EPA Report Number
October 1993

FINAL DRAFT

A GEOGRAPHIC INFORMATION SYSTEM FOR THE OHIO RIVER BASIN

by

Walter M. Grayman
W.M. Grayman Consulting Engineer
Cincinnati, Ohio 45229

Sudhir R. Kshirsagar
Global Quality Corp.
Cincinnati, Ohio 45255

Richard M. Males
RMM Technical Services, Inc.
Cincinnati, Ohio 45208

Number CR-817414-01

Project Officer

James A. Goodrich
Risk Reduction Engineering Laboratory
Cincinnati, Ohio 45268

This Study was Conducted in Cooperation with
Ohio River Valley Water Sanitation Commission
ORSANCO

Project Manager
Jason Heath

RISK REDUCTION ENGINEERING LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OHIO 45268

FOREWORD

In the past 25 years that computers have been actively utilized in the water quality management arena, their potential for assisting in a wide range of analysis and display tasks has been demonstrated. Technologies such as geographic information systems, database management systems, and mathematical modeling have been applied in the water quality management field and have been shown to be effective tools. However, for computers to achieve their full potential, it is necessary that they become integrated into the normal programmatic efforts of agencies and organizations in the planning, regulation and operational areas of water quality management.

In recognition of this need for routine utilization of computer based tools, the Ohio River Valley Water Sanitation Commission (ORSANCO) and the Risk Reduction Engineering Laboratory (RREL) of U.S. EPA commenced a study in 1990. The goals of the study included the adaptation, development and application of various modeling and spatial database management tools which could assist ORSANCO in its prescribed water quality management objectives. These goals were consistent with EPA's ongoing programs involving use of geographic information system and modeling technology.

ABSTRACT

The Ohio River Valley Water Sanitation Commission (ORSANCO) is an interstate water pollution control agency serving the Ohio River and its eight member states. ORSANCO entered into a Cooperative Agreement with the U.S. Environmental Protection Agency (EPA) Risk Reduction Engineering Laboratory (RREL) to develop and apply spatially based computer models and database systems for the purpose of evaluation and management of water quality related issues within the Ohio River Basin.

In this study, three computer based technologies have been applied and integrated: geographic information system (GIS), water quality/hydraulic modeling and database management. GIS serves as a mechanism for storing, using and displaying spatial data. The ARC/INFO GIS, EPA's agency-wide standard, was used in the study under which databases of land and stream information were assembled for the Ohio River Basin. Streams in hydrologic catalog units along the Ohio River mainstem were represented in the GIS using EPA's new detailed RF3 level Reach File System. The full Ohio River Basin was represented using the less detailed RF1 level Reach File. Modeling provides a means of examining the impacts of man-induced and natural events within the basin and exploring alternative strategies for mitigating these events. EPA's WASP4 water quality model, utilizing hydraulic information from the Corps of Engineers' FLOWSED model, was embedded in a menu driven spill management system in order to facilitate modeling of the Ohio River mainstem under emergency spill conditions. A steady state water quality modeling component was also developed under the ARC/INFO GIS in order to trace the movement and degradation of pollutants through any reaches in the RF1 representation of the full Ohio River Basin. Database management (DBM) technology relates to the storage, analysis and display of data. A detailed database of information on dischargers to the Ohio River mainstem was assembled under the PARADOX database management system using EPA's Permit Compliance System (PCS) as the primary source of data. Though these three technologies have been widely used in the field of water quality management, integration of these tools into a holistic mechanism provided the primary challenge of this study.

This report was submitted in fulfillment of contract number CR-817414-01 under the partial sponsorship of the U.S. Environmental Protection Agency. This report covers a period from October 15, 1990 to October 14, 1993 and work was completed as of October 14, 1993.

CONTENTS

Foreword		iii
Abstract		iv
Figures		vi
Tables		vii
Acknowledgments		viii
1.	Introduction	1
2.	Conclusions and Recommendations	3
	Conclusions	3
	Recommendations	3
3.	Project Methodology	6
	Methodology Overview	6
	Hardware Platform	9
4.	Geographic Information System	12
	Introduction	12
	GIS Concepts	12
	ARC/INFO	13
	Database Development	14
	Reach File System	17
	Coverages	24
	Conclusions	30
5.	Water Quality Spill Models	34
	Introduction	34
	Basinwide Network Modeling	35
	Ohio River Mainstem Modeling	42
6.	Discharger Database Management System	65
	Introduction	65
	Database	65
	Application Discharger DBMS	72
References		84
Appendices		
A.	GIS Coverages	85
B.	Data Import Methods	117
C.	Spill Modeling System User's Manual	121
D.	Structure, Programs and Files of the Spill Management System	133
E.	Discharger Database Structure	146
F.	Procedures for Transferring PCS Data to PARADOX	170
Attachments		
I.	PCS Retrieval Program Library	
II.	Published Project Papers	

TABLES

Number		Page
1	Hardware Components	11
2	Coordinate System Description	24
3	Basic Coverages	25
4	Derived Coverages	28
5	Location of ARC/INFO Coverages by Workspace	29
6	Ohio River Segmentation	45
7	Ohio River Flow Segments	49
8	PARADOX Tables	68
9	PCS Retrieval Files	71
10	Primary Forms Developed for DBMS	74
D-1	Spill Modeling System Program Summary	135
D-2	FLOWSED Ohio River Mainstem Nodes	138

ACKNOWLEDGMENTS

Many people contributed their professional expertise during the conduct of this study and their assistance is gratefully acknowledged. Drs. James Goodrich, Project Officer, Robert Clark and Lewis Rossman, of the Drinking Water Research Division of U.S. EPA, provided continuing encouragement and technical assistance during the study formulation and throughout the course of the study. Jeff Finkeldey, formerly of the Computer Sciences Corporation, provided invaluable computer assistance to the study. At ORSANCO, Jason Heath, Project Manager, Alan Vicory, Peter Tennant, and Richard Herd provided continuing technical and administrative assistance throughout the 3-year life of the project. In the modeling area, the contributions by Tim Wool of the ASCI Corporation, Robert Ambrose of the EPA Environmental Research Laboratory at Athens, Georgia, Stan Wisbeth and David Legg of the Corps of Engineers' Ohio River Division, and Dr. Dean Djokic of the University of Texas are acknowledged. Robert Pease of EPA-OIRM, Sue Hansen of Horizon Systems, Loren Hall of EPA-OTS, provided assistance in various aspects of the GIS database development and their contributions are gratefully acknowledged. The assistance of Tom Sheitlin and John Shirey, located at the EPA National Computer Center in Research Triangle Park North Carolina, in the specification, acquisition and operation of the workstation hardware is acknowledged.

FIGURES

Number		Page
1	Hardware Configuration	10
2	Diagram Illustrating Example Basic and Derived Coverages	15
3	Example ArcView Screen	16
4	RF1 Reaches in the Ohio River Basin	19
5	Hierarchical Reach Numbering System	22
6	Comparison of RF1 and RF3 Reach File Representation for Catalog Unit 05090203	23
7	States and County Coverage for States in the Ohio River Basin	31
8	Discharger, Water intake and Dam Coverages for Portions of the Ohio River Basin	32
9	River Mile and RF3 Coverages for Portion of the Ohio River Mainstem	33
10	Dialog Box for the Basinwide Network Spill Model	39
11	Graphical Output from the Basinwide Network Spill Model	41
12	Calculation of Flows for WASP4 Flow Segments	50
13	Observed and Modeled Toluene Concentrations in the Ohio River	53
14	WASP4 Sensitivity Under High Flow Conditions	55
15	WASP4 Sensitivity Under Low Flow Conditions	56
16	Schematic Representation of Spill Modeling System Processes	58
17	Spill Modeling System Menu	60
18	Spill Modeling System Input Form	61
19	Spill Modeling System Output Report	63
20	Spill Modeling System Graphical Output	64
21	Discharger Relational Database Structure	67
22	Schematic Representation of Transfer of Data from PCS to PARADOX	69
23	Standard Form for PCSLIM Table	75
24	Custom Form for PCSFAC Table	76
25	Custom Form for Viewing Facility and Pipe Data	77
26	Custom Form for Docket Table	78
27	Custom Form for Permit Limit Data	79
28	Custom Form for Viewing information from Value Table	80
29	Custom Form for Comparing ORSANCO and PCS Facility Data	81
30	Standard Report for PCSLIM Table	82
31	Custom Report for Permit Limits Data	83
C-1	Complete Spill Modeling System Menu Items	122
C-2	Example Spill Modeling System Input Form	124
C-3	Example Spill Modeling System Plot Input Form	127
D-1	Spill Modeling System Program and File Interaction	134
D-2	Example Cross-Section Data Diagram	142
F-1	Example Input File for Retrieving Facility Data from PCS	171
F-2	Example Output Summary File from PCS Retrieval	172
F-3	Example Output ASCII File from PCS Retrieval	172
F-4	Example Specification File for Converting to Comma Delimited Form	173
F-5	Example Comma Delimited Format File for insertion into PARADOX	173
F-6	Example Input File for Calculating Annual Loads from PCS	176
F-7	Example Annual Loading Report Generated by PCS	177

SECTION 1
INTRODUCTION

In 1990, the Ohio River Valley Water Sanitation Commission (ORSANCO) entered into a Cooperative Agreement with the U.S. Environmental Protection Agency (EPA) Risk Reduction Engineering Laboratory (RREL) to develop and apply spatially based computer models and database systems for the purpose of evaluation and management of water quality related issues within the Ohio River Basin. The goals of the cooperative agreement were consistent with both EPA's long standing support of computer based models in the area of water quality management and its active role in spatial analysis through the application of geographic information systems (GIS), and ORSANCO's historical mission to maintain and improve the water quality of the Ohio River and its tributaries through sound water quality management.

ORSANCO is an interstate water pollution control agency serving the Ohio River and its eight member states: Illinois, Indiana, Kentucky, New York, Ohio, Pennsylvania, Virginia and West Virginia. Established in 1948, ORSANCO works with these eight states as well as the federal government to implement water pollution control and abatement projects. Activities of ORSANCO include: water quality monitoring and assessment of conditions, coordination of interstate water pollution control activities, development and implementation of pollution control standards, spill response and notifications, and public information programs.

In order to fulfill its central role on the Ohio River, ORSANCO has long recognized the need for collecting and utilizing data relative to the characteristics and health of the river. Such data is used to characterize the condition of the river and to analyze the likely short term and long term response of the river to various water pollution control strategies. On the operational side, ORSANCO plays a key role in collecting, analyzing and dispensing information during emergency spill situations.

Historically, much of the emphasis has been on the collection of data on the water quality conditions of the Ohio River. Their innovative on-line monitoring system in the 1960's and 1970's provided continuous measurements of various parameters. In the 1980's they instituted an Organics Detection System (ODS) providing an early warning of organics spills to the river. Associated with these and other data collection efforts was the management of large amounts of data.

By 1990, ORSANCO recognized the needs for more advanced methods for storing, displaying and using a wide range of data related to the Ohio River and its tributaries. Key to this information was the spatial component of 'where' the data point was located. Techniques such as geographic information systems, database management systems and spatially sensitive models were rapidly expanding technologies that could provide the capabilities needed by ORSANCO.

At the same time, the Drinking Water Research Division of EPA's Risk Reduction Engineering Laboratory, was interested in expanding its knowledge and experience in spatial analysis and modeling as a tool for improving drinking water quality. Previous studies by this division had shown the viability of using stream models to identify the potential vulnerability of surface water drinking sources to upstream discharges of toxic chemicals. The Division had also long utilized spatial analysis and display tools to provide a better understanding of the impacts of pollution on drinking water quality.

With these shared viewpoints and needs, the RREL and ORSANCO entered into a 2-year cooperative agreement (later expanded to three years at no additional cost) to explore and implement these tools on the Ohio River. During the course of this study, EPA and ORSANCO have worked closely together to develop these capabilities to the future advantage of both groups. The result has been a fully operational system at ORSANCO's offices and increased capabilities in the GIS and modeling fields at the RREL in Cincinnati.

SECTION 2
CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

The application of computer based display, analysis and modeling tools in conjunction with geographic information system (GIS) technology proved to be an effective strategy in the area of water quality management. In this study, an existing GIS package and database management system was used in conjunction with existing water quality and hydraulic models. Primary emphasis was placed on assembling available spatial and relational databases and integrating the systems to provide a useable and effective tool.

RECOMMENDATIONS

The following recommendations address both specific suggested actions that ORSANCO should follow in the short and long term relative to the subject of this study and general recommendations for further development in the field of spatial analysis and modeling, and geographic information systems.

Recommendations for ORSANCO

1) The GIS, modeling, and database management technology (both hardware and software) associated with this study is extensive and complex. An ongoing training and support program is essential in order to assure the continued full use of these capabilities and integration of these capabilities into all appropriate ORSANCO programmatic areas. The training program should assure redundancy in experience and sufficient knowledge by individuals to perform their required tasks. Different levels of training are required based on specific responsibilities including those of the UNIX systems administrator, ARC/INFO GIS specialist, modeling specialist, database management specialist, data entry technicians, office-wide ArcView user training and system users.

2) The spatial analysis field is moving at a very rapid pace. In order to be aware of developments in this arena, ORSANCO personnel should attend industry-wide or regional conferences. These may include the annual ARC/INFO User's Conference, regional GIS or ARC/INFO meetings and national conferences on GIS applications in the environmental and water resources area.

3) Several software products were used in the study that are now in use by ORSANCO. In some cases, newer versions of the software are currently available, while in other cases, upgrades are forthcoming.

Though it is not always essential to utilize the latest versions of the software, generally it is a good practice to upgrade. Specific upgrades that should be considered are WASP 5.1, ArcView 2, and an improved relational database system for use with ARC/INFO. The WASP 5.1 software has been requested from EPA and would expand the capabilities if implemented. However, it may also require some reprogramming of the Spill Management System if it is to be incorporated in that system. Alternatively, it could be acquired and used as a stand alone model for future modeling efforts on the Ohio River. ArcView 2 is purported to be a major improvement over the current version. It is recommended that this software be upgraded on both the workstation and PC when it becomes available. The relational database capabilities of native INFO data base system associated with ARC/INFO are limited though adequate for the present system. If the extent and complexity of data to be stored in ARC/INFO is projected to increase, then acquisition of one of the external database systems that ARC/INFO supports (e.g. ORACLE, SYBASE, INGRES) should be considered.

4) Calibration of the WASP4 model on the Ohio River mainstem has been limited due to the scarcity of field data that may be used for calibration. It is recommended that ORSANCO work with state and federal agencies to develop and implement an extensive tracer (dye) study so that a suitable data set can be assembled for model calibration and verification. Additionally, a formal plan should be established by ORSANCO to perform detailed sampling in the river, at a fine spatial and temporal scale, when a sizable spill occurs to the river.

General Recommendations

1) There has been considerable efforts by various research groups around the world to combine GIS and water quality modeling technology. However, most of these studies have been site specific and the resulting software products have been non transferable or held in a highly proprietary manner. It is recommended that EPA in consort with other interested groups provide a forum and mechanism for implementing a robust, integrated GIS-modeling tool that may be freely and widely applied throughout the world.

2) Most GIS packages have evolved from the planning field and have been most responsive to the needs of city and transportation planners. Though it is frequently possible to adapt these applications for use in the field of water resources, such applications are generally difficult to adapt and not as responsive to the needs of water resources. Professionals in the area of water resources should attempt to exert greater influence upon the commercial GIS vendors to provide products that are more adaptable to use in the field of water resources.

3) Many existing databases (available from EPA and other sources) which have a geographic component (e.g., latitude-longitude) were developed before the general availability of geographic information systems. As a result, rigorous quality control on the location of features in these databases may be lacking. GIS provides mechanisms for performing various quality control checks and for correcting locational data. It is recommended that a program be started to check and correct the geographic component in such existing databases.

SECTION 3
PROJECT METHODOLOGY

METHODOLOGY OVERVIEW

In order to address the goals of this project, three basic technologies have been applied and integrated: geographic information system (GIS), water quality/hydraulic modeling and database management. GIS serves as a mechanism for storing, using and displaying spatial data. Modeling provides a means of examining the impacts of man-induced and natural events within the basin and exploring alternative strategies for mitigating these events. Database management (DBM) technology relates to the storage, analysis and display of data. Though these three technologies have been widely used in the field of water quality management, integration of these tools into a holistic mechanism provided the primary challenge of this study.

GIS Technology

The guiding principle in developing the GIS capability was to maximize the use of existing GIS technology and spatial databases. The ARC/INFO GIS, EPA's agency-wide standard, was used in the study. Initial work was performed using remote access of ARC/INFO on a VAX minicomputer. Subsequently, both PC ARC/INFO and a workstation based system were obtained.

EPA has developed an extensive spatial database related to water quality and demographic parameters. This served as the primary source of spatial data for the study. Following is a summary of spatial data used in this study:

- State and county boundaries
- City locations and characteristics
- Water supply locations and characteristics
- Locations and characteristics of dischargers to water bodies
- Toxic loadings to air, water and land
- Dam locations and characteristics
- Stream reaches and characteristics

The primary organizing concept for the water related information was U.S. EPA's Reach File System (Horn and Grayman, 1993). This system provides a common mechanism within EPA and other agencies for identifying surface water segments, for relating water resources data, and for traversing the nation's surface water in hydrological order within a computer environment. Each reach is uniquely identified by a hierarchical hydrologic code. Information available on each reach includes topological identification of adjacent reaches, characteristic information such as length and stream name, and streamflow and velocity estimates. The original Reach File (designated as RF1) was developed in the early 1980's and included approximately 70,000 reaches nationwide. The most recent version (RF3) includes over 3,000,000 reaches nationwide. As part of this project, an RF1 level database was established for the entire Ohio River Basin. The RF3 Reach File was implemented for the Ohio River mainstem and lower portions of tributaries. River miles along the Ohio River were digitized and established as an ARC/INFO coverage in order to provide a linkage between the Reach File and river mile indexing used by ORSANCO and other agencies along the river.

Several EPA sources of information on dischargers to water bodies were used. The Industrial Facility Discharger (IFD) file contains locational and characteristic data for NPDES permitted discharges. Detailed permit limits and monitoring information was accessed from the Permit Compliance System (PCS). The Toxic Release Inventory (TRI) System includes annual loading of selected chemicals to water, land, air, and sewer for selected industries based on quantity discharged. All water data is referenced to the NPDES permit number which is spatially located by reach and river mile, and by latitude/longitude.

Spill Modeling

An important role that ORSANCO fills on the Ohio River relates to the monitoring and prediction of the fate of pollutant spills. Typically, ORSANCO serves as the overall communications link between states during such emergency conditions. They coordinate and participate in monitoring, and serve as the information center in gathering data and issuing predictions as to the movement of the spill in the river. In the past, a series of time-of-travel nomographs, based on National Weather Service flow forecasts, Corps of Engineers flow-velocity relationships and previous experience, were used to predict the movement of a spill. In this project, a hydraulic model was combined with a water quality model to serve as a more robust method for making such predictions.

The Corps of Engineers' FLOWSED model was selected as the means of predicting daily flow quantities and water levels along the mainstem and portions of major tributaries near their confluence with the Ohio River (Johnson, 1982). FLOWSED is applied daily by the Ohio River Division of the Corps of Engineers as part of their reservoir operations program. Five day forecasts of stage and flow are generated for 400 mainstem and tributary segments and the results were made accessible to ORSANCO via phone lines.

EPA's WASP4 water quality model was selected for use in the project (Ambrose et al., 1990). WASP4 is a dynamic compartment model that can be used to analyze a variety of water quality problems in a diverse set of water bodies. Since the primary use of the model in this project is quick response under emergency situations, only the toxic chemical portion of the model with first order decay is being used. The FLOWSED and WASP4 models have been combined into a user friendly spatial decision support system framework described later in this report.

Discharger Database Management System

EPA's Permit Compliance System (PCS) and historical records maintained by ORSANCO furnish a rich source of data on discharge information for the Ohio River. In order to organize this data and make it available for analysis, a database was developed using the PARADOX database management system (DBMS).

The database was established using a relational structure with a series of related tables (2-dimensional flat files). Individual tables contain information on facilities, outfalls, permit limits, monitoring data, and codes used in the other tables. The NPDES permit number is used as the primary key in each of the data tables. A mechanism for downloading and reformatting data from the national PCS database has been developed along with custom forms for viewing and editing data, and custom reports for preparing hard copy summaries. Latitude and longitude values for each facility can provide the locational mechanism for use of this data in conjunction with the GIS.

Integration of GIS/Modeling/Database Technologies

A major objective of this study was the integration of the GIS, modeling and DBMS technologies into a holistic tool for use by ORSANCO. Several integration mechanisms have been implemented as summarized below and described in greater detail later in this report.

Steady State Spill Tracing--

The NETWORK component of the ARC/INFO GIS provides a steady state transportation oriented routing capability. This capability was used in an ARC Macro Language (AML) program to construct a routing procedure for determining downstream concentrations and travel times. The pollutant may be treated as a conservative element or represented by a first order exponential decay function. This capability has been implemented for use with the RF1 Reach File representation of the full Ohio River Basin. The user may select from six flow regimes: average flow, low flow and four multiples of average flow ranging from one-tenth to ten times average flow. This system provides ORSANCO with the capability of estimating the arrival time of a spill from any RF1 tributary to the Ohio River mainstem.

Spill Management System--

A PC based spatial decision support system (SDSS) was built as a Spill Management System to serve a quick response tool for analyzing and displaying the results of a pollutant spill into the Ohio River. The system is implemented in the C language using a commercial menuing system and a series of graphical display routines developed at EPA. Custom written routines have been used to read the output from the Corps of Engineers' FLOWSED model, to generate the input files for EPA's WASP4 model, to create output reports and output plots, and to provide an animated representation of the concentration profiles moving down the river. Additionally, the system generates a file in DBF format that may be read by ArcView (the companion software to ARC/INFO for user friendly viewing of spatial data).

HARDWARE PLATFORM

Within the study, the initial hardware platform was a combination of local PC's (in Cincinnati) and a remote access terminal to a VAX computer located at EPA's National Computer Center in Research Triangle Park, North Carolina. The final platform, and the one on which the completed system was installed, was composed of a UNIX based Data General workstation and a PC computer workstation. These systems are described below.

The full hardware configuration is shown schematically in Figure 1. Individual hardware components are enumerated in Table 1. The two major components, the workstation and the PC, are currently largely independent. When needed, data is transferred using a floppy disk. A future network between the two stations is under consideration.

FIGURE 1
HARDWARE CONFIGURATION

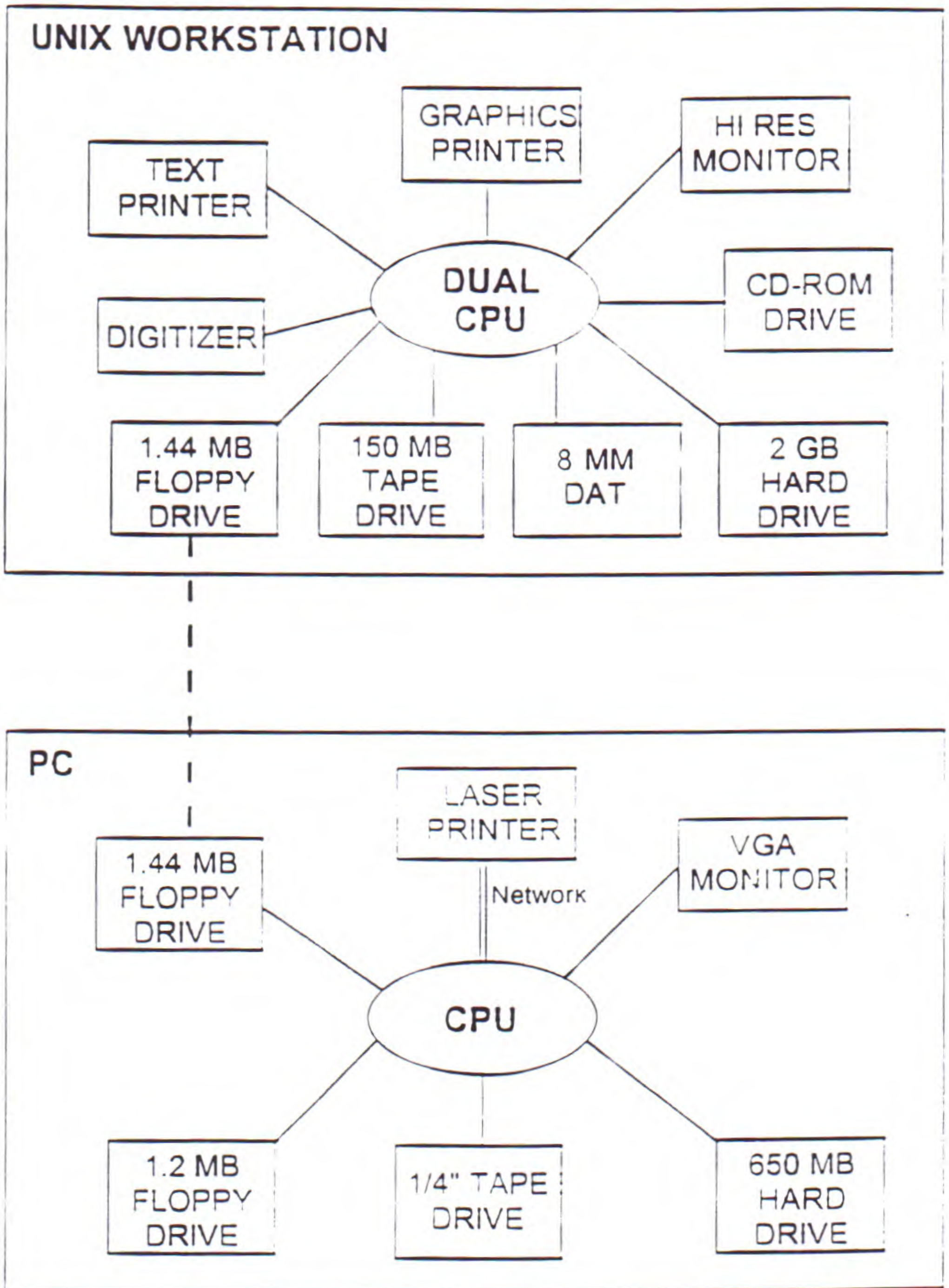


TABLE 1

HARDWARE COMPONENTS

UNIX WORKSTATION

Data General Aviion 412 Dual CPU

Data General 19" Color Monitor

Peripheral Housing Unit with:

2GB Hard Drive

150 MB Tape Drive

1.44 MB Floppy Drive

CD-ROM Drive

8 MM DAT

Seiko Colorpoint Postscript Printer

Panasonic Text Printer

Altek 36"x48" Digitizer with 16 Button Cursor

PC

CompuAdd 386/33, 8MB RAM, 650MB Hard Drive, Dual Floppy Drives

NEC 20" Multisync 5D SVGA Monitor

Tape Drive 150MB, 1/4"

External Modem, 9600 Baud

HP Laserjet 2 Printer on Lantastic LAN

SECTION 4
GEOGRAPHIC INFORMATION SYSTEM

INTRODUCTION

Many decisions in water resources depend upon data that has a spatial dimension. Historically, such data was mapped, and planners and engineers used the maps in developing plans and reviewing options. Keeping the maps current and attempting to visually integrate the information on the maps for decision making purposes were only two of the limitations associated with this approach. For example, performing suitability analysis in which several characteristics (e.g. land use, soils, etc.) were combined to determine areas that were most suitable for supporting a specific activity was very cumbersome. More complex tasks that involved using spatial data in support of hydrologic or other water resources models proved even more difficult.

The concepts of storing spatial data in a computerized form and using this data in mapping and analysis first appeared in the 1960's. In fact, much of the early development and many of the early applications of computerized spatial analysis were in the field of water resources (Males and Grayman, 1992). In the 1970's, the concept of use of computers in conjunction with geographic data was first referred to as geographic information systems or GIS (Tomlinson, 1972). GIS technology has evolved since then and grown to major proportions in the 1990's in terms of applications and popularity. GIS technology is now widely used in the field of water resources for a range of storage, display, analysis and modeling tasks.

GIS CONCEPTS

Within a geographic information system, all data has both a spatial component and a characteristic component. The spatial component identifies the location of the data (e.g. its latitude and longitude) while the characteristic data can be any information ascribed to the location (e.g. soil type, age of a building, owner of a property).

There are three basic types of spatial components which may be arranged to represent any spatial feature:

1) points: component identified by a single set of geographic coordinates.

2) lines: a component defined by two end points. This feature may also have a direction associated with it such as upstream to downstream direction of a stream.

3) area (polygon): a 2-dimensional zone identified by a closed boundary that surrounds it.

These basic components may be combined or grouped to form other features such as:

- arc: a series of line segments which together define a boundary or a piecewise linear feature. For example, an arc may represent a portion of a stream between two tributaries.

- network: a series of points (referred to as nodes) and lines (or arcs) which together define a set of pathways. For example, stream networks and road networks are frequently represented in a GIS. Networks may be directional, such as streams or one way streets, which specify a preferred direction defined by an upstream to downstream direction or they may be non-directional.

- grids: a collection of regularly shaped polygons. In a rectangular grid, each polygon is a square or rectangle of the same size organized into a series of rows. Many geographic information systems store all data into such a pattern. A triangulated irregular network (TIN) is a grid composed of a series of interconnected triangles of differing size. A TIN structure is frequently used to represent a surface since each triangle defines a plane.

ARC/INFO

There are several commercially packaged geographic information systems available that could have fulfilled the requirements of this project. The ARC/INFO geographic information system developed by ESRI was selected for use in this project because it is the GIS standard within the U.S. EPA and other federal and state agencies, and many of the databases used in the study were available from EPA or other sources in ARC/INFO format. ARC/INFO is available on a wide range of hardware platforms ranging from large VAX minicomputers, to UNIX workstations and DOS based personal computers.

ARC/INFO is actually an amalgam of two products that separately handle the two forms of data in a GIS. The 'ARC' product developed by ESRI deals with the geographic spatial data representing points, lines and areas. In fact, the name 'ARC' refers to the primary mechanism for storing spatial data with ARC/INFO in which connected line segments are stored as arcs and a series of arcs are used to represent polygonal areas. INFO is a proprietary database management system developed by HENCO, Inc. Characteristic data is stored and manipulated in the INFO portion of ARC/INFO. In recent years, the ARC/INFO product has moved away from the total reliance upon INFO as the embedded database management system. On the PC, characteristic data is stored in a DBF format while on workstations, the user has an option to replace or supplement INFO with other.

more powerful database management systems, such as ORACLE and INGRES. However, independent of the platform and the specific database system, the user should be aware of the presence of this dual system for storing data within ARC/INFO.

Within ARC/INFO, the basic database organization is the 'coverage'. A coverage may be best explained as a single layer of data. For example, within a database, there may be several coverages, one for land use, another for soil, another for streams, and another for wells.

Coverages may be considered as basic data (such as soil type) or may be derived from combining several coverages. For example, if one started with a coverage representing a stream network and another coverage representing wells, an ARC command could be used to generate a 'buffer' zone of 1 mile around all streams. This coverage could then be combined with the well coverage to contain all wells located within 1 mile of a stream. This process is shown graphically in Figure 2.

A second product developed by ESRI called ArcView was also used extensively in the project. ArcView is not a full featured GIS, but rather a user friendly package for viewing and querying an existing GIS database. It operates both on a PC (under Windows) or on a UNIX workstation. Figure 3 shows an ArcView screen in which both a display and table are in view.

DATABASE DEVELOPMENT

The development of a large GIS database can be very time consuming and expensive. Within this project, primary emphasis in database development was on the identification and acquisition of existing databases. The primary source of these databases was EPA.

Since ORSANCO's primary area of concern is the streams and rivers in the Ohio River Basin, a primary emphasis in the database development related to the representation of these features in the GIS. The stream network system was represented using EPA's Reach File System: a nationwide database of surface waters and their connectivity. Further details on this system are provided later in this section.

Initial work on the GIS portion of the study was conducted using an ARC/INFO system resident on a VAX computer at EPA's National Computer Center (NCC) in Research Triangle Park, North Carolina. This installation was accessed remotely by a 9600 baud line. Additionally, a PC ARC/INFO installation at EPA's Risk Reduction Engineering Laboratory (RREL) in Cincinnati was used for some work. Many of the coverages used in the study were set up on the VAX computer and, in a few cases, some editing and quality assurance

FIGURE 2
DIAGRAM ILLUSTRATING EXAMPLE BASIC AND DERIVED COVERAGES

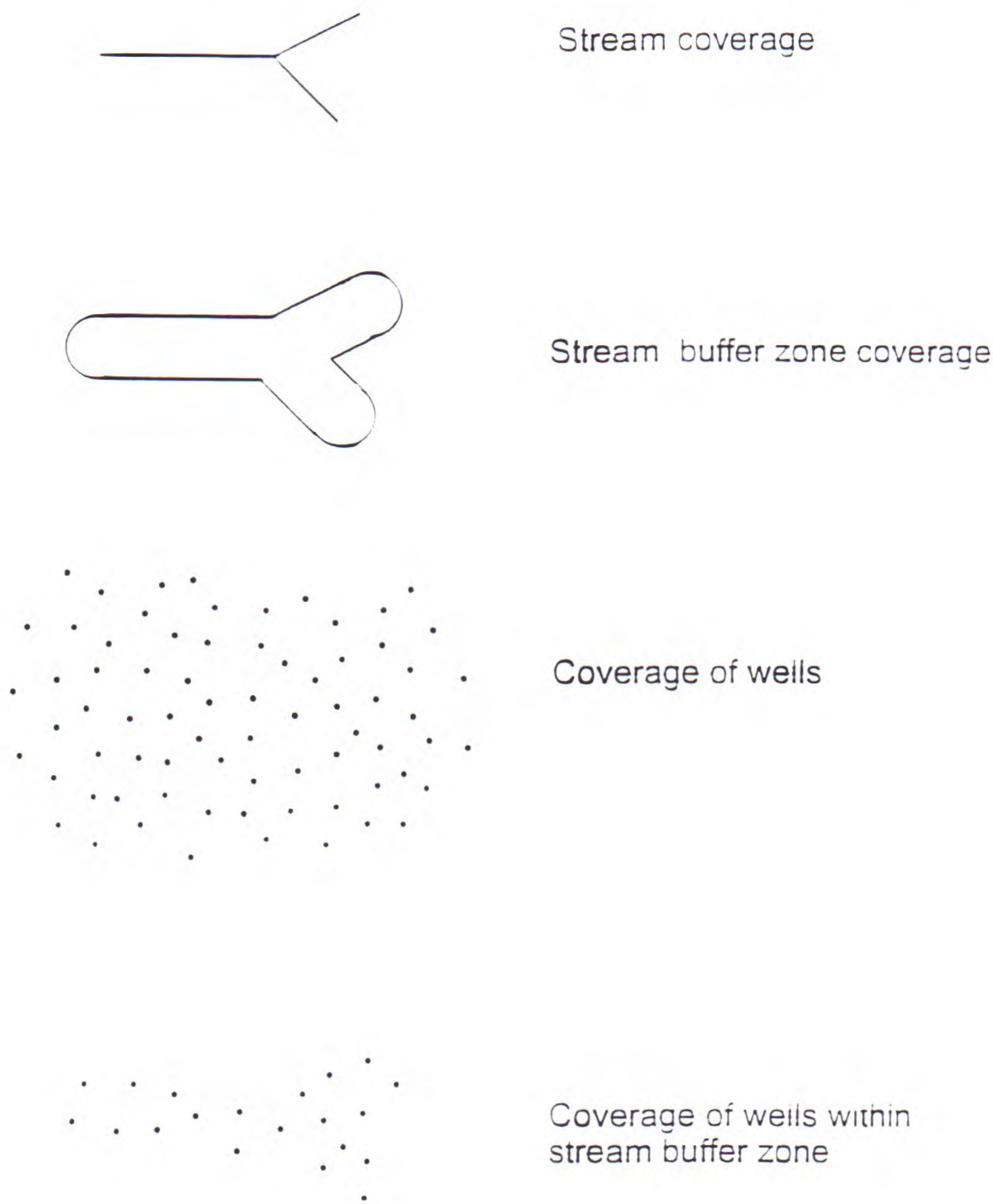
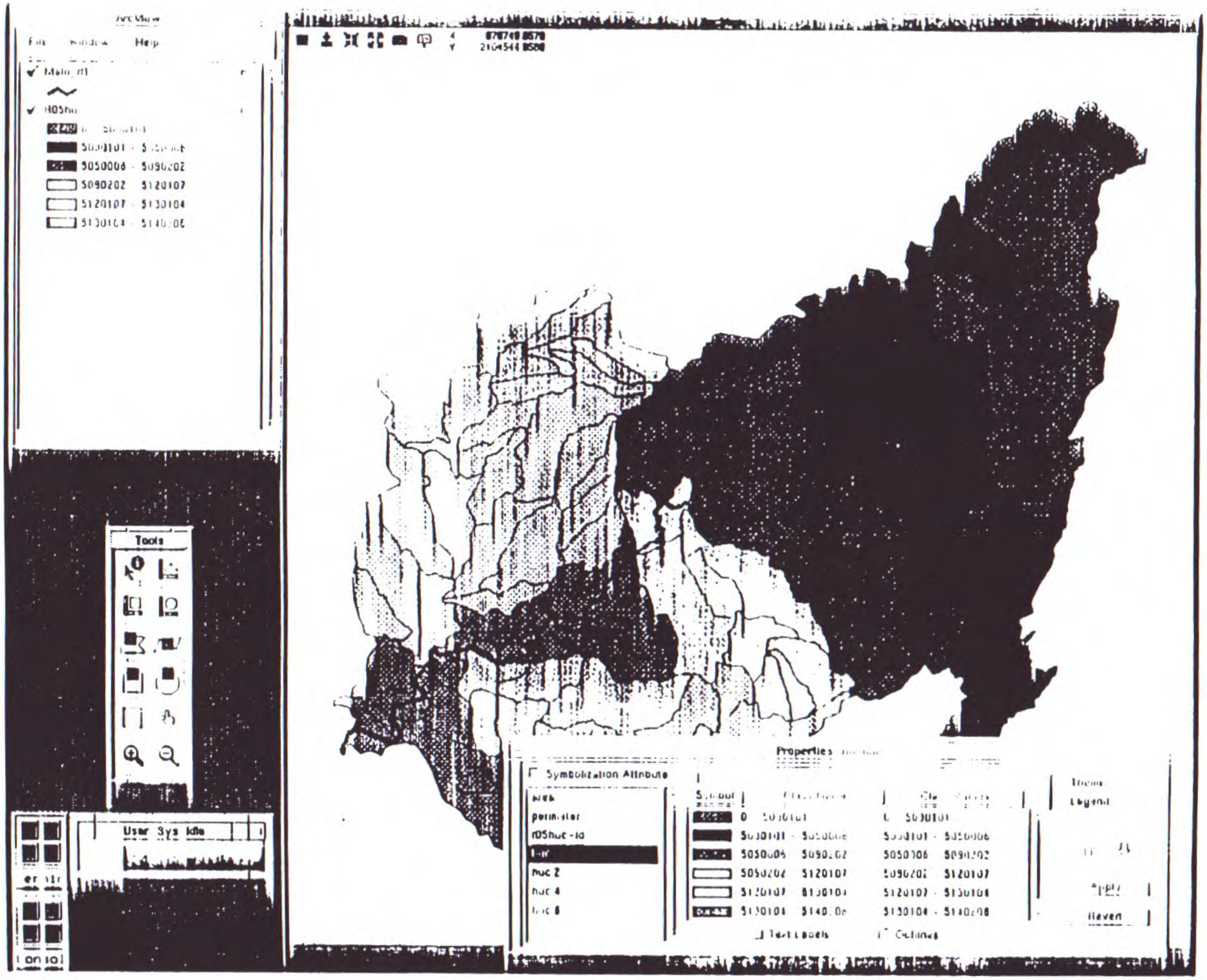


FIGURE 3
EXAMPLE ARCVIEW SCREEN



tasks were performed on this system.

Towards the end of 1992, a Data General workstation was acquired and became the major focus of the study. The coverages on the VAX computer were downloaded to a PC at the RREL in Cincinnati over a dedicated line and moved to the workstation. Other coverages were downloaded from EPA's IBM mainframe computer at the National Computer Center via phone lines. Still other coverages were acquired from EPA headquarters and Region V office on tapes. Additional coverages were digitized from maps and some created from lists of features with associated latitude and longitude coordinates.

Generally some level of quality assurance was performed locally on the coverages. However, for most of the large coverages obtained from EPA, the quality control was limited to visual checking or limited consistency checks. Unacceptable features were either corrected or identified for future correction and deleted from the production level database.

A complete listing of coverages and their origins is summarized in the following sub section and described in detail in Appendix A.

REACH FILE SYSTEM

Because of the central importance of the stream network system to much of the work in this study, a detailed description of the Reach File System used as the basis for the stream and river coverages will be provided. The following discussion draws heavily from a recently published paper on the Reach File System (Horn and Grayman, 1993).

The Reach File is a network-oriented, spatially referenced geographic database of surface water features of the United States. In the Reach File, streams are identified by a hierarchical, hydrologically referenced coding system. The connectivity of each reach to its adjacent (upstream and downstream) reaches is known, along with various characteristics of the reach. Additionally, the geographic trace for each reach is stored as a string of latitude - longitude coordinates. The Reach File has progressed through four versions: RF1A, RF1, RF2, and RF3. RF1A was designed in 1973, funded for development in 1974, and completed in 1975.

RF1A was digitized by EPA from the U.S. Geological Survey (USGS) Hydrological Map of the United States having a scale of 1:2,500,000. This first version of the Reach File was a single national coverage used for database design testing and demonstration. Its design was found to be well suited for integration of surface water databases and for hydrologically ordered retrievals of any data which might be indexed to it along river courses or through open water bodies. Spatial data were captured as latitude/longitude coordinate strings representing the traces of segments of streams and shorelines: these segments were called reaches. The topological relationships among reaches were derived from the digital traces using interactive graphics software prepared for the project. In 1978, EPA began developing the Reach File as a fully functional database. It was completed as RF1 in 1982, at which time it was adopted for use by several offices within EPA, the Fish and Wildlife Service, and some States. This RF1 version and all subsequent versions retained the basic topological design developed for RF1A, although new data types and data relationships were added for each new version of the Reach File .

RF1 was digitized by EPA from aeronautical charts prepared by the National Oceanographic and Atmospheric Administration (NOAA). These charts provided the most detailed hydrographic coverage of any complete set of maps available for the United States at that time. NOAA developed them from photographic copies of the USGS maps having a scale of 1:250,000, and except for updates for improved detail, the resultant maps retained all the hydrography shown on the 1:250,000 scale maps. These NOAA charts therefore became the maps of choice for the Reach File. They were optically scanned, edge-matched, and transformed into a single contiguous vector database by EPA during the period of 1978-1982. The resultant database, RF1, includes information on links representing stream centerlines (referred to as hydraulic transport reaches) and on shorelines for a limited number of wide rivers, lakes, open waters, and oceans (referred to as shoreline reaches). Figure 4 shows the RF1 traces for the Ohio River Basin.

An intermediate update (RF2) was constructed during the late 1980's to aid in building a much more detailed version (RF3). RF2 was developed using hydrographic feature names and spot coordinates obtained from USGS's Geographic Names Information System (GNIS). Overlay software was developed for the project on the EPA mainframe computer. This software included many specialized functions to speed final processing and to greatly enhance the quality control of the overlay process. For example, the software culled features by their estimated size to eliminate overlay clutter, and in combination with RF1, provided high level information on the relative significance of various matches of end points and intermediate points of GNIS features with RF1 features. The simplest types of matches were those which were found to be unambiguous and which involved the connection of streams in GNIS by their downstream end points to intermediate points along the traces of reaches in RF1. The overall process involved a wide variety of much more complex matches than these. Interactive graphics software and operator protocols were developed to validate the overlays and to provide quality control in the project. The RF2 development project essentially doubled the number of reaches in the Reach File.

The RF3 development project began in the Fall of 1988. RF3 includes approximately 3,200,000 miles

FIGURE 4
RFI REACHES IN THE OHIO RIVER BASIN

of streams and many thousands of reservoirs, lakes and ponds. The RF3 version was developed by augmenting the Reach File with data from newly developed digital cartographic files of the hydrography of the U.S.(Bondelid et al., 1990). This new source of digital cartographic data was the digital line graph (DLG) data for the 1:100,000 scale maps prepared by the USGS and the U.S. Bureau of the Census for the 1990 census. The DLG data were purchased by EPA from USGS in the form of approximately 54,000 files. These files, one for each 7-1/2 minute quadrangle within the 1:100,000 scale maps, were consolidated into a single contiguous file on the EPA-IBM mainframe computer. Software was developed for processing this very large database in preparation for overlaying it on the Reach File. The information content of the file thus prepared was increased over the DLG files by providing single direction progression of coordinates within geographic entities such that the occurrence of coordinates for a given entity began at one end of the entity and progressed to the other. In addition, this single contiguous file vastly improved the accessibility of entities within the database, a factor deemed essential for quality control in this enhancement of the Reach File. The data were then brought into relation with the Reach File to provide more extensive and detailed coverage of hydrographic features to the Reach File. Because the DLG database contained no names, the GNIS was used again to provide surface water names where available for the new reaches.

Topologic data, geographic data, and trace data are stored in the Reach File for each reach. The topologic information defines the connections (inter-reach linkages) between the reaches. This identifies the reaches attached to the upstream and downstream ends of each reach, and incorporates special data elements to facilitate rapid traversals through the database for data storage and retrieval purposes and for mathematical modeling. The geographic data include the names of streams and water bodies represented in the file and other data which help to geographically locate and identify these surface water features. The trace data are a series of connected points, in latitude / longitude coordinates, forming a line trace of each reach.

Each reach is identified by a 17-character number. Each reach number consists of the 8-character numeric identifier of the USGS cataloging unit within which the reach is located, a 4-character numeric identifier which is unique within the cataloging unit until the reach is subdivided for inserting new tributaries or other updates, and 5 additional characters for use when reaches are subdivided. Until a reach is subdivided, these 5 additional characters are zeroes.

Reaches may be subdivided to improve the file where needed. The resulting subdivisions are also referred to as reaches. The 5-character numeric extension corresponds to the mileage measured from the base of the original reach being subdivided to the base of a new smaller reach within it. To illustrate these points, the nationally unique reach number 04010203-0006 corresponds to stream reach 0006 in cataloging unit 04010203, whereas 04010203-0006-04.23 is the identifier of the reach whose downstream end is 4.23 miles above the downstream end of reach 0006 within cataloging unit 04010203. The diagram presented in Figure 5 illustrates the hierarchical features of the Reach File numbering system.

Using the reach numbers and the relative distance between downstream ends of reaches and points of interest along the reach, other major EPA databases have been indexed to the Reach File. This, in effect, imparts hydrological order to those databases and integrates them with each other. By traversing the Reach File in the upstream or downstream directions, the user can retrieve data from the Reach File and related files in hydrological order. This type of retrieval is made possible by inter-reach linkages provided in the Reach File itself. By using Reach File reach numbers as indexed variables in common among the several files, data in any of the files can be related to the data in any of the other files. Databases on pollutant dischargers, domestic water supply intakes, and stream gages are examples of EPA databases that have been referenced to the Reach File (Pandolfi, 1988). Data collected at water quality monitoring stations can be retrieved in hydrological order using Reach File retrieval procedures of EPA's STORET System (Taylor et al., 1988).

Other data in the Reach File can be used to control or focus user-defined retrievals. For example, stream names in the Reach File enable the user to select whole rivers or streams by their names. Various types of mileage data in the file facilitate selection of reaches by size of drainage system or position within drainage system. Other data in the Reach File that may be used this way include the hydrologic level or stream order of each reach, reach type, and other data that describe and locate streams, rivers, lakes, reservoirs, shorelines and related surface water features.

With these types of data provided by the Reach File coupled with Reach File reach numbers inserted as indexed variables in other files, it is possible, for example, to analyze pollutant discharges described in one file with respect to downstream domestic water supply intakes found in another file, and further to associate these analyses with data from water quality monitoring stations found in yet another file. Furthermore, hydrologically ordered retrievals from multiple water resource files are ideally suited for water quality modeling because they can provide needed data in the same sequence required by the models as the simulation progresses downstream from points of origin to final sinks.

In all, there are 2,111 cataloging units in the contiguous United States. In RF1, there are approximately 70,000 reaches; in RF3, there are approximately 3,500,000 reaches. Figure 6 illustrates the relative reach density between the RF1 and RF3 files for a single catalog unit, 05090203, representing the Ohio River in the vicinity of Cincinnati, Ohio.

The RF1 system was designed for use on the EPA-IBM mainframe computer. The RF3 version may be used

FIGURE 5
HIERARCHICAL REACH NUMBERING SYSTEM

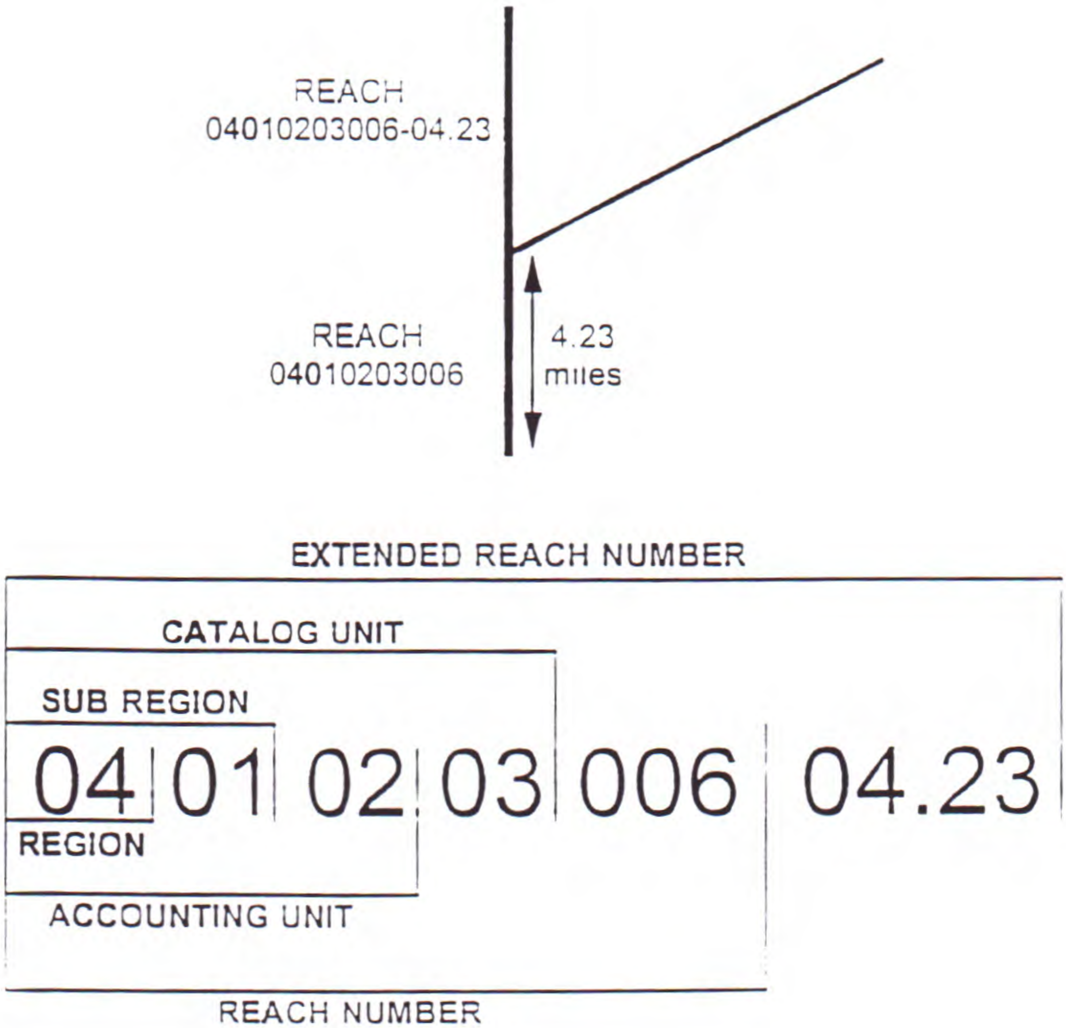
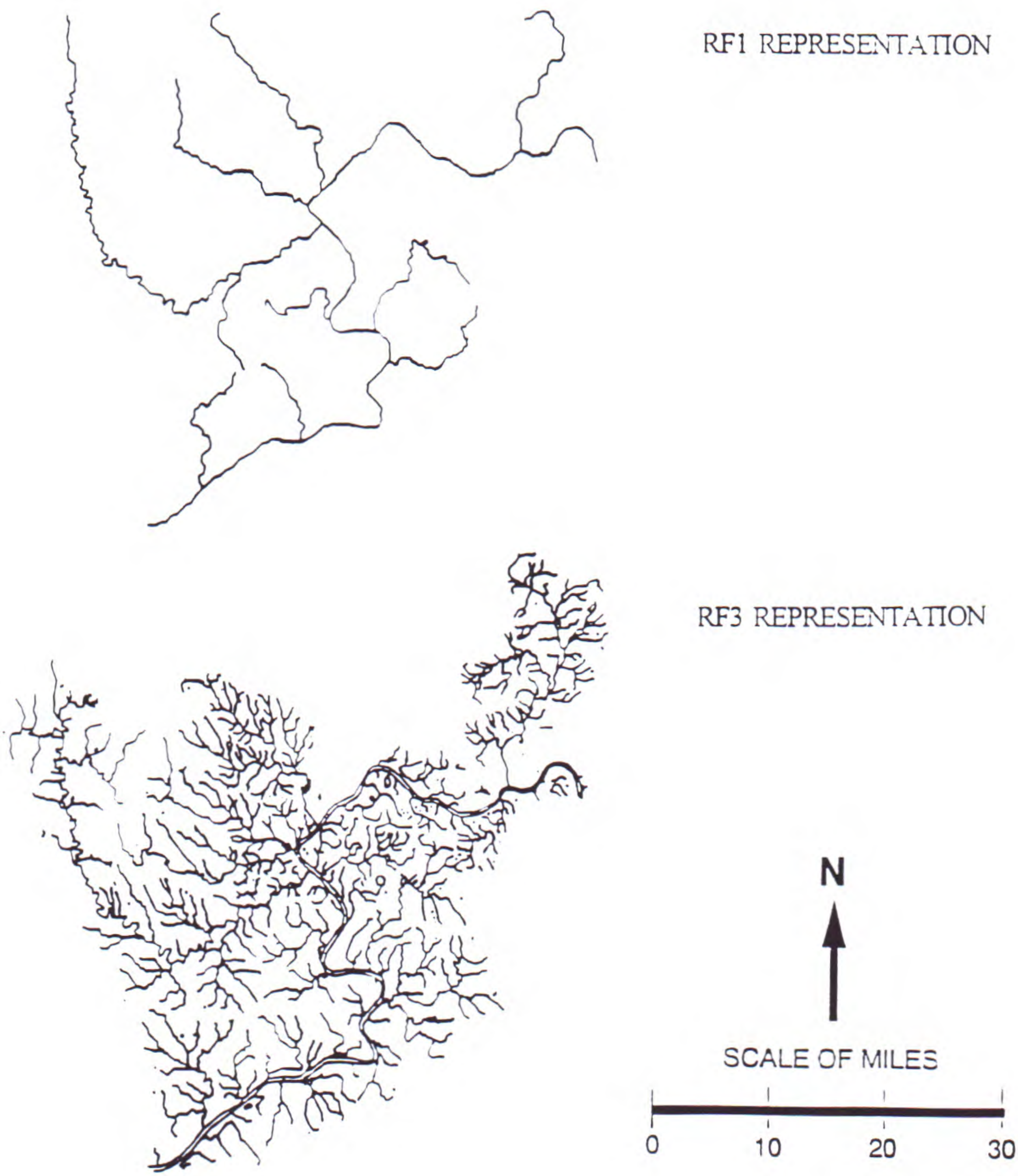


FIGURE 6
COMPARISON OF RF1 AND RF3 REACH FILE REPRESENTATION OF CATALOG UNIT 05090203



on the mainframe computer or, for a single cataloging unit at a time, on a PC. RF3 data are also available for use in commercial geographic information systems or in user constructed systems. The RF1 file is still used for many production efforts at EPA. However, RF3 files are available nationwide and efforts are underway to reference facilities to the RF3 reaches and to estimate streamflows for the RF3 reaches. Some quality assurance is still underway within EPA on the RF3 files. The RF3 files obtained in this study are considered interim products by EPA subject to some future modification and cleanup.

COVERAGES

Though various coverages were acquired in different geographic projections (e.g. latitude-longitude, Albers, etc.), all coverages were converted to a common Albers projection for comparability. A full description of the projection used in the study is provided in Table 2.

TABLE 2

COORDINATE SYSTEM DESCRIPTION

Projection	ALBERS
Units	METERS
Spheroid	CLARKE1866
Parameters:	
1st standard parallel	29 30 0.000
2nd standard parallel	45 30 0.000
central meridian	-96 0 0.000
latitude of projection's origin	0.000
false easting (meters)	0.000
false northing (meters)	0.000

A listing and description of all basic coverages is provided in Table 3. These coverages were obtained from various sources (as detailed in Appendix A) and, except for some quality assurance, have not been modified. Several other coverages were derived by combining basic coverages to produce a new coverage. These coverages are listed and described in Table 4. These coverages are all stored on the Data General workstation under workspaces contained in sub directory: /usr2/orsanco/data. Table 5 lists the workspace location for each of the coverages.

TABLE 3
BASIC COVERAGES

Coverage	*Description*
r05dams	dams in region 5
r05flo	flows for rf1 reaches(no arcs)
r05huc	catalog units in region 5
r05rf1	reach file 1 for region 5
mile_pt	mile points along the main stem of the Ohio River
waspsegs	WASP segments of the Ohio River
tri_dc	toxic release inventory for District of Columbia
tri_il	toxic release inventory for Illinois
tri_in	toxic release inventory for Indiana
tri_ky	toxic release inventory for Kentucky
tri_md	toxic release inventory for Maryland
tri_nc	toxic release inventory for North Carolina
tri_ny	toxic release inventory for New York
tri_oh	toxic release inventory for Ohio
tri_pa	toxic release inventory for Pennsylvania
tri_tn	toxic release inventory for Tenesee
tri_va	toxic release inventory for Virginia
tri_wv	toxic release inventory for West Virginia
city	cities from EPA GIS
county	counties from EPA GIS
ifd	industrial facilities dischargers from EPA GIS
nplsites	national priority list sites from EPA GIS
state	coverage of states from EPA GIS
supply	water supply intakes from EPA GIS
cso	combined sewer overflows

Table 3 (continued)

Coverage	*Description*
r05rr	rf1 coverage of region 5 that does not have all arcs from the flows file
r05rr75	rf1 arcs that were missing in r05rr
a5010006	rf3 coverage for catalog unit a5010006
a5010008	rf3 coverage for catalog unit a5010008
a5010009	rf3 coverage for catalog unit a5010009
a5020005	rf3 coverage for catalog unit a5020005
a5020006	rf3 coverage for catalog unit a5020006
a5030101	rf3 coverage for catalog unit a5030101
a5030104	rf3 coverage for catalog unit a5030104
a5030105	rf3 coverage for catalog unit a5030105
a5030106	rf3 coverage for catalog unit a5030106
a5030201	rf3 coverage for catalog unit a5030201
a5030202	rf3 coverage for catalog unit a5030202
a5030203	rf3 coverage for catalog unit a5030203
a5030204	rf3 coverage for catalog unit a503024
a5040004	rf3 coverage for catalog unit a5040004
a5050008	rf3 coverage for catalog unit a5050008
a5060002	rf3 coverage for catalog unit a5060002
a5060003	rf3 coverage for catalog unit a5060003
a5070102	rf3 coverage for catalog unit a5070102
a5070204	rf3 coverage for catalog unit a5070204
a5080002	rf3 coverage for catalog unit a5080002
a5090101	rf3 coverage for catalog unit a5090101
a5090102	rf3 coverage for catalog unit a5090102
a5090103	rf3 coverage for catalog unit a5090103
a5090104	rf3 coverage for catalog unit a5090104
a5090201	rf3 coverage for catalog unit a5090201

Table 3 (continued)

Coverage	*Description*
a5090202	rf3 coverage for catalog unit a5090202
a5090203	rf3 coverage for catalog unit a5090203
a5100101	rf3 coverage for catalog unit a5100101
a5100102	rf3 coverage for catalog unit a5100102
a510205	rf3 coverage for catalog unit a5100205
a5110005	rf3 coverage for catalog unit a5110005
a5120113	rf3 coverage for catalog unit a5120113
a5120114	rf3 coverage for catalog unit a5120114
a5120207	rf3 coverage for catalog unit a5120207
a5130205	rf3 coverage for catalog unit a5130205
a5140101	rf3 coverage for catalog unit a5140101
a5140102	rf3 coverage for catalog unit a5140102
a5140103	rf3 coverage for catalog unit a5140103
a5140104	rf3 coverage for catalog unit a5140104
a5140201	rf3 coverage for catalog unit a5140201
a5140202	rf3 coverage for catalog unit a5140202
a5140203	rf3 coverage for catalog unit a5140203
a5140204	rf3 coverage for catalog unit a5140204
a5140205	rf3 coverage for catalog unit a5140205
a5140206	rf3 coverage for catalog unit a5140206
a6040005	rf3 coverage for catalog unit a6040005
a6040006	rf3 coverage for catalog unit a6040006
a8010201	rf3 coverage for catalog unit a8010201

TABLE 4
DERIVED COVERAGES

Coverage	*Description*
hucbuf	8 km buffer of r05huc
hucohior	r05huc split by the Ohio River
hucstohr	r05huc split by the Ohio River and the state boundaries
r05city	cities within statebuf
r05ifd	industrial facilities dischargers within hucbuf
r05state	states in region 5 with boundaries aligned with the Ohio River. and no small islands with area less that 100 sq. km
r05suppl	water supply intakes within statebuf
state in /usr2/orsanco/data/basin	derived from /usr2/orsanco/data/epa/state with no islands with area less than 100 sq. km
statebuf	8km buffer of r05state
statline	line coverage from r05state (no polygons. just arcs)
main_buf	8 km buffer of main_rf1
main_rf1	main stem of the Ohio River obtained from r05rf1 by selecting arcs with pname = 'OHIO R'
main_rf3	main stem of the Ohio River obtained from rf3 coverages by selecting arcs with pname = 'OHIO R'
mile_buf	8 km buffer of mile_pt
mile_pt	mile points along the main stem of the Ohio River
ohr_city	cities within main_buf
ohr_dams	dams within main_buf
ohr_ifd	industrial facilities dischargers within main_buf
ohr_supp	water supply intakes within main_buf
rf3	RF3 catalog units along the main stem of the Ohio River
mainpoly	polygon coverage of /usr2/orsanco/data/oh_river/main_rf1 by buffering for 100 meters

TABLE 5
LOCATION OF ARC/INFO COVERAGES BY WORKSPACE

All **directory names** are shown in bold. Workspace names are obtained by preceding directory names by /usr2/orsanco/data/ e.g., /usr2/orsanco/data/basin

basin

hucbuf
hucohior
hucstohr
monitesa
r05city
r05dams
r05flo
r05huc
r05ifd
r05rf1
r05state
r05suppl
state
statebuf
statline
cso

oh_river

main_buf
main_rf1
main_rf3
main_buf
mile_buf
mile_pt
ohr_city
ohr_dams
ohr_ifd
ohr_supp
rf3
waspsegs

tooldata

mainpoly

old

r05rr
r05rr75

tri

tri_dc
tri_il
tri_in
tri_ky
tri_md
tri_nc
tri_ny
tri_oh
tri_pa
tri_tn
tri_va
tri_wv

epa

city
county
ifd
nplsites
state
supply

rf3

A5010006
A5010008
A5010009
A5020005
A5020006
A5030101
A5030104
A5030105
A5030106
A5030201
A5030202
A5030203
A5030204
A5040004
A5050008
A5060002
A5060003

rf3 (cont.)

A5070102
A5070204
A5080002
A5090101
A5090102
A5090103
A5090104
A5090201
A5090202
A5090203
A5100101
A5100102
A5100205
A5110005
A5120113
A5120114
A5120207
A5130205
A5140101
A5140102
A5140103
A5140104
A5140201
A5140202
A5140203
A5140204
A5140205
A5140206
A6040005
A6040006
A8010201

The general extent and detail of some of the coverages is illustrated in the following maps generated within ARC/INFO and ArcView. In Figure 7, the state and county coverages are displayed for 11 states within the Ohio River Basin. Points in the discharger (IFD) file coverage, the dams coverage and the water supply intake coverage within a portion of the States of Ohio, Kentucky and Indiana are shown in Figure 8. A portion of a coverage of river mile points along the mainstem of the Ohio River digitized from USGS 7 1/2 minute topo sheets is shown in Figure 9 along with the RF3 coverage.

CONCLUSIONS

An extensive spatial database has been assembled for the Ohio River Basin and implemented as a series of coverages under the ARC/INFO GIS package. A large majority of this database was accessed from existing files maintained by the U.S. EPA. The EPA databases were augmented by additional data that were digitized, manually input or transferred from other formats into ARC/INFO. Though some quality control has been performed as part of this study on the acquired databases, a comprehensive program for checking and correcting the data has not been done.

FIGURE 7
STATES AND COUNTY COVERAGE FOR STATES IN THE OHIO RIVER BASIN

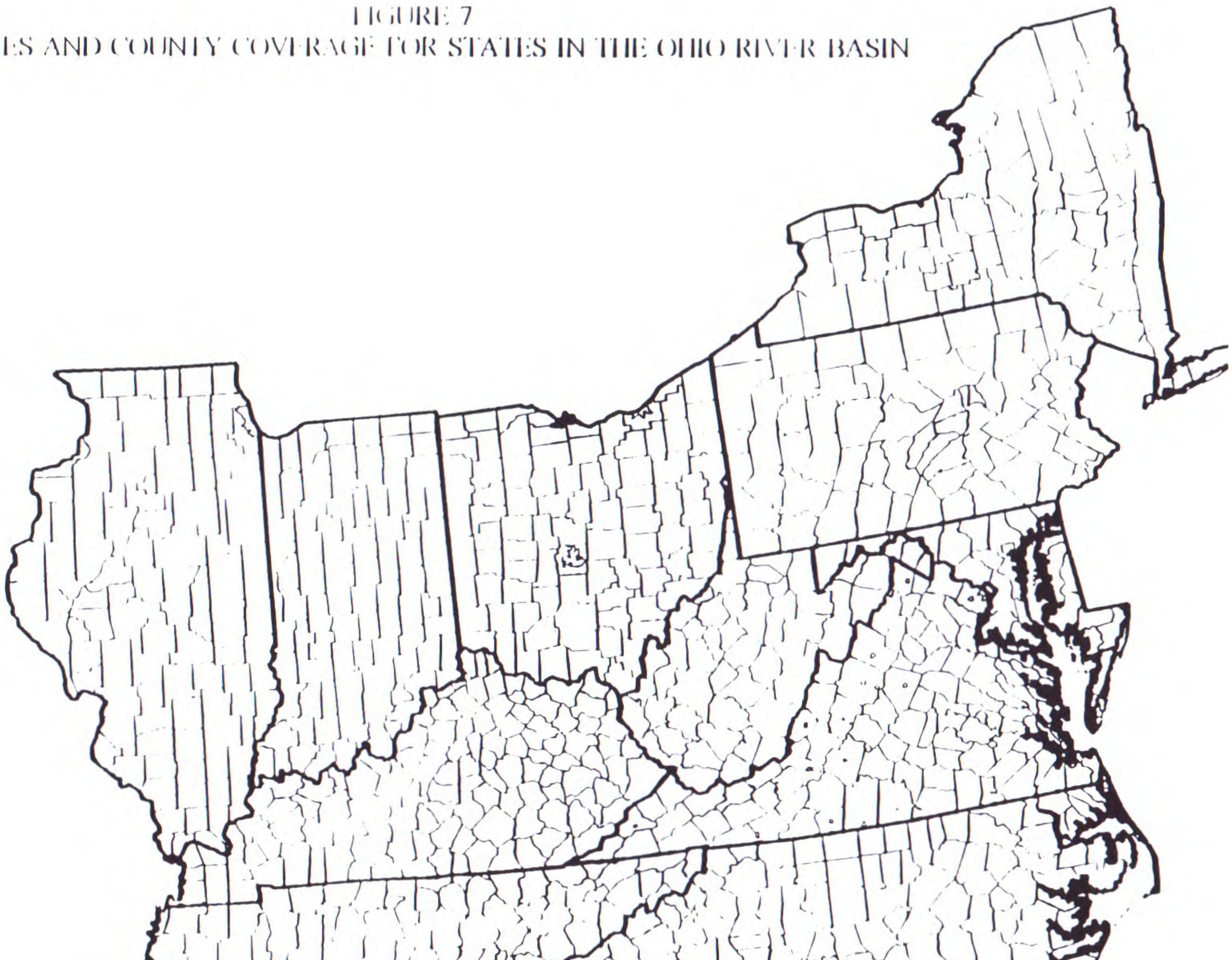


FIGURE 8
DISCHARGER, WATER INTAKE AND DAM COVERAGES FOR
PORTIONS OF OHIO RIVER BASIN

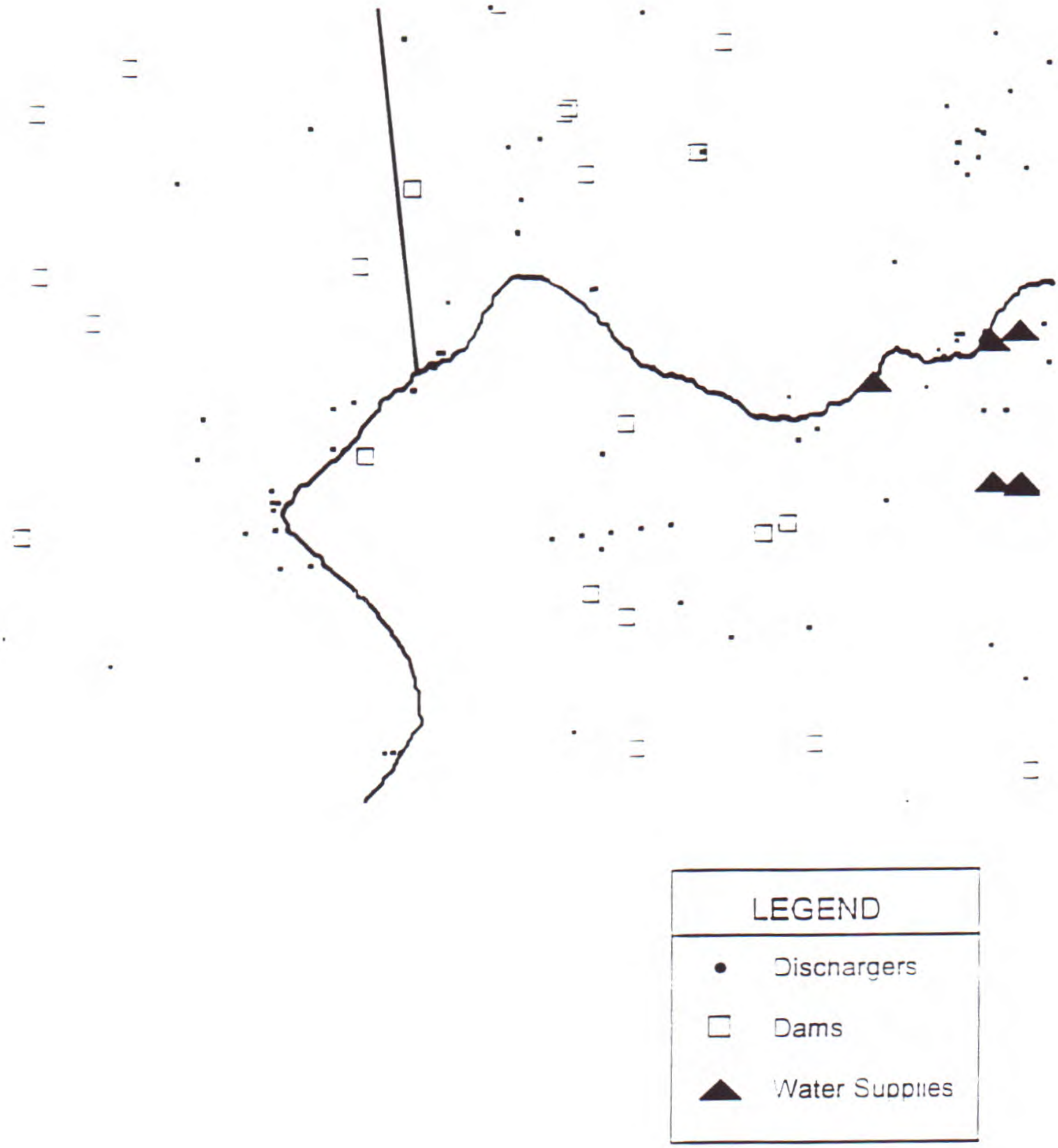
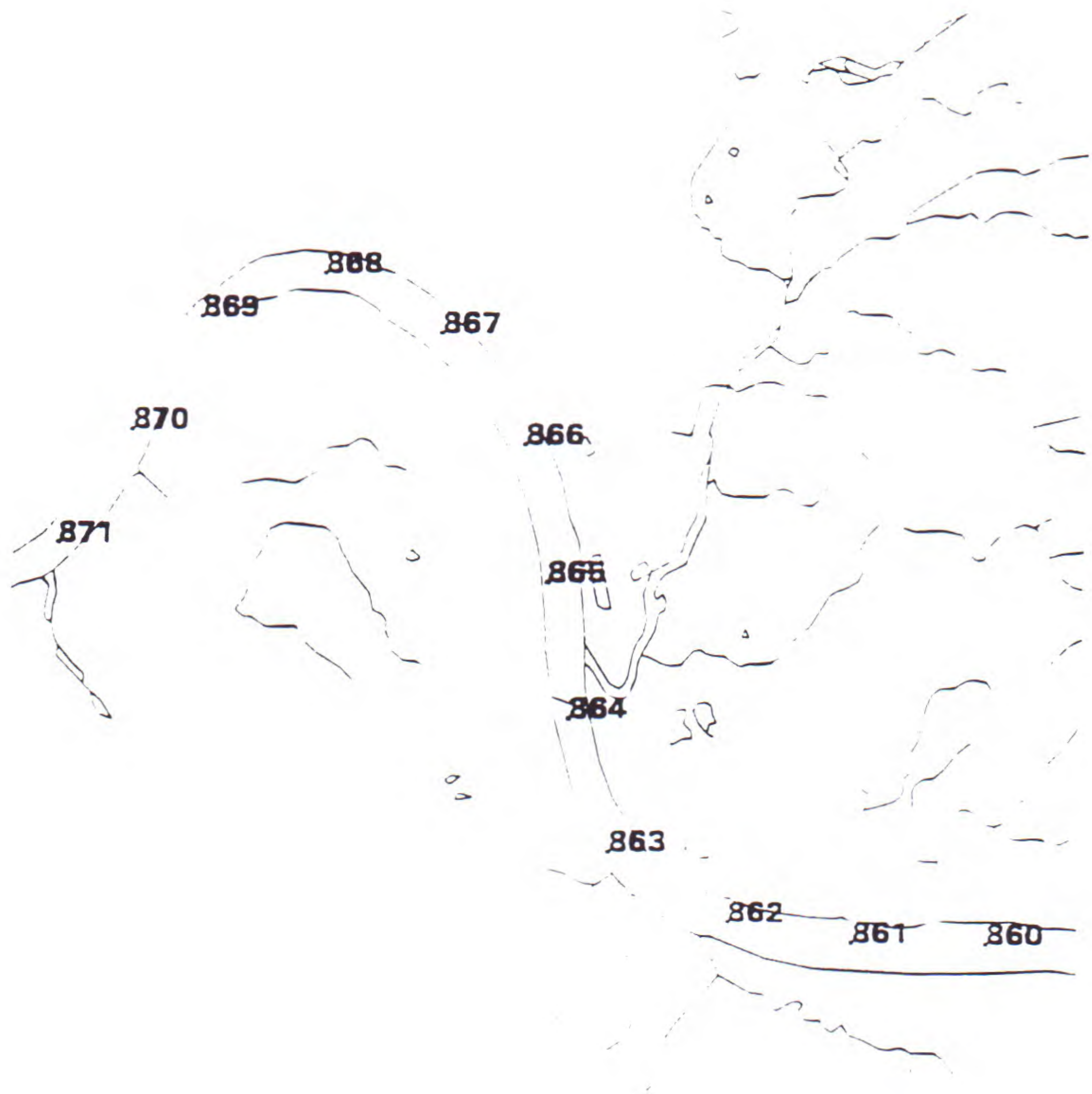


FIGURE 9
RIVER MILE AND RF3 COVERAGES FOR PORTION OF OHIO RIVER MAINSTEM



SECTION 5
WATER QUALITY SPILL MODELS

INTRODUCTION

Since its establishment in 1948, ORSANCO has had a primary interest in the impacts of discharges to the mainstem or tributaries on the water quality in the Ohio River. A major function of ORSANCO has been monitoring the river in order to document and assess the quality of the river. However, monitoring can only provide an historical picture of the water quality in the river as caused by past actions and events. Water quality models of rivers and streams, on the other hand, can provide the means to estimate the potential impact of discharges on the receiving stream. Such models may be used to study the effects of permitted discharges, proposed point discharges, nonpoint sources or the effects of accidental spills. Water quality models have been in use for many decades and are generally accepted as a potentially effective tool in the water quality management arena.

One function of ORSANCO is to serve as a central coordinating office during spill events on the Ohio River. In this role, it oversees the collection of sampling data, estimate the movement of the spill in the river and disseminate information on the estimated arrival, duration and concentration of the spill at water intakes and other locations along the river. In addition to this role, ORSANCO has programmatic interest in other areas in which water quality models can serve a role such as the permitting process and non point source analysis. Its interests include both the mainstem of the Ohio River and on tributaries which may eventually affect the quality in the mainstem. It is towards these various ends that two water quality modeling systems were developed as part of this study.

The two user friendly modeling systems include: 1) a wide scale simplistic model of the entire Ohio River Basin; and 2) a detailed model of the Ohio River mainstem that may be used under emergency spill conditions. Both models are built to interact with the GIS system for display and/or input generation. The wide scale model uses representative steady-state flow regimes and represents movement by simple travel time relationships and transformation by dilution and decay mechanisms. Pollutants are routed through the RF1 Reach File representation of the basin. The detailed mainstem model utilizes actual non steady-state flow patterns as input to EPA's WASP4 water quality model. The Ohio River mainstem is represented in the model by a series of segments ranging in size from two to ten miles in length.

The wide scale model is built as part of the ARC/INFO GIS system and utilizes information on the streams and rivers in the Ohio River Basin available in the GIS. It uses the NETWORK routing software component of ARC/INFO and is written in the native ARC/INFO Arc Macro Language (AML).

The Ohio River mainstem model is composed of a series of linked programs and uses EPA's WASP 4 model as the actual water quality modeling engine. Other components include linkage to the Corps of Engineer FLOWSED hydraulic model, a user friendly menuing system, graphic and report writers, and a tie to the ARC/INFO and ArcView GIS packages as a means of graphically displaying the results of the modeling. The system is intended to facilitate its use under emergency spill conditions.

The development, design and use of both of these systems is described in this section. Additional supporting information is provided in Appendices.

BASINWIDE NETWORK MODELING

The ARC/INFO NETWORK module provides a generic capability to model spatial networks (ESRI, 1992). The traditional applications of this technology have been in the area of transportation networks but recent applications have dealt with hydrological networks. Djokic and Maidment (1993) discuss the application of ARC/INFO NETWORK tools to drainage network verification and sizing using the rational method, watershed isochrone determination and steady state flow computations. The application discussed here deals with the spill modeling under steady state conditions. A formulation is developed that allows the routing of a pollutant discharge through the river network. The modeling can be performed for the entire Ohio River Basin by using the RF1 Reach File river network (Horn and Grayman, 1993) coverage under ARC/INFO . The pollutant may be treated as a conservative element or represented by a first order exponential decay function. The user can select from a number of flow regimes such as average flow, low flow and multiples of these flows.

The model utilizes flow and velocity information available as part of EPA's RF1 Reach File System. The flow estimates were made based on USGS gage information with interpolation and extrapolation methods used to estimate flow in ungaged reaches (Gates, 1982). Velocity is calculated based on a statistical relationship between velocity and streamflow and stream order based on over 2000 separate time-of-travel studies (Gates, 1982).

Network Modeling Terminology and Methodology

Since the ARC/INFO NETWORK module was written for transportation network problems, much of the terminology and concepts are foreign to the water resources field. Therefore, as a first step in network modeling the application specific entities associated with spill modeling must be mapped into the general network entities supported by the ARC/INFO NETWORK module. For the spill modeling application, the mapping was completed as follows. Each reach segment was assumed equivalent to a directed network link. Therefore, the confluence of two or more reaches is a network node. A link has two impedances which are its resistances for the two directions of flow. The fact that hydraulic flow cannot take place from the downstream node to the upstream node is represented in NETWORK by assigning a negative value for the corresponding impedance. The upstream-to-downstream impedance value is set equal to the time of travel in days for that reach segment. The time of travel is calculated by dividing the segment length by the velocity corresponding to a specified flow regime. In the spill model, either the segment length that ARC/INFO automatically calculates based on the geometry or the segment length that RF1 provides could be used. The RF1 lengths were chosen in order to maintain consistency with that data base. However, the two lengths for any segment did not differ by more than 10% and choosing the calculated length would have been equally acceptable.

Once the network description was complete, the next step was to define how the general NETWORK commands can be used to route the pollutants through the river network. NETWORK version 6.1 offers commands for pathfinding, allocation of resources, tracing and spatial interaction. The specific need in spill modeling was to find the spill path and the cumulative time of travel along that spill path. The ALLOCATE OUT command provided the means to accomplish that objective. The spill point was the network center for the ALLOCATE command from which the resource (the pollutant load) could be allocated out. Since this was not a situation where supply was matched with demand but a case where supply items traveled out and could decay during travel, the ALLOCATE command was executed without specifying supply and demand item amounts. The output of the command consisted of the route (an ordered list of links) that the pollutant would take while moving down the river network, and the cumulative impedance (time of travel) to each node in the route.

Having obtained the spill path, the final step was to compute the concentration of the pollutant along the spill path(route). Mass balance equations were used for this purpose. Essentially, the mass balance equations require that the total mass (pounds) of pollutant stays the same (conservative pollutant) or decays out exponentially (nonconservative pollutant) as it moves downstream. For a conservative pollutant such as mercury, the concentration decreases as it travels downstream because of the dilution effect. For a nonconservative pollutant such as BOD (Biochemical Oxygen Demand), both dilution and decay affect the concentration. In the RF1 ARC/INFO coverage, each reach segment is geometrically represented by an arc and all parameters of interest are stored as attributes of each arc. Given that there are 5142 RF1 reach segments in the Ohio River Basin, a fast method had to be devised to calculate concentrations along the route. The flow value for each segment for a given flow regime was available for each segment in the arc attribute table (.AAT table) within the ARC/INFO database. The ALLOCATE command creates another table of route attributes that provides each arc's database record number and the cumulative time of travel to the end-point of that arc (ESRI, 1991). Therefore the flow value for each arc (link) in the route can be quickly obtained through a relational join on the arc rec number item in both tables. It will be shown next that the concentration in any reach can be calculated directl from three quantities: the flow in the reach (cfs), the cumulative time of travel (days) to that reach and the pollutant load at the spill point (lbs/day). Having all three quantities available as columns in a single (related) table, it was possible to create a 'calculated column' for the concentration values. Thus all modeling operation were completely internalized within ARC/INFO.

Pollutant Decay Model

In applying a spill to the river system, it was assumed that the spill takes place at the upstream end c reach segment (arc). If further definition were required for spill location, it is always possible to split arcs us the ARCEDIT module of ARC/INFO. It was also assumed that the pollutant stream flow is negligible compar to the reach segment flow. First consider the first order decay model for a non-conservative pollutant. The conservative pollutant model can be then obtained as a special case.

Let Q_R (cfs) be the flow in reach R where M_R (lbs/day) of the pollutant enters at the upstream end. T concentration C_R at the downstream end of reach R is given by.

$$C_R \text{ (in mg/l)} = (M_R / Q_R) * \exp(-k * T_R) * 0.1855 \qquad (1)$$

where

k is the decay coefficient with a typical value of 0.5/day.

T_R is the time of travel(days) in reach R.

exp denotes the exponential function and

0.1855 is a units conversion factor.

Similarly for the reach segment S immediately downstream of segment R.

$$C_S = (M_S / Q_S) * \exp(-k * T_S) * 0.1855 \qquad (2)$$

The mass balance equation states that

$$M_S = C_R * Q_R / 0.1855 \qquad (3)$$

Combining (1), (2) and (3) and simplifying, results in

$$C_S = (M_R / Q_S) * \exp(-k * (T_R - T_S)) * 0.1855 \quad (4)$$

Denoting $T_R - T_S$ by CT_S, the cumulative time of travel to reach S, results in

$$C_S = (M_R / Q_S) * \exp(-k * CT_S) * 0.1855 \quad (5)$$

Applying the same logic in succession, we find that equation (5) is valid for any reach downstream of reach R. Thus the concentration in any reach can be calculated directly from three quantities: Q_S the flow in the reach (cfs); CT_S, the cumulative time of travel(days) to that reach; and M_R the pollutant load at the spill point(lbs/day).

The equation for conservative pollutants is obtained by setting k = 0.0 in equation (5), resulting in

$$C_S = (M_R / Q_S) * 0.1855 \quad (6)$$

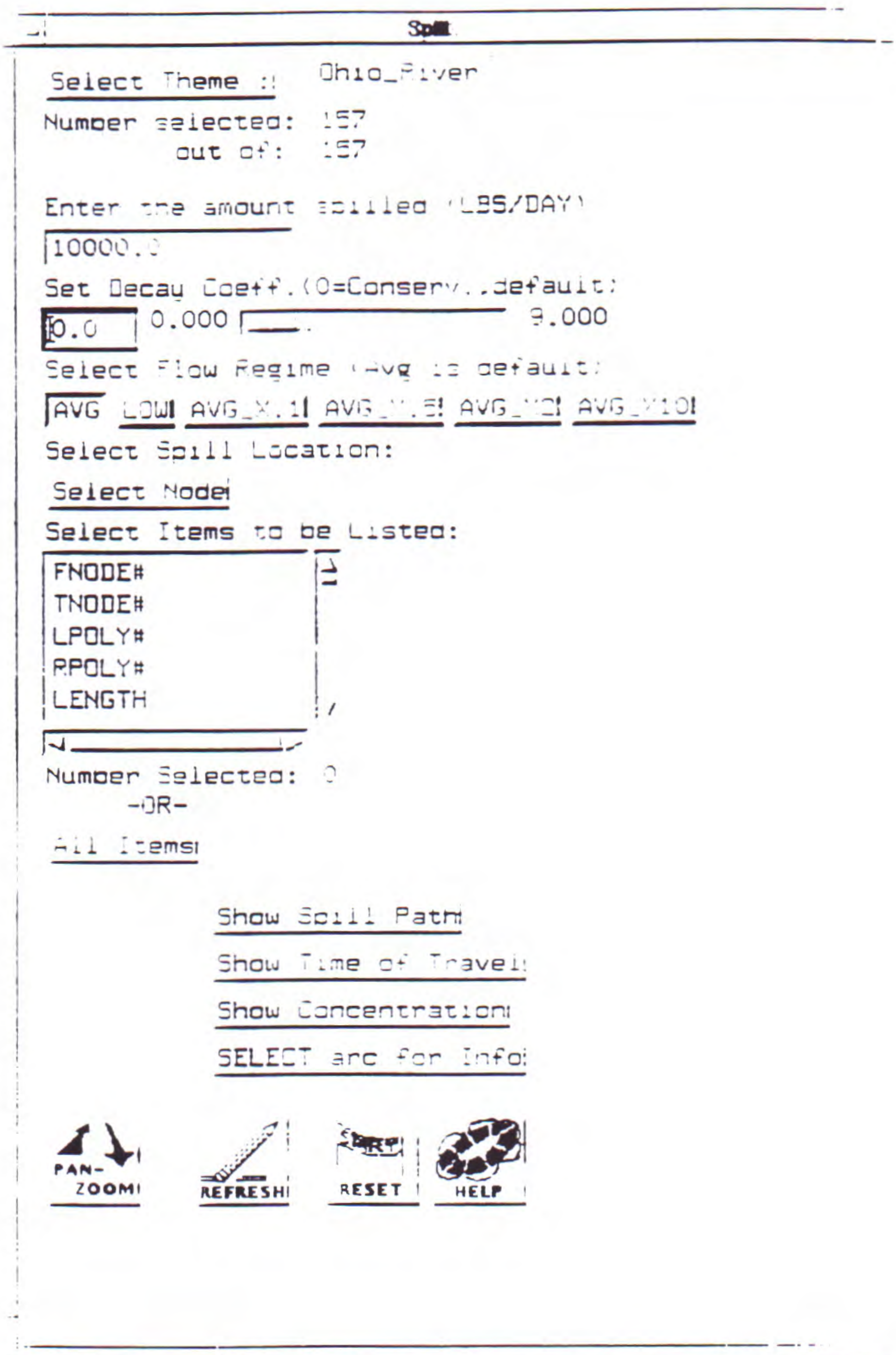
Equation (6) shows that the cumulative time of travel does not affect the pollutant concentration for conservative pollutants and that only dilution is important.

User Interface

The entire modeling procedure has been embedded into an easy-to-use graphical user interface (GUI). The GUI has been created using the Arc Macro Language that runs within ARC/INFO. Thus the spill modeling application has been seamlessly integrated into the ARC/INFO GIS. It has been created using the building blocks of the Generic Demo Interface (GDI) shareware available with ARC/INFO and provides several other query facilities such as statistics and measurements.

The dialog box for the Spill Application is shown in Figure 10. First the user selects the river network to be used from different available networks. For example, the entire Ohio River Basin network can be selected or just the Ohio River mainstem can be selected. Next, the amount of pollutant in lbs/day is entered into its entry box. The decay coefficient can be adjusted using the slider control or entered directly. The flow regime is selected by clicking on the appropriate button. The items of interest for the river segments can be chosen or all items can be viewed by default. The simulation is started by clicking on the 'Pick Spill Location' button. An interactive dialog helps the user pick a node using the mouse where the pollutant enters the river network. In less than 10 seconds, the program (running on a Data General Aviion 412 workstation) draws the spill path in red over the river network. The concentration calculations take an additional 30 seconds. The mouse cursor assumes

FIGURE 10
DIALOG BOX FOR THE BASINWIDE NETWORK SPILL MODEL

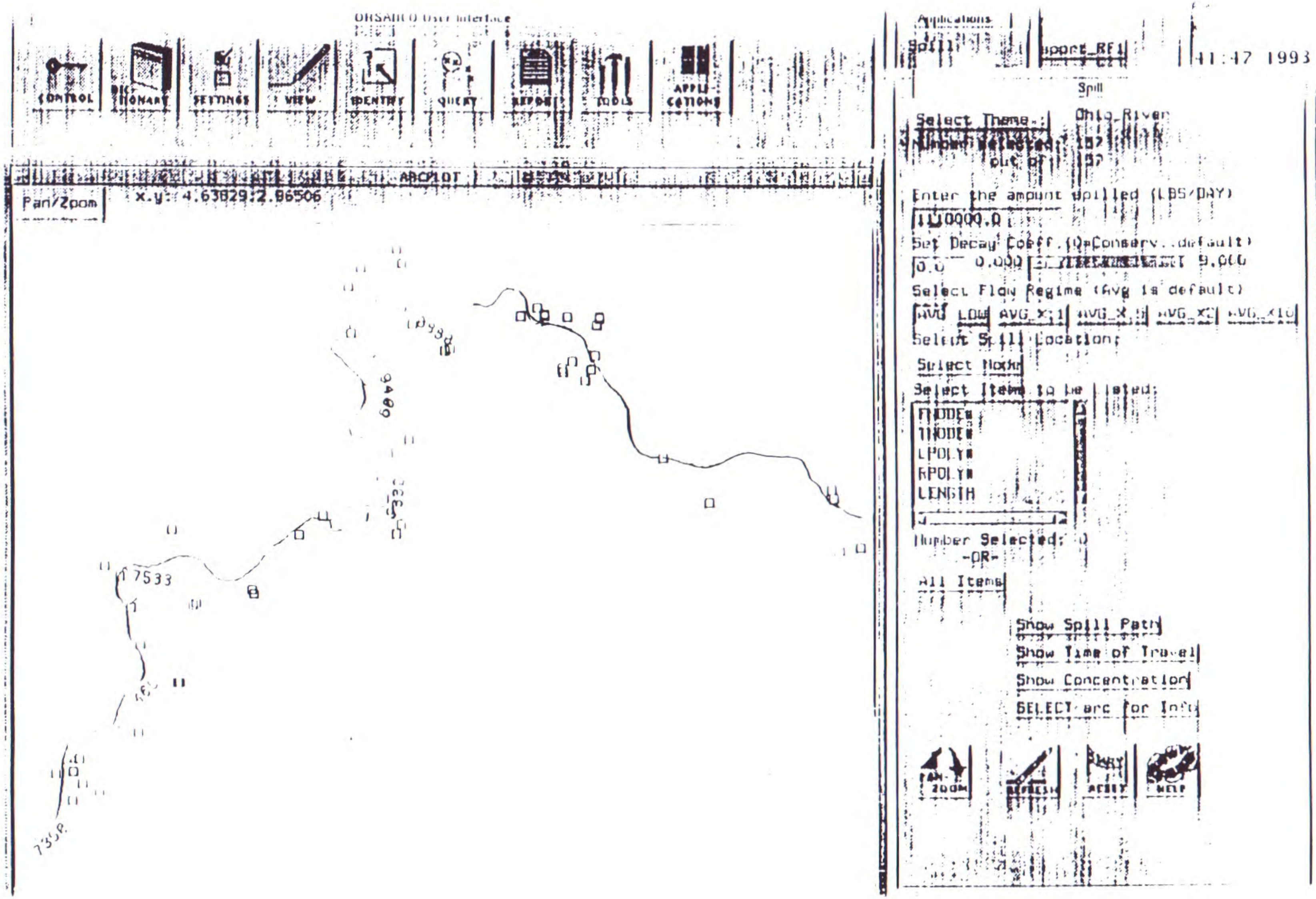


the shape of an hour-glass during the computations. Next, the user can display the concentration values along the reach segments as shown in Figure 11. A number of display control buttons such as pan and zoom are available to pinpoint the areas of interest. Online help is also available to explain the functions of the various buttons. Information about any reach segment (arc) can be obtained quickly by clicking on 'Select arc for info' button, and then interactively selecting arcs. The information for each arc is displayed in a list box that can be scrolled as needed.

Conclusions

An efficient implementation of a simplistic spill model was created using the generic capabilities for network modeling within the ARC/INFO geographic information system (GIS). Emphasis was placed on creating an easy-to-use graphical user interface for executing the spill simulation. The entire system is completely integrated with ARC/INFO through the use of the Arc Macro Language and enables the user to model numerous scenarios quickly and efficiently. The accuracy of model predictions of travel times and pollutant concentration is most sensitive to the flow regime and the decay coefficient. This system serves as another example of the emergence of GIS in hydrology.

FIGURE 11
GRAPHICAL OUTPUT FROM THE BASINWIDE NETWORK SPILL MODEL



OHIO RIVER MAINSTEM MODELING

Background

As an interstate water quality management agency, ORSANCO has long had an interest in the impacts of discharges to the mainstem or tributaries on the water quality in the Ohio River. By the early 1970's ORSANCO had developed and routinely were using computer models which simulated the movement and transformation of contaminants in the river. During the 1970's a more advanced model, QUAL II, was applied to the Ohio River by various agencies using data supplied by ORSANCO (Goodrich and Clark, 1984; Dames & Moore, 1975). In recent years, the computer model had been replaced by a series of nomographs based on travel times that was amenable to use during emergency spill situations (ORSANCO, 1988).

Model Selection

One objective of the present study was the selection and application of a water quality model that could be used by ORSANCO in conjunction with the GIS technology being developed in the study. Though the initial application was to be in the prediction of pollutant movement and transformation under emergency spill situations, the model was also required to be robust enough to allow for future expansion for use in studying the cumulative impacts of discharges to the river.

Several prospective models were identified and examined for potential use in the study. Two water quality models evolved as serious candidates for use in the study: EPA's WASP 4 model (Ambrose et al., 1990) and an Oil Spill Model developed for the Corps of Engineers' Cold Regions Research and Engineering Laboratory (Yapa et al., 1990). Though both models had pros and cons associated with them, the greater likelihood of continued support for WASP by EPA and the wider applicability of WASP for future modeling needs led to the selection of that model for use in this study.

WASP requires detailed hydraulic information as input to the model. Generally this information is provided to the model by linking it to a hydraulic model. Two such models were examined for potential use in the study: EPA's DYNHYD4 model (Ambrose et al., 1990) and the Corps of Engineers' FLOWSED model (Johnson, 1982). DYNHYD can provide detailed flow information, at a small time scale, to WASP and the linkages between DYNHYD4 and WASP have been developed and are routinely supported by EPA. However, DYNHYD4 also requires detailed inflow hydrographs from tributaries and is quite difficult and time consuming to apply. FLOWSED was specifically developed by the Corps of Engineers to provide daily flow and stage elevation predictions. It is used on a daily basis by the Corps' Ohio River Division for reservoir operation and the results of this routine application (5-day forecasts of flow and stage) are available to ORSANCO in electronic form and can be downloaded using a PC and modem at any time. Investigation of the programming requirements to link the output from FLOWSED with WASP indicated that such linkages could be established with relative ease. Since the primary modeling objective in the present study was development of a model that could be used quickly and easily under emergency spill conditions, it was felt that the availability of hydraulic data through the use of FLOWSED by the Corps of Engineers, far outweighed the high operational costs associated the more detailed hydraulic information that could be supplied by DYNHYD4.

Thus the decision was made to utilize EPA's WASP model and to build an interface which would allow the quick and easy use of hydraulic data generated by the Corps of Engineers using their FLOWSED model.

WASP Model

The WASP4 model (Water Quality Simulation Program-4) is an enhancement of the original WASP model (Di Toro et al. 1983). By the end of the project, newer versions of WASP (WASP 5.1) have been released but for continuity purposes, the project has continued with the WASP4 version. WASP4 is a dynamic compartment modeling program for aquatic systems, including both the water column and the underlying benthos. The time varying processes of advection, dispersion, point and diffuse mass loading, and boundary exchange are represented in the program (Ambrose et al., 1990). It has been used widely by EPA and other organizations to model lakes, estuaries and free flowing rivers. EPA's Environmental Research Laboratory in Athens, Georgia continues to enhance the model and provide support.

WASP4 provides a very flexible modeling environment. Within WASP, the aquatic environment is represented as a series of completely mixed compartments. These compartments may be arranged to structure one, two or three dimensional models. Similarly, the modeler may choose from a wide range of kinetic and reactive modules or the modeler may substitute their own modules. Complete details on the theory, structure and use of WASP4 are provided in the EPA documentation (Ambrose et al., 1990). Further details on the application of WASP4 are presented later in this Section.

Ohio River Representation

The WASP4 is a general water quality system with a great deal of flexibility in representing a wide range of aquatic environments. The initial step in using this model (or any other water quality model) is to represent the specific aquatic environment in the model. This includes representing the physical, hydrologic, chemical and biological characteristics of the actual system and the particular pollutant loading patterns being studied in the model. Each of these areas is discussed below.

Segmentation--

The Ohio River was represented in the WASP model by a series of compartments that were strung together in a one-dimensional chain. For the purposes of emergency spill simulation, lateral movement can generally be ignored (except in relatively near field situations) and vertical variation and movement, with possible interaction with the benthos, was considered to be infeasible. It was the general experience of ORSANCO personnel that the lock and dams on the Ohio River tended to result in relatively complete mixing (both laterally and vertically) throughout the river.

The mainstem was divided into 176 segments (compartments) ranging in length from 2 miles to 10 miles and averaging 5.5 miles in length. A minimum segment length of 2 miles was selected to improve the stability of the numerical scheme in WASP. The upstream and downstream locations of all segments corresponded to locations on the Ohio River where Corps of Engineers approximate cross sectional information was available and where flow and stage information was reported by FLOWSED. Table 6 contains a listing of segments and the locations of water intakes, dams and tributaries by segment.

A maximum of 60 segments can be simulated by WASP4. Thus, the maximum length of river that can be modeled, at a single time, is approximately 250 miles in the upper portion of the river to 450 miles in the lower portion of the river where there are longer segments.

Within WASP, the volume of the segment is calculated at the start of the simulation and is either assumed to remain constant throughout the run or is linearly adjusted with flow. Neither assumption is totally valid since the first method does not account for changes in water surface elevation while the second method assumes that velocity is directly related to the flow. The EPA support group at the Athens laboratory recommended the use of the first assumption. The volume of the segment is determined by: a) calculating the cross sectional area at the upstream and downstream ends of each segment for the water surface elevation at the time of the spill; 2) averaging these two areas; and 3) multiplying the average cross sectional area by the length of the segment.

TABLE 6
OHIO RIVER SEGMENTATION

SEG. NO.	U/S MILE	D/S MILE	DRINK. WATER INTAKES	DAMS (U/S End)	TRIBUTARIES
1	0.00	2.00			Alleg-Mon
2	2.00	4.00			
3	4.00	6.20	West View		
4	6.20	10.50	Robinson Twp.	Emsworth	
5	10.50	13.20	Moon Twp.		
6	13.20	16.00		Dashields	
7	16.00	21.00			
8	21.00	25.00			
9	25.00	28.50			Beaver R.
10	28.50	31.70			
11	31.70	35.00		Montgomery	
12	35.00	39.00	Midland		
13	39.00	42.50	E.Liverpool		
14	42.50	46.00			
15	46.00	50.50			
16	50.50	54.40			
17	54.40	57.50		New Cumberland	
18	57.50	61.90	Toronto		
19	61.90	67.20	Weirton, Steubenville		
20	67.20	72.20			
21	72.20	78.30			
22	78.30	84.20			
23	84.20	86.50		Pike Island	
24	86.50	90.62	Wheeling		
25	90.62	96.00	Bellaire		
26	96.00	101.62			
27	101.62	107.25			
28	107.25	112.37			
29	112.37	116.75			
30	116.75	120.75			
31	120.75	124.00			
32	124.00	126.40			
33	126.40	129.50		Hannibal	
34	129.50	133.50			
35	133.50	138.50	Sisterville		
36	138.50	143.50			
37	143.50	148.50			
38	148.50	153.00			
39	153.00	157.00			
40	157.00	161.70			
41	161.70	164.00		Willow Island	
42	164.00	168.80			
43	168.80	172.60			
44	172.60	178.00			Muskingum R.
45	178.00	184.10	Parkersburg		
46	184.10	189.00			Little Kanawha
47	189.00	193.50			
48	193.50	199.00			
49	199.00	203.90			Hocking R.
50	203.90	207.50		Belleville	
51	207.50	212.50			
52	212.50	217.50			
53	217.50	223.00			
54	223.00	228.50			
55	228.50	233.50			
56	233.50	237.50			
57	237.50	240.00		Racine	
58	240.00	244.00			

TABLE 6 OHIO RIVER SEGMENTATION (continued)

SEG. NO.	U/S MILE	D/S MILE	DRINK. WATER INTAKES	DAMS (U/S End)	TRIBUTARIES
59	244.00	250.00			
60	250.00	255.00			
61	255.00	260.00			
62	260.00	266.10			
63	266.10	270.50			Kanawha R.
64	270.50	276.00			
65	276.00	279.20			
66	279.20	282.50		Gallipolis	
67	282.50	287.00			
68	287.00	291.50			
69	291.50	298.50			
70	298.50	303.50			
71	303.50	305.70	Huntington		Guyandot R.
72	305.70	311.40	Huntington		
73	311.40	313.50			
74	313.50	317.60			
75	317.60	319.50			Big Sandy R.
76	319.50	322.50	Ashland		
77	322.50	327.00			
78	327.00	332.50	Ironton		
79	332.50	336.90			
80	336.90	341.00			
81	341.00	345.50		Greenup	
82	345.50	348.50			
83	348.50	352.00	Portsmouth		
84	352.00	354.50			
85	354.50	357.00			
86	357.00	360.00			Scioto R.
87	360.00	364.50			
88	364.50	369.50			
89	369.50	374.50			
90	374.50	379.50			
91	379.50	384.50			
92	384.50	390.00			
93	390.00	395.50			
94	395.50	400.50			
95	400.50	405.50			
96	405.50	408.70	Maysville		
97	408.70	412.00			
98	412.00	417.50			
99	417.50	422.50			
100	422.50	427.50			
101	427.50	432.50			
102	432.50	436.20			
103	436.20	438.40		Meldahl	
104	438.40	441.25			
105	441.25	447.00			
106	447.00	453.00			
107	453.00	458.00			
108	458.00	463.90	(Cincinnati, Kenton Co., Newport)		
109	463.90	470.50			Little Miami R.
110	470.50	475.25			Licking R.
111	475.25	483.50			
112	483.50	491.50			
113	491.50	498.55			Great Miami R.
114	498.55	506.55			
115	506.55	514.15			
116	514.15	521.25			
117	521.25	527.50			

TABLE 6 OHIO RIVER SEGMENTATION (continued)

SEG. NO.	U/S MILE	D/S MILE	DRINK. WATER INTAKES	DAMS (U/S End)	TRIBUTARIES
118	527.50	531.60			
119	531.60	535.00		Markland	
120	535.00	542.00			
121	542.00	546.30			
122	546.30	552.15			Kentucky R.
123	552.15	557.80			
124	557.80	561.50			
125	561.50	568.00			
126	568.00	576.00			
127	576.00	584.00			
128	584.00	595.00	Louisville		
129	595.00	603.00	Louisville		
130	603.00	606.80			
131	606.80	610.00	New Albany	McAlpine	
132	610.00	614.15			
133	614.15	623.50			
134	623.50	632.85			Salt River
135	632.85	642.20			
136	642.20	651.55			
137	651.55	660.90			
138	660.90	670.93			
139	670.93	680.97			
140	680.97	691.00			
141	691.00	701.03			
142	701.03	711.07			
143	711.07	716.08			
144	716.08	720.70			
145	720.70	730.33		Cannelton	
146	730.33	739.57			
147	739.57	748.80			
148	748.80	758.03			
149	758.03	767.27			
150	767.27	771.88			
151	771.88	776.10			
152	776.10	780.55		Newburgh	
153	780.55	782.57			
154	782.57	784.80			
155	784.80	794.90	Evansville		Green River
156	794.90	805.20	Henderson		
157	805.20	815.50			
158	815.50	825.80			
159	825.80	836.10	Mt.Vernon		
160	836.10	841.25	Morganfield		
161	841.25	846.00	Uniontown		
162	846.00	848.20		Uniontown	
163	848.20	858.81			Wabash River
164	858.81	868.83			
165	868.83	878.84	Sturgis		
166	878.84	888.86			
167	888.86	898.87	Rosiclaire		
168	898.87	908.89	Golconda		
169	908.89	918.50			
170	918.50	923.12		Smithland	
171	923.12	929.56			Cumberland R.
172	929.56	935.80	Paducah		
173	935.80	938.90			Tennessee R.
174	938.90	949.85		L&D 52	
175	949.85	960.50			
176	960.50	971.15			
177	971.15	979.60	Cairo		

Flow information--

When WASP is used in conjunction with a model such as FLOWSED which provides flow information on a daily basis, then flow is assumed to be spatially constant over long stretches of the river (flow segments). Due to limitations in the number of such flow segments that can be simulated by WASP (maximum of 4 in a single run of the model), flow segments were defined based on only the major tributaries. A list of flow segments and their relationship to tributaries is presented in Table 7. For each flow segment, the average flow is calculated (for each day) by averaging the flow at the upstream and downstream ends of the flow segment.

Within WASP, the flow increments resulting from each tributary are, in effect, modeled separately. Thus, for example, as shown in Figure 12, if there are three major tributaries, then there are four flow segments. This is represented by four "passes" of the model. In the first pass, the entire mainstem is modeled assuming a flow corresponding to the flow in flow segment A. In the second pass, the flow from the most upstream tributary is modeled through flow segments B, C and D. This flow corresponds to the difference in flow between flow segment B and flow segment A. This is repeated in the third and fourth pass for the other two tributaries. It is currently assumed that the flow in the river will always be greater downstream of a tributary. This would be expected to be true in the large majority of cases. However, if a flood wave was moving down the river and had not reached the point where the tributary enters the mainstem, the flow in the mainstem upstream of the tributary could exceed the flow downstream of the tributary.

Dispersion--

Dispersion in the river can be represented in WASP4 by dispersion coefficients. However, additionally, due to the solution technique, numerical dispersion is also introduced causing a pollutant plume to spread out in space and time as it moves down the river. Because of the relatively large numerical dispersion that can be introduced and the lack of actual field data on dispersion that can be gleaned from dye studies, no explicit dispersion was assumed in representing the river. The sensitivity of the model response in terms of numerical dispersion to changes in the model time step were investigated and will be discussed later in the Section.

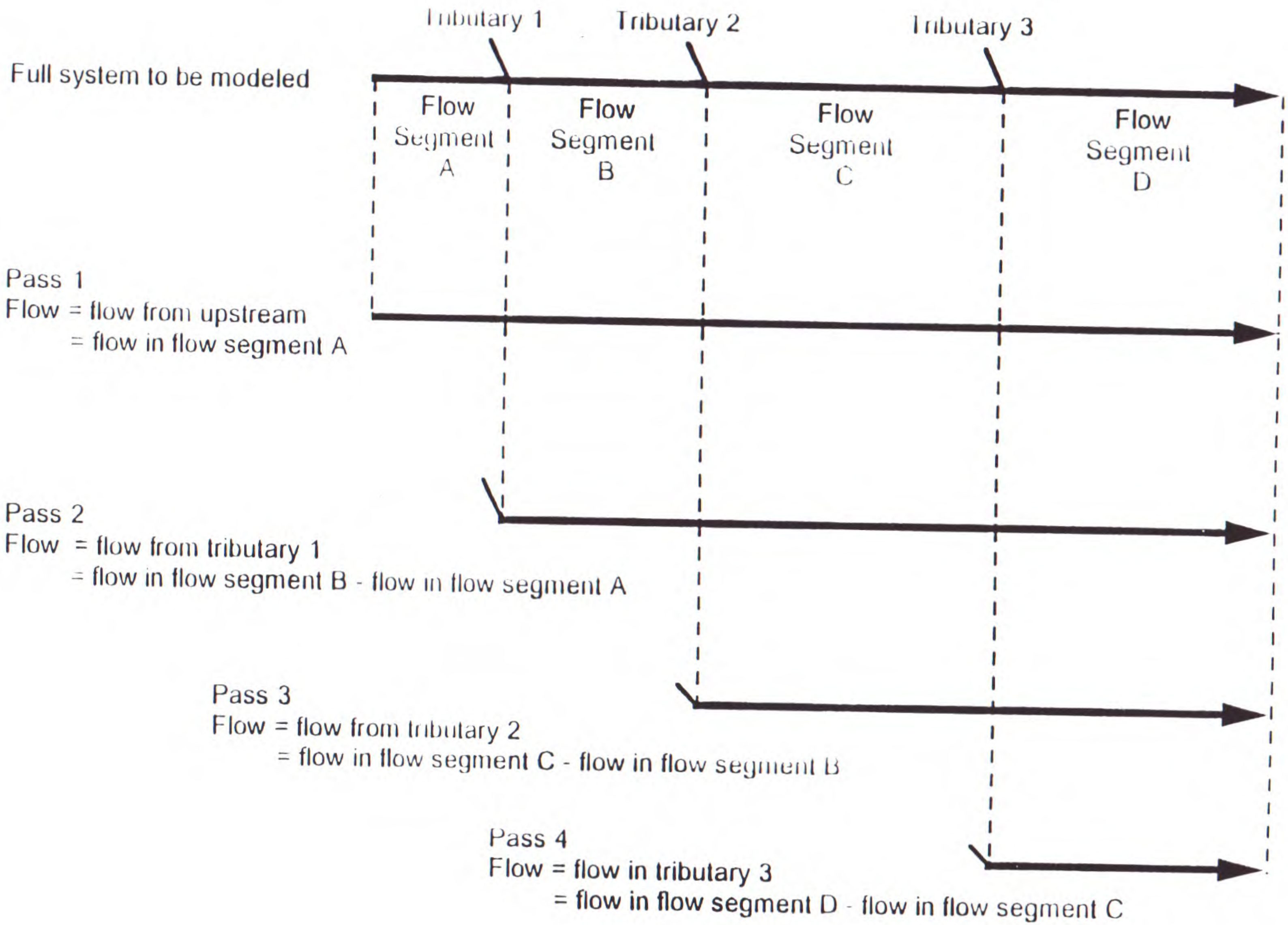
Reaction processes and rates--

WASP4 is capable of representing a wide range of kinetic reactions. For each constituent (pollutant) there may be different processes that occur (e.g. volatilization, bacterial degradation, oxidation, etc.), different pathways such as interaction with the benthos, and multiple rate coefficients for each process and path. Application of such specific kinetic processes is generally very time consuming, data intensive and pollutant, and sometimes, location specific. Such a process does not lend itself well to the prediction of a very wide range of pollutants under emergency spill conditions.

TABLE 7
OHIO RIVER FLOW SEGMENTS

UPSTREAM RIVER MILE	DOWNSTREAM RIVER MILE	UPSTREAM TRIBUTARY
0.00	171.60	ALLEGHENY-MONONGAHELA
172.60	265.10	MUSKINGUM
266.10	316.60	KANAWHA
317.60	490.50	BIG SANDY
491.50	545.30	GREAT MIAMI
546.30	847.60	KENTUCKY
848.20	922.72	WABASH
923.12	979.60	CUMBERLAND

FIGURE 12
CALCULATION OF FLOWS FOR WASP4 FLOW SEGMENTS



An alternative mechanism for representing a wide range of pollutants is the use of a first order decay function of the form:

$$c_t = c_o \exp(-kt) \qquad (7)$$

where c_t is the concentration at time t

c_o is the concentration at time 0

k is the rate coefficient (generally in $days^{-1}$)

Values of k may range from near zero representing conservative substances to values exceeding 100. A value of k of 1.0 means that after 1 day, the concentration of the substance will be only 36.8% of the original concentration. Various publications provide data that may be used to estimate values for k (Mabey et al. 1981; Bowie et al., 1985).

Pollutant loading--

In WASP4, pollutant loadings are specified by segments. Loads may be constant or may vary over time. In using WASP4 as an emergency spill model, only a single load is specified and this load is represented as a pulse of specified duration and rate (pounds). WASP4 is capable of accepting a time history of loadings at single or multiple points though this feature is not used in the emergency spill model.

Model Verification

Verification of a model generally requires detailed field data collected under different hydrologic conditions and pollutant loadings. Alternatively, controlled dye tracer experiments can be used to develop a data set that may be used to calibrate or verify a model. Unfortunately, there is relatively little detailed field data tracing spills on the Ohio River and, due to the high costs of dye studies, no systematic dye studies have been performed. As a result, there is only limited information that can be utilized to test the validity of WASP4.

One limited data set that is available resulted from a spill of toluene that occurred at river mile 609.4 in the Louisville, Kentucky area on March 10, 1993 at approximately 12:00 PM (noon) CST. This spill was estimated at a total of 470,000 pounds with an estimated duration of approximately 3 hours. Samples were taken on the following day at the Cannelton Dam sampling site and analyzed at the organics data station (ODS) at Evansville, Indiana. The WASP model was applied on March 11, 1993 to give some first cut estimates of the movement and transformation of the plume. Subsequently, the model was applied to study the calibration process and to verify the model results.

Flow in the river at the time of the spill was considered to be moderately high. The flow at the spill site was approximately 425,000 cfs. As a comparison, average annual flow at that location in the river is 115,000 cfs and average monthly flow for March (the statistically highest flow month) is 250,000 cfs.

Analysis of the toluene concentrations at Cannelton (see Figure 13), led to the following observations:

- the width of the plume remained remarkably narrow during the 33 hour, 110 mile travel between the spill location and Cannelton dam. At Cannelton, the time between the leading edge and the (extrapolated time of the) trailing edge is approximately 13 hours. The temporal width of the plume at a concentration of 50% of the peak concentration at Cannelton is approximately 4 hours.

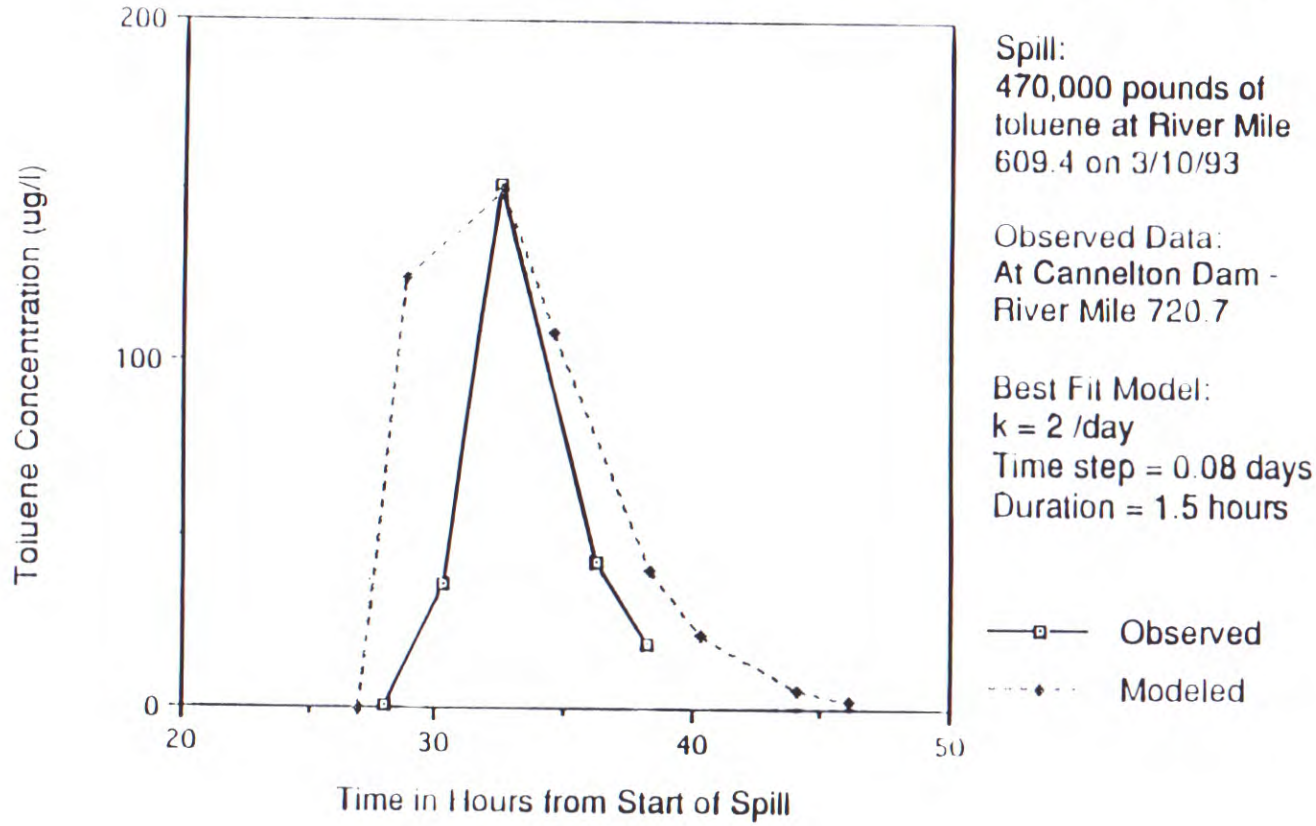
- The calculated mass of toluene passing Cannelton is approximately 66,000 pounds or around 14% of the estimated total spill.

In applying the WASP4 model, three parameters (duration, time step, and decay) were varied over a reasonable range and the best fit curve at Cannelton was found. Durations of 3 hours (the reported spill duration) and 1.5 hours were studied. Because of the narrow temporal width of the observed plume at Cannelton, the shorter duration of 1.5 hours resulted in a better fit. For the decay coefficient, values of 0.01 per day (essentially conservative) up to 2.5 per day were studied. When a simple mass balance between the reported spill load and the observed mass passing Cannelton is performed and an average travel time of 33 hours is applied, the calculated first order decay rate is 1.43 per day. Though toluene is considered quite volatile, no representative decay rates were found in the literature. A decay rate of 2.0 per day was found to give the best model results.

The third parameter, the WASP model time step, was varied between 0.02 days and 0.08 days. Generally, smaller time steps result in greater numerical dispersion in the model results. At the other end, if the time step is too large, the model will become unstable.

The narrowness of the plume at Cannelton suggests that there is little dispersion in the river under the observed flow conditions. Thus, the logical time step (and the one that gave the best results) was the largest time step that did not lead to model instability. A time step of 0.08 days was selected. In fact, this time step did lead to some instability in the model downstream of Cannelton but appeared to be stable in the area of interest. Instability is easily recognized by large fluctuations in concentrations between segments and sometimes, increasing concentrations as you move downstream.

FIGURE 13
OBSERVED AND MODELED TOLUENE CONCENTRATIONS IN THE OHIO RIVER



Model results based on the selected values for the three parameters are shown in Figure 13 along with the sampled concentrations. Based on this single set of data and parameterization, the observed and modeled results seem to be in relatively close agreement. However, one cannot make any statement that the model is verified based on this small sample. At best, one can claim that, for this set of data, the parameters of the model can be adjusted, while being kept within a reasonable range, so that observed and modeled results are within an acceptably close level of agreement. Obviously, a much larger and diverse field data set is needed to provide more experience in calibrating the model and in developing the confidence to accept the model as verified. Such a data set can be developed from a controlled field study using an environmentally acceptable dye or other tracer or by sampling an actual spill event. The controlled study is preferable because of the time provided for planning and preparing for the study.

Model Sensitivity

Sensitivity analysis is the process of systematically varying model parameters through a logical range of values to determine the response of the model to these changes in values. Such a procedure can be used to provide a better understanding of the model and serve as a guide in determining where best to put one's resources in gathering data.

A limited sensitivity analysis was performed on the WASP4 model in order to gain a better understanding of its operation. The response of the model was determined for two parameters: decay rate and the time step, under two different flow conditions. The first flow condition was the moderately high flow that accompanied the toluene spill in March 1993. The second flow condition was a moderately low flow condition in February 1992. While the average flow at Louisville is approximately 115,000 cfs, the March 1993 flow was 425,000 cfs and the February flow was 58,000 cfs.

For each flow condition, a loading comparable to the spill loading (470,000 pounds) was input into the model over a duration of 24 hours. Decay rates of 0.1, 1.0, and 2.5 per day were studied for the high flow condition and decay rates of 0.01, 0.1, and 0.5 per day were tested for the low flow condition. The results for the two flow conditions are presented in Figure 14-a and Figure 15-a. In both cases, the lower decay rate results in a significantly higher concentration. At the lower flow rate, the velocities in the river are much lower (thus longer travel times) resulting in greater attenuation. As one would expect, the decay rate significantly affects the predicted peak concentration. However, it has little impact upon the timing; generally only shifting the time to peak slightly forward with increasing decay rate.

FIGURE 14: WASP SENSITIVITY UNDER HIGH FLOW CONDITIONS

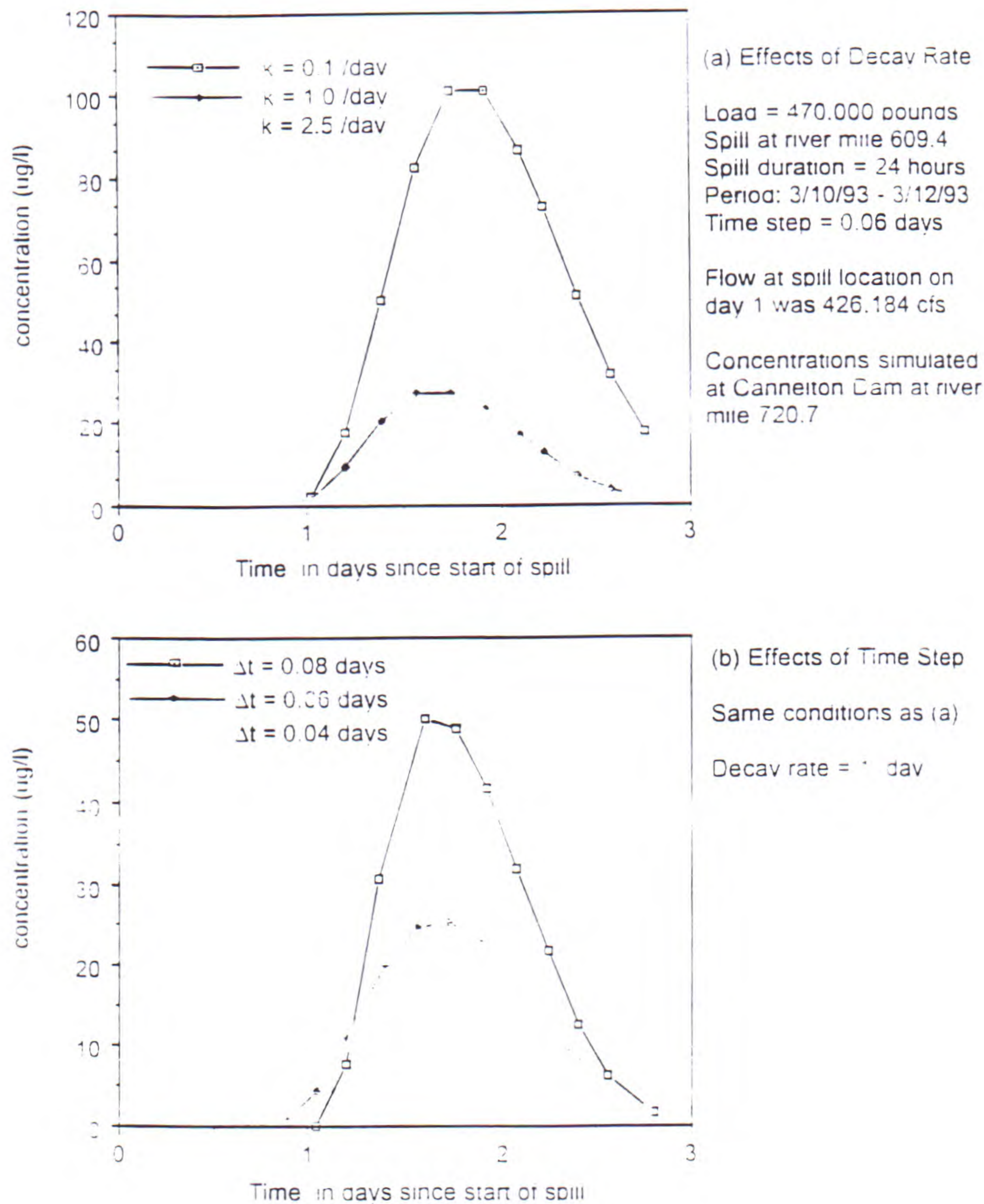
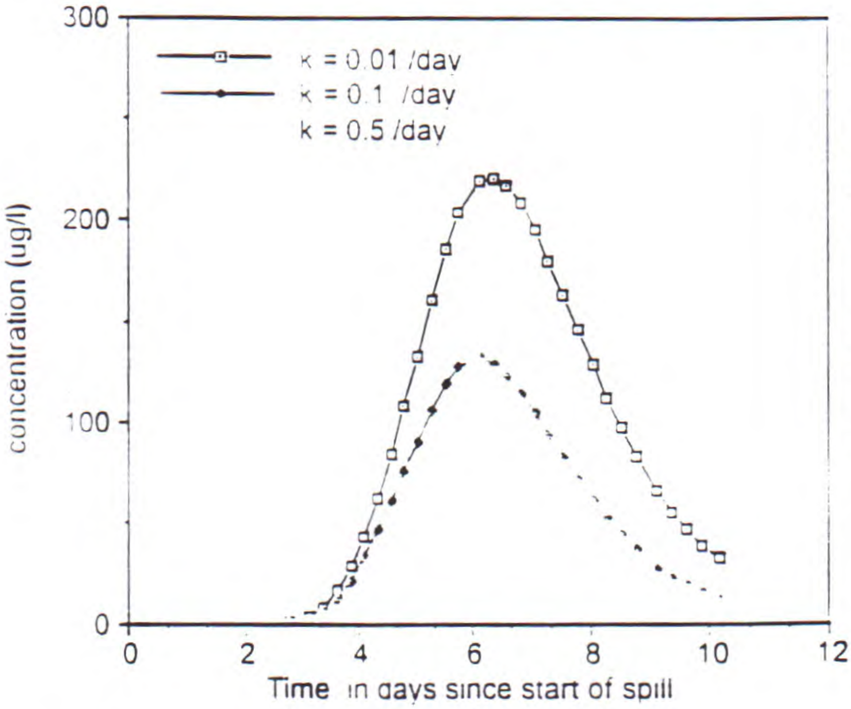


FIGURE 15: WASP SENSITIVITY UNDER LOW FLOW CONDITIONS

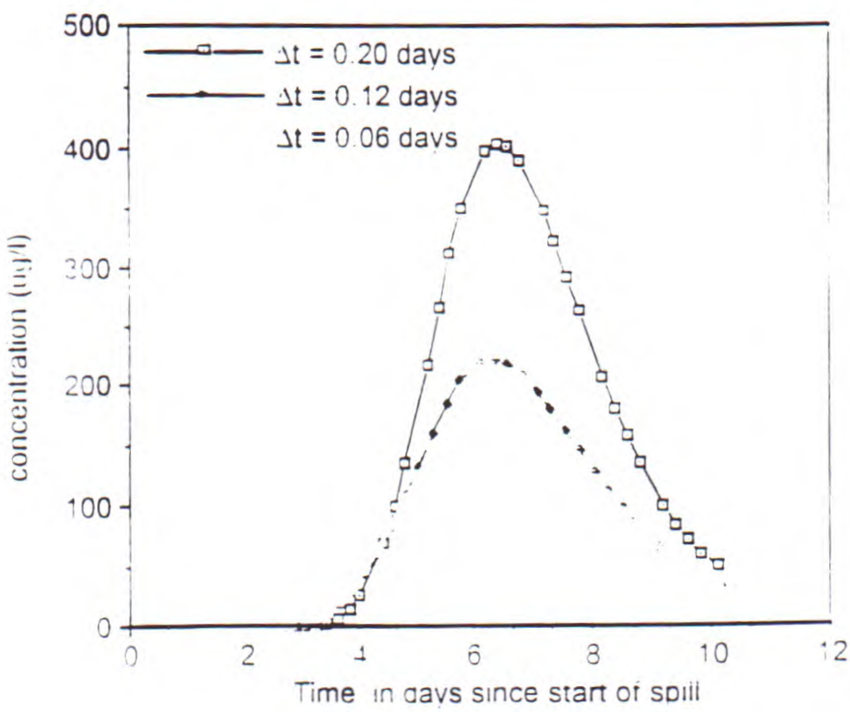


(a) Effects of Decay Rate

Load = 470,000 pounds
Spill at river mile 609.4
Spill duration = 24 hours
Period: 2/6/92 - 2/15/92
Time step = 0.12 days

Flow at spill location on day 1 was 57,690 cfs

Concentrations simulated at Cannelton Dam at river mile 720.7



(b) Effects of Time Step

Same conditions as (a)

Decay rate = 0.01 /day

Figures 14-b and 15-b illustrate the model response to different time steps under high and low flow conditions respectively. In both cases, the largest value for the time step, 0.08 days and 0.20 days for high and low flow respectively resulted in some model instability and also significantly higher peak concentrations. The model was relatively insensitive to the lower values for time steps. With decreasing flows, the allowable time step values increase. The reasons and consequences of this fact are described later in this section.

Selection of Model Time Steps

The WASP4 User's Manual (Ambrose, 1990) provides some guidance on the selection of the model time step. The model is most susceptible to becoming unstable if the time step approaches or exceeds the residence time (flow through time) in any segment. The residence time in a segment is defined as:

$$\text{residence time} = \text{Segment Volume} / \text{Flow} \qquad (8)$$

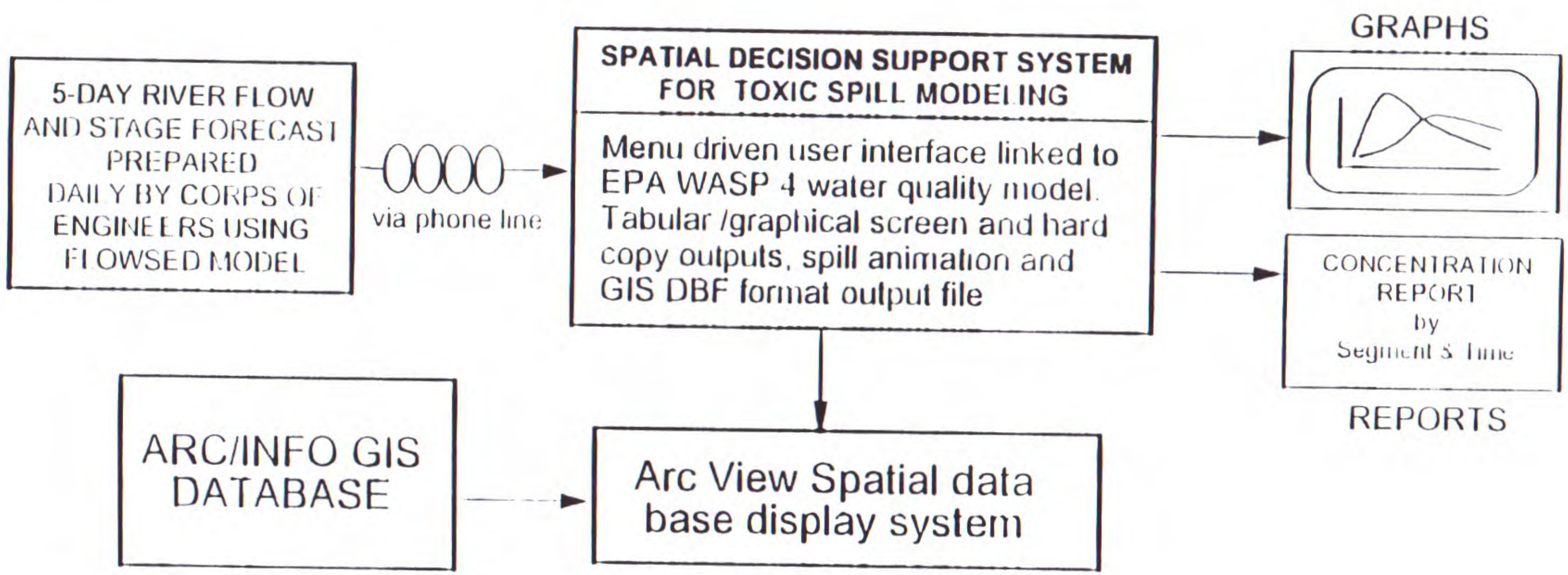
Generally, if the time step is set so that it is approximately 80 to 90% of the minimum residence time then instability will be avoided and numerical dispersion is minimized. In order to achieve the best results, it is optimal to set the segmentation so that the residence times in segments are approximately equal.

Spill Management System

The WASP4 model is a relatively complex model that requires an input file that is quite lengthy. Furthermore, the input file format mimics the fixed format, card image form that was popular on mainframe computers with card readers. In order to facilitate its use under emergency spill conditions, a user friendly 'Spill Management System' was developed. A complete user's manual for the Spill Management System is presented in Appendix C. Details on the structure, programs and files associated with the Spill Management System are presented in Appendix D.

The approach selected for the Spill Management System was to utilize well accepted and tested, existing models and to embed those models into a user friendly 'shell'. The system is implemented on a PC-based work station. The primary elements in the Spill Management System are shown in Figure 16. Five day flow and stage predictions along the Ohio River are downloaded from the Corps of Engineers via phone lines and used as input to the System. The Spill Management System itself, is composed of a menu driven interface which uses the flow information and guides the user through a series of operations centered around the WASP4 model and produces several graphical and tabular displays.

FIGURE 16
SCHEMATIC REPRESENTATION OF SPILL MODELING SYSTEM PROCESS



The Spill Management System consists of a series of programs and files which interact and are invoked by the user from a menuing system. All of the programs are written in the C language with the exception of EPA's WASP4 model which is written in FORTRAN. Interaction between the program modules is provided by files which are written by one program and used as input to other programs.

The primary functions of the Spill Management System are as follows:

1) Construct an input file for the WASP4 model based on a short form into which the user enters information on the spill.

2) Run the WASP4 model.

3) Generate a report summarizing the WASP4 results which may be viewed on the screen or directed to a printer.

4) Generate x-y plots showing the results of the WASP4 run in the form of concentration vs. time at particular sites or concentration vs. river mile at particular times. The plots may also be animated to show the movement of the spill plume down the river over time.

5) Generate a summary file which may be used by the ARCVIEW program to view the results of the WASP4 run in relationship to other GIS coverages.

The main menuing system is composed of a main menu bar and a series of 'pull-down' menus (pillars) as shown in Figure 17. The five pull-down menus allow the user to perform a series of tasks:

1) The System pillar provides the user access to several utilities and access to the DOS prompt.

2) The Control File pillar is used to construct a 'control file' containing the characteristics of the spill to be simulated and to construct the input file for the WASP4 model. The general characteristics of the spill to be simulated are entered into a form as shown in Figure 18.

3) The Wasp pillar is used to run the WASP4 model.

FIGURE 17
SPILL MODELING SYSTEM MENU

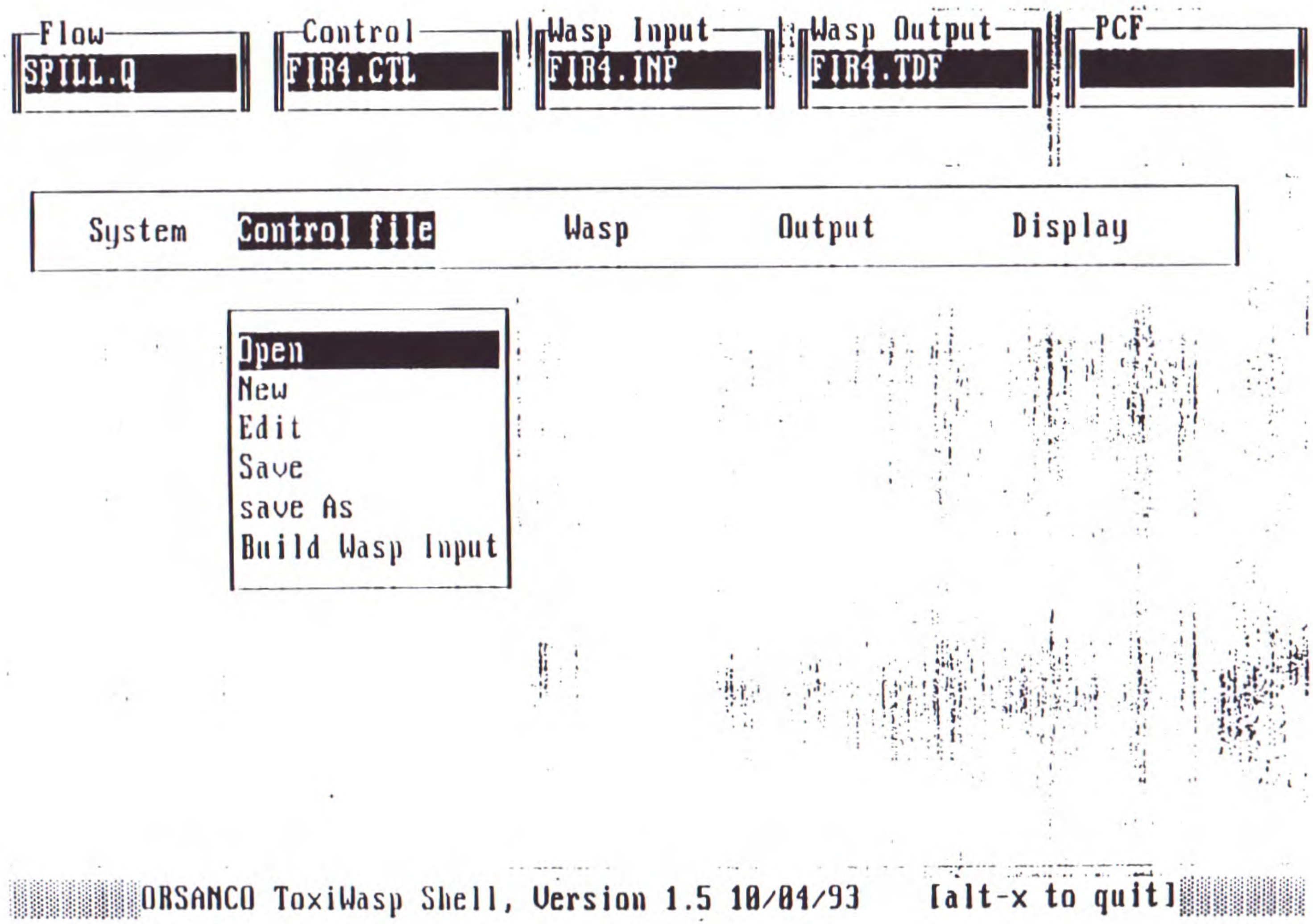
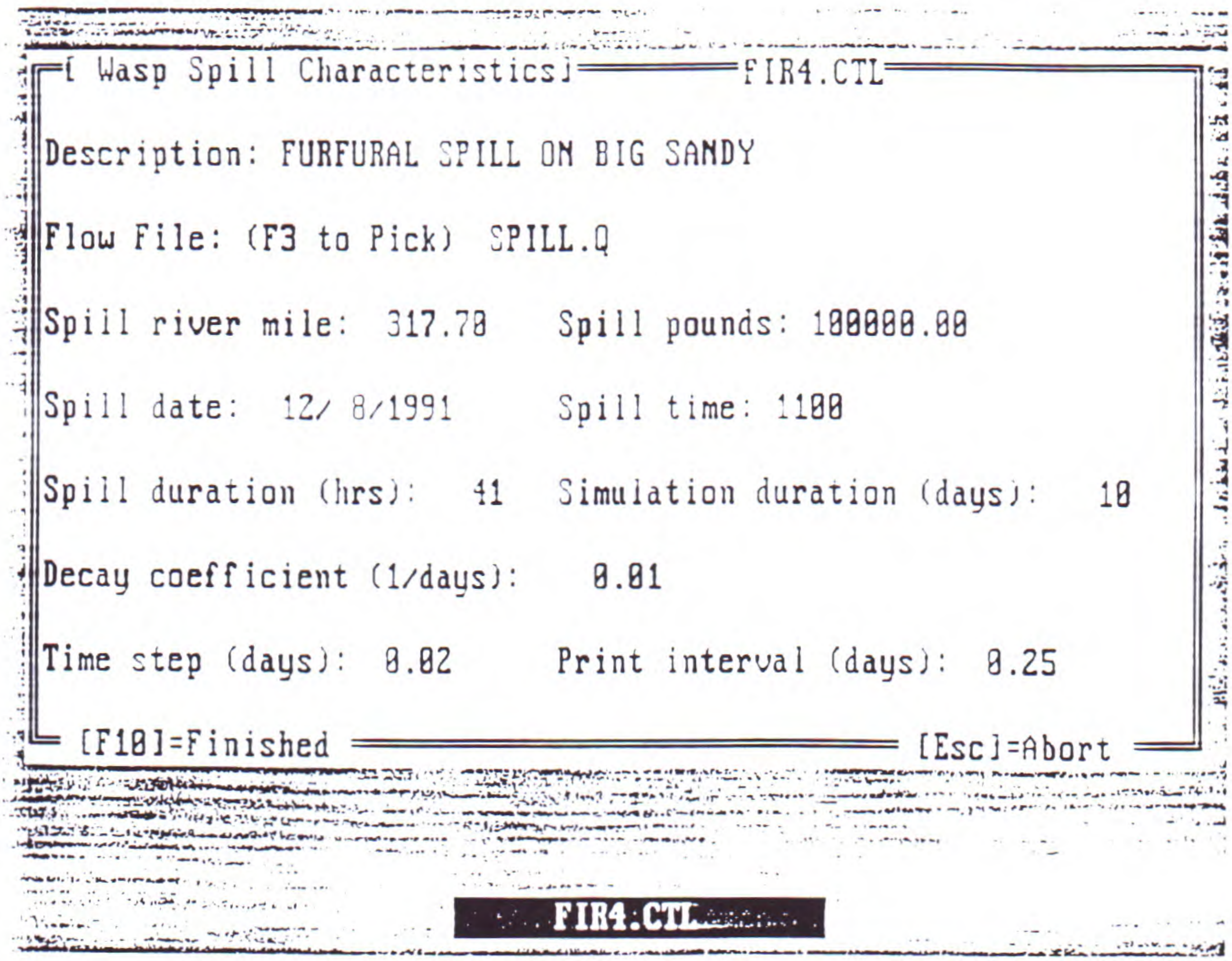


FIGURE 18
SPILL MODELING SYSTEM INPUT FORM



4) The Output pillar is used to view and print the summary report, to construct control files specifying the characteristics of the x-y plots, and to construct an output file that can be used by ARCVIEW to view the results of the simulation. A portion of the summary table is shown in Figure 19.

5) The Display pillar is used to view the screen plots and to generate plots on a printer. An example of an x-y plot is shown in Figure 20.

In actual operation, the user generally moves through the pillars from left to right, specifying the spill, running WASP4 and constructing and viewing reports and plots of the results of the model application. It is recommended that after running the WASP4 model that the animation option be used to examine the output for instability and, if unusual fluctuations are found, that the time step be modified and the WASP4 model rerun.

Conclusions

The Spill Management System provides a fast, user friendly mechanism for tracing a spill to the Ohio River mainstem under emergency conditions. It uses a widely accepted water quality model, WASP4, that is supported and updated by EPA. Flow and stage predictions may be acquired daily from the Corps of Engineers via phone lines and used as input to the model. Additionally, the results of the model application may be viewed in graphical or tabular form from within the Spill Management System or may be exported and viewed within the ARCVIEW system in relationship to other GIS coverages.

The model was successfully adjusted to simulate, with reasonable accuracy, an actual spill to the Ohio River. However, actual field information on the Ohio River is relatively sparse and future applications using either data collected during spill events to the Ohio River or, preferably, during a controlled dye or tracer study, will lead to better calibration information for the model. The WASP4 model may also be used to simulate the impacts of multiple discharges to the Ohio River.

FIGURE 10
SPILL MODELING SYSTEM OUTPUT REPORT

SEGMENT CONCENTRATION REPORT Page 1 of [illegible]
TOLUENE SPILL NEAR LOUISVILLE
9- 9-1993 [illegible]:20 spill at rm 609.40

Time(days)	Seg CH137	Seg CH138	Seg CH139	Seg CH140	Seg CH141	Seg CH142
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.18	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.36	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.54	55.0000	16.8000	2.3500	0.0000	0.0000	0.0000
0.72	119.0000	77.7000	37.4000	13.4000	3.5800	0.4170
0.90	137.0000	123.0000	91.5000	58.0000	32.7000	13.5000
1.02	129.0000	130.0000	115.0000	89.8000	64.1000	36.8000
1.20	102.0000	118.0000	124.0000	117.0000	104.0000	80.2000
1.38	69.4000	91.3000	108.0000	117.0000	117.0000	109.0000
1.56	37.1000	59.2000	80.9000	97.5000	108.0000	113.0000
1.74	15.7000	31.2000	51.0000	69.9000	84.6000	97.9000
1.92	5.5700	13.6000	26.9000	42.5000	57.0000	73.4000
2.10	1.7600	5.2000	12.2000	22.1000	32.9000	47.6000
2.22	0.7830	2.5700	6.7100	13.3000	21.1000	32.9000
2.40	0.2200	0.8400	2.5400	5.6600	9.7900	17.0000
2.58	0.0592	0.2580	0.8890	2.2000	4.1000	7.8700
2.76	0.0154	0.0758	0.2940	0.7990	1.5900	3.3300
2.94	0.0039	0.0216	0.0933	0.2750	0.5790	1.3100
3.12	0.0010	0.0060	0.0287	0.0908	0.2010	0.4870
3.24	0.0004	0.0026	0.0129	0.0428	0.0974	0.2460
3.42	0.0001	0.0007	0.0039	0.0136	0.0322	0.0862
3.60	0.0000	0.0002	0.0012	0.0043	0.0105	0.0295
3.78	0.0000	0.0001	0.0003	0.0013	0.0034	0.0099
3.96	0.0000	0.0000	0.0001	0.0004	0.0011	0.0033
4.14	0.0000	0.0000	0.0000	0.0001	0.0003	0.0009
4.26	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003
4.44	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001
4.62	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4.80	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4.98	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
[illegible]	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

FIGURE 20
SPILL MODELING SYSTEM GRAPHICAL OUTPUT