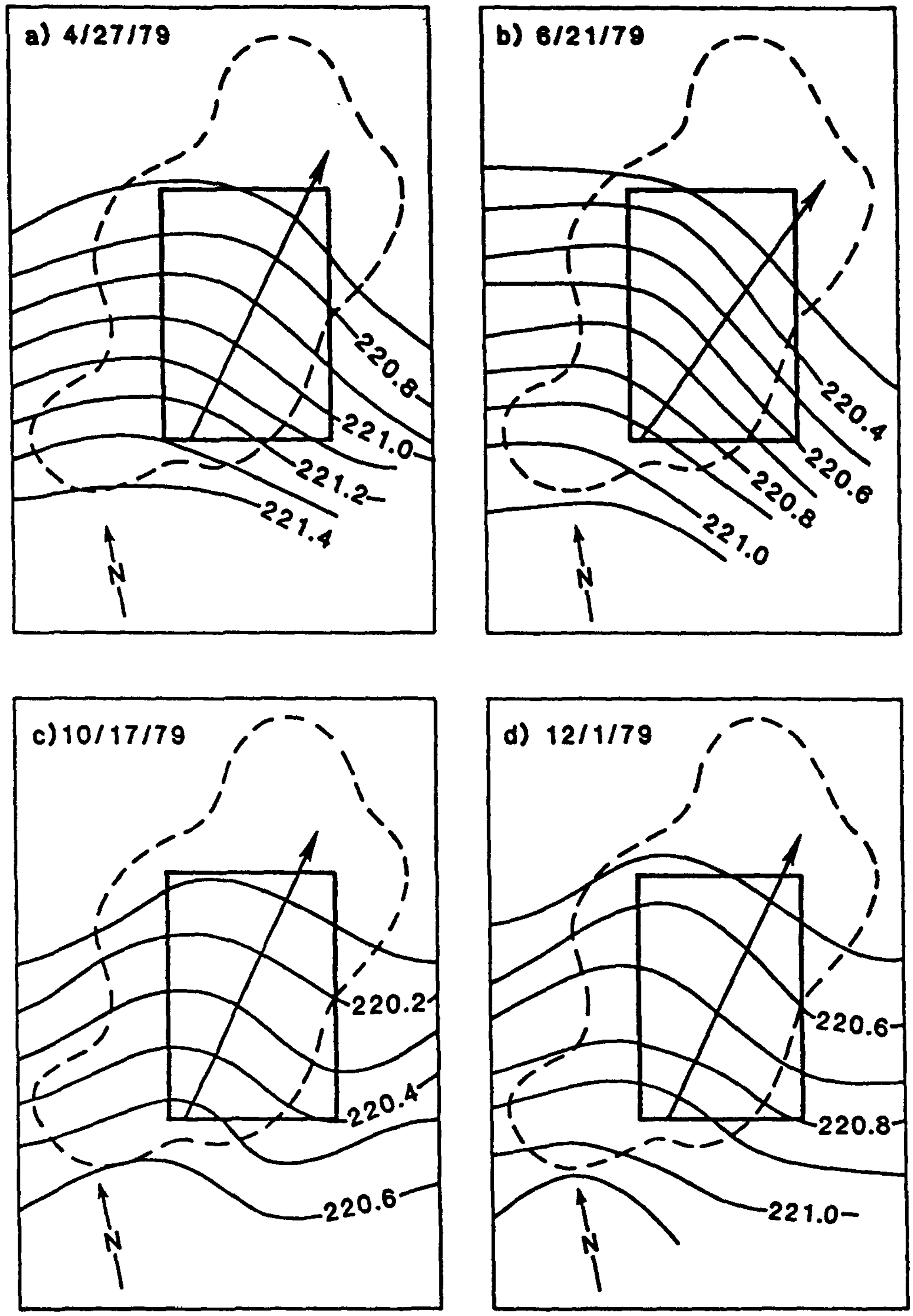

Figure 11. Water table maps for the tracer experimental site and vicinity (after Roberts and Mackay 1986)

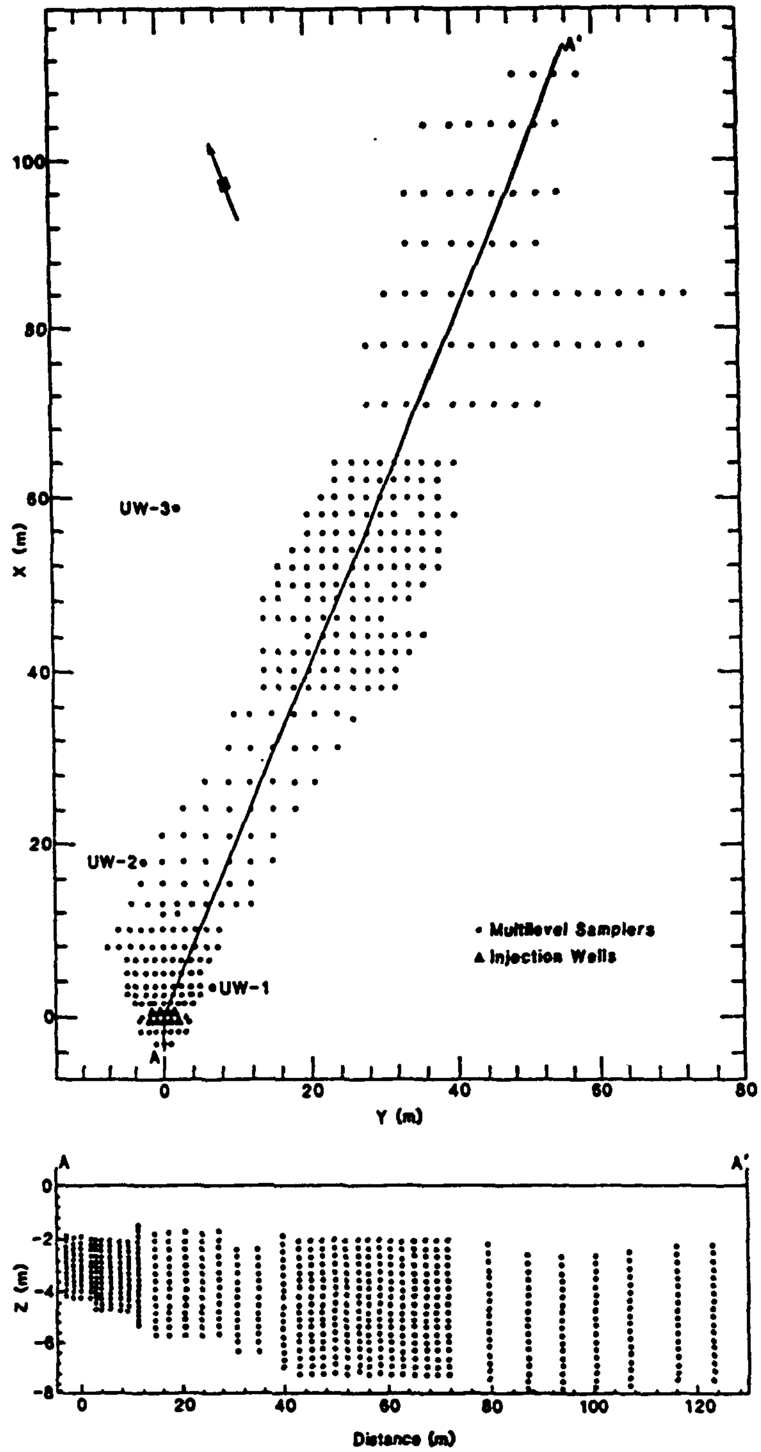

Figure 12. Location of multilevel samplers and injection wells as of January 1986: plan view (top); vertical distribution of sampling points (bottom) projected onto cross-section AA' (vertical exaggeration = 4.6) (after Roberts and Mackay 1986)

ANALYSIS

The time between collection (including filtering and storing in containers) and analysis of the water quality samples (including transportation) was approximately three weeks. Total analyses time for a single sampling period (up to 1800 samples) was less than two weeks. Samples were trucked from the site to the University of Waterloo and air-freighted to Stanford University for analysis. Inorganic samples were analyzed with an automated ion chromatograph and organic samples were analyzed with two gas chromatographs.

QUALITY ASSURANCE

Duplicate samples were collected from one sampling point at each multilevel device (bundle piezometer). A prototype sampling station was tested in the laboratory and field tests were performed to determine the vertical spacing integrity and appropriate volume of sample and purging requirements. The injection system, field sampling equipment and protocols, and analytical methodologies for the project were documented in detail. Monitoring and sampling materials, containers, and techniques were selected with consideration to sample representativeness, especially with respect to the volatile organic tracers. The criteria for the injection concentration levels are as follows: levels were to exceed 100 times the background or analytical detection limit, whichever was greater; the conservative tracer concentrations were to be kept as low as possible to minimize the density contrast between the injected pulse and the native groundwater; and the organic solute concentrations were set to yield roughly equal peak areas in the gas chromatographic analysis in order to achieve similar sensitivity for all compounds. Several groups of data with questionable validity are discussed in the documentation (Roberts and Mackay 1986).

MOVEMENT OF THE PLUME

To derive measures of mass, mean velocity, and dispersion, the zeroth-, first-, and second-order spatial moments were defined over the volume of the plume (Freyberg 1986). The zeroth-order spatial moment measures the mass of the respective solutes. The first-order spatial moments measure the location, movement, and velocity of the center of mass of the solute plume. The vertical displacement of the plume is small. The vertical component of the mean solute velocity vector is negligible. The second-order spatial moments define the spatial covariance tensor, which measures the spread of a plume about its center of mass. This is a measure for the dispersion of the plume. Evidence was found of what commonly is called "scale-dependent" dispersion. The role of growth of the covariance over time is probably not linear, as would be predicted by the classic advection-dispersion equation with constant effective parameters. Also, the effect of heterogeneities in aquifer characteristics on dispersion could be distinguished.

RETARDATION OF THE PLUME

Due to sorption of the halogenated organic solutes, the different plumes of these organic solutes are retarded in varying degrees (Roberts et al. 1986). The retardation factors increased with elapsed time in a manner that suggests deviation from local equilibrium.

The retardation estimates obtained from time series sampling at particular points agreed well with estimates based on periodic high-resolution spatial sampling.

SORPTION OF ORGANIC SOLUTES

Batch laboratory experiments with core samples have been used to determine sorption-based retardation factors of the halogenated organic solutes (Curtis et al. 1986).

The isotherms measured were all essentially linear throughout the concentration ranges studied and appeared nearly reversible. The retardation factors inferred from the observed distribution coefficients were generally in good agreement with observed temporal and spatial field data.

DATASET CONTENTS

The dataset of the Stanford/Waterloo natural gradient tracer experiment at the Borden site is currently available from the International Ground Water Modeling Center (IGWMC) on an "as-is" basis. The tape being distributed contains:

- sampling dates
- x,y,z coordinates of sample location
- concentrations of the seven constituents of the injected tracer solution.

An extensive abstract and additional information on QA/QC-procedures, data formats, field conditions, and bibliographic references, etc., are available from the IGWMC's SATURN database (see Section 5). The documentation consists of a report from the data collection team (Roberts and Mackay 1986). This report includes an appendix on the experimental database tape documentation.

SCREENING RESULTS

After analyzing the extensive literature documenting the experiment and the resulting dataset, the Data Center's prototype dataset annotation form was updated and finalized (see Appendix C). This form functioned both as a checklist for the Center's dataset evaluation procedure and to evaluate the test dataset.

The referral database descriptors do not address the variability of parameters in detail. The dataset description allows only for describing whether or not a parameter was measured, and whether or not information is available on procedures, equipment, and methods used, and on documentation. This means that for most measurements no indication exists regarding:

- number of measurements
- spatial variability of parameters
- time dependence of parameters
- standard deviation of parameters
- actual parameter values

The "Remarks" field was used to indicate that information on this issue was available.

For large, detailed datasets, the SATURN dataset description format (dataset annotation) might prove insufficient to fully cover the extent of the available data. In that case, the modular, flexible program structure of SATURN allows for easy modification and updating of the data files by incorporating additional fields in the dataset record. Future use of the database should make clear whether or not additional descriptive information of datasets should be incorporated.

DISCUSSION

Analyzing the Stanford/Waterloo natural gradient tracer experiment provided important feedback regarding the referral database descriptors. The main corrections concerned the level of detail in the descriptor list that was necessary to concisely describe the available data and the manner in which to allow for inclusion of additional information not represented by descriptors.

In the future, updates to the current descriptor list may prove necessary as the interests of the research and modeling community change.

SECTION 7

DATA CENTER PROCEDURES

INTRODUCTION

In order to obtain the best possible operation of the Data Center, procedures have been established to ensure proper (referral) database and dataset management, acquisition and annotation of dataset information, acquisition and evaluation of selected datasets, dissemination of information, and distribution of datasets. These procedures are discussed in the following sections. The QA/QC procedures for these operation are discussed in Section 8.

DATABASE AND DATASET MANAGEMENT

Computer data management procedures should emphasize data integrity and security, whether for referral information or actual datasets. This can be done by developing and enforcing strict data processing procedures that include authorization rules specifying that certain tasks be performed only by a specified group of users. External users, for example, are not allowed to write, modify, or delete data from the database or have direct access to the files containing the actual data of distributable datasets. Another important procedure would require routine data backup, thus allowing recovery of the database or the dataset files in the event that their content is corrupted, destroyed, or lost.

Although database management covers many topics, the present section discusses security, integrity, and recovery.

Database security is the protection of the database or its software against unauthorized or illegal access, either intentional or accidental. Database security can be provided in three basic ways (Kroenke 1977):

- Encryption: storing data in an encrypted format which will only be 'translated' by the system when authorized users work with the database system

- Definition of subschemas or views: referring to the information that a user is allowed to see

- Authorization rules: limiting access privileges to the database, both internal and external, allowing only certain tasks (e.g., database design, programming, adding, updating, and deleting data, performing searches and selecting information for display or reporting, and implementation of an accounting system), using passwords and audit logs

Integrity of the database management system (dbms) is the mechanism that ensures that both the data in the database and changes to the data follow certain rules. For example, transmissivity, permeability and aquifer thickness should not be less than zero. Another integrity-assuring method is the use of spelling checkers for text-field data entry (e.g., bibliographic information). Either the dbms or the accompanying software should enforce integrity constraints. Integrity is also important when data are transferred between computer systems. For example, when tapes containing data files have been through many undocumented manipulations, their integrity might not have been preserved. Data transfer integrity might be preserved by comparing the transferred data with the original data, by returning a copy of the tape to the author for verification; this can be done by computer, using existing software.

Recovery is the mechanism for recovering the database in the event that it is damaged in any way. Generally, two mechanisms assure this: backups and journaling.

In backup procedures, hard copies of both the database and the software are stored at a physical location away from that of the database, as in a fire-resistant storage place in the same building and one somewhere else. If calamity should befall the host computer, or if the data files are corrupted (as through operator error, hardware failure, or a so-called computer virus), the backup version assures that no valuable information will be lost. In order to maintain an up-to-date version of the database in each of these backup locations, they must be replaced at regularly scheduled intervals. Such a schedule is determined by the frequency of updating the database or the data files and the cost of replacing the information lost between backup and time of data loss. Measures need to be in place to check the integrity of the backup immediately after it has been prepared.

For the IGWMC Data Center, internal backups of datasets will be replaced after each new dataset that has been downloaded on the computer system. In addition, the internal backup procedure calls for backup as part of the computer facility's own backup procedure (e.g., every month). Backups will be stored for at least three years as will the backups of every third year prior to this period.

To prevent program alteration, additional measures can be taken, including separation of program source codes from the user-accessible environment, using only executable images of the software. In addition, periodically refreshing the operational software (both database management software and application programs) by reloading from backup might prevent unauthorized program alteration.

As the SATURN referral database will be widely distributed, complete with search and report-generating software, procedures are adopted to prevent data corruption off-site while maintaining the integrity of database content. To protect this content, only compiled versions of the search and report-generating software will be widely distributed, and database users will be provided periodically with updated data files and application programs to replace their previous versions. On request, database users will receive the complete SATURN version, as it is public domain software. In that case, they need to assure themselves of the integrity of the database content. Specific software under development by the Data Center will allow user-made modifications of the database content to be checked against the master database at the Data Center and when accepted to be incorporated in the master database.

Journaling is keeping a log or journal (preferably automatically) of all the activities that update or modify the database. In case of a computer or database system crash, the system manager can always restore the latest correct version of the database. At the IGWMC, journaling is part of QA/QC procedure and is currently being automated.

INFORMATION MANAGEMENT

Information Acquisition and Processing

Procedures have been established for collecting and processing information on field research sites which might provide datasets suitable for model validation. These procedures have been tested using the information collected during the first phase of the Data Center project. The procedures have drawn on the experience of the IGWMC in modeling and information processing.

Identification of Potential Datasets—
The IGWMC Data Center staff continuously collects and analyzes information on datasets they have identified. The initial information may come from open literature or from presentations and discussions at conferences, workshops, and other meetings, or may be obtained directly from researchers.

Once a dataset of interest is located, additional information is collected from the research team that collected the data, and from pertinent literature, to enable the Center's staff to include the dataset in the SATURN database. In selecting a dataset for inclusion in the referral database, special attention is given to the quality of the available data, to whether or not standard procedures were used during data collection, to the extent of the dataset, to the kind of chemicals and processes involved, and to existing restrictions on data transfer and documentation distribution.

To assure consistency in the evaluation of the site and dataset information and in the data entered in the referral database, a standardized form, the SATURN data entry form, has been designed (see Appendix D). A complete dataset annotation includes comments made by the original research team and the IGWMC staff, as well as bibliographic references regarding its collection, interpretation, and primary and secondary use. After detailed evaluation of the field site and the data collected by the Center's staff, a SATURN data entry form is filled out, and the data is entered in SATURN (see Figure 1 in Section 5).

To ensure that the dataset description is correct and complete, a full report of the stored information is verified with the dataset contact person, if identified.

To collect information on field sites where large data collection projects have taken place, or where a large number of separate investigations have been conducted, the Center might choose a project approach in cooperation with outside experts.

Once all the information describing a dataset is entered in the referral database, the information is checked for errors (see also related QA/QC in Section 8).

Periodically, the database content will be reviewed and updated where necessary, following the procedures described above.

Information Dissemination—

The Center distributes the information stored in the SATURN database in three ways: through published overviews, distribution of database files and search and report programs, and in response to field support requests. Regularly, the Center will publish, in the open literature, reviews and overviews of the datasets described in the referral database. Major outlets include journals, conference proceedings, and IGWMC's own publication series. For those who need assistance in locating detailed datasets for model validation, the IGWMC staff is available for consultation by telephone, correspondence, or visits to the Center.

To provide rapid, direct access to the database, remote access by telephone will be made available; or, the user may obtain a distributed version of the database system. This distributed version is available on a subscription basis, to allow the Center to provide the user with updated content and application software.

The experience of the Center in disseminating model information indicates that persons contacting the Center for such information are not familiar with the variety of models available, the detail in which these models are described, or with the selection process. In such cases, IGWMC staff provides assistance in helping define user needs and determining the most efficient search and report strategy to meet those needs. The same situation is expected to occur with requests for dataset-related information.

If the search is to be performed by an IGWMC staff member, care should be taken that the proper descriptors are used as a search criterion. A retrieval form will be filled out by the requestor or by the Center's staff. A report containing information on the search and a description of the selected datasets will be checked for errors by the Center's staff before is sent to the requestor. From this information the user might select a dataset for a specific problem. Eventually IGWMC staff members can play an advisory role in this decision, as they are familiar with the usefulness of the different datasets for specific testing and validation purposes. If the desired dataset and additional information are available from the Data Center, they can be sent to the user on request.

To evaluate the efficiency of its information dissemination process, the Center's staff will maintain a record of the information requests received and of the follow-up by the Center.

DATASET MANAGEMENT

Selection, Acquisition, and Evaluation Procedures

The Center may receive unsolicited datasets, or it may actively seek out a dataset for incorporation in its distribution package.

Selection of datasets to be distributed by the Center is determined by distribution criteria, priorities, and availability of staff time and funding. As discussed in Section 2, the dataset selection criteria adopted by the Center include:

- its significance to current groundwater pollution or model validation problems
- accessibility of information regarding the dataset and its collection
- presence of sufficient documentation (might be expanded and improved by Data Center)
- completeness of dataset (as raw data and/or processed data)
- quality of the data
- current format (electronic format or hard copy); if data are important and only available as hard (paper) copy, the Data Center might transfer them to an electronic storage medium
- timeliness of the data
- potential to obtain permission for distribution (e.g., if data collection has been publicly funded)

The first step in the selection process is screening descriptive information available in the SATURN database. When the Center considers a dataset for possible distribution, the staff contacts the dataset author (or custodian) to obtain detailed information regarding the dataset status, its availability, restrictions in distribution, its format, and pertinent documentation. If the Center decides to distribute a dataset, it will obtain written permission from the dataset author, custodian, or responsible agency or institute.

The Center will ensure that any dataset to be obtained is uncorrupted. Whenever possible, an integral copy of the electronically transferred dataset will be returned to the author for automatic verification of its content, or the dataset provider will be asked to prepare a duplicate for verification at the Center. Upon receiving the dataset, the Center will copy the tape or disk content into a computerized master-directory and will store the original tape or disk in a safe place. An identical procedure will be followed for the documentation. In any case, the Data Center will acknowledge the receipt of the material and follow the QA/QC procedures detailed in Section 8.

Potential Problems in Acquiring Datasets

Various problems, limiting data availability and accessibility, can occur in acquiring datasets. Such problems include the presence of mixed formats, data "tied-up" in projects, and data available only in hard-copy form ("shoebox filing"). Obtaining data from research projects often means spending significant effort coercing the data custodian to release the data and to provide the necessary data documentation. Another problem is that the datasets pertinent to a particular site often result from different investigations and thus reside with different custodians.

Furthermore, researchers often are reluctant to share data, even when their collection was made possible through public funding, because they want to make absolutely sure that their data contain no errors. During the discussions at the Data Center workshop (see Appendix A) it was noted that "you must cover yourself completely before the release of data." After data is collected during a research project, much effort is required to document, reformat, and acknowledge the data before they are released. There may be internal authorship battles which inhibit the release of data and documentation. Furthermore, a researcher often wants to get as much out of the data as possible. It is difficult to collect data and analyze it at the same time—the data may not be available for a long time, as when they are not released for two to three years after their collection. Even after data is made available in principle, its accessibility may be limited by any of the aforementioned factors. Specifically, a problem occurs when researchers responsible for the initial data collection claim that secondary data use, as in model validation, constitutes an infringement of their authorship rights; some researchers consider the results of such validation as their own research finding even if they have not contributed to the theoretical development of the model. The Data Center encourages model developers in search of field datasets to validate their models, to team up with the data generators to ensure optimal and correct use of the data. Dealing with these issues, the Data Center follows the general IGWMC approach and does not take a confrontational position, but will accomodate the data generators' concerns where possible. In any case, secondary users are asked to give full credit to the source of the data and inform the data generators of this secondary use. Arrangement of financial support for data generators documenting their data before release for secondary use and for their participation in the secondary use should be made in advance.

The Data Center will develop a protocol for secondary data use as discussed in the last paragraph of Section 3. This protocol will address such issues as protecting the integrity of the dataset, prevention of misuse or misinterpretation of the data, adhering to the original intent of the data collectors, the role of data collectors in secondary use, and proper acknowledgment of the data generator's work. The protocol must be accepted by the data user before the Data Center will release the data for secondary use.

Dataset Preparation and Distribution

When the dataset is acquired, the Center's final evaluation will be made of its quality and completeness and the completeness of its documentation. The outcome of the Center's final evaluation will be added to the distributed documentation.

As part of the distribution process, the computerized data files will be prepared for easy reproduction. The transfer medium will be a tape or disk. If the dataset has been altered by the Center, the distributed files will contain both the original "unaltered" data and the modified data. Each tape or disk will be labeled as to its content and format specifications and will be accompanied by a listing of the files and their content. Furthermore, the Center will distribute with the dataset a listing of all pertinent information regarding the site and the dataset (from the SATURN database) and a copy of the pertinent (core-) documentation.

Initially, each dataset will be distributed on an as-is basis. The dataset requestor should check the integrity of the dataset received. For primary datasets, those that IGWMC judges to be of high quality, the Center will eventually summarize the error-checking conducted by data collectors and perform its own extensive error-checking. However, such evaluations are time-consuming and require additional funding. If additional funding becomes available, the dataset formats and documentation might be improved and additional data restructuring, analysis, and reporting might be provided.

Additional information, such as field forms, notes, etc., are not likely to be distributed by the Data Center. However, if these are available, the same strict regulations and procedures will apply as for the other distributed materials.

In distributing datasets the Center will apply the same safeguards as it does in acquiring datasets from researchers. Therefore, the Center will advise dataset requestors of proper verification procedures.

SERVICES OFFERED BY THE DATA CENTER

To cover the costs of the operation of the IGWMC Research Data Center, seven types of activities need to be distinguished:

- dataset identification and annotation
- dataset selection and acquisition
- dataset preparation (for distribution) and documentation
- dataset transfer
- technical assistance (in quality assessments, dataset use evaluations, reformatting, etc.)
- information dissemination
- internal software maintenance

Two of these activities directly involve secondary data users: (1) information dissemination on research datasets, and (2) distribution of selected datasets. Information dissemination will be two-fold: (1) by responding directly to written or telephone requests; and (2) through the distribution on a subscription basis of the dataset referral system SATURN. The fee structure for these services will be based on the costs related to the preparation of the information or the datasets in requestor-specified format and the costs related to the transfer of the materials to the requestor. ASCII flat files will be available for most systems without the need to reformat, and the data will be available on magnetic tape or diskette. For standard preparation and transfer activities a fixed fee will be determined.

The cost of information analysis and storage in the SATURN system and the maintenance of SATURN software will be covered in part from overhead on the services directly related to the use of this system, and to a large extent from future contracts and agreements with funding agencies, especially the U.S. EPA, focused on its maintenance and utilization.

To fulfil one of the major objectives of the Center, that of distributing existing datasets efficiently, separate funding is required to cover identification, acquisition, and preparation of datasets for distribution. This is especially important if the selected dataset is not yet available on an electronic medium, if documentation is lacking or incomplete, or if extensive interaction with the dataset collecting team is required in readying the dataset for secondary use. In the first place, such funding needs to be sought from the agency that funded the original research or from agencies interested in the secondary use of these datasets. One possibility is for the Data Center to act as subcontractor to and in cooperation with the data collectors to complete the transfer and final documentation of the data as the last phase of a funded research project. If this funding mechanism cannot be realized, much of the potential of the Data Center concept will be unfulfilled.

SECTION 8

INTERNAL QUALITY ASSURANCE/QUALITY CONTROL

INTRODUCTION

As discussed in Section 4, quality assurance at the Research Data Center is a functional methodology for the documentation, filing, and control of technical reports, data, and computer programs prepared and/or distributed by the Center. The data handled by the Center can be divided into two groups: (1) numerical data generated in field and laboratory bench studies and managed by the IGWMC Groundwater Research Data Center, and (2) informational data stored in the Center's referral database SATURN.

This section describes the safeguards taken at the Center to insure that the quality standards adopted are applied through a variety of mechanisms. The applicable standards and procedures have been described in Sections 4 and 7, respectively.

The primary objective in implementing QA practices at the IGWMC Data Center is to concurrently document all data-related activities in order to provide evidence that standards of quality have been maintained. This objective implies the concepts of integrity, traceability, and accountability.

QA ORGANIZATION

The QA organization of the IGWMC Data Center is integral with the QA organization of the IGWMC as a whole. There are two levels in the IGWMC QA framework: (1) a permanent organization complete with QA management policies, goals, and objectives, and (2) project QA organization where general QA policies and assignments are detailed toward project objectives. A separate IGWMC QA Manual presents the charter of the Center's permanent QA organization, defining each element of the organization and outlining its responsibility. Where possible, the persons responsible for QA are independent from those responsible for its operational activities such as data management and software development and maintenance.

The hierarchy of responsibility for Data Center activities within the IGWMC is shown in Figure 13 and is designed with routine QA activities embedded in the technical chain-of-command. This approach differs from possible structures separating all QA activities. The QA coordinator has a key role outside other QA responsibilities; this role is currently assumed by the IGWMC Office Directors, whereas the daily QA activities of the staff are an integral part of the technical work. This implies that QA is not something new or additional, but rather a part of a permanent, comprehensive approach to file maintenance and reporting through the chain-of-command. The QA coordinator is brought into a project or task as needed to clarify QA issues, audit project (or task) files for QA compliance, and coordinate to a certain extent any inter-division activities.

Figure 13. Data Center QA organization

```
        office director      ---------------------->      QA coordinator
               |                                                |
               v                                                v
      division coordinator                              internal and external
               |                                           review (ITAC)
               v
     project or task leader
```

Responsibilities:

- Project (or task) leader is responsible for adequate QA files; or, if project is the direct responsibility of the division coordinator, the latter should maintain the QA files.

- QA coordinator audits all QA files and forms; at the close of the project the coordinator ensures that all folders have been filed (archived).

- After project ends QA control takes over (e.g., adequacy of documentation, form of filing or archiving, establishing retention period)

- The International Technical Advisory Committee (ITAC) provides advice and technical and scientific review of the IGWMC programs, services, and products.

QA TRACKING

The internal QA tracking is an important aspect of quality control (QC). The Center uses a series of forms to ensure that the QC is consistent for all models, information systems, and datasets it maintains and to provide staff guidance with respect to the QA/QC procedures to be followed. Each form will be subjected to final QA review by the IGWMC Director before the QA file is completed.

<u>Referral Database</u>

Data Entry—

To analyze a dataset for the SATURN referral database a standard annotation form is completed (see Section 7). To ensure consistent and adequate completion of this form and subsequent data entry and filing of the pertinent materials, an annotation-processing tracking form is used (see Table 13). Through this form the Data Center tracks the completion of the annotation, the entering of the data in the database, the internal verification of the entered data, the verification of the dataset analysis with the dataset author(s) or custodian, and the filing of the pertinent materials. For each of the checked activities, the operator, the date; and the initials of the checking person are required.

Table 13. SATURN ANNOTATION PROCESSING FORM

HD is Head, IGWMC Data Division

GENERAL INFORMATION

Site Name: ____

Site Number: ____________________ Study Number: ____________________

ANNOTATION FORM / DATA ENTRY INFORMATION

Encircle: New / Update Log (Up) date: ____________________

Task	Operator	Date	Check
Form completed:			
Annotation entered:			
Printouts checked:			
Sent to Author:			
Received from author:			
Modifications made:			
Hard copy filed:			

Comments: ____

____

____

QA Review (IGWMC Director): __________ (date) __________ (initials)

INTERNATIONAL GROUND WATER MODELING CENTER
Holcomb Research Institute, Butler University
Indianapolis, Indiana, USA

Information Retrieval—

To document the execution of a search request and the subsequent IGWMC response, a search processing form is used (see Table 14). The major elements of this form are the primary and secondary search criteria requested (criteria that have to be met and criteria that would be useful to the requestor, respectively), and an administrative check-off list. For each of the checked activities, the operator, the date, and the initials of the checking person are required.

Dataset Distribution

Acquisition—

To document the acknowledgment of the receipt of datasets and proper internal response, the Center uses a dataset tracking form (see Table 15). For each of the checked activities, the operator, the date, and the initials of the checking person are required.

Evaluation—

To ensure that a dataset received at the Center is evaluated according to established procedures (see Section 7), a dataset evaluation form is used. This form documents completion of the major evaluation steps and key findings of the Center staff member performing the particular evaluation (see Table 16). The main categories are:

- the extent of the data, their completeness and usefulness
- the dataset documentation
- file structure and data formats
- QA/QC documented
- Assessment of the required level of support (e.g. explanation, additional documentation, reformatting)

The form also documents the follow-up of the evaluation and the Center's decision with respect to distribution to secondary users. For each of the checked activities, the operator, the date, and the initials of the checking person are required.

Distribution—

To document the transmittal of a dataset and appropriate documentation the Center uses a dataset distribution form (Table 17). This form is both a check-list to ensure that all pertinent materials have been included, and a proof of transmittal. Therefore, each dataset distributed by the Center will have a dedicated transmittal form, listing the form in which the dataset has been distributed and the specific items (e.g., number of disks, type of tape, list of files, complete listing of printed documentation, etc.). For each of the checked activities, the operator, the date, and the initials of the checking person are required.

Each dataset distributed by the Center will be accompanied by the statement, "Enclosed are the items you have requested and which are described below. Please, report to us your acceptance of the materials, or any problem encountered during installation."

Table 14. SATURN SEARCH PROCESSING FORM

HD is Head, IGWMC Data Division

GENERAL INFORMATION

Information requested by:

Name: ______________________________
Organization: ______________________________
Department: ______________________________
Address: ______________________________
City: ____________________ State/Country: ____________________
Telephone: (___)____________________ ZIP/Postal Code: ____________________
Date of request: ____________________ Form of request: Written / Telephone

SEARCH REQUEST DESCRIPTION

Main Search Criteria/Problem Statement (First Search Level):

____
____
____
____
____
____
____

(continued)

TABLE 14. (continued)

Other Search Criteria (Second Search Level):

____

____

____

____

____

Requested Reporting Format: Summary / Full

Task	Operator	Date	Check
Search performed:			
Reporting checked:			
Sent to requestor:			
Copy of report filed:			

Comments: ________________________________

____

____

QA Review (IGWMC Director): __________ (date) __________ (initials)

INTERNATIONAL GROUND WATER MODELING CENTER
Holcomb Research Institute, Butler University
Indianapolis, Indiana, USA

TABLE 15. IGWMC DATASET TRACKING FORM

HD is Head, IGWMC Data Division

Dataset Name/Acronym: ______________________________ IGWMC Key: ___________

Form (encircle): tape / disk(s) Number of tapes/disks: ______

Documentation Included (encircle): yes / no

IGWMC Addressee: ___________________________________ Date: ____________

Task	Operator	Date	Check
Passed on to HD:			
IGWMC Director notified:			
Transmittal letter filed:			
Internal review completed:			
Authors informed of results:			
Dataset baselined/archived:			
QA file updated:			

Comments: ____

____

____

QA Review (IGWMC Director): _________(date) ______(initials)

INTERNATIONAL GROUND WATER MODELING CENTER
Holcomb Research Institute, Butler University
Indianapolis, Indiana, USA

TABLE 16. IGWMC DATASET EVALUATION FORM

HD is Head, IGWMC Data Division

GENERAL INFORMATION

Dataset Name/Acronym: ______________________________ IGWMC Key: ________
Site Name: ____
Number: __________ Study Number: __________
Received for evaluation: __________(date)
Contact Person: ____
Organization: ____
Address: ____
____
City: ____________________ State: ______ ZIP: ________

EXTENT OF DATA/COMPLETENESS/USEFULNESS

Evaluation performed by: ______________________________ Date: __________
Evaluation: ____
____
____
____
____
____
____
____
____
____

DOCUMENTATION

Evaluation performed by: ______________________________ Date: __________
Evaluation: ____
____
____
____

(continued)

TABLE 16. (continued)

FILE STRUCTURE/DATA FORMATS

Evaluation performed by: ______________________________ Date: __________

Number of files: ______________ Total storage required: __________Mbytes

Evaluation: ____

____

____

____

QA/QC APPLIED

Evaluation performed by: ______________________________ Date: __________

Evaluation: ____

____

____

____

____

REQUIRED LEVEL OF SUPPORT

Evaluation performed by: ______________________________ Date:__________

Evaluation: ____

____

____

____

____

General Comments: ____

____

____

____

____

(continued)

TABLE 16. (continued)

ADDITIONAL INFORMATION REQUESTED

Topic	Date Requested	Date Received	HD-Check
Documentation:	____________	____________	____________
QA/QC:	____________	____________	____________
Error checking:	____________	____________	____________
(Re-)formatting:	____________	____________	____________
Author support:	____________	____________	____________
Restrictions:	____________	____________	____________

DISTRIBUTION STATUS

	Date	HD-Check
To be distributed by IGWMC: yes / no	____________	____________
Documentation distribution ready:	____________	____________
Packing list prepared:	____________	____________
Annotation completed/updated:	____________	____________
Notification Indianapolis Office:	____________	____________
Notification Delft Office:	____________	____________
Announcement IGWMC Newsletter:	____________	____________

Restrictions: ____
____
____

QA Review (IGWMC Director): ____________(date) ______(initials)

INTERNATIONAL GROUND WATER MODELING CENTER
Holcomb Research Institute, Butler University
Indianapolis, Indiana, USA

TABLE 17. IGWMC DATASET DISTRIBUTION FORM

HD is Head, IGWMC Data Division

To:

Name: ________________________________

Organization: ________________________________

Department: ________________________________

Address: ________________________________

City: ____________________ State/Country: ____________

Telephone: (___)____________ ZIP/Postal Code: ____________

Date of Request: ____________ *Form of Request: Written / Telephone*

Name of *Staff* Person Handling Request: ________________________________

Enclosed materials:

o Magnetic Tape (9-track)
Unlabeled o / *other:* ____________
1600bpi o / other: ____________
ASCII o / EBCDIC o

o Floppy Disk

#		
___	5 1/4"	Double-sided/double density (360Kb)
___	5 1/4"	Double-sided/high density (1.2Mb)
___	3 1/2"	Double-sided/double density (720Kb)
___	3 1/2"	Double-sided/high density (1.44Mb)

Task	Operator	Date	Check
Tape/diskette(s) generated:	____________	______	______
Documentation prepared:	____________	______	______
Packing list checked:	____________	______	______
Invoice prepared:	____________	______	______
Final check:	____________	______	______

Comments: ________________________________

____

____

QA Review (HD): ________(date) _____(initials)

INTERNATIONAL GROUND WATER MODELING CENTER
Holcomb Research Institute, Butler University
Indianapolis, Indiana, USA

QA FILING

The IGWMC QA filing system is a three-tiered administrative structure consisting of archiving pertinent materials (e.g., baselined computer codes, datasets, documentation, and other products, either in electronic form or as printed material), QA folders, and QA control files. To organize the QA folders, a hierarchical filing system is used (e.g., division->project or task->folder), linking the files to projects or divisions of the Center. For example, in addition to the separate document folders prepared for each record in the IGWMC databases (e.g, containing all the past and present versions of a model or dataset annotation and related correspondence and notes), a separate QA tracking form is completed for filing in the QA task folder for the database of concern (MARS or SATURN) together with relevant QA information (e.g., audit reports, form reviews, etc.).

For each model or dataset received by the IGWMC, the Center maintains a separate folder (or group of folders) containing technical information specific to the model or dataset, including pertinent documentation, references to related technical reports, copies of technical memorandums and technical letters, internal and external technical reviews of published or delivered reports, and a hard copy of the computer codes, code input/output, or raw data received from external sources (complete with transfer correspondence).

In addition, a separate QA filing system is in place using QA control files for each of the models or datasets. These files contain the QA forms, QA audits, problem reports and subsequent remedial action, etc. In the near future this QA control system will be implemented on a microcomputer using a database management system (see Johnson et al. 1987).

For easy identification each model or dataset has its own identification number (IGWMC key), date, and short description.

SECTION 9

CONCLUSIONS AND RECOMMENDATIONS

Sharing data resulting from detailed field experiments and laboratory bench studies is important to the furtherance of scientific research. It stimulates interdisciplinary use of data and it enables verification, refutation or refinement of original research results. Added to this is an economical factor as many experimental studies, especially those carried out in the difficult-to-access subsurface environment, require complex facilities, elaborate instrumentation and equipment, and are often of prolonged duration.

Increasingly, large, multi-disciplinary field studies are conducted to enhance our understanding of the complex processes that govern the transport and fate of contaminants in the subsurface. These studies are often accompanied by laboratory experiments and laboratory testing of field samples. Such studies often are the result of cooperation between various funding agencies and research groups, each with its own missions and objectives. Increasingly, one objective of these studies is to promote their secondary use, and funding agencies are often prepared to provide additional financial support to this purpose. Moreover, various national and international forums have indicated the need for high quality datasets in order to perform testing and validation of subsurface contaminant transport and fate models, and to validate theoretical concepts on which these models are based.

As most research projects and their supporting organizations are not set up to distribute the elaborate datasets resulting from current field experimental research, the International Ground Water Modeling Center (IGWMC) in response to this need has established a comprehensive referral facility for selected, well-documented research datasets. Through this service the IGWMC hopes to prevent situations where datasets of value to many potential users go unrecognized (often, along with the researchers who generated the data) and therefore unused.

The datasets of interest to the Center should satisfy various selection criteria, such as accessibility of pertinent information, availability in automated format, adequate level of documentation, adequate quality of the data, acceptability of dataset distribution restrictions, and their significance for the main objectives of potential secondary use.

The secondary user of environmental measurements must have access to information relevant for their assessment of data quality. Therefor, the evaluation and documentation of the level of quality assurance applied during data acquisition, data handling and data storage will be a primary concern for the newly established IGWMC Research Data Center.

A key element of the Data Center's activities is the referral database SATURN. This database, which will be distributed, is designed to provide IGWMC staff and other users efficient access to important descriptive information on the content and quality of groundwater research datasets, available either through the Center or directly from the data originators. In order to design an efficient data directory type of data base, a list of major dataset descriptors has been developed, facilitating both rapid analysis and characterization of datasets by the Center's staff and fast search and retrieval its users. To allow future modification of database structure, descriptors and software, the database (programmed in Pascal and Assembly languages) has a flexible, transparent design.

In the future, additional detail in describing dataset quality and a system of dataset quality levels might prove an useful improvement. This extra effort should focus on the documentation by the data originators of information on project QA not included in publications.

Because often data originators and their supporting organization don't have the facilities or institutional framework to actually distribute datasets and their documentation routinely, the IGWMC Research Data Center also distributes selected datasets. Wherever possible, this distribution will be done in consultation with and with the consent of the data originators. The Data Center will encourage publicly funded data generators to share their data with other researchers for scientific use.

Although currently, the Center only distributes datasets on a "as-is" basis, expanding its services in the future might be considered to provide processed data according to user's specification. These additional services might include data reformatting, providing graphics, performing (geo)statistic analysis, and providing additional analysis regarding the utility and quality of datasets for user's specific objectives.

As the quality of the research data often is of great importance to the end-user, QA/QC aspects of data collection, handling, transfer, and storage are not only an issue for the data originators, but also for the Research Data Center. Internal QA/QC procedures and related institutional organization has been tailored to the existing, highly successful QA/QC program in model information and software distribution of the IGWMC. It is expected, that in the near future many of the internal QA/QC procedures of the IGWMC will be automated.

For this new facility to optimally function, groundwater research groups and their funding agencies should provide it with the necessary information, and funding should be made available to the Center to maintain and update the database. It is hoped that this new activity of the IGWMC will provide groundwater modelers with a new efficient means to improve the confidence placed by regulators and decision-makers in modeling as a powerful analytic tool.

SECTION 10

REFERENCES

American Society of Agronomy (ASA) and Soil Science Society of America (SSSA). Methods of Soil Analysis. 2 Vols., Agronomy Monograph #9, Madison, Wisconsin, 1986.

American Society of Testing and Materials (ASTM). Standard Definitions of Terms Relating to Water. In: 1987 Annual Book of ASTM Standards, Vol. 11.01, Water (1), Philadelphia, Pennsylvania, 1987a.

American Society of Testing and Materials (ASTM). Standard Practices for Sampling Water. In: 1987 Annual Book of ASTM Standards, Vol. 11.01, Water (1), *Philadelphia, Pennsylvania, 1987b.*

American Society of Testing and Materials (ASTM). Standard Practice for Determination of Precision and Bias of Applicable Methods. In: 1987 Annual Book of ASTM Standards, Vol. 11.01, Water (1). Philadelphia, Pennsylvania, 1987c.

American Society of Testing and Materials (ASTM). 1988 Annual Book of ASTM Standards, Vol. 4.08 Soil and Rock, Building Stones; Geotextiles. Philadelphia, Pennsylvania, 1988.

Anderson, M.P. Field Studies in Groundwater Hydrology—A New Era. Reviews of Geophysics 25(2):141-147, 1987.

Armentano, T.V., and O.L. Loucks. Ecological and Environmental Data as Under-Utilized National Resources: Results of the TIE/ACCESS Program. The Institute of Ecology (TIE), Indianapolis, Indiana, 1979. 98 pp.

Barcelona, M.J., J.P. Gibb, J.A. Helfrich, and E.E. Garske. Practical Guide for Ground-Water Sampling. EPA 600/S2-85/104, U.S. Environmental Protection, Agency, R.S. Kerr Environmental Research Laboratory, Ada, Oklahoma, 1985.

Barcelona, M.J., and J.P. Gibb. Development of Effective Ground-Water Sampling Protocols. In: A.G. Collins and A.I. Johnson, eds. Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 17-26.

Barcelona, M.J., and J.A. Helfrich. Well Construction and Purging Effects on Ground-Water Samples. Environmental Science and Technology, 20(11):1179-1184, 1986.

Barcelona, M.J., J.A. Helfrich, E.E. Garske, and J.P. Gibb. A Laboratory Evaluation of Ground Water Sampling Mechanisms. *Ground Water Monitoring Review, 4(2):32-41, 1984.*

Barker, J.F., G.C. Patrick, L. Lemon, and G.M. Travis. Some Biases in Sampling Multilevel Piezometers for Volatile Organics. Ground Water Monitoring Review, 7(2):48–54, 1987.

Barth, D.S., and B.J. Mason. Soil Sampling Quality Assurance User's Guide. EPA-600/4-84-043. U.S. Environm. Protection Agency, Environm. Monitoring Systems Lab., Las Vegas, Nevada, 1984.

Benson, R.C., R.A. Glaccum, and M.R. Noel. Geophysical Techniques for Sensing Buried Wastes and Waste Migration. National Water Well Assoc., Dublin, Ohio, 1982.

Benson, R.C., M. Turner, P. Turner, and W. Vogelsong. In Situ, Time Series Measurements for Long-Term Ground-Water Monitoring. In: A.G. Collins and A.I. Johnson, eds. Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 58–72.

Bissell, V.C., P.A. Pasteris, and D.G. Bennett. The Standard Hydrologic Exchange Format (SHEF) for Operational Hydrological Data, Hydrological Applications of Remote Sensing and Remote Data Transmission, IAHS Publ. 145. Internat. Assoc. for Hydrological Sciences, Exeter, UK, 1988. pp. 181–188.

Bohn, H.L., B.L. McNeal, and G.A. O'Connor. Soil Chemistry. John Wiley and Sons, New York, 1979. 329 pp.

Bruner, R.J., III. A Review of Quality Control Considerations in Soil Sampling. In: C.L. Perket (ed.), Quality Control in Remedial Site Investigation, ASTM STP 925, Am. Sc. for Testing and Materials, Philadelphia, Pennsylvania, 1986. pp. 35–42.

Bryden, G.W., W.R. Mabey, and K.M. Robine. Sampling for Toxic Contaminants in Ground Water. Ground Water Monitoring Review, 6(2):67–72, 1986.

Campbell, M.D., and J.H. Lehr. Water Well Technology. McGraw-Hill, New York, New York, 1973. 681 pp.

Campbell, J.A., and W.R. Mabey. A Systematic Approach for Evaluating the Quality of Ground Water Monitoring Data. Ground Water Monitoring Review, 5(4):58–62, 1985.

Cardin, C.W., J.E. Moore, and J.M. Rubin. "Water Resources Division in the 1980's." A Summary of Activities and Programs of the U.S. Geological Survey's Water Resources Division 1986, USGS Circular 1005, 1986. 79 pp.

CDM Federal Programs Corporation. Data Quality Objectives for Remedial Response Activities; Development Process. EPA/540/G-87/003, U.S. Environm. Protection Agency, Off. of Solid Waste and Emergency Response, Washington, D.C., 1987.

Cherry, J.A. (ed.). Migration of Contaminants in Groundwater at a Land Fill: A Case Study. Journal of Hydrology, 63(1/2): 1983.

Chu, W., E.W. Strecker, and D.P. Lettenmaier. An Evaluation of Data Requirements for Groundwater Contaminant Transport Modeling. Water Resources Research, 23(3):408–424, 1987.

Clarke, D. Microcomputer Programs for Groundwater Studies. Elsevier Science Publishers, Amsterdam, The Netherlands, 1987.

Collins, A.G., and M.E. Crocker. Laboratory Protocol for Determining Fate of Waste Disposed in Deep Wells. EPA Project Summary, EPA/600/S8–88/008, U.S Environmental Protection Agency, CERI, Cincinnati, Ohio, 1988.

Collins, A.G., and A.I. Johnson (eds.). Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988.

Committee on Data for Science and Technology (CODATA). International Council of Scientific Unions, Project Report, Paris, France, 1982.

Cordes, E.H. Ground Water Instrumentation Program of the U.S. Geological Survey. Ground Water Monitoring Review, 4(4):103–114, 1984.

Crocker, M.E., and L.M. Marchin. Adsorption and Degradation of Enhanced Oil Recovery Chemicals. In: A.G. Collins and A.I. Johnson (eds.), Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 358–369.

Curtis, G.P., P.V. Roberts, and M. Reinhard. Sorption of Organic Solutes and Its Influence on Mobility. In: P.V. Roberts and D.M. Mackay (eds.), A Natural Gradient Experiment on Solute Transport in a Sand Aquifer, Technical Report 292, Department of Civil Engineering, Stanford University, Stanford, California, 1986. pp. 124–144.

Date, C.J. An Introduction to Database Systems, Vol. 1. Addison-Wesley Publishing Co.

Davis, S.N., D.J. Campbell, H.W. Bentley, and T.J. Flynn. Groundwater Tracers. National Water Well Association, Worthington, Ohio, 1985.

Driscoll, F.G. Ground Water and Wells. 2nd Edition, Johnson Division, Saint Paul, Minnesota, 1986. 440 pp.

Dunbar, D., H. Tuchfeld, R. Siegel, and R. Sterbentz. Ground Water Quality Anomalies Encountered During Well Construction, Sampling and Analysis in the Environs of a Hazardous Waste Management Facility. Ground Water Monitoring Review, 5(3):70–74, 1985.

Dunlap, W.J., J.F. McNabb, M.R. Scalf, and R.L. Cosby. Sampling for Organic Chemicals and Microorganisms in the Subsurface, EPA 600/2-77-176, U.S. Environmental Protection Agency, R.S. Kerr Environmental Research Laboratory, Ada, Oklahoma, 1977.

Edwards, M.D. NAWDEX: A Key to Finding Water Data. U.S. Geological Survey, National Water Data Exchange, Reston, Virginia, 1978. 15 pp.

Edwards, M.D. Water Data and Information Exchange Programs of the U.S. Geological Survey. Proceedings of Conference on Improving Access to On-line Water Information, Cornell University, Center for Environmental Research, Ithaca, New York, 1983. pp. 30-39.

Edwards, M.D., A.L. Putnam, and N.E. Hutchinson. Conceptual Design for the National Water Information System. U.S. Geological Survey Bulletin 1792, Washington, D.C., 1987.

Eggenberger, L.M. Establishing an Environmental Quality Assurance Program. In: J.K. Taylor and T.W. Stanley (eds.), Quality Assurance for Environmental Measurements, ASTM STP 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1985. pp. 391-406.

Electric Power Research Institute (EPRI). Field Evaluation of Instruments for the Measurement of Unsaturated Hydraulic Properties of Fly Ash, EPRI EA-5011, Palo Alto, California, 1987. 86 pp.

Environmental Systems Research Institute (ESRI). Final Report, San Gabriel Basin Geographic Information System. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada, 1986.

Everett, L.G. Monitoring in the Vadose Zone. Ground Water Monitoring Review, 1(2):44-51, 1981.

Everett, L.G., E.W. Hoylman, L.G. Wilson, and L.G. McMillion. Constraints and Categories of Vadose Zone Monitoring Devices. Ground Water Monitoring Review, 4(1):26-32, 1984b.

Everett, L.G., L.G. Wilson, and E.W. Hoylman. Vadose Zone Monitoring for Hazardous Waste Sites. Pollution Technology Review No. 112, Noyes Data Corporation, Park Ridge, New Jersey, 1984a. 358 pp.

Fang, H.Y., and J.C. Evans. Long-Term Permeability Tests Using Leachate on a Compacted Clayey Liner Material. In: A.G. Collins and A.I. Johnson, eds. Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 397-406.

Ferris, J.G., D.B. Knowles, R.H. Browne, and R.W. Stallman. Theory of Aquifer Tests. Water-Supply Paper 1536-E, U.S. Geological Survey, Reston, Virginia, 1962.

Fetter, C.W., Jr. Potential Sources of Contamination in Ground-Water Monitoring. Ground Water Monitoring Review, 3(2):60-64, 1983.

Finkelstein, P.L., D.L. Mazzarella, T.J. Lockhart, W.J. King and J.H. White. Quality Assurance Handbook for Air Pollution Measurement Systems, Volume IV: Meteorological Measurements, EPA 600/4-82-060, Office of Air and Radiation, Research Triangle Park, North Carolina, 1983.

Fienberg, S.E., M.E. Martin, and M.L. Straf, eds. Sharing Research Data. National Academy Press, Washington, D.C., 1985. 240 pp.

Flatman, G.T. Design of Soil Sampling Programs: Statistical Considerations. In Quality Control in Remedial Site Investigation—Hazardous and Industrial Solid Waste Testing, Cary L. Perket (ed.), ASTM STP 925, Philadelphia, 1986. pp. 43–56.

Francis, C.W., M.P. Maskarinec, and D.W. Lee. Physical and Chemical Methods for the Characterization of Hazardous Wastes, Oak Ridge National Lab., Tennessee, 1988. 27 pp.

Freeze, R.A., and J.A. Cherry. Groundwater. Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1979. 604 pp.

Freyberg, D.L. Spatial Moments and the Advection and Dispersion of Non-Reactive Tracers. In: P.V. Roberts and D.M. Mackay, eds., A Natural Gradient Experiment on Solute Transport in a Sand Aquifer, Technical Report 292, Department of Civil Engineering, Stanford University, Stanford, California, 1986. pp. 50–92.

Friedman, D. Waste Testing and Quality Assurance. ASTM STP 999, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988.

Friedman, L.C., L.J. Schroder, and M.G. Brooks. Recovery of Several Volatile Organic Compounds from Simulated Water Samples: Effect of Transport and Storage. Environ. Sci. Technol., 20(8):826–829, 1986.

Frind, E.O., and G.E. Hokkanen. Simulation of the Borden Plume Using the Alternating Direction Galerkin Technique. Water Resources Research, 23(5):918–930, 1987.

Fritterman, D.V. Examples of Transient Sounding for Ground-Water Exploration in Sedimentary Aquifers. Ground Water, 25(6):685–692, 1987.

Gaspar, E. (ed.). Modern Trends in Tracer Hydrology, 2 volumes, CRC Press, Inc., Boca Raton, Florida, 1987. 312 pp.

Gelhar, L.G. Stochastic Prediction of Dispersive Transport at Hazardous Waste Sites. U.S. Environmental Protection Agency. R.S. Kerr Environmental Research Laboratory, Ada, Oklahoma, 1986.

Hammill, L., and F.G. Bell. Groundwater Resource Development. Butterworths, London, UK, 1986.

Hern, S. and S. Melancon. Vadose Zone Modeling of Organic Pollutants. Lewis Publishers, Inc., Chelsea, Michigan, 1986.

Hillman, J.J. Certification of Local Laboratories Analyzing Drinking Water in Nonprimacy States. In: J.K. Taylor and T.W. Stanley (eds.), Quality Assurance for Environmental Measurements, ASTM STP 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1985. pp. 103-109.

Hoffman, J.L. A Computer Databse for Overview of Ground Water Pollution Investigations. Ground Water Monitoring Review, 6(1):76-79, 1986.

Hopke, P.K., and D.L. Massart. Environmental Data Management; Directions for Internationally Compatible Environmental Data. Proceedings CODATA Workshop, International Council of Scientific Unions, Paris, France, 1986.

Huyakorn, P.S., A.G. Kretschek, R.W. Broome, J.W. Mercer, and B.H. Lester. Testing and Validation of Models for Simulating Solute Transport in Ground-Water. GWMI 84-13, International Ground Water Modeling Center, Holcomb Research Institute, Butler University, Indianapolis, Indiana, 1984.

HYDROCOIN. Progress Report No. 4., January - August 1986. Swedish Nuclear Power Inspectorate, Stockholm, Sweden, 1986.

HYDROCOIN. The International HYDROCOIN Project; Background and Results. Nuclear Energy Agency, Organisation for Economic Co-operation and Development (OECD/NEA) Paris, France, 1987.

INTRACOIN. Final Report Levels 2 and 3, model validation and uncertainty analysis. Report SKI 86-2, Swedish Nuclear Power Inspectorate, Stockholm, Sweden, 1986.

Jenkins, T.F., and P.W. Schumacher. Comparison of Methanol and Tetraglyme as Extraction Solvents for Determination of Volatile Organics in Soil. CRREL-SR-87-22, Cold Regions Research and Engineering Lab., U.S. Army Corps of Engineers, Hanover, New Hampshire, 1987. 36 pp.

Johnson, G.L., J.S. Ford, and L.E. Michalec. PC-QTRAK: An Automated Tracking System for Environmental Quality Assurance Activities. EPA/600/D-87/138, U.S. Environmental Protection Agency, Air and Energy Engineering Laboratory, Research Triangle Park, North Carolina, 1987.

Kaplan, E., J. Naidu, M. Hauptman, and A. Meinhold. 1985. Guidebook for the Assembly and Use of Diverse Ground Water Data. BNL-37356, Brookhaven Nat. Lab., Dept. of Applied Science, Brookhaven, New York, 1985.

Kashef, A-A. I. Groundwater Engineering. Mc-Graw Hill Inc., New York, New York, 1986.

Keely, J.F. Chemical Time-Series Sampling. Ground Water Monitoring Review, 2(4):29-38, 1982.

Keith, L.H., W. Crumlett, J. Deegan, Jr., R.A. Libby, J.K. Taylor, and G. Wentler. Principles of Environmental Analysis. Analytical Chemistry, 55:2210–2218, 1983.

W.K. Kellogg Biological Station. Data Management at Biological Field Stations. Workshop Report, prepared for National Science Foundation, Michigan State University, Hickory Corners, Michigan, 1982. 46 pp.

Kerfoot, W.B. Darcian Flow Characteristics Upgradient of a Kettle Pond Determined by Direct Ground Water Flow Measurement. Ground Water Monitoring Review, 4(4):188–192, 1984.

Kerfoot, W.B. Monitoring Well Construction, and Recommended Procedures for Direct Ground-Water Flow Measurements Using a Heat-Pulsing Flowmeter. In: A.G. Collins and A.I. Johnson, eds., Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 146–161.

Kerfoot, W.B., and V.A. Massard. Monitoring Well Screen Influences on Direct Flowmeter Measurements. Ground Water Monitoring Review, 5(4):74–77, 1985.

Keys, W.S., and L.M. MacCary. Application of Borehole Geophysics to Water Resources Investigations. U.S.G.S. Techniques of Water Resources Investigations, Book 2, Chapter E1, U.S. Geol. Survey, Washington, D.C., 1971.

Kroenke, D. Database Processing. Science Research Associates, Inc., Chicago, Illinois, 1977.

Kruseman G.P., and N.A. De Ridder. Analysis and Evaluation of Pumping Test Data. Bulletin 11, Internat. Inst. for Land Reclamation and Improvement, Wageningen, The Netherlands, 1970.

Leap, D.I. A Simple Pneumatic Device and Technique for Performing Rising Water Level Slug Tests. Ground Water Monitoring Review, 4(4):141–146, 1984.

Leap, D.I., and P.G. Kaplan. A Single-Well Tracing Method for Estimating Regional Advective Velocity in a Confined Aquifer: Theory and Preliminary Laboratory Verification. Water Resources Research, 24(7):993–998, 1988.

Locke, J.W. Laboratory Accreditation. Environmental Science and Technology, 21(4):332–333, 1987.

Lockhart, T.J. Quality Assurance of Meteorological Measurements. In: J.K. Taylor and T.W. Stanley, eds., Quality Assurance for Environmental Measurements, ASTM STP 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1985. pp. 253–259.

Lorenzen, D., R.A. Conway, L.P. Jackson, A. Hamza, C.L. Perket, and W.J. Lacy, eds. Hazardous and Industrial Solid Waste Testing and Disposal; 6th Volume. ASTM STP 933, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1986.

Mackay, D.M., W.P. Ball, and M.G. Durant. Variability of Aquifer Sorption Properties in a Field Experiment on Groundwater Transport of Organic Solutes: Methods and Preliminary Results. J. of Contaminant Hydrology, 1:119-132, 1986a.

Mackay, D.M., D.L. Freyberg, and P.V. Roberts. A Natural Gradient Experiment on Solute Transport in a Sand Aquifer, 1. Approach and Overview of Plume Movement. Water Resources Research, 22(13):2017-2029, 1986b.

Marble, D.F., and D.J. Peuquet. Geographic Information Systems and Remote Sensing. In: Manual of Remote Sensing, 2nd Edition, Volume I., American Society of Photogrammetry, Falls Church, Virginia, 1983. pp. 923-958.

Mercer, M.W., and C.O. Morgan. Storage and Retrieval of Ground-Water Data at the U.S. Geological Survey. Geological Survey Circular 856, Reston, Virginia, 1982. 9 pp.

Mercer, J.W., S.D. Thomas, and B. Ross. Parameters and Variables Appearing in Repository Siting Models. NUREG/CR-3066, U.S. Nuclear Regulatory Commission, Washington, D.C., 1982.

Melville, J.G., F.J. Molz, and O. Güven. Laboratory Investigation and Analysis of a Groundwater Flowmeter. Ground Water, 23(4):486-495, 1985.

Merrill, D. Practical Applications of Integrated Data Approaches. In Proceedings of the 1983 Integrated Data Users Workshop. ORNL, CONF-831117, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1985. 94 pp.

Meyer, P.D., and E.D. Brill, Jr. A Method for Locating Wells in a Groundwater Monitoring Network under Conditions of Uncertainty. Water Resources Research, 24(8):1277-1282, 1988.

Mills, W.B., D.B. Porcella, M.J. Ungs, S.A. Gherini, K.V. Summers, D.A. Haith, L. Mok, G.L. Rupp, and G.L. Bowie. Water Quality Assessment: A Screening Procedure for Toxic and Conventional Pollutants-Part II. EPA/600/6-85/002b, U.S. Environmental Protection Agency, Cincinnati, Ohio, 1985.

Mitchell, P.J., R.S. Argo, S.L. Bradymire, and C.A. Newbill. Hanford Ground-Water Data Base Management Guide and User's Manual. Battelle Pacific Northwest Laboratory, Richland, Washington, 1985.

Moffett, T.B. Where Have All the Ground Water Data Gone—Into the Files Never to Return? In Proceedings of the Sixth National Ground Water Quality Symposium, Sept. 22-24, 1982, Atlanta, Georgia. National Water Well Association, Dublin, Ohio, 1983. pp. 116-126.

Molz, F.J., J.G. Melville, O. Guven, R.D. Crocker, and K.J. Matteson. Design and Performance of Single Well Tracer Tests at the Mobile Site. Water Resources Research, 21(10):1497-1502, 1985.

Montgomery, R.H., and T.G. Sanders. Uncertainty in Water Quality Data. In: A.M. El-Shaarawi and R.E. Kwiatkowski (eds), Statistical Aspects of Water Quality Monitoring, Elsevier, Amsterdam, The Netherlands, 1986. pp. 17–29.

Moran, M.S., and L.J. Mezgar. Evaluation Factors for Verification and Validation of Low-Level Waste Disposal Site Models. DOE/OR/21400–T119, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1982.

Murarka, I.P., and D.A. McIntosh. Solid-Waste Environmental Studies (SWES) Description, Status, and Available Results. EPRI-EA-5322-SR, Electric Power Research Inst., Palo Alto, California, 1987.

NWQL (National Water Quality Laboratory). Quality Assurance in the National Water Quality Laboratory, NWQL and Canada Centre for Inland Waters, Burlington, Ontario, Canada, 1986. 98 pp.

NWQL (National Water Quality Laboratory). Brochure, Canada Centre for Inland Waters, Burlington, Ontario, 1986.

Nacht, S.J. Ground-Water Monitoring System Considerations. Ground Water Monitoring Review, 3(2):33–39, 1983.

National Research Council (NRC). *Sharing Research Data: A Scientific* Responsibility. News Report (6):15–18, 1985.

National Water Quality Laboratory (NWQL). Quality Assurance in the National Water Quality Laboratory. Environment Canada, Canada Centre for Inland Waters, Burlington, Ontario, Canada, 1986.

Nebraska Natural Resources Commission (NNRC). Nebraska Natural Resources Data Bank Information System, Data Access Manual, 2nd ed. Nebraska Natural Resources Commission, Lincoln, Nebraska, 1981. 348 pp.

Nelson, M.R. Index to EPA (Environmental Protection Agency) Test Methods. EPA/901/3-88/001, U.S. Environm. Protection Agency, Region 1, Boston, Massachusetts, 1988.

Nelson, J.D., and R.C. Ward. Statistical Considerations and Sampling Techniques for Ground-Water Quality Monitoring. Ground Water, 19(6):617–625, 1981.

Nicholson, T.J., T.J. McCartin, P.A. Davis, and W. Beyeler. NRC Experiences in HYDROCOIN: An International Project for Studying Ground-Water Flow Modeling Strategies. Proceedings GEOVAL 87, Stockholm, Sweden, April 4-9, 1987. Swedish Nuclear Power Inspectorate, Stockholm, Sweden, 1987.

Nielson D.M. Developing Standards for Ground Water Monitoring–Improving the State–of–the–Science. Editorial. Ground Water Monitoring Review, 7(4):4–5, 1987.

Nimmo, J.R., J. Rubin, and D.P. Hammermeister. Unsaturated Flow in a Centrifugal Field: Measurement of Hydraulic Conductivity and Testing of Darcy's Law. Water Resources Research, 23(1):124-134, 1987.

Office of Information and Regulatory Affairs (OIRA). Procedure for Preparation of Abstracts for Public Use of Statistical Machine-Readable Data Files. Proposed Voluntary Standard. Office of Management and Budget, Washington, D.C., 1983. 49 pp.

Olhoeft, G.R. Geophysics Advisor Expert System. EPA/600/X88/257. U.S. Environm. Protection Agency, Environm. Monitoring Systems Lab., Las Vegas, Nevada, 1988.

Olsen, H.W., T.L. Rice, and R.W. Nichols. Measuring Effects of Permeant Composition on Pore-Fluid Movement in Soil, In: A.G. Collins and A.I. Johnson, eds., Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 331-342.

Olson, R.E., and D.E. Daniel. Measurement of the Hydraulic Conductivity of Fine-Grained Soils In: T.F. Zimmie and C.O. Riggs (eds.), Permeability and Ground Water Contaminant Transport, ASTM STP 746, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1981. pp. 18-64.

Olson, R.J. Review of Existing Environmental and Natural Resources Data Bases. ORNL/TM-8928, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1984.

Olson, R.J., and N.T. Millemann, eds. Proceedings of the 1983 Integrated Data Users Workshop. USGS National Center, Reston, VA. ORNL, CONF-831117, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1985. 94 pp.

Parker, D.G., and S.C. Parker. Local Water Resource Information Management System. Publication No. 120, Arkansas Water Resources Research Center, University of Arkansas, Fayette, Arkansas, 1986.

Paulsen, S.G., C.L. Chen, K.J. Stetzenbach, and M.J. Miah. Guide to the Application of Quality Assurance Data to Routine Survey Data Analysis. EPA/600/4-88/010, U.S. Environm. Protection Agency, Environm. Monitoring Systems Lab., Las Vegas, Nevada, 1988.

Pearsall, K., and D.A.V. Eckhardt. Effects of Selected Sampling Equipment and Procedures on the Concentration of Tricloroethylene and Related Compounds in Groundwater Samples. Ground Water Monitoring Review, 7(2):64-73, 1987.

Perket, C.L., ed. Quality Control in Remedial Site Investigation: Hazardous and Industrial Solid Waste Testing; 5th Volume. ASTM STP 925, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1986.

Powers, J.P. Construction Dewatering. John Wiley & Sons, New York, New York, 1981.

Roberts, P.V., M.N. Goltz, and D.M. Mackay. Retardation Estimates and Mass Balances for Organic Solutes. In: P.V. Roberts and D.M. Mackay, eds., A Natural Gradient Experiment on Solute Transport in a Sand Aquifer, Technical Report 292, Department of Civil Engineering, Stanford University, Stanford, California, 1986. pp. 93-111.

Roberts, P.V., and D.M. Mackay, eds. A Natural Gradient Experiment on Solute Transport in a Sand Aquifer. Technical Report 292, Department of Civil Engineering, Stanford University, Stanford, California, 1986. 153 pp.

Ross, D., D. Sharma, and W.W. Shuster. Modeling the Fate of Toxic Organic Materials in Aquatic Environments. EPA-600/S3-82-028, U.S. Environmental Protection Agency, Environmental Research Laboratory, Athens, Georgia, 1982.

Scalf, M.R., J.F. McNabb, W.J. Dunlop, R.L. Cosby, and J.S. Fryberger. Manual of Ground-Water Quality Sampling Procedures. National Water Well Association, Dublin, Ohio 1981. 93 pp.

Schiegg, H.O., and J.F. McBride. Laboratory Setup to Study Two-Dimensional Multiphase Flow in Porous Media. PNL-SA-14988, Battelle Pacific Northwest Lab., Richland, Washington, 1987.

Schweitzer, G.E., and S.C. Black. Monitoring Statistics. Environmental Science and Technology, 19(11):1026-1030, 1985.

Semprini, L., P.V. Roberts, G.D. Hopkins, and D.M. Mackay. A Field Evaluation of In-Situ Biodegradation for Aquifer Restoration. EPA/600/2-87/096, U.S. Environm. Protection Agency, R.S. Kerr Env. Res. Lab., Ada, Oklahoma, 1987. 53 pp.

Sexton, K.G., H.E. Jeffries, J.R. Arnold, T.L. Kale, and R.M. Kamens. Validation Data for Photochemical Mechanisms. EPA/600/3-87/003, Atmospheric Sciences Res. Laboratory, Research Triangle Park, North Carolina, 1987.

Shacham, S., and H. McLellan. Abstracts of the Data Base—Wisconsin Power Plant Impact Study. Preliminary draft, Environmental Research Laboratory, U.S. Environmental Protection Agency, Duluth, Minnesota, 1979. 109 pp.

Shirley, P.A. Use of STORET as a Database for Ground-Water Quality Management. In: Proceedings of the Sixth National Ground-Water Quality Symposium, Sept. 22-24, Atlanta, Georgia, National Water Well Association, Dublin, Ohio, 1982. pp. 131-141.

Slawson, G.C. Jr., K.E. Kelly, and L.G. Everett. Evaluation of Ground Water Pumping and Bailing Methods-Application in the Oil Shale Industry. Ground Water Monitoring Review, 2(1):27-32, 1982.

Smart, P.L., and I.M.S. Laidlaw. An Evaluation of Some Fluorescent Dyes for Water Tracing. Water Resources Research, 13(1):15-33, 1977.

Smith, S., ed. Ground Water Hydrology for Water Well Contractors, National Water Well Association, Dublin, Ohio, 1982. 288 pp.

Stanley, T.W., and S.S. Verner. The U.S. Environmental Protection Agency's Quality Assurance Program. In: J.K. Taylor and T.W. Stanley, eds., Quality Assurance for Environmental Measurements, ASTM STP 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1985. pp. 12-19.

Steele, T.D. Strategies for Water Quality Monitoring. In: J.C. Rhoda, ed., Facets of Hydrology, Vol. II. John Wiley & Sons, Chichester, United Kingdom, 1985.

Stierman, D.J., and L.C. Ruedisili. Integrating Geophysical and Hydrogeological Data: An Efficient Approach to Remedial Investigations of Contaminated Ground Water. In: A.G. Collins and A.I. Johnson, eds., Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Penn., pp. 331-342, 1988.

Sudicky, E.A., J.A. Cherry, and E.O. Frind. Migration of Contaminants in Groundwater at a Landfill: A Case Study, 4. A Natural Gradient Dispersion Test. J. of Hydrology, 63:81-108, 1983.

Summers, K.V., and S.A. Gherini. Sampling Guidelines for Groundwater Quality. EPRI EA-4952, Electric Power Research Institute, Palo Alto, California, 1987.

Sykes, J.F., S.B. Pahwa, R.B. Lantz, and D.S. Ward. Numerical Simulation of Flow and Contaminant Migration at an Extensively Monitored Landfill. Water Resources Research, 18(6):1687-1704, 1982.

Sykes, J.F., S.B. Pahwa, D.S. Ward, and R.B. Lantz. The Validation of SWENT, A Geosphere Transport Model. In: R. Stepleman et al. (eds.), Scientific Computing. IMACS/North-Holland Publishing Company, Amsterdam, The Netherlands, 1983.

Taylor, J.K. What is Quality Assurance?. In: J.K. Taylor and T.W. Stanley (eds.), Quality Assurance for Environmental Measurements, ASTM STP 867, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1985. pp. 5-11.

Texas Natural Resources Information System (TNRIS). File Description Report, TNRIS-002, Texas Natural Resources Information System, Austin, Texas, 1982. 33 pp.

Texas Natural Resources Information System (TNRIS). TNRIS-A Progress Report, Texas Natural Resources Information System, Austin, Texas, 1983. 59 pp.

Tompson, A.F.B., and W.G. Gray. A Second-Order Approach for the Modeling of Dispersive Transport in Porous Media, 3. Application to Two Porous Media Problems. Water Resources Research, 22(13):1959-1971, 1986.

U.S. Department of Energy (DOE). Research Plan Subsurface Transport Program. DOE/ER-01563. Office of Health and Environmental Research, Ecological Research Division, Washington, D.C., 1985. 43 pp.

U.S. Environmental Protection Agency (EPA). Quality Assurance Handbook for Air Pollution Measurement Systems: Volume I - Principles. EPA 600/9-76-005, Office of Air and Radiation, Research Triangle Park, North Carolina, 1976.

U.S. Environmental Protection Agency (EPA). Procedures Manual for Groundwater Monitoring at Solid Waste Disposal Facilities. SW-611, Office of Solid Waste, Washington, D.C., 1981.

U.S. Environmental Protection Agency and National Center for Ground Water Research (EPA/NCGWR). A Ground Water Research Strategy. Ground Water Research Branch, R.S. Kerr Environmental Research Laboratory, Ada, Oklahoma, 1982. 56 pp.

U.S. Environmental Protection Agency (EPA). Handbook for Sampling and Sample Preservation of Water and Wastewater. EPA-600/4-82-029. Environm. Monitoring and Support Lab., Cincinnati, Ohio, 1982a.

U.S. Environmental Protection Agency (EPA). Guidelines for Assessing and Reporting Data Quality for Environmental Measurements. Environm. Monitoring and Support Lab., Cincinnati, Ohio, 1983.

U.S. Environmental Protection Agency (EPA). Ground Water Protection Strategy. Office of Ground Water Protection, Washington, D.C., 1984a. 73 pp.

U.S. Environmental Protection Agency (EPA). Review of the EPA Ground Water Research Program. Office of the Administrator, Science Advisory Board, Washington, D.C., 1985a. 55 pp.

U.S. Environmental Protection Agency (EPA). Methods for the Storage of RCRA (Resource Conservation and Recovery Act) Ground-Water Monitoring Data on STORET - Users Manual (Revised). Office of Solid Waste, Washington, D.C., 1985b.

U.S. Environmental Protection Agency. Ground-Water Data Requirements Analysis for the Environmental Protection Agency. Office of Ground-Water Protection and Office of Information Resources Management, 1987a.

U.S. Environmental Protection Agency (EPA). On-site Meteorological Program Guidance for Regulatory Modeling Applications. EPA-450/4-87-013, Office of Air and Radiation, Research Triangle Park, North Carolina, 1987b.

U.S. Geological Survey (USGS). National Handbook of Recommended Methods for Water-Data Acquisition, Chapter 2: Groundwater (revised). Office of Water Data Coordination, Reston, Virginia, 1977. 177 pp.

U.S. Geological Survey (USGS). Scientific and Technical, Spatial, and Bibliographic Data Bases and Systems of the U.S. Geological Survey, 1979. Geological Survey Circular 817, Reston, Virginia, 1979.

U.S. Geological Survey. ESIS User Manual. USGS, Information Systems Division, Reston, Virginia, 1982.

van der Heijde, P.K.M. A.I. El-Kadi, and S.A. Williams. Groundwater Modeling: An Overview and Status Report. GWMI 88-10, International Ground Water Modeling Center, Holcomb Research Institute, Butler University, Indianapolis, Indiana, 1988.

van Ee, J.J., and L.G. McMillion. Quality Assurance Guidelines for Ground-Water Investigations: The Requirements. In: A.G. Collins and A.I. Johnson (eds.) Ground-Water Contamination: Field Methods. ASTM STP 963, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1988. pp. 27-34.

Violette, P. Surface Geophysical Techniques for Aquifer and Wellhead Protection Area Delineation. EPA/440/12-87/016, U.S. Environm. Protection Agency, Office of Ground-Water Protection, Washington, D.C., 1987.

Walton, W.C. Computer Programs for Pumping Test Analysis, Groundwater Pumping Tests; Design and Analysis. Lewis Publishers, Inc., Chelsea, Michigan, 1987.

Way, S.C., C.R. McKee, and H.K. Wainwright. A Computerized Ground-Water Monitoring System. Ground Water Monitoring Research, 4(1):21-25, 1984.

White, P.A. Measurement of Ground-Water Parameters Using Salt-Water Injection and Surface Resistivity. Ground Water, 26(2):179-186, 1988a.

Wilson, L.G. Monitoring in the Vadose Zone: A Review of Technical Elements and Methods. EPA 600/7-80-134, Environm. Monitoring Systems Lab., Las Vegas, Nevada, 1980.

Zimmie, T.F., and C.O. Riggs, eds. Permeability and Groundwater Contaminant Transport. ASTM STP 746, Am. Soc. for Testing and Materials, Philadelphia, Pennsylvania, 1980.

Zohdy, A.A.R., G.P.Eaton, and D.R. Mabey. Application of Surface Geophysics to Ground-Water Investigations. Techn. of Water Resourc. Investig., Chapter D1, Book 2, U.S. Geological Survey, Washington, D.C., 1974.

October 1988

APPENDIX A

SUMMARY OF "WORKSHOP ON THE ESTABLISHMENT OF A GROUNDWATER RESEARCH DATA CENTER FOR VALIDATION OF GROUNDWATER MODELS"

SUMMARY OF "WORKSHOP ON THE ESTABLISHMENT OF A GROUNDWATER RESEARCH DATA CENTER FOR VALIDATION OF GROUNDWATER MODELS"

August 9-10, 1988

1. INTRODUCTION

On August 9-10, 1988, a workshop was held at the Holcomb Research Institute, Indianapolis, Indiana, on the establishment of a groundwater research data center for validation of groundwater models. This workshop is part of the EPA-funded two-year agreement with the Holcomb Research Institute, "Development and Operation of a Groundwater Research Data Center for Model Validation," funded through the R.S. Kerr Environmental Research Laboratory, Ada, Oklahoma. The workshop was attended by eleven persons, representing three EPA environmental research laboratories (Ada, Las Vegas, and Athens), the USGS (local and national), academic institutions, and Holcomb Research Institute.

The workshop goals were to discuss and evaluate Data Center procedures and the results of the work performed under the two-year agreement, with special focus on the formal description of datasets, quality assurance, the structure of the database (SATURN), and the role of research data in model validation. The two-day workshop involved presentations, group discussions, and written comment on provided questionnaires. A workshop notebook included graphs and peripheral information. Questions and comments concerning all aspects of the Data Center were addressed during the workshop.

The following summary of the workshop describes the presentations given and the issues, concerns, and suggestions expressed by the participants.

2. PURPOSE AND SCOPE

2.1. INTRODUCTORY COMMENTS (Paul van der Heijde)

Many researchers have found that data for the validation of complex models are not sufficiently accessible. A recent EPA study (U.S. EPA 1987) revealed a need to improve access to groundwater data and to lower the transaction costs associated with obtaining and using such data. The EPA study stressed the need to coordinate and standardize the large volume of groundwater data (and its storage) presently being generated and stored by many organizations in many different locations, files, and formats. The high costs of field research studies make it vital to share data and avoid duplication of research (U.S. EPA 1985).

To address these needs, the Holcomb Research Institute (HRI) is developing a Groundwater Research Data Center, with the support of the U.S. Environmental Protection Agency. The Data Center activities are taking place within the framework of the International Ground Water Modeling Center (IGWMC) and HRI's Data Center. The main tasks of the Groundwater Research Data Center will be the collection, conditioning, documentation, storage, and distribution of datasets resulting from research on groundwater pollution, and the establishment of a referral service for those datasets not managed by the Data Center. The Center will meet the need for better documentation of, information on, and access to existing groundwater research data, and will concentrate on those datasets that might be useful for groundwater model validation studies.

2.2. DISCUSSION

The following issues and concerns were expressed by the workshop participants after the presentation and in response to the questionnaire for this section of the workshop.

2.2.1. Model Validation

Model validation may be defined from many different viewpoints, e.g., management, legal, research, and model development. Validation involves determining and evaluating the differences between model predictions and field observations under a variety of conditions and stresses. It was pointed out that model validation can be seen as an iterative process, or at least a process of continued learning during the course of model development. Generally, a model developer will test the model to a certain extent and will often claim partial validation. Many professionals feel that reliable field data are essential for model validation, backed up by process-related laboratory data.

One practical purpose of model validation is to convince potential users that the performance of a model is acceptable for the conditions tested. Furthermore, a model's validity is often construed on the basis of the number of successful applications. Frequent use enhances credibility and reliability as model users find existing coding "bugs" and as limitations in applying the code to various situations are explored.

Model development often takes a long time; consequently, continuity during this phase is often a concern. Funding may be discontinued during the development process, personnel come and go, and interest may wane or change focus. Frequently, model documentation is delayed because of poor planning or a lack of funds allocated to the task. Due to these factors, a model often becomes available before its validity is established.

To illustrate some of the obstacles delaying validation, the validation of MINTEQ (a hydrochemical speciation model) was discussed. During the course of validating this model, the participants in the validation process were of different backgrounds (statisticians, geohydrologists, laboratory bench scientists, etc.) and were hard to convince of the model's overall validity because each was concerned with a different aspect of its performance. However, despite disagreement on the final validity or MINTEQ, it was generally agreed that it is the best model currently available for this type of problem.

Management generally has a limited understanding of models and model validation issues. Therefore, it would be useful to present a range of reviewed and tested models to assist in the analysis of management alternatives. Management in turn should have procedures in place to assure that a model's application is within the range of its established validity. For example, the USGS has a review system in place for applications of models: any application model must be checked with the original developer or another authority on that model. A "band of error" must be defined, meaning that the outcome of validation calculations should be within a predefined range from independently measured values. It should be noted that, although the USGS prefers not to use the term "validation," determining model reliability is a key concern of the USGS.

2.2.2. Purpose of the Referral Database

A number of possible uses of the database are:

(1) datasets for model validation
(2) datasets as example for other types of field studies
(3) datasets for education in model use
(4) comparison studies of research datasets
(5) guidance for collection of research data.

The participants discussed and weighed the relative importance of each of these uses. The consensus was that the Center should focus on data that can be used in relation to modeling activities. The use of datasets as examples for other research data acquisition projects is specially important because of the lack of good data collectors. It was noted that the first type of use should be the Center's primary focus, and that uses 2 and 3 are of secondary importance. Point 4 was rated important by some who thought that it might be useful for gathering statistical information. Most participants felt that the last type of use is least important.

2.2.3 Sharing Research Datasets

Different views were expressed concerning the scientific value of the storage and hence "third-party" use of data. The sharing of datasets is essential for cooperative studies and will facilitate the testing of hypotheses, and will prevent unnecessary duplication of effort and redundant data. Workshop participants agreed that these shared datasets will not lead to new groundwater research "breakthroughs," as these will have already occurred in the initial studies.

2.3. CONCLUSION

The participants agreed that the Data Center can function, as part of the International Ground Water Modeling Center, to facilitate and promote effective model validation and to improve understanding of the limitations of field validation of models. Another role the Data Center might play is that of providing a reference base for model parameters to be used in model application studies. Furthermore, the Center's products might be useful for teaching purposes and as examples for future data studies. Although these services and functions might be important spin-offs of the Center, its primary responsibility is providing a referral service for datasets. Of lower priority is its role in the distribution of research datasets. Finally, through its accumulated expertise, the Center might provide dataset analysis services to third parties. The workshop participants felt that initially, the Center should focus on a national audience, leaving acquisition of datasets from abroad for a later phase in its development.

3. DATA REQUIREMENTS FOR SELECTED GROUNDWATER MODELS

3.1. INTRODUCTORY COMMENTS (Stan Williams)

As the focus at the Data Center is on groundwater quality, an overview was given of (1) processes that influence transport and fate of contaminants in groundwater, and (2) related data requirements. Attention was given to factors that influence model selection, governing equations, mathematical solution techniques, and specific data requirements for the various types of models. A discussion of model limitations and current research developments was included.

Model selection is influenced by factors such as availability of the model, extent of its use, its user-friendliness, hardware dependence, level of review and testing applied to the model, and availability and quality of model documentation.

Generally, two governing equations are employed in a model for solute transport: one for groundwater flow and one for solute transport. The principal parameters in the groundwater flow equation are the hydraulic conductivity or transmissivity, specific storage or storativity, and sometimes fluid or matrix compressibility, fluid density, and dynamic viscosity. The solute transport equation usually includes terms for advection, dispersion, and solute sinks or sources. Nonconservative transport of species is often controlled by chemical processes such as adsorption, ion exchange, dissolution-precipitation, and biochemical transformation. Radioactive decay may take place. The parameters that enter into the solution of a solute transport equation include groundwater velocity, longitudinal and transverse dispersivity, distribution coefficient, and chemical half-life. Initial conditions for these solutions consist of values for the respective dependent variables; i.e., piezometric surface or solute concentration. Boundary conditions for the flow equation may include a combination of specified head, specified flux, and head-dependent flux boundaries. Boundary conditions for transport are usually specified concentration, specified solute flux, and concentration-dependent flux.

Some issues limiting the usability of available solute transport models include lack of flexibility for spatial discretization, accuracy limitations caused by vertical averaging (used in many 2-dimensional models), inability to handle multiple fluid problems adequately, and, in general terms, the lack of understanding of basic processes related to describing the behavior of real-world groundwater systems. New developments in groundwater transport modeling include modeling of multiphase systems, incorporation of biological decay mechanisms in existing transport formulations, coupling of transport models with geochemical models, refining the theory of dispersion, and increased capability in describing groundwater flow and transport in variable saturated media and in fractured rock or macroporous soil.

To illustrate the type of data required, handouts were presented to workshop participants.

3.2. DISCUSSION

The following issues and concerns were expressed by the workshop participants in discussion after the presentation and in response to the questionnaire for this section of the workshop.

3.2.1. Use of Available Data

There is a wide range of quality and testing of the many variables used in the more complex solute transport models. To apply such models, investigators usually take "a potpourri of different data types" from more than one source. Even for a single type of parameter, major differences exist in the quality of data. For example, some equilibrium constants are NBS standards, while many others frequently used are not (yet) accepted as a standard. Furthermore, there is a range of possible resolutions and accuracy in data used to test models. This requires that data consistency must be addressed in the context of model validation.

Data currently collected are not always suitable for model testing because of incompatible spatial or temporal resolution, sufficient accuracy, or because the wrong type of data has been collected. In many cases, existing measurement technologies are inadequate to provide the necessary level of detail. Recognizing the limitations posed by existing measurement technologies is important in determining the usefulness of datasets for the purpose of model validation.

Currently, and in the future as more sophisticated models are developed, greater quantities of certain data types will be needed in much more detail, these include data on adsorption and other geochemical processes, biodegradation, multiphase flow and transport, and mineralogical and biological characteristics of aquifer material. As an increased emphasis on risk assessment is foreseen, statistical descriptions of system properties become increasingly important. Therefore, more attention should be given to the natural spatial variability of parameters and characterization of correlation scale by increasing the spatial/temporal resolution of data. In the meantime, despite these severe limitations, we must use existing or measurable data for any validation.

3.2.2. Limitations in Field Measurement Technology

Theoretical and experimental scientists often work separately and isolated from each other. This frequently causes problems, as system properties needed for a successful model aren't always specified by the model developer. For example, some of the parameters required by theory are not accessible in the field. Furthermore, problems may be encountered with the scale of measurements and the dimensionality of the processes involved. It was stated during the workshop that an "overzealous" approach often is taken to model development, and that in such cases, more planning and thought regarding the development of parameter estimation methods is justified. Among other considerations, future data collection technologies should focus on better control of the aquifer volume for which water samples are collected and on better methods for in-situ measurement of chemical constituent concentrations.

4. CASE STUDY: HISTORY OF STANFORD/WATERLOO TRACER STUDY AND DATASET FOR THE BORDEN SITE

4.1. INTRODUCTORY COMMENTS (Doug MacKay)

The workshop overview of the Stanford/Waterloo Tracer Study focused on data collection, project documentation, and undocumented problems/snags during the project. Of particular of interest to participants were the candid statements concerning the quality assurance/quality control (QA/QC) actually applied. There was much concern during the design stage of the project about the loss of the volatile solutes during tracer injection. One incident of poor sampling technique where the purging rate was too high, resulted in discarding 1600 samples. This error was detected during review of laboratory analysis results when low concentrations of the volatile organics were noticed. Another problem concerned the time taken for sample labeling during the first sampling episode; as a result, a simpler labeling system was adopted for the next sampling sessions.

As part of the project, twelve core samples were taken (described in a thesis by Meredith Deranth, University of Waterloo) for experimental laboratory studies. Meanwhile, Barker and Patrick launched a BTX experiment at the source site while the field coring for the tracer study was still in progress. As a consequence, the coring sequence-timing was changed and delayed. The hydraulic conductivity values used for the tracer study were taken from previously published studies by Sudicky (University of Waterloo). Detailed head data were collected with the monitoring network but these were never published. Gary Hopkins (Stanford University) was in charge of technical facilities needed for the project. All project data was entered onto sheets, punched into the computer, and subsequently checked by Gary Hopkins. Stanford University was also involved in other studies at the Borden Site. Publications relatinggarding to these experiments are available through Paul Roberts of Stanford University. For the University of Waterloo studies, Sudicky can be contacted.

In practice, field experiments are often done before one knows how to do them. To fulfill the EPA QA project plan requirements, researchers often adopt an attitude of "you say what you need to say," but during the actual study "you do what you need to do," as most data collection procedures in research studies are still in an experimental stage or even under development as part of the study objectives.

4.2. DISCUSSION

The following suggestions, issues, and concerns were expressed by the workshop participants after the presentation and in response to the questionnaire for this section of the workshop.

4.2.1 Information Usually Not Included in Documentation

Many aspects of field data collection projects are rarely included in the project reporting and QA. Typical omissions are changes in approach, preliminary conclusions, changes in funding, changes in workplan, changes in staff, a brief written chronology of project events, key contacts for future project information, any unforseen project changes, site closures or other site activities, site anthropology, reasons for project changes, previous contaminant exposures, actual QA/QC procedures, and explanation of how the monitoring system was designed. Other omissions may include statistical techniques used for sampling design, problems that may have occurred in sampling, previous experimental work in the area that might have impacted physical properties, biases of the individuals performing the experiment, qualifications of the field team and laboratory analytical team, and lists of available project data. The type and amount of such information should be site- and project-specific.

4.2.2. Thoroughness in Reporting Field Activities

Workshop participants expressed a wide range of opinions concerning thoroughness in the reporting of field activities. Most participants felt that it is generally easier to accept data from studies in which they are personally involved, but that other datasets must have enough peripheral information to be acceptable for their own use. Much crucial information is often lacking in the actual project reporting and data documentation. As one participant said, "It is just not possible to include all of the potentially pertinent facts in most publications when editors and reviewers are constantly seeking terseness," and that descriptive, subjective information is often edited out. There was agreement that researchers must always be cautious in using third-party data.

4.2.3. Difference between What Occurs in the Field and What Is Reported

EPA and other agencies have many requirements and guidelines for studies performed by or for them. These requirements often include extensive QA procedures in preparation for administrative audits. QA is the project manager's responsibility, starting with development of workplans and QA plans to be submitted to the funding agency. However, during collection of field data, researchers often depart from these stringent plans, and major differences frequently occur between planning and practice. This is not necessarily bad, as field studies are conducted in order to allow unbiased observation of phenomena; if a test is too constrained by expectations, observations may be of limited value and important phenomena missed. In any case, such deviations from QA plans should be documented to allow assessment of the reasons for and validity of the deviations, and thus the value of the data acquisition effort. Recognizing that such deviations are probably more the rule than the exception in conducting field experiments, a practical level of QA should be defined and field activities should be honestly reported.

5. IDENTIFICATION OF POTENTIAL AND EXISTING DATASETS

5.1. INTRODUCTORY COMMENTS (Rachel Miller)

Different types of field research studies, endeavoring to describe different processes, were presented to workshop participants. A literature search was performed to gather information on datasets that might be of interest for model validation purposes. Information obtained was entered into a preliminary database containing approximately 150 such datasets. Several reports generated from the database were presented to the workshop participants for review. Various searches were conducted, using this database to identify datasets with certain characteristics (e.g., fractured rock, tracer tests, certain organizations involved, and so forth). At this stage of Data Center development, decisions must be made now to proceed with the gathering of detailed information. Issues involved include the means by which additional information should be collected, the considerations and criteria to be used in selecting datasets for incorporation in the Center's database, and identification of other possible sources of information on datasets, sites, and research projects. It should be noted that although a number of potentially interesting research sites have been identified, little is known about the accessibility and availability of the actual datasets.

5.2. DISCUSSION

The following issues, concerns, and suggestions were expressed by the workshop participants in discussion after the presentation and in response to the questionnaire for this section of the workshop.

5.2.1. Field Datasets versus Laboratory Bench Datasets

Some participants expressed their opinion that the Data Center should not compile data resulting from independent laboratory studies. As reasons for this advice, they pointed to the large number of laboratory bench studies, the variable quality of these studies ("fraught with problems"), and that the experts in this field are already aware of the really good datasets. However, many model developers have expressed their interest in this type of data for model evaluation purposes. They contend that a model cannot be properly validated without validation of its individual components. Laboratory datasets might prove useful for such partial evaluation. It was generally concluded that laboratory bench data should have low priority; rather, the Center's prime concern should be field research datasets followed by laboratory bench studies connected to research sites. Initially, the Center should be concerned only with lab bench studies from high quality field sites with which Center personnel are familiar.

5.2.2. Dataset Selection Criteria

It is important to establish criteria for both the selection of datasets, with which annotated descriptions will be included in the referral database, and for datasets that will be distributed by the Center. It was suggested that the Center consider only field datasets that have been documented in peer-reviewed publications. In an initial screening, the Center should select only those datasets that have been reviewed by others, e.g., by such agencies as the USGS, having their own internal review procedures for data collection.

It is important to focus on those datasets collected primarily for examination of scientific hypotheses (research data). Datasets for which no information is available on data quality should not be considered. Time and resources should initially be concentrated on those sites on which the most comprehensive and accessible field work have been performed and for which extensive and timely data are available. If collecting information on a specific site proves cumbersome, the effort should receive a lower priority. Datasets already used in modeling studies and datasets with extensive first-glance data should be included. The willingness of a data generator to have his/her data included in the referral database should be a major requirement for the selection of datasets. However, datasets for selected distribution should pass certain predefined QA criteria. In screening the datasets the Center should concentrate on the high quality material that is too voluminous to be published in the peer literature. Logistically, dataset availability will be one of the Center's first screening criteria.

5.2.3 Data Needed

Because each researcher or modeler has a special area of interest, a wide range of data types should be available for groundwater quality model validation. Some of the data identified include natural and forced gradient tracer tests, especially with organic or inorganic contaminants, and controlled experiments with remedial action technology. The data needs of the model designed for such experiments will dictate the type of data to be collected. Datasets should include groundwater elevations, spatial hydraulic conductivity and transmissivity, well locations, pumpage rates, contamination source characterization, and porosity, among other components. Of frequent interest is information on land use and population. Data resolution and quantity will be a function of the scale of study.

5.2.4. Additional Existing Datasets

The following are datasets that the participants felt might be of interest for the validation of groundwater models.

- Rocky Mountain Arsenal–current UCLA-conducted experiments; however, the data will not be available for a while.

- University of Waterloo experiments on NAPL behavior in the subsurface

- Picatinny Arsenal, NJ–a pollution site where heavy metals, TCE, and other degreasers are present. (Contact Tom Imbrogiotta, USGS District Office, NJ.)

- Wentsmith AF Base–a JP4 (jet fuel) spill (Contact T. Ray Cummings, USGS District Office, MI.)

- Hazardous Waste Ground-Water Task Force database–(EPA, Washington, DC) currently containing data from about 58 RCRA facilities.

6. REFERRAL DATABASE STRUCTURE AND SEARCH CRITERIA CRITERIA

6.1. INTRODUCTORY COMMENTS (Wilbert Elderhorst)

If model testers/validators are to use the Center's referral database to identify datasets suited to their particular purposes, the database must be well structured. Data entry, data retrieval, and maintainability and portability of software and data content must all be considered. The database will be structured using interrelated tables; its design will reflect efficient use of computer memory storage requirements and software execution speed.

For information analysis and data entry purposes, forms will be designed containing all descriptors identical to the database use interfaces. These forms will be partitioned into separate sections covering different topical areas of field measurements, site characterization, and operational characteristics of the datasets. This partitioning will facilitate collection of information scattered among various studies related to the same site.

Major search strategies have been determined and considered in the design of the database. Additional criteria are being used to develop dataset reports from the database information.

The referral database will contain information on a number of datasets collected at different experimental sites. This information is collected using a standardized descriptive analysis and reporting system focused on the major characteristics of the datasets (e.g., indicating which parameters are measured or determined, what information is available concerning methods and equipment used, organizations involved in the project data quality, geohydrological conditions present, etc.).

Each dataset will be described in a uniform way by a hierarchical set of descriptors, thus assuring consistency among the dataset descriptions anticipated. The selection of descriptors is a subjective process. The referral database should contain enough information to allow a searcher to determine the usefulness of a dataset for a specific application. The database should not contain information that is not crucial and that will be used only sparingly or not at all. It is evident that time and resources are related to the complexity of the system designed.

In a model validation example presented, it was shown that a well documented field experiment (pumping test) may yield an elaborate and detailed list of descriptors. However, the use of such detailed descriptors for all research data categories of potential interest would make the database too elaborate. Therefore, limitations must be imposed and criteria chosen to describe the datasets efficiently. A well-balanced list of descriptors for each of the specific research areas related to field experimentation on groundwater quality is essential for effective dissemination of information to a wide variety of users.

As the Data Center becomes operational, the staff will regularly evaluate its progress. Based on operational experience gained and user feedback expected, the database design may be modified. To facilitate such maintenance, the database should have a flexible structure that is efficiently programmed and well documented.

A draft dataset annotation format with about 400 descriptors was presented and comments were solicited.

6.2. DISCUSSION

The following issues, suggestions, and concerns were expressed by the workshop participants after the presentation and in response to the questionnaire for this section of the workshop.

6.2.1. Structure

The selection of the database descriptors and structure is intimately related to the main purpose of the database. Therefore, establishing the level of detail for the descriptors follows from a careful analysis of that purpose. Because input data are prerequisite to modeling, descriptors reflecting model data requirements should receive major attention. Model developers need detailed, process-oriented information. Local regulatory analysis requires GIS-type surveys; on a national level, there is a regulatory need for characterization of uncertainty to be used in generic Monte-Carlo-type simulations. With this information present, the Center's referral database could be used by regulators to compare sites for similar hydrogeologic environments in order to evaluate such issues as network design and/or remedial action alternatives. However, model validation should be the primary purpose of the datasets to be included.

In developing its database descriptors, the Center should be concerned with consistency in terminology across disciplines and should consider standard data formats. It should address terminology problems such as the existing disagreement over the use of such terms as "field capacity" or "water retention." The same applies to the units used in different disciplines for the same parameters. Selection of units may involve the description of different techniques or procedures used in determining the values of the specific parameter. The development and publication of a data dictionary would be helpful.

The complex datasets collected at many sites require a database structure that handles a variety of data types such as the different bulk densities for different soil areas within a site, complex aquifer systems varying in nature from site to site, or parameters whose values are by definition dependent on the method of measurement.

In designing the database, a basic structure should be defined along with a top-level set of 15 to 20 basic descriptors. Further detailing of such levels can be based on feedback from users. The structure should be easy to understand and flexible enough to facilitate future modifications; the level of detail must not stand in the way of its main purpose. Database reporting should include a one-page abstract of the researcher's original intent and use of the data. Descriptions of the region, type of site, and type of contaminants included in the first general level of the database would be useful. The reports generated should give all information of interest to the individual user and should be intelligible (not cryptic or in symbolic language).

The design criteria for database structure should:

(1) achieve completeness with respect to relevant information
(2) ensure balance of information between sections
(3) facilitate efficient searches
(4) provide efficient storage (minimizing storage requirements)
(5) provide useful and complete reporting.

Doug Mackay proposed the following list of basic descriptors/categories as essential for the first level of the database structure:

General (abstract, original purpose, duration of project)

Data Collectors
- principal contact(s)
- organizations involved

Model(s) tested or testable
- fluid flow
- solute transport/transformation
- biologic exposure
- chemical speciation
- exposure/risk prediction

Site Location
- area

general location

Site Characteristics
- Climate
- Surface
  - hydrology
  - geography
  - botany
- Unsaturated zone - hydrogeology
- Saturated zone

Primary (Field) Experimental Data
- spatial network
- parameters measured
- sampling methods
- analytical methods
- QA/QC procedures
- dataset description

Related (Field) Data
- pump tests
- geophysical methods
- meteorological measurements
- other

The category for experimental data should include descriptors that give a general impression of the dataset/site, including the water quality, engineering aspects, and any problems at the site.

The following are basic descriptors suggested for the laboratory datasets:

Single phase tests	Process-related lab tests
2-phase system	Property-related tests
partition coefficients	
degradation rates, etc.	

6.2.2. Level of Detail

The EPA Data Requirements study (U.S. EPA 1987) included a meeting to develop a "Required List of Data Elements." Many iterations were needed to determine the most important descriptors from a list of thousands of suggested terms. A main guideline in determining which descriptors to use is that the level of detail must be sufficient to quickly access datasets of interest.

Care must be taken so that the database will not exceed the "critical mass." An overly complex database might cause problems such as lack of personnel familiar enough with the database to use it, user manuals too complex to create and update, and certain basic information obscured by detail.

7. QA/QC: EVALUATION AND DOCUMENTATION OF DATASET QUALITY

7.1. INTRODUCTORY COMMENTS (Rachel Miller)

A major concern in the use of experimental data for model validation is data quality. There is need for high quality datasets for use in model testing and validation. Quality assurance for groundwater data collection projects is increasingly receiving attention. Of special importance is the long duration of these projects and the large scale of other activities. The EPA requires QA project plans for projects carried out by or for them to be prepared during the design of the study. In collecting data, different QA approaches are taken by the various organizations involved. These differences may be in the form of various standards set and applied, different regulatory requirements for collection and documentation, or in technical guidance provided. Currently, various organizations (e.g., ASTM) are formulating standards for groundwater field measurements. Quantitative measures of data quality, such as precision, accuracy, and detection limits, are used mainly in laboratory analysis. Data synthesis and reporting also may affect quality of reported data, and QA of these activities must be considered. (The Data Center is preparing a document, "Development of Criteria for the Evaluation and Annotation of Research Datasets.") Examples of existing QA procedure documentation were presented for various data acquisition activities including well logs, well construction diagrams, and analyzed-synthesized data (aquifer tests, geophysical data). Points for further discussion are how the Center should incorporate QA descriptors into the referral database and dataset distribution, and what minimal level of QA/QC documentation is minimally required.

7.2. DISCUSSION

The following issues, concerns, and suggestions were expressed by workshop participants in discussion and in response to the questionnaire for this section of the workshop.

7.2.1. Evaluation of Dataset Quality

Many organizations such as the USGS and EPA already have extensive internal/external review procedures in place to ensure the quality of research data collection and synthesis. The USGS also has a standard procedure for the release of data; sometimes data are reviewed and modified after being released. This is the case if a subsequent user detects inconsistencies; this would be of concern to the Center, and a mechanism should be in place to keep track of such incidents. It was noted that a distinction should be made between data quality for datasets in the referral database, and datasets distributed by the Center. In the latter case, the additional QA should be applied by the Center. For datasets included in the referral database, at least a rudimentary QA check should be performed. However, the user should have the responsibility for determining if the data is of sufficient quality to meet the user's needs. The extent of documentation available will be a major factor in a judgment of dataset quality. The database should contain information concerning all procedures followed and equipment used. This is only feasible if these procedures have been documented by the original researcher.

In evaluating the quality of datasets, the Center should be flexible, since many different levels of QA/QC are practiced for collection of different data types. Above all, practicality and common sense should be employed in establishing QA criteria. For example, a good way to test geophysical contractors is to have a test course for geophysical measurements (e.g., buried targets). A test such as this may ensure better quality than reviewing the documented use of Standard Operating Procedures (SOPS) by the field teams.

7.2.2. Documentation of Dataset Quality

Workshop participants shared the opinion that a classification system using a limited number of data quality levels should be employed in the referral database. The EPA discerns six levels of QA, from a basic research category to a 16-point QA documentation (including chain of custody procedures). The flagging of data elements should be considered and a system of qualitative descriptors (from rigorous to superficial QA) adopted. Again it should be noted that QA classifications for the referral database differ from those used in dataset distribution.

Somehow, the practice of field data collection should be evaluated by the Center. The Center should collect information on the innovative, unplanned, and often undocumented procedures used during data collection--especially when data collection is part of a research project. This information is crucial to datasets planned for distribution by the Center. Although the database should not include too much detail on data quality, information on the QA/QC is an essential part of the datasets to be distributed.

7.2.3. Field Notes and Forms

The Center should document field forms and procedures, but not field notes. It is more important to document these for distributing high quality datasets, but is of lower priority for the referral database. A method of evaluation might be to select some example forms and check them for accuracy and consistency. The user could then track these forms down if needed. Field notes would not be of much value to the Center; they are really of value only to the original researcher because the level of detail and the form in which they normally exist require that the original researcher interpret them.

7.2.4. Field Practice

QA problems often occur in field projects; errors and omissions are made on a regular basis. Difficulties inherent to field experimental work are frequently related to rush and fatigue (for example, the 1600-sample set that Mackay's group had to scrap). It is clear that actual data collection differs markedly from the "standard practice" procedures described in text or workplans. Furthermore, there is a lack of statistics on the amount and spatial arrangements of the data collected. It was agreed that more emphasis should be placed on the QA/QC of data collection from the beginning of research activities, including the statistics on the type of data collected.

7.2.5. Role of the Center in QA/QC Practice

The Center might provide researchers/data collectors with guidelines for understanding QA concerns. The Center could compare QA procedures in studies identified as generally good to procedures in questionable studies; however, rigorous QA/QC may be too cumbersome, leading to few data and an uninterpretable overall study. The Center could compare QA/QC plans with QA/QC actually done; as one participant stated, "Perhaps we should stop the charade and home in on the really important matters."

In evaluating applied QA/QC, the Center might take an administrative view, focusing on datasets that have been collected in projects that met formal EPA QA/QC requirements, or a more personal view, providing information on the investigators, methods used, and past reputations among peers so that the user could judge standards for the datasets.

8. DATA CENTER PROCEDURES

8.1. INTRODUCTORY COMMENTS (Wilbert Elderhorst)

In order to obtain effective and efficient operation of the Data Center, procedures should be established ensuring proper database management, acquisition of data, distribution of data, and description of datasets. Database management procedures should focus on assuring data integrity and security. This can be done by adopting authorization rules specifying that certain tasks can be performed only by a specific group of users. External users, for example, are not allowed to write or delete data from the database. Another important procedure would require routine data backup, allowing recovery of the database in the event that its content is corrupted.

Acquisition procedures must be established in order to obtain proper data transfer from data providers to the Data Center. This involves the design and use of standard terminology and forms. The actual data entry can be performed only by authorized Center personnel. Once all data that describe a dataset are entered in the referral database, they should be checked for errors. Furthermore, the data transfer media (floppy, disks, and tape), and optional reformatting procedures, must be determined.

Documentation should describe various report-generating options available in the database system. This would allow the user to obtain printed information from the referral database to the extent and detail desired. In addition to printing procedures, additional information should be available concerning user guides, availability of additional data, and reports describing the database structure.

Distribution procedures should contain restrictions on distribution such as proprietary rights or liabilities. When a dataset is accepted by the Data Center for distribution, these restrictions should be taken into account to avoid future problems. The Data Center should acquire and distribute only datasets that are public domain and do not have restrictions imposed on their general use.

8.2. DISCUSSION

The following issues, concerns and suggestions were expressed by the workshop participants in discussion after the presentation and in response to the questionnaire for this section of the workshop.

8.2.1. Dataset Integrity

The incorporation of procedures for maintaining data integrity is essential. In the discussion, an example was brought up in which certain data tapes were not decipherable. The tapes had been through many nondocumented manipulations and their integrity had not been preserved; hence, the Center must ensure that data is in its original form. A method for assuring integrity might be to compare the data on tape with the original data by sending completed forms and data output back to the originator to check for error; this can be done by computer, using existing software. Communication procedures should be established between the data source and the Center, and these procedures should allow the Center's staff to discover any inconsistencies in the data. In this context, problems within the STORET database system of EPA were discussed, particularly in relation to ranges of pH-related ion concentration values found to be unreasonably high.

These kinds of problems are comparable to those encountered by the IGWMC staff in the distribution of models. In such cases the Center's policies require that, if a user complains, the Center analyzes the problem and takes some course of corrective action, or takes the model out of distribution and informs the user of specific problems.

It was suggested that initially, the buyer should check data integrity; but for primary datasets, those that IGWMC judges to be of high quality, IGWMC should eventually summarize error-checking conducted by data collectors and conduct its own error-checking. However, extensive evaluations are time-consuming and require additional funding. For information entered in the referral database, or for datasets containing information entered by IGWMC, error-checking should be mandatory.

8.2.2. Services Offered by the Data Center

The Center would provide related services as funding or cost-recovery allows and as deemed necessary for the distribution of datasets and for information concerning them. These services might include guiding the deciphering of automated formats, providing a level of reliability for datasets, providing a summary table of the dataset, and reformatting at cost. ASCII flat files will be available for most systems without the need to reformat, and the data will be available on magnetic tape or diskette. It was generally agreed that although interpretation of datasets is not the responsibility of the Center, use could be made of standardized machine tests or screening (the Center must be cautious, as this approach has caused problems in other systems such as STORET). It was thought that variability of different dataset parameters would be of interest to data users. Summary tables could include extreme values, means, SDs, outliers, . . . (ciding with QA claims). The Data Center should include user support and the Center should incorporate procedures to keep track of the status, updating, and user needs relating to the datasets it refers or distributes.

8.2.3. Cost of Services

Several logistical issues were identified in the areas of funding and data transfer. The Center needs funds to distribute data efficiently and effectively, especially if extra formatting or documentation is required. Ideally, the funding agencies for data collection should provide budgets for dataset formatting, documentation, and preparation for distribution (as EPA did for the Stanford/Waterloo tracer study). Datasets should be distributed at cost to the requestor (after formatting, and documentation by the Center). If the original data-provider (e.g., a funding agency) subsequently requests the reformatted data, and when this is not covered by an existing agreement, the provider must pay for these additional services.

8.2.4. Problems with Acquiring Datasets

Various problems can occur in acquiring datasets. Personal experiences with data availability and accessibility involve problems with mixed formats, data "tied-up" in projects, and much data that is available only in hard-copy form ("shoebox filing"). To obtain results from research studies, a secondary user is often forced to spend significant effort identifying the source and in coercing the data custodian to provide the necessary information, often an exercise in sheer perseverance. For EPA projects, datasets should be obtained through the project officer of the data acquisition study. A knowledge of the reporting mechanisms of different agencies might be helpful in this respect.Datasets for model testing and validation often require data from different sources, since most datasets are incomplete in themselves or are a part of a large group effort.

Researchers often are reluctant to share data because they want to make absolutely sure that it contains no errors. "You must cover yourself completely before the release of data." After data is collected, much effort is required to document, reformat, and acknowledge the data. There may be internal authorship battles which inhibit the release of data and documentation. Furthermore, a researcher often wants to get as much out of the data as possible. It is difficult to collect data and analyze it at the same time—the data may not be available for a long time, as when they are not released for two to three years after their collection. There are many ways that researchers can hang on to data; basically, "Researchers just aren't going to give data until they are ready." Even after data is made available in principle, its accessibility may be limited by any of the aforementioned factors.

For the referral database, initially qualified Center personnel should complete the forms using existing dataset documentation, followed up by confirmatory interviews with or reviews by the data collectors. A team approach will probably be required, including a database manager, hydrogeologist, and modeler. As there are legal limitations on questionnaire mailings in studies performed with government funding, telephone contacts may be the best way to perform these surveys.

9. SUMMARY OF THE DATA CENTER WORKSHOP

Workshop participants agreed that there might very well be a "cascade" of research datasets in the future. Currently, many USGS research studies are being conducted at contamination sites, some in cooperation with DOD (such as the Picatinny site). Also, DOE, NSF, EPA, and other agencies are funding extensive field studies. It was estimated that the number of quality field datasets of eventual interest to the Center might be on the order of 50-100.

The workshop provided an excellent opportunity for the Center's staff to clarify understanding of the Center's objectives and the best ways to implement them. The workshop was held at a critical stage of the project, thus providing the necessary feedback for the final design of the Center's operation. As the Data Center becomes operational, information on new datasets will become available for the Center to process. The Center must decide on a practical system of gathering site information and balancing these activities with the resource made available to it.

Overall, the workshop members offered much encouragement toward the growth and future value of the Center. The Data Center would have limited value as a separate entity, but its affiliation with the IGWMC makes it viable and highly useful. Continued professional contacts with other organizations must be maintained and pursued.

The workshop served well in refocusing on problem areas in the Center's development. Many key issues have been addressed and the Center's staff is highly motivated to continue its work.

10. REFERENCES

U.S. Environmental Protection Agency. Report on the Review of the Environmental Protection Agency's Ground Water Research Program, Ground Water Research Review Committee, Science Advisory Board. U.S. EPA, Washington, D.C., 1985.

U.S. Environmental Protection Agency (EPA). Ground-Water Data Requirements Analysis. EPA Office of Ground-Water Protection, Washington, D.C., 1987.

APPENDIX B

DATASET SURVEY - February 1989

SITE NAME:	735 MW MICHIGAN CITY
DESCRIPTION:	coal-fired utilities, groundwater contamination
LOCATION:	north-central Indiana
CONTAMINANT OR TRACER:	fly-ash, coal storage pile leachates, heavy metals
GEOLOGY:	dunal region
AQUIFER MATERIAL:	sand
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	Univ. of Notre Dame
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	ALKALI LAKE
DESCRIPTION:	chemical waste disposal
LOCATION:	south-central Oregon
CONTAMINANT OR TRACER:	chlorophenolic compounds
GEOLOGY:	fractured porous soil, playa
AQUIFER MATERIAL:	gravel, sand, silt, clay
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	OGC, Univ. of Waterloo
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	ALLEN SITE
DESCRIPTION:	fly ash disposal ponds
LOCATION:	Gason County, North Carolina
CONTAMINANT OR TRACER:	fly ash leachate
GEOLOGY:	weathered bedrock, intrusive dikes
AQUIFER MATERIAL:	sand, silt, diorite
SOURCE(S) OF FUNDING:	EPRI, EPA
ORGANIZATIONS INVOLVED:	TetraTech Incorp., Arthur D.Little(ADL)
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ARMY CREEK
DESCRIPTION:	landfill leachate migration
LOCATION:	Wilmington, Delaware
CONTAMINANT OR TRACER:	landfill leachate
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	USGS, EPA
ORGANIZATIONS INVOLVED:	USGS, Del. DNREC
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ATLANTIC COAST
DESCRIPTION:	contamination of heterogeneous near shore aquifer
LOCATION:	Atlantic Coastal Plain of South Carolina
CONTAMINANT OR TRACER:	Various
GEOLOGY:	fluvial to marginal marine
AQUIFER MATERIAL:	sand, clay
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	Savannah River Lab
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	BABYLON LANDFILL
DESCRIPTION:	landfill contaminant plume
LOCATION:	south-east New York City, Long Island
CONTAMINANT OR TRACER:	landfill leachate
GEOLOGY:	glacial outwash plain
AQUIFER MATERIAL:	sand, gravel, silt
SOURCE(S) OF FUNDING:	USEPA
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	3

SITE NAME:	BAILLY GENERATING STATION
DESCRIPTION:	fly-ash ponds, dewatering plan
LOCATION:	north-west Indiana, lakeshore
CONTAMINANT OR TRACER:	fly-ash pond leachate
GEOLOGY:	Cowles Unit, dunal region
AQUIFER MATERIAL:	sand
SOURCE(S) OF FUNDING:	Nat. Park Service(NPS)
ORGANIZATIONS INVOLVED:	NIPSCO, USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	BARNWELL SITE
DESCRIPTION:	evaluate potential for deep percolation of radionuclides
LOCATION:	Barnwell, South Carolina
CONTAMINANT OR TRACER:	radionuclides
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	BATTELLE NWLAB1
DESCRIPTION:	adsorption of aromatic nitrogen bases in subsurface, laboratory study
LOCATION:	Battelle NW Labs
CONTAMINANT OR TRACER:	aromatic nitrogen bases
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Battelle PNL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	BAYVIEW PARK
DESCRIPTION:	landfill, fractured porous bedrock
LOCATION:	Burlington, Ontario
CONTAMINANT OR TRACER:	trichloroethane, landfill leachates
GEOLOGY:	Queenston Fm.-Paleozoic
AQUIFER MATERIAL:	fractured shale
SOURCE(S) OF FUNDING:	EPA, NSERC-Canada
ORGANIZATIONS INVOLVED:	OGC, Univ. of Waterloo
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	BEALE AFB
DESCRIPTION:	monitoring of waste sites, geophysical survey
LOCATION:	Marysville, California
CONTAMINANT OR TRACER:	varied wastes, fuels, landfills, photo, battery
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	AeroVironment, Inc.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	BEATTY SITE
DESCRIPTION:	low-level radioactive waste site
LOCATION:	Beatty, Nevada
CONTAMINANT OR TRACER:	low-level radioactive waste
GEOLOGY:	alluvial fan, flood-plain depos
AQUIFER MATERIAL:	coarse-grained sand, gravel
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	BEMIDJI
DESCRIPTION:	crude oil contamination in groundwater, multi-phase
LOCATION:	Bemidji, Minnesota
CONTAMINANT OR TRACER:	crude oil, hydrocarbons
GEOLOGY:	glacial outwash, till
AQUIFER MATERIAL:	sand, silt, clay, gravel
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	BISCAYNE AQUIFER
DESCRIPTION:	9 hazardous waste sites on EPA NPList
LOCATION:	Florida
CONTAMINANT OR TRACER:	landfill and industrial waste, hydrocarb, solvents, sewage
GEOLOGY:	sedimentary sequence-Florida
AQUIFER MATERIAL:	limestone, sandstone
SOURCE(S) OF FUNDING:	EPA-Superfund
ORGANIZATIONS INVOLVED:	EPA, CH2M Hill
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	BORDEN LANDFILL
DESCRIPTION:	contaminant plume, tracer tests
LOCATION:	NW of Toronto
CONTAMINANT OR TRACER:	liquid radioactive waste, (strontium)
GEOLOGY:	Quaternary
AQUIFER MATERIAL:	sand, clay, silt
SOURCE(S) OF FUNDING:	Nuc. Fuel Waste Man. Pr.
ORGANIZATIONS INVOLVED:	Chalk R. Nuc. Lab., Nat. Hy. Res. Inst. Canada
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	BORDEN SITE
DESCRIPTION:	tracer test
LOCATION:	NW of Toronto, Canada
CONTAMINANT OR TRACER:	chloride, bromide, 5 halogenated organic chemicals
GEOLOGY:	Quaternary
AQUIFER MATERIAL:	MF Sand
SOURCE(S) OF FUNDING:	EPA-RSKERL
ORGANIZATIONS INVOLVED:	Stanford, Waterloo
NUMBER OF MODEL APPLICATIONS:	9

SITE NAME:	CAMP LEJUENE
DESCRIPTION:	gas spill, Navy-Air Force Base, microbial action
LOCATION:	-0-
CONTAMINANT OR TRACER:	gas, hydrocarbons
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of North Carolina
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	CANADA URL
DESCRIPTION:	in situ geotechnical experiments, plutonic rock body, fract.flow
LOCATION:	Lac du Bonnet, Manitoba
CONTAMINANT OR TRACER:	radioactive waste disposal
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	Can. Nuclear WMP
ORGANIZATIONS INVOLVED:	Atomic Energy of Canada Limited
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	CANON CITY SITE
DESCRIPTION:	uranium ore processing, tailings disposal, groundwater contamination
LOCATION:	Canon City, Colorado
CONTAMINANT OR TRACER:	TDS, sulfate, molybdenum, uranium, selenium
GEOLOGY:	Tertiary, Cretaceous
AQUIFER MATERIAL:	alluv, SS, SH
SOURCE(S) OF FUNDING:	USEPA
ORGANIZATIONS INVOLVED:	USGS, USNRC, Alliance Technologies Corp.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	CAPE COD
DESCRIPTION:	sewage treatment plant, colloidal trans. of pollutants
LOCATION:	Otis AFB, SE Massachusetts
CONTAMINANT OR TRACER:	sewage
GEOLOGY:	Pleistocene
AQUIFER MATERIAL:	sand, gravel, silt, clay
SOURCE(S) OF FUNDING:	USGS, MDEQEng., RSKERL
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	CASTLEGUARD KARST
DESCRIPTION:	tracer study in karst hydrology
LOCATION:	Alberta, Canada
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of Western Ontario, London, Ontario
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME: CEDARSAUK SITE

DESCRIPTION: flyash landfill, groundwater contamination
LOCATION: southeastern Wisconsin
CONTAMINANT OR TRACER: sulfates, calcium, magnesium
GEOLOGY: glacial drift
AQUIFER MATERIAL: sand, gravel, clay, dolomite
SOURCE(S) OF FUNDING: USDI-WR Cen., Uni. Wis
ORGANIZATIONS INVOLVED: Univ. of Wisconsin, Wisc. Elec. Power Co.
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: CENTER COAL SITES

DESCRIPTION: environmental effects of fly ash disposal and FGD wast
LOCATION: Center, North Dakota
CONTAMINANT OR TRACER: fly ash, FGD waste
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: DOE
ORGANIZATIONS INVOLVED: MMRRI
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: CHEM-DYNE

DESCRIPTION: solvent reprocessing site
LOCATION: Hamilton, Ohio
CONTAMINANT OR TRACER: solvents
GEOLOGY: glacial
AQUIFER MATERIAL: fill, sand, silt, clay, till
SOURCE(S) OF FUNDING: EPA-Reg. 5-Chicago
ORGANIZATIONS INVOLVED: EPA, RFWESTON Inc., E&E, CH2MHill, OGC
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: CHROMIUM LAB1

DESCRIPTION: chemical attenuation of chromium in soils, solution in groundwater
LOCATION: -0-
CONTAMINANT OR TRACER: chromium
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: EPRI
ORGANIZATIONS INVOLVED: EPRI
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME:	CONCORD STATION
DESCRIPTION:	investigations for hazardous waste disposal
LOCATION:	Concord, California Naval Weapons
CONTAMINANT OR TRACER:	hazardous waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Army Eng. Waterways Exp. Station, Vicksburg, MS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	CONROE SITE
DESCRIPTION:	aerobic biodegradation, creosote plume
LOCATION:	Conroe, Texas
CONTAMINANT OR TRACER:	chloride, organic compounds, creosote
GEOLOGY:	unconsolidated Pliocene
AQUIFER MATERIAL:	sand, sandy clay, clay
SOURCE(S) OF FUNDING:	EPA, NCGWR, AMOCO Fnd.
ORGANIZATIONS INVOLVED:	Rice University
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	CRATER LAKE
DESCRIPTION:	hydrological effects of test drilling
LOCATION:	Winema National Forest, Crater Lake
CONTAMINANT OR TRACER:	drilling fluid/mud
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	LBL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	CREUX DE CHIPPIS SITE
DESCRIPTION:	tracer study in a stony field soil
LOCATION:	Sierre, Switzerland
CONTAMINANT OR TRACER:	chloride, bromide tracers
GEOLOGY:	colluvial, glaciofluvial
AQUIFER MATERIAL:	sand, gravel, clay
SOURCE(S) OF FUNDING:	Swiss Fed. Inst. Tech.
ORGANIZATIONS INVOLVED:	-0-
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	DAWSONVILLE SITE
DESCRIPTION:	fractured flow, contaminant transport
LOCATION:	Dawsonville, Georgia
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	Georgia State University, USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	DEAF SMITH
DESCRIPTION:	candidate site for radioactive waste disposal
LOCATION:	Texas
CONTAMINANT OR TRACER:	radioactive waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	DOE
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	DOUGHERTY PLAIN
DESCRIPTION:	pesticides trans. and fate, agricultural non-point source
LOCATION:	Albany, Georgia
CONTAMINANT OR TRACER:	pesticides, aldicarb, metachlor, bromine
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	EPA-Athens, USGS
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	ELRAMA SITE
DESCRIPTION:	scrubbing waste disposal, coal strip-mine spoils
LOCATION:	Washington County, west-central Pen
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	alluv., Pennsylvanian rocks
AQUIFER MATERIAL:	shale, siltstones
SOURCE(S) OF FUNDING:	EPRI, EPA
ORGANIZATIONS INVOLVED:	Tetra Tech Incorp., Arthur D. Little (ADL)
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ETIWANDA FIELD SITE
DESCRIPTION:	unsaturated zone dispersion experiment
LOCATION:	-0-
CONTAMINANT OR TRACER:	tracers: chloride, boron, nitrate, bromacil
GEOLOGY:	field soil
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA, EPRI
ORGANIZATIONS INVOLVED:	EPRI, Univ. of CA Riverside
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	EUROPE SOIL
DESCRIPTION:	nitrite migration through a loess soil
LOCATION:	northwestern Europe
CONTAMINANT OR TRACER:	nitrite
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Catholic Univ. of Louvain, Belgium
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	FANAY-AUGERES
DESCRIPTION:	uranium mine, nuclear waste storage research, fracture flow
LOCATION:	Limousin, France
CONTAMINANT OR TRACER:	radioactive waste disposal
GEOLOGY:	-0-
AQUIFER MATERIAL:	granite
SOURCE(S) OF FUNDING:	DOE, AFEE
ORGANIZATIONS INVOLVED:	LBL, French Bureaude Rechercles Geologiqueset Minieres
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	FEDERAL PIONEER
DESCRIPTION:	chemical spill, transformer manufacturing
LOCATION:	Regina, Saskatchewan
CONTAMINANT OR TRACER:	Polychlorinated Biphenyls(PCBs), TCB, multi-phase
GEOLOGY:	recent unconsolidated, glacial, Cret. bedrock
AQUIFER MATERIAL:	soil, clay, silt, till, sand, grav.
SOURCE(S) OF FUNDING:	NRCC, Saskatchewan
ORGANIZATIONS INVOLVED:	Univ. of Waterloo, Univ. of Alberta
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	FIELD COAL WASTES
DESCRIPTION:	study leaching of advanced coal process waste
LOCATION:	-0-
CONTAMINANT OR TRACER:	advanced coal process waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE, METC
ORGANIZATIONS INVOLVED:	NDEMRC, MMRRI
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	FLY-ASH LAB1
DESCRIPTION:	attenuation capacity, fly-ash disposal site
LOCATION:	North Dakota
CONTAMINANT OR TRACER:	fly ash, FGD waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	NDEMRC, MMRRI, Univ. of ND
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	FLYING J REFINERY
DESCRIPTION:	hydrocarbon contam., assessment and remediation
LOCATION:	Williston, ND
CONTAMINANT OR TRACER:	hydrocarbons
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	NDEMRC, MMRRI, EES
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	FORT UNION SITES
DESCRIPTION:	spoils from uraniferous lignite mines
LOCATION:	North Dakota
CONTAMINANT OR TRACER:	uranium in spoils piles
GEOLOGY:	-0-
AQUIFER MATERIAL:	lignite beds
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	FOSSUM SITE
DESCRIPTION:	evaluate brine and oil-gas drilling fluid disposal site
LOCATION:	North Dakota
CONTAMINANT OR TRACER:	oil brine, oil-gas drilling fluid
GEOLOGY:	glacial till
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	NDWRRI
ORGANIZATIONS INVOLVED:	NDGS, Univ. of ND, MMRRI
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	FRESNO SITE
DESCRIPTION:	aldicarb contamination, saturated and unsaturated soil
LOCATION:	central California
CONTAMINANT OR TRACER:	aldicarb
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Union Carbide- Research Triangle
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	GLIL YAM SITE
DESCRIPTION:	anthropogenic anoxification of a deep phreatic aquifer
LOCATION:	15 km. north of Tel Aviv, Israel
CONTAMINANT OR TRACER:	oxygen, organics
GEOLOGY:	Pleistocene, Pliocene
AQUIFER MATERIAL:	SS, siltstone, soils, clay
SOURCE(S) OF FUNDING:	Nat. Council for R&D
ORGANIZATIONS INVOLVED:	Technion Israel Inst. of Tech., Weizmann Inst. of Science
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	GLOBE LAB1
DESCRIPTION:	lab column exper. in alluv., copper mining contam.
LOCATION:	Globe, Arizona
CONTAMINANT OR TRACER:	dissolved metals, sulfate
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS-AZ-CO
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	GLOBE SITE
DESCRIPTION:	acidic groundwater contam. from copper mining
LOCATION:	Globe, Arizona
CONTAMINANT OR TRACER:	dissolved metals, sulfate
GEOLOGY:	alluv., Gila conglomerate
AQUIFER MATERIAL:	sand, gravel, silt, clay
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS-AZ-CO
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	GRAND ISLAND
DESCRIPTION:	subsurface contam. from disposal of munition wastes
LOCATION:	Grand Island, Nebraska
CONTAMINANT OR TRACER:	munition wastes, RDX, TNT
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of Nebraska
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	HAIFA BAY EXPERIMENTAL FIELD
DESCRIPTION:	tracer study for porosity and permeabilty
LOCATION:	Israel
CONTAMINANT OR TRACER:	radioactive tracer, Co60
GEOLOGY:	-0-
AQUIFER MATERIAL:	M sand, SS, Sh, C sand, clay
SOURCE(S) OF FUNDING:	Israel Gov., UN SpFun
ORGANIZATIONS INVOLVED:	TAHAL Water Planning for Israel Ltd.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	HANFORD SITE
DESCRIPTION:	radiocontaminants, sat. and unsat.
LOCATION:	south-central Washington State
CONTAMINANT OR TRACER:	radiocontaminants, plutonium
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	At. Richfield, DOE
ORGANIZATIONS INVOLVED:	Battelle, USGS, Boeing Comp. Services
NUMBER OF MODEL APPLICATIONS:	9

SITE NAME:	HESKETT STATION
DESCRIPTION:	site selection for long-term disposal of coal fly-ash
LOCATION:	Mandan, ND
CONTAMINANT OR TRACER:	coal fly-ash
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	Montana Dakota Util.
ORGANIZATIONS INVOLVED:	NDEMRC, MMRRI
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	HOE CREEK
DESCRIPTION:	underground coal gasification experimental site, contam
LOCATION:	south of Gillette, Wyoming
CONTAMINANT OR TRACER:	organic contam., phenols
GEOLOGY:	sedimentary sequence
AQUIFER MATERIAL:	sandstone, mudstone, coal seams
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	DOE, LLNL, Western Research Inst.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	HYDE PARK LANDFILL
DESCRIPTION:	immiscible organic transport, chemical waste disp. site
LOCATION:	Niagara Falls, New York
CONTAMINANT OR TRACER:	chemical waste
GEOLOGY:	glaciolacustrine, Lockport Fm.
AQUIFER MATERIAL:	silt, clay, dolomite
SOURCE(S) OF FUNDING:	Nat. Sci. & Eng.Council
ORGANIZATIONS INVOLVED:	University of Waterloo, EPA.
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME: IDAHO NEL SITE1

DESCRIPTION: radioactive waste, Nat.Reactor Testing Station
LOCATION: SE Idaho
CONTAMINANT OR TRACER: liquid low-level rad. waste, dilute chem., sewage
GEOLOGY: Quaternary, Snake R. Group
AQUIFER MATERIAL: basalticvol., interbd. seds.
SOURCE(S) OF FUNDING: DOE
ORGANIZATIONS INVOLVED: INEL, USGS, Atomic Energy Com. (AEC)
NUMBER OF MODEL APPLICATIONS: 3

SITE NAME: IDAHO NEL SITE2

DESCRIPTION: purgeable organic compounds in gw
LOCATION: Idaho National Engineering Laborato
CONTAMINANT OR TRACER: purgeable organic compounds
GEOLOGY: Quaternary, Snake R. Group
AQUIFER MATERIAL: basalticvol., interbd. sed.
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: INEL, USGS, DOE
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: IDAHO SPRINGS SITE

DESCRIPTION: measurement of stress in rocks near mined repositories
LOCATION: Idaho Springs, Colorado
CONTAMINANT OR TRACER: -0-
GEOLOGY: metamorphic rock
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: USGS
ORGANIZATIONS INVOLVED: USGS
NUMBER OF MODEL APPLICATIONS: 99

SITE NAME: INDUST SITE

DESCRIPTION: VOC contamination
LOCATION: -0-
CONTAMINANT OR TRACER: VOCs
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: EPA-RSKERL
ORGANIZATIONS INVOLVED: EPA, Arizona State
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME:	JEFFREY CITY
DESCRIPTION:	uranium tailings contam., sat. and unsat., lab studies
LOCATION:	central Wyoming, 2.4 km northeast
CONTAMINANT OR TRACER:	radionuclides, toxic elements
GEOLOGY:	alluv, dune, sed. BR, cryst. BR
AQUIFER MATERIAL:	sand, gravel, sandstone, granite
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	D'Appolonia Inc., Earth Sciences Consultants, Inc.
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	JET FUEL LAB
DESCRIPTION:	fate of jet fuel in aquatic sediments, lab study
LOCATION:	-0-
CONTAMINANT OR TRACER:	jet and missile fuel
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	ERL, Gulf Breeze, FL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	KELLY AFB
DESCRIPTION:	field-testing of in-situ biological degradation
LOCATION:	Kelly Air Force Base
CONTAMINANT OR TRACER:	organic contaminants
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Science Applications International Corp.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	KOPPERS COKE PLANT
DESCRIPTION:	groundwater contamination
LOCATION:	St. Paul, MN
CONTAMINANT OR TRACER:	toxic organic substances
GEOLOGY:	St. Peter aq., glacial
AQUIFER MATERIAL:	sand, gravel, silt, clay
SOURCE(S) OF FUNDING:	USEPA, MPCA
ORGANIZATIONS INVOLVED:	USGS, USEPA, MPCA
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LAB3
DESCRIPTION:	chemical transformation processes in groundwater
LOCATION:	-0-
CONTAMINANT OR TRACER:	organic chemicals
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-RSKERL
ORGANIZATIONS INVOLVED:	EPA
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LAB4
DESCRIPTION:	biodegradation of xenobiotics in subsurface environment
LOCATION:	-0-
CONTAMINANT OR TRACER:	organic pollutants
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-RSKERL
ORGANIZATIONS INVOLVED:	EPA, Univ. of Dayton
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LAC DU BONNET SITE
DESCRIPTION:	fractured flow studies, granitic batholith
LOCATION:	Lac du Bonnet, Manitoba Canada
CONTAMINANT OR TRACER:	radioactive waste disposal
GEOLOGY:	granitic batholith
AQUIFER MATERIAL:	granite, crystalline rock
SOURCE(S) OF FUNDING:	DOE, CNFWM program
ORGANIZATIONS INVOLVED:	Atomic Energy of Canada Limited(AECL), Battelle
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	LAKE HAMILTON
DESCRIPTION:	aldicarb residue study, citrus groves on Florida ridge
LOCATION:	central ridge area of Florida, near
CONTAMINANT OR TRACER:	aldicarb residues
GEOLOGY:	-0-
AQUIFER MATERIAL:	soil, C sand
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	Union Carbide, Univ. of Florida, FDER, FDACS
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	LANSING SMITH SITE
DESCRIPTION:	ash disposal pond
LOCATION:	Panama City Florida, coastal plain
CONTAMINANT OR TRACER:	fly ash leachate
GEOLOGY:	marine dep., Floridian Aquifer
AQUIFER MATERIAL:	silt, sand, clay, frac. limestone
SOURCE(S) OF FUNDING:	EPRI, EPA
ORGANIZATIONS INVOLVED:	Tetra Tech Inc., Arthur D. Little (ADL)
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LATUQUE SITE
DESCRIPTION:	waste pulp liquor contam., disposal pit
LOCATION:	LaTuque, Quebec
CONTAMINANT OR TRACER:	tannin, lignin, organic compounds
GEOLOGY:	fluvial
AQUIFER MATERIAL:	sand
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	University of Waterloo
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LBL FRACTURE SITES
DESCRIPTION:	organic contaminant migration in fractured-porous rock
LOCATION:	Lawrence Berkley Lab, California
CONTAMINANT OR TRACER:	organic pollutants
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-RSKERL
ORGANIZATIONS INVOLVED:	EPA, LBL
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME: LIPARI LANDFILL

DESCRIPTION: landfill leachate, site remediation
LOCATION: adjacent to Pitman and Glassboro, New Jersey
CONTAMINANT OR TRACER: numerous organic compounds, ether
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: EPA
ORGANIZATIONS INVOLVED: EPA Reg. 2, Radian Corp., Camp Dresser & Mckee, Inc.
NUMBER OF MODEL APPLICATIONS: 1

SITE NAME: LIQUIFIED FUEL SPILL TEST SITE

DESCRIPTION: study accidental release of various hazardous liquids
LOCATION: Nevada Test Site
CONTAMINANT OR TRACER: various hazardous liquids, liq. gaseous fuels
GEOLOGY: dry lake bed
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: DOE
ORGANIZATIONS INVOLVED: LLNL
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: LIVINGSTON SITE

DESCRIPTION: aldicarb contamination, sat. and unsat.
LOCATION: central California
CONTAMINANT OR TRACER: aldicarb
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: Union Carbide- Research Triangle
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME:	LLNL SITE
DESCRIPTION:	VOCs, groundwater migration
LOCATION:	Livermore, California
CONTAMINANT OR TRACER:	seven VOCs
GEOLOGY:	alluvial sed.
AQUIFER MATERIAL:	sand, gravel, clay, silt
SOURCE(S) OF FUNDING:	LLNL
ORGANIZATIONS INVOLVED:	LLNL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LLNL SITE 300
DESCRIPTION:	tritium contam. in vicinity of landfills and exp. facil
LOCATION:	LLNL, California
CONTAMINANT OR TRACER:	tritium
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	LLNL
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	LONG ISLAND
DESCRIPTION:	aldicarb residue in soil and groundwater, nonpoint contam.
LOCATION:	Long Island, New York
CONTAMINANT OR TRACER:	aldicarb, nonpoint contaminants
GEOLOGY:	glacial moraine, outwash
AQUIFER MATERIAL:	sand, gravel, clay
SOURCE(S) OF FUNDING:	Union Carbide, EPA
ORGANIZATIONS INVOLVED:	Cornell Univ., Union Carbide, USGS, EPA
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LOS ALAMOS
DESCRIPTION:	environmental monitoring studies
LOCATION:	LANL
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	LANL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LOVE CANAL
DESCRIPTION:	landfill
LOCATION:	Niagara Falls, New York
CONTAMINANT OR TRACER:	mixed landfill wastes
GEOLOGY:	glacial till, Lockport Dolomite
AQUIFER MATERIAL:	sand, silt, clay, till
SOURCE(S) OF FUNDING:	EPA-Res. Triangle Pk.
ORGANIZATIONS INVOLVED:	GCA Corp., GeoTrans Inc.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	LOVIISA POWER STATION
DESCRIPTION:	tracer meas. by radioactive isotopes
LOCATION:	Finland
CONTAMINANT OR TRACER:	reactor wastes
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Helsinki, Finland
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	MADE
DESCRIPTION:	macrodispersion exper., solid wastes, utility disposal
LOCATION:	Columbus AFB, Mississippi
CONTAMINANT OR TRACER:	seven conservative tracers
GEOLOGY:	alluv., Pleist. terrace
AQUIFER MATERIAL:	sand, gravel, clay, silt
SOURCE(S) OF FUNDING:	EPRI
ORGANIZATIONS INVOLVED:	Tenn. Valley Authority, MIT
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MALIGNE KARST
DESCRIPTION:	field fluorometry in a karst aquifer system
LOCATION:	Alberta, Canada
CONTAMINANT OR TRACER:	field fluorometry tracing
GEOLOGY:	faulted rock, Can. Rocky Mtns.
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of Western Ontario
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MANTECA SITE
DESCRIPTION:	aldicarb contamination, sat. and unsat.
LOCATION:	central California
CONTAMINANT OR TRACER:	aldicarb
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Union Carbide- Research Triangle
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MAXEY FLATS SITE
DESCRIPTION:	low-level radioactive waste burial
LOCATION:	Maxey Flats, Kentucky
CONTAMINANT OR TRACER:	low-level radioactive wastes
GEOLOGY:	fractured sedimentary rocks
AQUIFER MATERIAL:	shale, sandstone
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MEREDOSIA SITE
DESCRIPTION:	fertilizer contam. plume, geochem., remediation
LOCATION:	south of Meredosia, Illinois
CONTAMINANT OR TRACER:	ammonia, sulfate, potassium, chloride
GEOLOGY:	alluv., glacial, Penn. BR
AQUIFER MATERIAL:	sand, gravel, sed. BR
SOURCE(S) OF FUNDING:	ISWS
ORGANIZATIONS INVOLVED:	ISWS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	MOBILE SITE
DESCRIPTION:	tracer tests, dispersion
LOCATION:	north of Mobile, Alabama
CONTAMINANT OR TRACER:	sodium bromide
GEOLOGY:	Quaternary terrace
AQUIFER MATERIAL:	sand, clay
SOURCE(S) OF FUNDING:	USEPA-RSKERL
ORGANIZATIONS INVOLVED:	Auburn Univ.
NUMBER OF MODEL APPLICATIONS:	9

SITE NAME:	MOFFETT AIR STATION
DESCRIPTION:	trichloroethylene plume, biotransformation
LOCATION:	San Francisco Bay, California
CONTAMINANT OR TRACER:	trichloroethylene
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	RSKERL-EPA, Stanford
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MONTOUR TEST CELL
DESCRIPTION:	leachate chemistry study, fly ash
LOCATION:	Batelle, Pacific NW
CONTAMINANT OR TRACER:	fly ash leachate
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA, EPRI
ORGANIZATIONS INVOLVED:	EPRI SWES, Battelle NW
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MONTPELIER SITE
DESCRIPTION:	coal flyash landfill, ground water contamination
LOCATION:	near Montpelier, Iowa
CONTAMINANT OR TRACER:	sulfate, Se, As, flyash leachate
GEOLOGY:	Mississippi R. bluffs
AQUIFER MATERIAL:	loess, till, SS
SOURCE(S) OF FUNDING:	USGS, Univ. of Iowa
ORGANIZATIONS INVOLVED:	IL. SWS, Univ. of Iowa
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	MORENO FIELD SITE
DESCRIPTION:	unsaturated-zone field exper., ponded intro. tracers
LOCATION:	-0-
CONTAMINANT OR TRACER:	tracers
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA, EPRI, S. CA Edison
ORGANIZATIONS INVOLVED:	Univ. of CA Riverside, EPRI
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NDBRINE SITES
DESCRIPTION:	oil-and-gas brine leachate
LOCATION:	north-central North Dakota
CONTAMINANT OR TRACER:	oil-and-gas brine
GEOLOGY:	glacial, Fox Hills Fm.
AQUIFER MATERIAL:	sand, gravel, clay, sandstone
SOURCE(S) OF FUNDING:	NDWRRI
ORGANIZATIONS INVOLVED:	NDMMRRI, NDGS, ND State Univ.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NDCOAL MINE SITES
DESCRIPTION:	determine hydrogeochemistry of spoils settings
LOCATION:	North Dakota
CONTAMINANT OR TRACER:	coal mining spoils
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	US Bureau of Mines
ORGANIZATIONS INVOLVED:	MMRRI
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	NEBRASKA NON-POINT
DESCRIPTION:	nonpoint pollution for agricultural chemicals, 6 areas
LOCATION:	High Plains Aquifer Nebraska
CONTAMINANT OR TRACER:	agricultural chemicals
GEOLOGY:	High Plains Aquifer
AQUIFER MATERIAL:	unconsolidated sed.
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NEVADA TEST SITE
DESCRIPTION:	monitoring radioactive waste
LOCATION:	Nevada
CONTAMINANT OR TRACER:	radioactive waste, munitions waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	DOE
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NEW JERSEY ARSENAL
DESCRIPTION:	disposal of metal-plating wastes at an arsenal
LOCATION:	north-central NJ
CONTAMINANT OR TRACER:	chlorinated solvents, TCE
GEOLOGY:	glacial drift, till, outwash
AQUIFER MATERIAL:	sand, gravel, silt, clay
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS-NJ
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NM DESERT SITE
DESCRIPTION:	unsat. flow in stratified soils, effective hydraul. conductivity
LOCATION:	New Mexico
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	MIT
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	NMSU EXP STATION
DESCRIPTION:	mes. of variability of hydraulic param. in unsat. zone
LOCATION:	40 km northeast of Las Cruces, New Mexico
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	het. layered soil-alluvial
AQUIFER MATERIAL:	unsat. soil
SOURCE(S) OF FUNDING:	USNRC
ORGANIZATIONS INVOLVED:	MIT, NM State Univ., PNL, USNRC
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	NORTH SWISS RIVER
DESCRIPTION:	accidental tritium release, infil. of river to gw
LOCATION:	northern Switzerland
CONTAMINANT OR TRACER:	tritium
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Swiss Federal Institutes
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	OGCFRACTURE STUDY
DESCRIPTION:	organic contaminant migration in fractured porous rock
LOCATION:	-0-
CONTAMINANT OR TRACER:	organic pollutants
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-RSKERL
ORGANIZATIONS INVOLVED:	OGC, EPA
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	OHIO CORN
DESCRIPTION:	pesticide leaching in agricultural soils
LOCATION:	Ohio
CONTAMINANT OR TRACER:	aldicarb
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-ERL-Athens
ORGANIZATIONS INVOLVED:	EPA-Athens, Union Carbide-RTP-NC
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	ORACLE SITE
DESCRIPTION:	movement of fluids in low-perm. fractured rock
LOCATION:	Oracle, Arizona
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	granite
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ORNL FIELD FAC
DESCRIPTION:	research field facility for subsurface transport
LOCATION:	Oak Ridge, Tennesee
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	ORNL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ORNL FIELD SITES
DESCRIPTION:	monitoring of HW facil., macropores, spatial variability
LOCATION:	Oak Ridge, Tennessee
CONTAMINANT OR TRACER:	chemical pollutants
GEOLOGY:	Cambrian, Ordovician sequence
AQUIFER MATERIAL:	shale, limestone, sandstone
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	ORNL, Geraghty & Miller
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ORNL HYDROFRACTURE FACILITY
DESCRIPTION:	radioactive waste disposal, fractured rock system
LOCATION:	Oak Ridge, Tennessee
CONTAMINANT OR TRACER:	radioactive waste
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	ORNL
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	OTIS AFB
DESCRIPTION:	natural gradient tracer study
LOCATION:	Cape Cod, Massachusetts
CONTAMINANT OR TRACER:	bromide, lithium bromide, fluoride, molybdenum
GEOLOGY:	glacial outwash, drift
AQUIFER MATERIAL:	sand, gravel, clay, silt
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	OVIEDO SITE
DESCRIPTION:	aldicarb contamination, sat. and unsat.
LOCATION:	northeast of Orlando, Florida
CONTAMINANT OR TRACER:	aldicarb
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	State of FL
ORGANIZATIONS INVOLVED:	EPA-Athens, Univ. of FL, State of FL
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	PALO DURO BASIN
DESCRIPTION:	regional, high-level radioactive waste disposal
LOCATION:	north-east Texas
CONTAMINANT OR TRACER:	site evaluation, radioactive waste disposal
GEOLOGY:	Ogallala, Penn, Perm, Triassic
AQUIFER MATERIAL:	sand, clay, gravel, sed sequence
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	INTERA, TBEG, SWEC, ONWI-Battelle
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	PALOS FOREST PRESERVE
DESCRIPTION:	low-level radioactive waste burial, monitoring, remedial
LOCATION:	north-east Illinois
CONTAMINANT OR TRACER:	tritium, radioactive waste, metallurgical waste
GEOLOGY:	glacial, Silurian dolomite
AQUIFER MATERIAL:	clay, silt, sand
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS, Argonne Nat. Lab
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME: PANTANO SITE

DESCRIPTION: virus tracer studies
LOCATION: Pantano, Washington
CONTAMINANT OR TRACER: viruses
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: EPA
ORGANIZATIONS INVOLVED: EPA-RSKERL
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: PARADOX BASIN

DESCRIPTION: regional, high-level radioactive waste, Gibson Dome
LOCATION: south-east Utah
CONTAMINANT OR TRACER: evaluation of potential repository site
GEOLOGY: Hermosa & Cutler Groups
AQUIFER MATERIAL: marine carb, some clastics, salt
SOURCE(S) OF FUNDING: DOE
ORGANIZATIONS INVOLVED: WCC, INTERA, ONWI-Battelle
NUMBER OF MODEL APPLICATIONS: 2

SITE NAME: PARRIS ISLAND

DESCRIPTION: core study of potential denitrification rates
LOCATION: Parris Island, southeastern coastal
CONTAMINANT OR TRACER: nitrate
GEOLOGY: uncon. Pleist., Floridan aquifer
AQUIFER MATERIAL: sand, silt, clay, limestone
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: Univ. of SC, USGS
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: PEASE AFB

DESCRIPTION: analyses of screening parameters for water quality
LOCATION: Pease Air Force Base, New Hampshire
CONTAMINANT OR TRACER: TOX, TOC, O & G, phenols, hyanide, organic pollutants
GEOLOGY: -0-
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: Weston Inc.
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: PENSACOLA

DESCRIPTION: phenolic contamination
LOCATION: NW FL-western panhandle
CONTAMINANT OR TRACER: creosote, pentachlorophenol
GEOLOGY: fluvial, deltaic sediments
AQUIFER MATERIAL: Sand, clay and gravel
SOURCE(S) OF FUNDING: USGS
ORGANIZATIONS INVOLVED: USGS, Am. Creosote Works, Pensacola-City, FL. Dept. Env. R
NUMBER OF MODEL APPLICATIONS: 1

SITE NAME: PICEANE BASIN

DESCRIPTION: oil-shale development
LOCATION: north-west Colorado
CONTAMINANT OR TRACER: oil-shale site evaluation
GEOLOGY: -0-
AQUIFER MATERIAL: SS, SH, LS, marlstone
SOURCE(S) OF FUNDING: DOE, EPA
ORGANIZATIONS INVOLVED: USGS
NUMBER OF MODEL APPLICATIONS: 3

SITE NAME: PLAINS SITE

DESCRIPTION: movement and fate of agricultural chemicals
LOCATION: Plains, Georgia
CONTAMINANT OR TRACER: pesticides, nitrogen compounds
GEOLOGY: unconsolidated
AQUIFER MATERIAL: sand, gravel, clay
SOURCE(S) OF FUNDING: USGS
ORGANIZATIONS INVOLVED: USGS-GA, USDA-ARS, EPA
NUMBER OF MODEL APPLICATIONS: 99

SITE NAME: POWERTON SITE

DESCRIPTION: fly ash disposal (6-site study)
LOCATION: north-central Illinois
CONTAMINANT OR TRACER: fly ash leachate
GEOLOGY: alluv., glacial, Penn. BR
AQUIFER MATERIAL: sand, gravel, clay
SOURCE(S) OF FUNDING: EPRI, EPA
ORGANIZATIONS INVOLVED: Tetra Tech, Incorp., Arther D. Little (ADL)
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME:	PRICE'S LANDFILL
DESCRIPTION:	landfill, mixed contamination
LOCATION:	Atlantic City, NJ
CONTAMINANT OR TRACER:	acetone, acid chloroform, hexane, cesspool waste, oil, xylene
GEOLOGY:	CohanseySnd, Kirkwood FM, Tertiar
AQUIFER MATERIAL:	sand, clay
SOURCE(S) OF FUNDING:	EPA, ACMUA
ORGANIZATIONS INVOLVED:	Paulus, Sokolowski and Sartor(PSS), NJDEP, Pinder and B
NUMBER OF MODEL APPLICATIONS:	2

SITE NAME:	PRINCE EDWARD ISLAND
DESCRIPTION:	pesticide contam. in groundwater below potato fields
LOCATION:	Prince Edward Island, Ontario
CONTAMINANT OR TRACER:	aldicarb
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Natl. Research Inst., Burlington, Ontario
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	RADIAN FIELD SITES
DESCRIPTION:	collect field data to validate geochemical models
LOCATION:	-0-
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPRI
ORGANIZATIONS INVOLVED:	Radian Corp.
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	REGINA SITE
DESCRIPTION:	transformer manufacturing, PCB spill
LOCATION:	Regina, Saskatchewan Canada
CONTAMINANT OR TRACER:	PCBs, TCBs, Inerteen 70-30
GEOLOGY:	lacustrine, glacial till
AQUIFER MATERIAL:	clay, silt, fine sand
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	-0-
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	RICHTON DOME
DESCRIPTION:	regional, high-level radioactive waste
LOCATION:	south-east Mississippi
CONTAMINANT OR TRACER:	evaluation of potential repository site
GEOLOGY:	quaterary-alluv., Tertiary
AQUIFER MATERIAL:	sand, gravel, silt, clay, SS, LS, SH
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	INTERA, Batelle-ONWI
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	RIFLE SITE
DESCRIPTION:	definition of reservoir and hydraulic fracture system
LOCATION:	Rifle, Colorado
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Sandia Ntl. Labs
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	ROCKY MTN. ARSENAL
DESCRIPTION:	dissolved chemical transport, unlined disposal ponds
LOCATION:	north-central Colorado
CONTAMINANT OR TRACER:	liquid industrial wastes, Cl
GEOLOGY:	alluvial
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	R Mnt. A, Core of Eng., CO Dept. of Health
NUMBER OF MODEL APPLICATIONS:	3

SITE NAME:	SAN JOAQUIN VALLEY
DESCRIPTION:	regional study
LOCATION:	San Joaquin Valley, CA
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of CA
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	SAND RIDGE STATE FOREST
DESCRIPTION:	dye tracer study
LOCATION:	west-central Illinois
CONTAMINANT OR TRACER:	fluorescent dyes
GEOLOGY:	Wisconsin glacial
AQUIFER MATERIAL:	FM sand, C sand, gravel
SOURCE(S) OF FUNDING:	USEPA-OSU
ORGANIZATIONS INVOLVED:	IL SWS & ENR
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	SAVANNAH RIVER
DESCRIPTION:	migration of radionuclides in soil
LOCATION:	Aiken, South Carolina
CONTAMINANT OR TRACER:	radionuclides
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	DOE
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	SEYMOUR SITE
DESCRIPTION:	hazardous waste site
LOCATION:	Seymour, Indiana
CONTAMINANT OR TRACER:	solvents, metal finishing wastes
GEOLOGY:	alluvial, glacial
AQUIFER MATERIAL:	sand, gravel, clay, silt
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	EPA
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	SHEFFIELD SITE
DESCRIPTION:	low-level radioactive-waste burial site
LOCATION:	Sheffield, Illinois
CONTAMINANT OR TRACER:	low-level radioactive waste
GEOLOGY:	glacial till, outwash
AQUIFER MATERIAL:	sand, clay, silt
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	SHERCO SITE
DESCRIPTION:	sludge/flyash waste pond
LOCATION:	Sherburne County, Minnesota
CONTAMINANT OR TRACER:	flyash leachate
GEOLOGY:	glacial, PreCambrian rock
AQUIFER MATERIAL:	sand, gravel, granite
SOURCE(S) OF FUNDING:	EPRI, EPA
ORGANIZATIONS INVOLVED:	Tetra Tech Inc., Arthur D. Little (ADL)
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	SHULLSBURG SITE
DESCRIPTION:	zinc-lead mine contamination
LOCATION:	Shullsburg, Wisconsin
CONTAMINANT OR TRACER:	sulfate
GEOLOGY:	Galena-Platteville Fm.
AQUIFER MATERIAL:	dolomite, limestone, shale
SOURCE(S) OF FUNDING:	WI-WRC, Univ. WI, AGU
ORGANIZATIONS INVOLVED:	Univ. of Wisconsin
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	SOIL MOISTURE
DESCRIPTION:	soil moisture diffusivity char. of loamy to silty soil
LOCATION:	Ardeche Basin, France
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of Utrecht, Netherlands
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	SOILS LAB 1
DESCRIPTION:	sorption and transport of toxic organic substance
LOCATION:	-0-
CONTAMINANT OR TRACER:	organic chemicals
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA-RSKERR
ORGANIZATIONS INVOLVED:	EPA, Univ. of Florida
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	STRIPA SITE
DESCRIPTION:	geochemistry of trace elements for radioactive waste disposal
LOCATION:	Stripa Mine in Sweden
CONTAMINANT OR TRACER:	radioactive wastes
GEOLOGY:	fract. crystalline rock
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	STRIPMINE SITES
DESCRIPTION:	environmental effects of stripmining sites
LOCATION:	Montana, Wyoming, North Dakota
CONTAMINANT OR TRACER:	mining spoils
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	USEPA
ORGANIZATIONS INVOLVED:	MMRRI
NUMBER OF MODEL APPLICATIONS:	99

SITE NAME:	STROUDSBURG SITE
DESCRIPTION:	abandoned illuminating gas plant
LOCATION:	Stroudsburg, Pennsylvania
CONTAMINANT OR TRACER:	coal tar wastes
GEOLOGY:	-0-
AQUIFER MATERIAL:	gravel, fine silty sand
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	-0-
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	SWES PROJECTS
DESCRIPTION:	EPRI studies on trans. and fate of utility wastes
LOCATION:	-0-
CONTAMINANT OR TRACER:	utility wastes
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPRI
ORGANIZATIONS INVOLVED:	EPRI
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	TEXAS TECH SITE
DESCRIPTION:	tracing migration of chemicals from nonpoint sources
LOCATION:	Lubbock, Texas
CONTAMINANT OR TRACER:	salt solution injection
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	TX Wat. Dev. Board
ORGANIZATIONS INVOLVED:	Texas Tech. Univ.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	TOKYO SITE
DESCRIPTION:	sewage sludge applications to soils
LOCATION:	60 km. northeast of Tokyo, Japan
CONTAMINANT OR TRACER:	sewage sludge application
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Natl. Inst. for Env. Studies Tsukuba, Ibaraki
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	TRACER RESEARCH SITE
DESCRIPTION:	shallow soil gas study, TCE plume
LOCATION:	-0-
CONTAMINANT OR TRACER:	TCE, gaseous phase
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Tracer Research Corp., Tucson, AZ
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	TRAVERSE CITY
DESCRIPTION:	gasoline spill-underground storage
LOCATION:	Traverse City, Michigan- US Coast G
CONTAMINANT OR TRACER:	gasoline
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	EPA-RSKERL, Rice University
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	TWIN-CITIES
DESCRIPTION:	army ammunition plant
LOCATION:	New Brighton, MN
CONTAMINANT OR TRACER:	toxic organic substances
GEOLOGY:	glacial, PrarieduChienJordan
AQUIFER MATERIAL:	sand, gravel, clay, silt, dolomite
SOURCE(S) OF FUNDING:	USGS, USEPA, MPCA
ORGANIZATIONS INVOLVED:	MN Poll. Control Agency (MPCA), USEPA, USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	UNITED CHROME
DESCRIPTION:	chromium plating waste, United Chrome Products
LOCATION:	Washington State
CONTAMINANT OR TRACER:	chromium
GEOLOGY:	Quaternary alluv.
AQUIFER MATERIAL:	clay, silt, sand, gravel
SOURCE(S) OF FUNDING:	EPA Reg. 10-Seattle
ORGANIZATIONS INVOLVED:	OGC, E&E
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	VESICOL SITE
DESCRIPTION:	hazardous waste site
LOCATION:	western Tennessee
CONTAMINANT OR TRACER:	chemical wastes from pesticide manufacturing
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	EPA, Vesicol Chem. Corp.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	VICTORIA PROVINCE
DESCRIPTION:	hydrogeological study in crystalline rock, fractures
LOCATION:	Victoria Province, United Kingdom
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	crystalline basement rock
AQUIFER MATERIAL:	granite, gneiss
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Hydrotechnica Shrewsbury, UK
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	VOLK AIR BASE
DESCRIPTION:	in-situ soil washing, removal of hydrocarbons
LOCATION:	Volk Air National Guard Base, WI
CONTAMINANT OR TRACER:	hydrocarbons, chlorinated hydrocarbons
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	EPA
ORGANIZATIONS INVOLVED:	EPA, USAF, Mason and Hanger-Silas Mason Co., Inc.
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME: WASTE ISOLATION PLANT

DESCRIPTION: evaluate geohydrology of a setting for nuclear waste disposal
LOCATION: 40 km east of Carlsbad, New Mexico
CONTAMINANT OR TRACER: nuclear waste
GEOLOGY: Permian rock, faulted area
AQUIFER MATERIAL: sed. rocks, salt
SOURCE(S) OF FUNDING: DOE
ORGANIZATIONS INVOLVED: DOE, USGS
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: WELDON SPRING SITES

DESCRIPTION: radioactive waste disposal sites
LOCATION: 48 km west of St. Louis, Missouri
CONTAMINANT OR TRACER: radioactive waste
GEOLOGY: carbonate rocks
AQUIFER MATERIAL: -0-
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: USGS
NUMBER OF MODEL APPLICATIONS: -0-

SITE NAME: WESTERN INDIA

DESCRIPTION: injection well tests in an alluvial aquifer
LOCATION: western India
CONTAMINANT OR TRACER: -0-
GEOLOGY: het. alluvial mat., layered
AQUIFER MATERIAL: sand, gravel, clay
SOURCE(S) OF FUNDING: -0-
ORGANIZATIONS INVOLVED: Univ. of Birmingham, Central GW Board-Ahmedabad Indi
NUMBER OF MODEL APPLICATIONS: 1

SITE NAME: WESTERN PROCESSING

DESCRIPTION: industrial hazardous waste, 26 priority pollutants
LOCATION: Kent, Washington
CONTAMINANT OR TRACER: TCE, 26 priority pollutants
GEOLOGY: Green R. Flood Plain
AQUIFER MATERIAL: sand, silt, peaty silt, clay
SOURCE(S) OF FUNDING: EPA-MERL Cinc.
ORGANIZATIONS INVOLVED: Battelle
NUMBER OF MODEL APPLICATIONS: 1

SITE NAME:	WINDERL SITE
DESCRIPTION:	evaluate brine and oil-gas drilling fluid disposal site
LOCATION:	North Dakota
CONTAMINANT OR TRACER:	oil brine, oil-gas drilling fluid
GEOLOGY:	glaciofluvial, till
AQUIFER MATERIAL:	sand, clay, silt
SOURCE(S) OF FUNDING:	NDWRRI
ORGANIZATIONS INVOLVED:	NDGS, Univ. of ND, MMRRI
NUMBER OF MODEL APPLICATIONS:	1

SITE NAME:	WISCONSIN FRACTURE SYSTEMS
DESCRIPTION:	anisotropy, directional connec., porosity
LOCATION:	60 sites throughout Wisconsin
CONTAMINANT OR TRACER:	-0-
GEOLOGY:	-0-
AQUIFER MATERIAL:	-0-
SOURCE(S) OF FUNDING:	-0-
ORGANIZATIONS INVOLVED:	Univ. of Wisconsin
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	WOOD RIVER JUNCTION
DESCRIPTION:	contaminant plume, uranium recovery plant
LOCATION:	Wood River Junction, Rhode Island
CONTAMINANT OR TRACER:	strontium 90, radionuclides, chemical solutes
GEOLOGY:	glacial till, outwash
AQUIFER MATERIAL:	sand, gravel, gneiss
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	YAKIMA SITE
DESCRIPTION:	processes study of gasoline and diesel oil contamination
LOCATION:	Yakima, Washington
CONTAMINANT OR TRACER:	gasoline, diesel oil
GEOLOGY:	unconsol.-Holocene age
AQUIFER MATERIAL:	sand, gravel
SOURCE(S) OF FUNDING:	USGS
ORGANIZATIONS INVOLVED:	USGS, OGC
NUMBER OF MODEL APPLICATIONS:	-0-

SITE NAME:	YUCCA MOUNTAIN
DESCRIPTION:	nuclear waste storage invest., fractured flow
LOCATION:	Nevada Test Site
CONTAMINANT OR TRACER:	nuclear waste
GEOLOGY:	Topopah Spring Member
AQUIFER MATERIAL:	ash-flow tuffs
SOURCE(S) OF FUNDING:	DOE
ORGANIZATIONS INVOLVED:	Sandia Nat. Lab., LBL
NUMBER OF MODEL APPLICATIONS:	3

APPENDIX C

ENTRY FORM FOR SATURN DATABASE

LEVEL 1: GENERAL SITE-SPECIFIC INFORMATION

Site Identification
Site Information
Geographic Information
Geohydrological Site Data
Other Site Data Availability
Model Applications/Testing

LEVEL 2: INFORMATION ON PARTICULAR STUDIES AT A SITE

Study Identification/Information
Type of Measurements Made
Type of Measurements/Tasks Performed

LEVEL 3: INFORMATION ON INDIVIDUAL TASKS (OR INVESTIGATIONS)

Task Identification
Dataset Information
General Information/Documented QA/QC

Tasks:

- Field - Chemistry - In-situ Monitoring
- Field - Chemistry - Sampling and Analysis
- Task: Field - Chemistry - Tracer Test
- Field - Hydrogeology - Water Level Observation
- Field - Hydrogeology - Aquifer Test (Aquifer Parameters)
- Field - Soils - Water Content/Soil Moisture Tension
- Field - Soils - Infiltration/Moisture Movement
- Field - Soils - Measurement of Soil Parameters
- Task: Laboratory Experiments (Y/N)

PUBLICATIONS

USER REFERENCES

****

Instructions:

The form asks for three types of information: numerical (#), text (text), or Boolean (Y/N). The variable type on the form is consistent with that used in the SATURN database. If a field is not filled in, it is assumed that the value is "no" or "0"; if no information is available the field should be left blank.

LEVEL 1: GENERAL AND SITE-SPECIFIC INFORMATION

Site Identification

Site Name (text): ______________________

General Site Research Coordination (text)

Organization: ______________________

Department: ______________________

Address: ______________________

____

____

Contact Person(s): ______________________

Site Owner/Manager (text)

Organization: ______________________

Contact Person: ______________________

Site Information

Contamination Present (Y/N): ___

Source of Contamination if Present (Y/N)

Waste Disposal

Industrial Impoundment: ___

Illegal Dumping: ___

Septic Tanks: ___

Deep Well Injection: ___

Industrial/Municipal Landfill: ___

Wastewater Treatment: ___

Radioactive Waste Disposal: ___

Oil and Gas Field Brines: ___

Accidental Pollution

Surface Spills: ___ Leaking Storage Tanks: ___
Pipeline Leaks: ___ Acid Mine Drainage: ___

Non-Point Pollution

Agricultural Chemicals: ___ Feedlots: ___
Irrigation: ___ In-situ Mining: ___
Salt-water Intrusion: ___
Industrial Impoundment: ___ Industrial/Municipal Landfill: ___
Illegal Dumping: ___ Wastewater Treatment: ___
Septic Tanks: ___ Radioactive Waste Disposal: ___
Deep Well Injection: ___ Oil and Gas Field Brines: ___

Accidental Pollution

Surface Spills: ___ Leaking Storage Tanks: ___
Pipeline Leaks: ___ Acid Mine Drainage: ___

Non-Point Pollution

Agricultural Chemicals: ___ Feedlots: ___
Irrigation: ___ In-situ Mining: ___
Salt-water Intrusion: ___

Other sources (text): ______________________________

Kind of Pollutants (Y/N)

Organics

Aromatic: ___ Oxygenated: ___ Halogenated: ___

Anorganics

Heavy Metals: ___ Other metals: ___ Nitrates: ___
Phosphates: ___ Sulfates: ___ Cyanides: ___
Chlorides: ___ Radionuclides: ___ Other: ___

Tracers: ___

Major Pollutants (text):

_______________ _______________ _______________
_______________ _______________ _______________
_______________ _______________ _______________

Pollution History (pollution amounts, rates; time period sources were active):

____
____
____
____

Geographic Information

Coordinates

Coordinate System (text): ________
Latitude (#): ________
Longitude (#): ________

Size of Site (#): ________ Units (text): ________

Location (text)

Nearest city/town: ______________________
County: ______________________
State/Province: ______________________
Country: ______________________

Elevation

Maximum Elevation (#): ________ Units (m or ft): ___

Minimum Elevation (#): ________ Units (m or ft): ___

Average Slope Gradient (#): ________ (fraction)

Topographical Setting (Y/N)

Flatland: ___ Depression: ___ Valley Bottom: ___

Stream Channel: ___ Hillside: ___ Hilltop: ___

Terrace: ___ Plateau: ___ Marshland: ___

Dunes: ___ Polder: ___

Other (text): ______________________________________

Geohydrological Site Data

Lithology (Y/N)

Sand: ___ Sand and Gravel: ___ Sandstone: ___

Sands and Silts: ___ Clay: ___ Till: ___ Shale: ___

Limestone: ___ Igneous/Metamorphic Rock: ___

Other (text): ______________________________________

Short Geological Description (lithology, layer thickness, major formations, karst, fractures):

____

____

____

____

____

____

____

____

____

____

Hydrogeology (Y/N)

Single Aquifer: ___
Confined: ___
Semi-Confined: ___
Unconfined (shallow water-table): ___
Unconfined (deep water-table): ___
Porous: ___
Fractured: ___

Multi-Aquifer: ___

Top-Layer

Spatial Continuous: ___
Semi-Confined: ___
Unconfined (shallow water-table): ___
Unconfined (deep water-table): ___
Porous: ___
Fractured: ___

Lower Layers

Spatial Continuous: ___
Porous: ___
Fractured: ___

Short Hydrogeological Description (schematization, aquifer names, aquifer and aquitard thickness, karst, fracture density and orientation):

____
____
____
____
____
____
____
____

Soil Types (Y/N)

Thin or Absent: ___

Non-Shrinking Clay: ___
Clayey-Loam: ___
Sandy Loam: ___
Gravel: ___
Peat: ___

Shrinking Clays: ___
Silty Loam: ___
Sand: ___
Muck: ___

Other (text): ____

Short Soil Description (layer names, thickness, soil type, presence of macro-pores, root density, etc.):

____
____
____
____
____
____
____
____

Surface Water Presence at or near Site (Y/N)

Streams/Creeks: ___
Lakes/Reservoirs: ___
Sea/Ocean: ___

Canals/Ditches: ___
Impoundments/Ponds: ___
Fresh-Water Estuary: ___

Short Description of Surface Water Features (relative importance, how does it relates to site/plume, flow characteristics, perennial, ephemeral, etc.):

____

____

____

____

____

____

____

____

____

Groundwater Levels

Data Availability (Y/N)

Long-term: ___ Short-term: ___

Continuous: ___

Observation Wells (#): _____

Average Depth to Water-table (#): _____ m or ft

Average Regional Groundwater Velocity (#): _____ m/d or ft/d

Is Vertical Flow Present? (Y/N): ___

Sinks/Sources in Area (Y/N)

Municipal Wellfield: ___ Industrial Discharge Well: ___

Private Wells: ___ Shallow Recharge Wells: ___

Deep Waste Injection Wells: ___ Pumping/Injecting Remediation: ___

Short Description of Groundwater Flow Characteristics (location and discharge rates of wells, groundwater flow direction and rates, etc.):

____
____
____
____
____
____
____

<u>Other Site Data Availability</u>

Meteorological Data (Y/N)

Precipitation: ___	Evaporation: ___
Evapotranspiration: ___	Air Temperature: ___
Relative Humidity: ___	Wind Velocities: ___

Other (text): ____

Land Use (Y/N)

Grassland: ___	Crops: ___	Forest: ___
Rangeland: ___	Prairie: ___	Desert: ___
Build Area: ___	Roads: ___	Waste Disposal: ___

Other (text): ____

Comments: ____
____
____
____
____

Previous Model Testing and Modeling Applications

Site Is Known for Model Applications (Y/N): ___
Study Data Have Been Used for Model Testing (Y/N): ___

Model(s) Used at Site/Study (text)

Acronym and IGWMC-Key: ______________________________
____
____

(Add references describing application/testing to user category)

Characterization of Applications/Testing (Y/N)

Saturated Flow: ___
Variably Saturated Flow: ___
Infiltration/Recharge: ___
Buoyancy Flow: ___
Multi-phase Flow: ___
Advective Transport: ___
Dispersion: ___
Soil/Water/Solute Interaction: ___
Vapor Diffusion: ___
Gas Transport in Soils: ___
Biodegradation/Bioremediation: ___
Flow/Transport in Freezing/Thawing Soils: ___
Flow and Consolidation/Subsidence: ___
Salt Water Intrusion: ___

Other (text): ______________________________
____

LEVEL 2: STUDY SPECIFIC INFORMATION

Study Identification/Information

Study Title (text): ______________________________

Study Coordination (text)

Organization: ______________________________

Department: ______________________________

Address: ______________________________

____

____

Coordinator/PI: ______________________________

Contact Person(s): ______________________________

Study Period/Dates (text)

Start: ________ End: ________ IGWMC Check-Date: ________

Major Measurements Made (Y/N)

Solute Concentrations: ___	Groundwater Levels/Drawdowns: ___
Groundwater Velocities: ___	Groundwater Flow/Dicharge Rates: ___
Soil Moisture Content: ___	Soil Tension/Suction: ___
Infiltration/Recharge: ___	Porosity: ___
Soil/Rock Density: ___	Precipitation: ___
Evaporation: ___	Evapotranspiration: ___

Other (Text): ______________________________

____

____

Type of Measurements/Tasks Performed

Primary Tasks (See Level 3 for Detail)

Field Experiments (Y/N): ___

Chemistry: ___

In-situ Groundwater/Soil-Water Monitoring: ___
Groundwater/Soil-Water Sampling and Analysis: ___
Tracer Test: ___

Hydrogeology: ___

Water Level Observation: ___
Aquifer Test (Aquifer Parameters): ___

Soil Physics: ___

Water Content/Soil Moisture Tension: ___
Infiltration/Moisture Movement: ___
Measurement of Soil Parameters: ___

Laboratory Experiments (Y/N): ___

Secondary Tasks (No further detail provided)

Laboratory Tests (Y/N): ___

Rock Mechanical Properties: ___
Rock Hydraulic Properties: ___
Soil-Physical Properties: ___
Soil/Rock Chemical Properties: ___

Borehole (Rock) Sampling

Number of Sampling Locations (#): _____
Number of Sampling Points (#): _____

Total Number of Samples Taken (#): _____

Soil Sampling

Number of Sampling Locations (#): _____
Number of Sampling Points (#): _____
Total Number of Samples Taken (#): _____

Geophysical Surveys (Y/N): ___

Surface Methods (Y/N): ___

Surface Resistivity: ___
Reflection Seismics: ___
Low-Induction Conductivity: ___
Electromagnetic Sounding: ___

Other (text): ____
____

Borehole Methods (Y/N): ___

Spontaneous Potential: ___
Electric Resistivity: ___
Gamma-Gamma: ___
Groundwater Velocity: ___
Temperature: ___
Dip: ___

Natural Gamma: ___
Neutron: ___
Flow Rate: ___
Acoustic Logs: ___
Caliper: ___
Video: ___

Geodetic Surveying (Y/N): ___

Remote Sensing (Y/N): ___

Other (Text): ____

LEVEL 3: TASK-SPECIFIC INFORMATION

Task Identification

Task Title (text): ____

Task Coordination (text)

Organization: ____

Department: ____

Address: ____

____

____

Coordinator/PI: ____

Contact Person(s): ____

Task Period/Dates (text)

Start: ________ End: ________ IGWMC Check-Date:________

Dataset Information

Dataset Name/Acronym: ____

Dataset Structure: ____

____

____

Spatial Data Type (Y/N)

Point: ___ Line: ___ Cell/Element: ___ Grid: ___

Other: ____________________________

Timespan Dataset Coverage; Dates (text)

Start: _______ End: _______ IGWMC Check-Date: _______

Status of Dataset (Y/N)

Operational: ___ Under Development: ___
Under Modification: ___ Closed (Not Further Maintained): ___

Availability for Use/Access (Y/N)

Computerized: ___ Report Only: ___ Data Sheets Only: ___
Unrestricted: ___ Proprietary: ___ Conditional Use: ___

Access Method (Y/N)

On-Line: ___ Batch: ___ Indirect: ___

Transfer Media (Y/N)

Magnetic Tape: ___ Diskette: ___ Printout: ___
Electronic File Transfer: ___

Resident Storage Media (computer, mass storage device):

____

____

DBMS (if applicable; name, version): ___________________________________

Size of Dataset (#)

Logical Records: ________
Bytes per Record: ________
Expected Yearly Growth (# of records): ________

Datatype (Y/N)

Raw: ___ Conditioned: ___ Summarized: ___ Derived: ___

General Information/Documented QA/QC

Available Documentation (Y/N)

Data Report:	___	Equipment Used:	___
Procedures Applied:	___	Field Conditions:	___
Personnel:	___	Data Conditioning:	___
Data Conversions:	___	Data Analysis:	___
Data Transfer and Storage:	___	Applied QA/QC:	___

Documented Dataset Assessment by Source (Y/N)

Consistency:	___	Outliers:	___	Accuracy:	___
Precision:	___	Completeness:	___	Reproducibility:	___
Procedures:	___	Equipment:	___	Personnel:	___

IGWMC Dataset and Documentation Assessment (Y/N)

Consistency:	___	Outliers:	___	Accuracy:	___
Precision:	___	Completeness:	___	Reproducibility:	___
Procedures:	___	Equipment:	___	Personnel:	___
Field Conditions:	___	Data Handling:	___	Data Analysis:	___
Data Report:	___	QA/QC Applied:	___		

Note: IGWMC evaluations, if present, are separately documented.

Level 3 Task: Field - Chemistry - In-situ Monitoring

Soil-Water (Y/N): ___ Groundwater (Y/N): ___

Measurements (Y/N)

Specific Conductance:	___	pH:	___
Temperature:	___	Eh:	___

Other (text): ______________________________________

Monitoring Scenario and Network Design (text): ______________________

Procedures (text): ______________________

Equipment (text): ______________________

Level 3 Task: Field - Chemistry - Sampling and Analysis

Soil-Water (Y/N): ___

Groundwater (Y/N): ___

Sampling Locations (#): _____

Sampling Points (#): _____

Total Number of Samples Taken (#): _____

Sampling Frequency (text): ________________________________

Primary Constituents (text): ________________________________

Sampling Scenario and Network Design (text): ________________________________

Procedures (text): ________________________________

Equipment (text): ________________________________

Level 3 Task: Field - Chemistry - Tracer Test

Sampling Locations (#): _____
Sampling Points (#): _____
Total Number of Samples Taken (#): _____
Sampling Frequency (text): ____________________

Tracer(s) (text): ____________________

Derived Parameters (text): ____________________

Sampling Scenario and Network Design (text): ____________________

Procedures (text): ____________________

Equipment (text): ____________________

Level 3 Task: Field - Hydrogeology - Water Level Observation

Observation Locations (#): _____
Observation Points (#): _____
Total Number of Observations Made (#): _____
Observation Frequency (text): ______________________________

____

Observation Scenario and Network Design (text): ______________________________

____

____

____

____

____

Procedures (text): ______________________________

____

____

____

____

____

Equipment (text): ______________________________

____

____

____

____

____

Level 3 Task: Field - Hydrogeology - Aquifer Test (Aquifer Parameters)

Aquifer Test Locations (#): _____
Total Tests Made (#): _____
Number of Observation Wells (#): _____

Type of Test (Y/N)

Pump Test: ___ Bailer: ___ Slug: ___ Recovery: ___

Other (text): ______________________________

____

Derived Parameters (Y/N)

Transmissivity: ___
Horizontal Hydraulic Conductivity: ___
Vertical Hydraulic Conductivity: ___
Leakage Factor: ___
Storage Coefficient: ___

Other (text): ______________________________

____

Aquifer Test Scenario and Network Design (text): ______________________________

____

____

____

____

____

Equipment (text): ______________________________

____

____

____

____

____

Method of Analysis (text: ______________________________

____

____

____

____

____

Level 3 Task: Field - Soils - Water Content/Soil Moisture Tension

Observation Locations (#): _____

Observation Points (#): _____

Total Number of Observations Made (#): _____

Observation Frequency (text): ______________________________

____

Procedures (text): ______________________________

____

____

____

____

____

Equipment (text): ______________________________

____

____

____

____

____

Level 3 Task: Field - Soils - Infiltration/Moisture Movement

Observation Locations (#): _____

Observation Points (#): _____

Total Number of Observations Made (#): _____

Observation Frequency (text): ____

____

Procedures (text): ____

____

____

____

____

____

Equipment (text): ____

____

____

____

____

____

Level 3 Task: Field - Soils - Measurement of Soil Parameters

Measurement Locations (#): _____

Total Measurements Made (#): _____

Type of Test (text): ______________________________

____

Derived Parameters (text): ______________________________

____

____

Aquifer Test Scenario and Network Design (text): ______________________________

____

____

____

____

____

Equipment (text): ______________________________

____

____

____

____

____

Method of Analysis (text: ______________________________

____

____

____

____

Level 3 Task: Laboratory Experiments (Y/N)

Moisture Movement in Variably Saturated Soils: ___
Water Movement in Saturated Porous Rock: ___
Water Movement in Saturated Fractured Rock: ___
Advective-Dispersive Solute Transport: ___
(Bio)chemical Transformation/Degradation: ___
Immiscible/Multiphase Flow: ___
Transport of VOCs: ___

Type of Information Collected (text): ______________________________

____

____

____

Technique(s) (text): ______________________________

____

____

____

____

____

PUBLICATIONS

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Continue on new page if necessary

USER REFERENCES

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Title: ______________________________

____

____

Author(s): ______________________________

____

Year: ________

Journal: ______________________________

Volume/Issue/Pages: ______________________________

Book/Report Publisher: ______________________________

____

Continue on new page if necessary

☆U.S GOVERNMENT PRINTING OFFICE. 1989-648-163/00316