

**Final Technical Report**

**Lubbock County Pilot Study for Development of a Hydrogeologic Geographic Information  
System (HGIS) to Support TNRCC Implementation of Risk-Reduction Rules**

**by**

**Susan D. Hovorka, Robert E. Mace,  
Susanne Porterfield, and Sandra P. Parra**

**Prepared for**

**Texas Natural Resource Conservation Commission  
under Interagency Contract 72-000000-38**

**Bureau of Economic Geology  
Noel Tyler, Director  
The University of Texas at Austin  
Austin, Texas 78713-8924**

**August 1997**

## CONTENTS

Abstract .....	1
Introduction .....	1
Methods.....	2
Software .....	3
GIS System Design .....	3
Data Sources and Data Entry .....	5
Ambient environment—surface .....	12
Ambient environment—ground water and vadose zone .....	12
LPST-site data .....	14
Assigning spatial coordinates to well locations .....	16
Data tables from LPST sites.....	18
Other contaminated-site data .....	20
Documentation and Metadata .....	20
Accuracy of Site Locations .....	20
Metadata .....	22
Demonstration of the Uses of GIS for LPST Evaluation .....	22
Site Location and Land Use .....	23
Ambient Heterogeneity in the Unsaturated Zone and Aquifer .....	27
Demonstration of the uses of SSURGO soils data for characterizing heterogeneity .....	29
Heterogeneity in the aquifer .....	33
Hydraulic-head gradient.....	39
Risk-Based Corrective-Action (RBCA) Calculations for Soil Vapor and Ground Water .....	40
Statistical description of site-specific measurements.....	41
Plan A equations .....	41
Plan B equations.....	51

Analysis.....	53
Contaminant-Plume Evolution.....	61
Potential Receptors .....	62
Costs.....	66
Evaluation and Recommendations.....	70
Spatial Data .....	70
Data Entry .....	71
Impact of Heterogeneity on RBCA Assessments. ....	71
Plume Analysis .....	72
Conclusions.....	73
Acknowledgments.....	73
References.....	74

## Figures

1. ArcView point coverage showing location of LPST sites in Lubbock County compiled for this study.....	13
2. ArcView point coverage showing LPST site monitoring wells in the city of Lubbock, compiled for this study.....	24
3. ArcView polygons showing Land use/Land Cover and point coverage of LPST site locations compiled for this study .....	25
4. Histograms of land use/land cover classification at LPST sites in the Lubbock east and west 7.5-minute quadrangle inventoried for this study.....	26
5. Depths at which geotechnical soil samples were collected.....	28
6. ArcView polygons showing soil classification and point coverage of LPST boreholes at which geotechnical data was collected .....	30
7. Histograms of representative geotechnical data sorted by soil group.....	31
8. Cross plots of geotechnical soil parameters .....	32
9. Hydraulic conductivity measured at LPST sites and percent sand and gravel in the aquifer in Lubbock County .....	34
10. Histogram and cumulative distribution function for hydraulic conductivity.....	35

11.	Digitized ArcView projection of the regional 1990 water elevation in feet above sea level and LPST-monitoring wells .....	37
12.	Arc GRID calculation of depth to water in the city of Lubbock calculated from USGS DEMs and 1990 water level .....	38
13.	Hand-contoured interpretation of water level in the first quarter of 1993 and the extent of representative benzene plumes during the same time period in the city of Lubbock .....	42
14.	Histogram and cumulative distribution function for organic-carbon content .....	44
15.	Histogram and cumulative distribution function for moisture content .....	45
16.	Histogram and cumulative distribution function for dry bulk density .....	46
17.	Histogram and cumulative distribution function for hydraulic-head gradient .....	47
18.	Histogram and cumulative distribution function for the ratio of expected air concentration to bulk soil concentration .....	55
19.	Histogram and cumulative distribution function for the target soil concentration for residential ingestion and inhalation of volatiles and particulates .....	56
20.	Histogram and cumulative distribution function for the target soil concentration for worker ingestion and inhalation of volatiles and particulates .....	57
21.	Histogram and cumulative distribution function for the bulk soil concentration that coincides with the saturation limit of the soil .....	58
22.	Histogram and cumulative distribution function for the target soil concentration protective of ground water for depths to water of less than 15 ft .....	59
23.	Histogram and cumulative distribution function for ground-water velocity .....	60
24.	Time series of contaminant concentrations at LPST site 91944 through its history .....	63
25.	Time series of contaminant concentrations at LPST site 93526 through its history .....	64
26.	ArcView grid used to calculate distances from LPST sites inventoried in this pilot project to water wells .....	67
27.	Cumulative number of water wells within measured distance of LPST sites .....	68

## Tables

1.	Software used for pilot study .....	4
2.	Contents of the hydrogeologic geographic information system (HGIS) .....	6
3.	Table names, content, measurement units, and source of data in tables generated for this report .....	9

4. Available data .....	15
5. Values for chemical-dependent constants for benzene, toluene, ethyl benzene, and xylene .....	21
6. Default values for soil properties .....	49
7. Default values for calculating volatilization factors .....	52
8. Labor costs for each task of pilot project .....	69

## ABSTRACT

Regional hydrologic, geologic, soils, and cultural background information and data from 217 leaking petroleum storage tank (LPST) sites in Lubbock County, Texas, are used to construct an ARC/INFO Geographic Information System (GIS). The study evaluates the uses of this technology to provide context information for new site evaluation and risk assessment as well as to evaluate the effectiveness of past site-characterization, risk-assessment, and remediation strategies. Methods and costs of producing the data base in this pilot study are described.

Several analyses of these data are presented as a demonstration of the uses of this tool. The heterogeneity within the unsaturated zone is characterized spatially and statistically. Hydrologic variables including water level and hydraulic conductivity from well tests are mapped. The effect of the observed variability on Risk-Based Corrective Action (RBCA) calculations is assessed. The evolution of contaminant plumes can be viewed and relationships between plumes and water-supply wells quantified.

This pilot study demonstrates an application of GIS technology to a moderate-size data set of contaminated-site information. The demonstration is intended not only to provide information about the Lubbock County study area but to serve as a prototype and feasibility study for the application of this technology to other large contaminated-site data sets, including LPST-site data in other major urban areas, other types of contaminated sites, and industry applications.

## INTRODUCTION

In the past decade, a great deal of experience and information have been collected by the Texas Natural Resource Conservation Commission (TNRCC) as part of the leaking petroleum storage tank (LPST) and ground-water protection program. The pilot study presented here demonstrates a methodology for building a geographic information system (GIS) compiling such information and experience into retrievable hydrogeologic data sets. The LPST-site data are

superimposed on published and digital cultural and geotechnical data, forming a hydrogeologic geographic information system (HGIS) that places them into a hydrologic, geologic, and spatial context and optimizes their usefulness. The pilot study also demonstrates some of the various types of analyses that can be done with the GIS and data base. The purpose of the GIS data base is not to replace the TNRCC LPST file system and other existing recordkeeping but to test the benefits of combining data from a variety of sources in a format that facilitates retrieval, analysis, and added-value usage. The effects of ambient geologic and hydrologic conditions on risk-assessment calculations is analyzed.

Lubbock County, Texas, was the area selected for the pilot study. This is an area of moderately shallow ground water, urban development, and public concern about water issues. The pilot study is intended to demonstrate not only the use of the methodologies in the Lubbock County study area but also the potential for application in other areas, such as the metropolitan Dallas–Fort Worth area, greater Houston/Harris County, and the urban corridor along I-35 between Waco and San Antonio. These three regions account for 50 percent of the recorded ground-water contamination incidents in Texas (TNRCC, 1996). Therefore, the methods developed, costs of producing the data base, and the representative types of analyses are described in detail to evaluate feasibility of use in other areas.

## METHODS

Four concepts drove the selection of the technologies, methods, and approaches used in this study:

- The GIS system should be compatible with the GIS system implemented at TNRCC at the time of this pilot study.
- The data should meet TNRCC digital-data guidelines.
- The GIS system should require only minor adjustment to apply this methodology elsewhere in Texas.

- The pilot project is intended as a demonstration, and its scope is designed to test a variety of approaches rather than attempting to be encyclopedic or exhaustive.

### Software

We determined that the GIS will use Environmental Systems Research Institute Incorporated (ESRI) products and be based around ARC/INFO and ArcView software. The strengths of ARC/INFO and ArcView GIS technologies are to

- overlay one coverage on another and look at spatial relationships;
- use the relational data base for both labeling and numerical analyses such as contouring, gridding, and data extractions;
- merge maps of different projections and reproject them into a common projection and scale;
- create continuous surfaces and contoured surfaces with new Spatial Analyst extension;
- query maps and data bases; and
- examine spatial relationships and geographic features at different scales (countywide, citywide, and site scale) in one map.

Software used for this study, platform, purposes, and rationale for its selection are listed in table 1.

### GIS System Design

GIS data are stored as coverages (ARC/INFO) or themes (ArcView). Each theme consists of points, which are locations defined by geographic coordinates; arcs, which are lines defined on the surface; and polygons, which are areas (Bonham-Carter, 1994). Data tables containing attributes (numerical and descriptive data) are linked with each of these types of data.

Conceptually, the data used for this study can be grouped into four types:

- Ambient environment—surface

Table 1. Software used for pilot study.

<u>Software</u>	<u>Platform</u>	<u>Purposes</u>	<u>Rationale for selection</u>
ARC/INFO Environmental Systems Research Institute (ESRI), Redlands, Calif.	UNIX Sun, PC	Data tables; project downloaded digital data and digitized spatial data; onscreen contouring, GRID processing to create surfaces, data analysis	Widely used for environmental applications and used at TNRCC
ArcView	UNIX Sun, PC	Data analysis, onscreen contouring	Compatible with ARC/INFO, widely used
Microsoft Excel	Macintosh	Data entry, plotting	Available at BEG. Microsoft Access was not used because of time and budget constraints.
KaleidaGraph, Cricketgraph	Macintosh	x-y plots, histograms	Available at BEG, fast
Data Thief © Kees Huyser and Jan vander Laan, National Institute for Nuclear Physics and High Energy Physics, P.O. Box 4395 1009AJ Amsterdam, The Netherlands	Macintosh	Fast onscreen digitizing of nongeographically registered spatial data (site maps)	Free and available

- Ambient environment—ground water and vadose zone
- Leaking petroleum storage tank (LPST) site data
- Other contaminated-site data

“Ambient environment—surface” includes cultural data, land surface elevation, land use and land coverage, and digital soils data. “Ambient environment—ground water and vadose zone” incorporates major sources of aquifer data from digital and published sources. LPST-site data are the focus of the compilation effort in this pilot project. TNRCC Plan A risk-assessment worksheets and attachments are used as the basic data source. Other contaminated-site data such as Permitted Industrial and Hazardous Waste sites, Superfund sites, and landfills can be used to complement and extend the LPST data. However, review of these data types in the Lubbock County pilot area led to meager results. The contents of the HGIS are listed in table 2.

The data compiled for this study fall into several categories, depending on the format in which we found them. Spatial (map) data digitized or created for this pilot study by the methods described in this section are formally included as deliverables and include Federal Geographic Data Committee (FGDC) compliant metadata in read-me files. Data tables created for this project are included as Excel (.xls), GIS input (.dbf), and comma-delimited text (.csv) files. The text provided in this contract report and in the read-me files in digital versions provides documentation for this data. Digital data from Internet and other public sources are not included as formal deliverables for this report. We edited and projected these files using standard procedures. Digital copies of these modified files are included for the user’s convenience; however, for quality control issues, the source is the originator. Metadata from the source are included where available, as listed in table 2.

### Data Sources and Data Entry

Existing cultural, hydrologic, and geologic data available in digital or published form provide context for the large data base extracted from TNRCC LPST files from Central Records. Table names, sources of data, content, and units used in the tables are detailed in table 3.

Table 2. Contents of the hydrogeologic geographic information system (HGIS).

<u>Layer name</u>	<u>Source</u>	<u>Content</u>	<u>Format</u>
<b>Ambient Environment—Surface</b>			
Digital Elevation Models 1:24,000	USGS EROS Data Center	Land surface elevation, Lubbock urban area	Grid
Digital Elevation Models 1:250,000	USGS EROS Data Center	Land surface elevation	Grid
Land Use/ Land Cover	EPA GIRAS ftp	Land use and vegetation coverages for conterminous U.S.A. based on census and other data sources from 1979	DLG
SSURGO	Natural Resource Conservation (SCS) Site <a href="http://www.fhw.nrcs.usda.gov">www.fhw.nrcs.usda.gov</a>	Soil mapping and properties	DLG
Cultural data	TNRIS	Streets and highways, other cultural data	DLG
<b>Ambient Environment—Ground Water and Vadose Zone</b>			
Major Aquifers	TNRIS	Name and extent of aquifer	Arc
Minor Aquifers	TNRIS	Name and extent of aquifer	Arc
Water Well Data	Texas Water Development Board ftp on request		
Water Wells	Well ID •Water level •TDS	Point	Point
	TNRCC Water Utilities (vap loc) ftp on request	Location of water wells	
	High Plains Underground Water Conservation District No. 1 Contoured potentiometric surface map	Arc, Metadata	
	High Plains Underground Water Conservation District No. 1 Contoured structure map	Arc, Metadata	
Approximate Altitude of the Water Table in the Ogallala Aquifer, 1990	Regional page-size map (Seni, 1980)	Percent of high-permeability strata in aquifer	Arc, Metadata
Paper maps (Wyatt and others, 1992)			
Approximate Altitude of the Base of the Ogallala Formation			
Paper maps (Wyatt and others, 1992)			
Percent sand/gravel			

Table 2. (cont.)

<u>Layer name</u>	<u>Source</u>	<u>Content</u>	<u>Format</u>
<b>LPST site data</b>			
Monitor-well locations	Generated for this report as described in methods Purchased by TNRCC	Decimal latitude and longitude for representative sites Geocode locations for LPST site addresses	Point coverage, metadata
Site locations	GPS data collected for this study		Point coverage
Monitor-well locations	TNRCC digital files	Decimal latitude and longitude for representative sites •LPST ID •Responsible party •Facility name •Street address •City •County •Hydrocarbon type •Number of monitor wells •Impervious cover (%) •Organic carbon •Water content •Dry bulk density •Effective porosity •Intrinsic permeability	Point coverage, metadata
ID and Setting			Table
Geotechnical soil parameters	LPST files, TNRCC Central Records		Table
Site-specific aquifer data	LPST files, TNRCC Central Records	•LPST-ID •Geologic formation •Formation texture •Aquifer name •Hydraulic conductivity •Notes on measurement •Transmissivity •Storativity •Minimum TDS from nonaffected well •Potential beneficial use category •Hydraulic gradient •LPST ID •Well ID •Measurement date •Ground-water elevation •Depth to PSH •Depth to water •PSH thickness	Table
Site-specific water-level data	LPST files, TNRCC Central Records		Table; temporal subdivisions

Table 2. (cont.)

<u>Layer name</u>	<u>Source</u>	<u>Content</u>	<u>Format</u>
Soil contaminant concentrations	LPST files, TNRCC Central Records	•LPST ID •Depth •Benzene •Toluene •Ethyl benzene •Total Xylenes •Total BTEX •Total Hydrocarbons •MTBE •	Table
Ground-water contaminant concentrations	LPST files, TNRCC Central Records	•LPST ID •Well ID •Benzene •Toluene •Ethyl benzene •Total Xylenes •Total BTEX •Total Hydrocarbons •MTBE •Temperature conductivity •pH •DO	Table, temporal data
Interpreted plume geometry	Contoured for this project	•LPST ID •Date •Component • Geometry of plume	Arc, metadata
Remediation history	LPST files, TNRCC Central Records	•Date of discovery • Remediation start date •Remedial actions •Free product? •Source abated?	Table
Vapor contaminant concentrations	Not compiled, minimal data in pilot area, available from LPST files, TNRCC Central Records		
Surface-water contaminant concentrations	Not compiled, minimal data in pilot area, available from LPST files, TNRCC Central Records		
<b>Other data types</b>			
Permitted and Industrial Hazardous Waste Sites	Not compiled, minimal information in county, available from TNRCC Central Records		
Superfund sites	None in county		
Landfills	Location TNRCC GIS, data available from TNRCC landfill records, minimal useful information in pilot study area		

Table 3. Table names, content, measurement units, and source of data in tables generated for this report.

Database Name	Content	Units	Description source
<b>1M LPST ID &amp; Setting</b>			
	LPST-ID	none	TNRCC database
	Unique well-ID	none	BEG assignment
	Priority	none	TNRCC database
	Responsible party	none	TNRCC database
	Facility ID	none	TNRCC database
	Facility name	none	TNRCC database
	Facility address	none	TNRCC database
	Facility city	none	TNRCC database
	Facility county	none	TNRCC database
	County code	none	TNRCC database
	Hydrocarbon type	none	TNRCC database
	Site map?	(Y/N)	TNRCC files
	Impervious cover	(%)	TNRCC Plan A reported
	Number of monitor wells	none	TNRCC Plan A reported
<b>2M Geotechnical Soil Parameters</b>			
	Unique well-ID	none	BEG assignment
	Well ID	none	TNRCC Plan A reported
	Sample depth	(ft)	TNRCC Plan A reported
	Organic carbon	(g/g)	TNRCC Plan A reported
	Water content	(cm <sup>3</sup> /cm <sup>3</sup> )	TNRCC Plan A reported
	Dry bulk density	(g/cm <sup>3</sup> )	TNRCC Plan A reported
	Effective porosity	(%)	TNRCC Plan A reported
	Intrinsic permeability, k	(cm <sup>2</sup> )	TNRCC Plan A reported
<b>3M Aquifer</b>			
	LPST-ID	none	TNRCC database
	Unique well-ID	none	BEG assignment
	Geologic formation	none	TNRCC Plan A reported
	Formation texture	none	TNRCC Plan A reported
	Aquifer	none	TNRCC Plan A reported
	Well ID	none	TNRCC Plan A reported
	Saturated hydraulic conductivity, K	(ft/d)	TNRCC Plan A reported
	Well ID	none	TNRCC Plan A reported
	Saturated hydraulic conductivity, K	(ft/d)	TNRCC Plan A reported
	Well ID	none	TNRCC Plan A reported
	Saturated hydraulic conductivity, K	(ft/d)	TNRCC Plan A reported
	Notes on saturated hydraulic conductivity	none	TNRCC Plan A reported
	Transmissivity, T	(ft <sup>2</sup> /d)	TNRCC files
	Notes on transmissivity	none	TNRCC files
	Storativity, S	none	TNRCC files
	Notes on storativity	none	TNRCC files
	Minimum TDS of groundwater from unaffected well	(mg/l)	TNRCC Plan A reported
	Potential beneficial use category	none	TNRCC Plan A reported
	Direction of groundwater flow	none	TNRCC Plan A reported
	Hydraulic gradient	none	TNRCC Plan A reported

Table 3. (cont.)

Database Name	Content	Units	Description source
<b>4M Groundwater Contaminant Concentrations</b>			
	Input date	none	BEG Input Date
	LPST-ID	none	TNRCC database
	Well ID	none	TNRCC files
	Unique well-ID	none	BEG assignment
	Measurement date	none	TNRCC files
	QTR-ID	none	BEG assignment
	Benzene	(ppb)	TNRCC files
	Lead	(mg/kg)	TNRCC files
	Toluene	(ppb)	TNRCC files
	Ethyl benzene	(ppb)	TNRCC files
	Total xylenes	(ppb)	TNRCC files
	Total BTEX	(ppb)	TNRCC files
	Total hydrocarbon	(ppm)	TNRCC files
	MTBE	(ppb)	TNRCC files
	TDS	(ppm)	TNRCC Plan A reported
	Temperature	(°F)	TNRCC files
	Conductivity	(1/μΩ)	TNRCC files
	pH	none	TNRCC files
	dO	(ppm)	TNRCC files
<b>5M Water Levels</b>			
	Input date	none	BEG Input Date
	LPST-ID	none	TNRCC database
	Well ID	none	TNRCC files
	Unique well-ID	none	BEG assignment
	Measurement date	none	TNRCC files
	QTR-ID	none	BEG assignment
	Groundwater elevation	(ft)	TNRCC files
	Casing height	(ft)	TNRCC files
	Depth to PSH	(ft)	TNRCC files
	Depth to water	(ft)	TNRCC files
	PSH thickness	(ft)	TNRCC files
<b>6M Remediation</b>			
	LPST-ID	none	TNRCC database
	Unique well-ID	none	BEG assignment
	Date of release, discovery, or report.	none	TNRCC files
	Have remedial actions occurred?	(Y/N)	TNRCC files
	If yes, start date.	none	TNRCC files
	Is remediation still in operation? Stop date	(Y/N)	TNRCC files
	What are/were remedial actions?	none	TNRCC files
	P&T	none	TNRCC files
	SVE	none	TNRCC files
	AS	none	TNRCC files
	PPS	none	TNRCC files
	BIO	none	TNRCC files
	IT	none	TNRCC files
	Has source been abated?	(Y/N)	TNRCC files
	Tank	(Y/N)	TNRCC files
	Soil	(Y/N)	TNRCC files
	Date of abatement	none	TNRCC files
	Site closure	(Y/N)	TNRCC files
	Free product?	none	TNRCC files
	Date of the last PSH recovery	none	TNRCC files

Table 3. (cont.)

Database Name	Content	Units	Description source
<b>7M Notes &amp; Comments</b>			
	LPST-ID	none	BEG assignment
	ID & Setting	none	BEG assignment
	Soil data	none	BEG assignment
	Water data	none	BEG assignment
	Plume data	none	BEG assignment
	Remedial actions	none	BEG assignment
	General comments	none	BEG assignment
<b>8M Soil Contaminant Concentrations</b>			
	Input date	none	BEG Input Date
	LPST-ID	none	TNRCC database
	Well ID	none	TNRCC files
	Unique well-ID	none	BEG assignment
	Depth	(ft)	TNRCC files
	Measurement date	none	TNRCC files
	Benzene	(ppb)	TNRCC files
	Lead	(mg/kg)	TNRCC files
	Toluene	(ppb)	TNRCC files
	Ethyl benzene	(ppb)	TNRCC files
	Total xylenes	(ppb)	TNRCC files
	Total BTEX	(ppb)	TNRCC files
	Total hydrocarbon	(ppm)	TNRCC files
	MTBE	(ppb)	TNRCC files

## Ambient environment—surface

A digital street and highway map of Lubbock County based on Texas Department of Transportation county highway mapping was downloaded from the Texas Natural Resources Information System (TNRIS) Web site. This map, which shows streets and highways and other cultural features, is used to geographically reference monitoring wells and soil borings shown on LPST-site maps (fig. 1). Other sources of cultural base maps identified for this project include scanned U.S. Geologic Survey (USGS) topographic maps showing cultural features, paper County Tax Assessor plat maps showing property lines, and a variety of proprietary maps. These other possible maps were not used for the pilot project because of cost considerations and uncertainty about whether the accuracy would be superior to the TNRIS map.

Land surface elevation was acquired from USGS as digital elevation models (DEM) at two scales. The entire county is available from USGS at a scale of 1:250,000. The 1:24,000-scale Lubbock East and Lubbock West 7.5-minute quadrangles, in which most LPSTs are located, were purchased from USGS. This information was used to calculate surface elevation at sites where it was not reported and as one of the digital surfaces used to create a regional map of depth to water. Land use/land cover data downloaded from the Environmental Protection Agency (EPA) are out of date (1979 census) but were used in the pilot to test the utility of a GIS approach to characterizing land use around a site. Digital soils data (SSURGO) provide a large amount of information in a GIS format. For this project, a subset of information was extracted and used to analyze statistical variation in the geotechnical parameters of soils from various geomorphic settings.

## Ambient environment—ground water and vadose zone

Coverages of major and minor aquifers from TNRIS provides information on aquifers threatened by contamination. In the pilot area, these coverages are superfluous because of the regional dominance of the Ogallala aquifer; however, they will be useful for other areas.

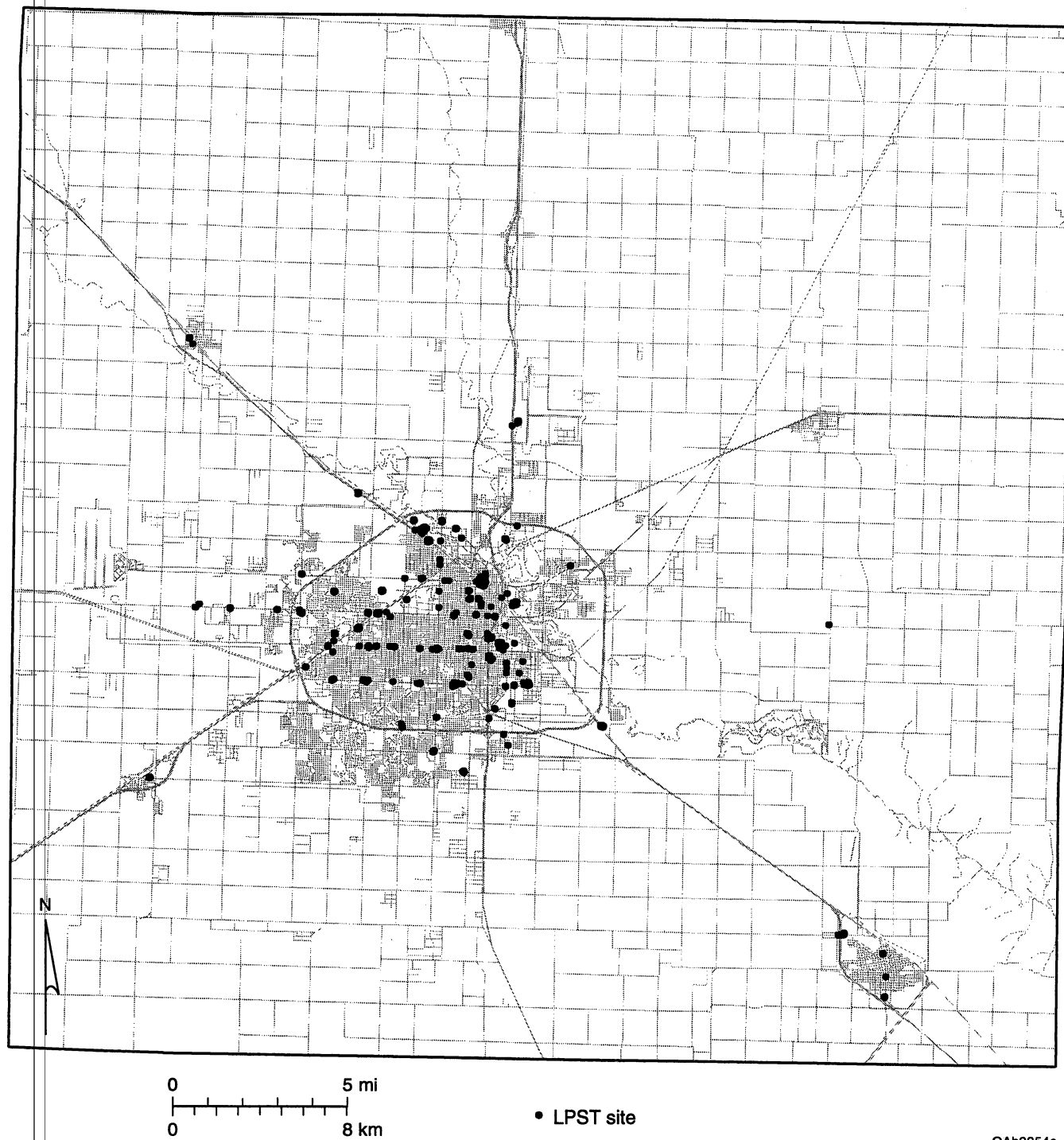


Figure 1. ArcView point coverage showing location of LPST sites in Lubbock County, compiled for this study. A digital version of the Texas Department of Transportation (TXDOT) county highway map downloaded from TNRIS serves as the cultural base. At this scale, all the monitoring wells and soil borings at each site merge into a single dot. Countywide maps are shown at this scale.

GIS could readily be used to search information on potential receptor wells. This type of work is done by consultants in the pilot study area, and their reports in LPST files suggest that they have made improvements by field checking the well data from TNRIS paper files. Rather than generating a new digital file of questionable quality using data in TNRIS files, we used digital water-supply well data from the TNRCC Water Utilities section to demonstrate the methods that could be used to enumerate and describe potential receptor wells.

A number of maps from the literature provide information about the aquifer at the county or regional scale (table 4). We selected the water level and base of the Ogallala aquifer (Wyatt and others, 1992) and percent sand and gravel (Seni, 1980) as examples. Other potentially useful maps, such as aquifer specific yield and permeability (Knowles and others, 1984), net sand and gravel in the Ogallala (Seni, 1980), tritium concentration in ground water, thickness of the vadose zone, and precipitation (Nativ, 1988), are available but were not digitized for the pilot study because of time constraints.

#### LPST-site data

This data set is the major new compilation resulting from this study. Plan A risk-assessment worksheets and attachments submitted by various environmental consultants to TNRCC were used as the basic data source. The data extracted, source, destination data table, column name, and units are listed in table 3. In case files where the Plan A assessment was not available at the time of data review, equivalent data were extracted from other available reports and letters. The data collection effort was designed to include most of the available hydrogeologic data so that the interrelationships among various types of data can be evaluated. Eight data tables have been generated (table 3): site identification and setting, geotechnical soil parameters, site-specific aquifer data, ground-water contaminant concentrations, site-specific water-level data, remediation history, notes and comments, and soil contaminant concentration. Surface-water contaminant concentration and vapor contaminant concentration were not compiled because in the pilot area these data were collected at very few sites.

Table 4. Available data.

Data	Scale	Source	Data extracted	Units	Format
Leaking Petroleum Storage Tanks (LPST) list	Digital	TNRCC Information Resources	LPST ID and location by county, zip code, and street address		Data base on 3.5-inch diskettes
GeoCode LPST locations	Digital	Purchased by TNRCC (Tom Lewis) from Geographic Data Technology	Location matching street address or 5-digit Zip code centroid	Latitude and longitude in decimal degrees	Digital
Soils					
Surface geology	Digital	Eifler and others, 1993			Paper plate
Water level, 1990	1:250,000 1 in = 2 miles	Wyatt and others, 1990	Altitude of the water table, 1990 Lubbock County	Surface geologic units ft	Paper map was digitized
Base of the Ogallala aquifer	1 in = 2 miles	Wyatt and others, 1990	Base of the Ogallala Formation, Lubbock County	Elevation (ft) MSL	Paper map was digitized
Saturated thickness of the Ogallala aquifer	1 in = 2 miles	Wyatt and others, 1990	Lubbock County	ft	Paper map
Water levels, city of Lubbock	1:115,200	Chen and others, 1988	1981, 1987 water levels, change in water levels, TDS	Elevation (ft) MSL	Page size paper maps
Water levels, city of Lubbock	Not shown	Rainwater and Thompson, 1994	1987-1991-1992 water levels, change in water levels	Elevation (ft) MSL	Page size paper maps
Percent sand and gravel	1 in = 30 miles	Seni, 1980	Distribution of coarse-grained materials in the Ogallala aquifer, regional	Percentage of Quaternary digitized	Page-size paper map was digitized
Net sand and gravel	1 in = 30 miles	Seni, 1980	Thickness of coarse grained materials in the Ogallala aquifer, regional	ft	Page-size paper map
Specific yield	1 in = 20 miles	Knowles and others, 1984	Specific yield	Percent	Paper plate
Permeability	1 in = 20 miles	Knowles and others, 1984	Hydraulic conductivity	ft/day	Paper plate
Dissolved solids	1 in = 20 miles	Knowles and others, 1984	TDS	mg/l	Paper plate
Dissolved solids	1 in = 20 miles	Knowles and others, 1984	TDS	mg/l	Paper plate
Natural tracers	1 in = 40 miles	Nativ, 1988	Tritium, oxygen isotopes, arsenic concentrations, showing areas of rapid recharge to aquifer	Page-size paper maps	
Thickness of the vadose zone	1 in = 40 miles	Nativ, 1988	ft	Depth to water	Page-size paper maps
Precipitation	1 in = 40 miles	Nativ, 1988	recharge	in	Page-size paper maps

Each monitoring well was assigned a unique identification number (UN) created from the LPST identification number (LPSTID) and the monitoring well (MW) number:

$$UN = 10 \times LPSTID + MW \quad (1)$$

For example, at LPST site 99999, monitoring well 1 is assigned unique number 999991. Recovery wells (RW) and soil borings (SB) are assigned consecutively higher numbers than the monitoring wells. The contractor's well identification number can be obtained by querying the LPST ID and Setting table.

#### *Assigning spatial coordinates to well locations*

Three procedures were used to determine latitude and longitude for monitoring well locations: (1) site map to TNRIS digital street map (site map method), (2) Geocode locations, and (3) field location referencing using Global Positioning System (GPS).

The site-map method developed for this study follows nine steps:

1. Photocopy site maps provided in LPST reports.
2. Identify the center point of a street intersection to use as an origin (if the center of a street intersection was not shown on the site map, a street width of 40 ft was assumed).
3. Orient the map with respect to north, plot x and y axis (northing and easting) through the origin.
4. Scale the axis using the scale marked on the map, with positive values to the north and east and negative to the south and west.
5. Scan marked photocopy on a flat-bed scanner to create a PICT file.
6. Calculate northing and easting for each monitoring well or soil boring in feet from the origin using the Data Thief software.
7. Import the resulting files of x and y values into a Microsoft Excel spreadsheet. Identify the street intersection on the digital street and highway map from TNRIS and extract UTM (Universal Transverse Mercator) coordinates.

8. Calculate UTM coordinates for each monitoring well or soil boring.
9. Create an ARC/INFO point coverage.

The well location nearest the former tank pit, generally monitoring well 1, was used as a general site location. The accuracy of the locations generated was limited by the accuracy of the street and highway map downloaded from TNRIS, the accuracy of the site map, and the ability to locate an identifiable street intersection in the site maps. This method was successful for locating 90 percent of the sites.

Digital Geocode location data created by Geographic Data Technology (GDT) for TNRCC were tested as an alternative to using the TNRIS digital street map for general site locations. Geocode technology uses proprietary digital maps and software to extract latitude and longitude for each street address. Where street address matches fail, U.S. Post Office Zip codes are used to create a centroid to approximate the location. GDT reported that 73 percent of the addresses were located for 43,366 LPST sites throughout Texas (GDT, written communication, 1997). Projecting Geocode locations on the TNRIS street map showed a systematic displacement of streets relative to Geocode locations, apparently because of minor base map misalignment, as discussed in the evaluation section.

To obtain high-quality location information, a representative subset of monitoring wells at LPST sites were located using a digitally corrected GPS. Accurate GPS location data allows calculation of representative error for other, less accurate location methods. GPS data collection for this project involved two principal tasks: (1) organizing the site files to facilitate efficient data collection and (2) collecting locational data in the field. During mission planning, we organized the data by street. Each record was reviewed to ensure that the address and facility name were complete. Using the city road map, we organized the locations geographically in a binder to be taken to the field, where an attempt was made to locate each facility. Thirty-five sites in several parts of the city were selected to visit. Of these, 10 could not be surveyed because either the site closed and the wells had been removed, or because the wells were inside buildings or fenced enclosures. GPS locations for 35 wells at 25 sites were measured in two days of field work.

Collection of the GPS data was accomplished with a Trimble Navigation Pathfinder Basic Plus unit. This receiver is a hand-held, battery-powered, six-channel receiver. It can track up to eight satellites simultaneously and has 256 Kbytes of nonvolatile memory. Positions can be calculated at a rate of one per second and stored for later transfer to a personal computer for processing with Trimble software, P-FINDER. The accuracy of the Pathfinder is rated by Trimble at  $\pm 2$  meters horizontally, based on the average of 180 data points from a differentially corrected file.

Real-time differential correction was achieved using an Omnistar Model 6300A receiver. The Omnistar system is based on 11 base stations in North America that monitor and send corrections to a network control center in Houston, Texas, where the data are uplinked to a geostationary satellite. The satellite then broadcasts the corrections to clients having Omnistar receivers. The accuracy of the Omnistar is rated at  $\pm 1$  m.

The GPS data were collected at two levels of accuracy:  $\pm 2$  m CEP (circular error probable, meaning 50 percent of the collection points are within a 5-m radius circle on a horizontal) and  $\pm 5$  m CEP. Of the 35 sites that were located during this project, 25 were at the  $\pm 2$  m accuracy and 10 were at the  $\pm 5$  m accuracy. In order to achieve  $\pm 2$  m accuracy, 180 data points must be collected and averaged. It takes about 4 to 7 minutes to collect 180 data points with the Trimble unit. With just one data point it is still possible to maintain acceptable accuracy and save time for more data collection. The location data were placed in coverages created in ESRI's ARC/INFO, and the field records were downloaded as metadata.

#### *Data tables from LPST sites*

An LPST listing on floppy disc from TNRCC Information Resources was used to inventory the sites in the Lubbock County study area. Additional sites not on this list were identified on the Geocode list and by TNRCC personnel. Sites were selected to sample the spectrum of available data, including sites on the edge and outside of the city of Lubbock, sites from various geologic environments, old sites with low LPST numbers, and recently identified sites with high numbers.

Not all of the more than 300 sites in the county could be inventoried in the scope of this pilot study.

Setting, soil data, site-specific aquifer data, site-specific water-level data, soil contaminant concentration, ground-water contaminant concentration, and remediation history were extracted from the Plan A risk-assessment worksheets and attachments or from monitoring, closure, or other types of reports in the TNRCC files (table 3). The Plan A assessment, the most recent site map, the most recent compilation of water-level records, and the most recent compilation of ground-water contaminant concentration records from the attachments were photocopied. Data were manually typed into a standardized Excel spreadsheet following the same format developed by Mace and others (1997). Units were standardized as shown in table 3. Obvious errors were corrected. Water levels calculated from an arbitrary datum were normalized to sea-level elevation using the DEM to find the elevation at the reference location used by the contractor. Inconsistencies in the data bases reflect both the complex history of site monitoring and the unavailability of some reports at the time of data entry.

Organization of data tables is needed to meet GIS requirements for determinant relationships (one-to-one relationships between attribute cells and locations). At LPST sites, ground-water elevation and contaminant concentration were sampled multiple times at each monitoring well. For input into the GIS, we identified these data temporally by quarters of each year, designated by two-digit year and quarter, starting with January (for example, 92-1). ArcView could then be used to plot in map view the measured parameters (water level, phase-separated hydrocarbon [PSH], thickness, benzene, toluene, ethyl benzene, total xylenes, total BTEX, and MTBE) during each quarter at all the wells for which data were reported. Changes in water level or contaminant concentration through time can be examined by plotting data from sequential temporal tables.

A more elegant approach allowing greater flexibility but requiring more time for table construction would be to use a data base such as Microsoft Access to key additional subsets of the data that meet GIS requirements for determinance. Examples of other sorting are to average and

group data more coarsely (annually) or view plume evolution by site history (before and after remediation).

Soil contamination was measured at several different depths in each boring. Viewing these data in GIS requires separating the data into several attribute cells by elevation, averaging the values, or plotting only the highest values.

The structure and linked columns that permit display and analysis of these data bases are shown in table 5.

#### Other contaminated-site data

Permitted Industrial and Hazardous Waste sites, Superfund sites, and landfill sites are potential sources of data to complement the LPST data in the GIS. No Superfund and only three permitted Industrial and Hazardous (I&H) Waste sites are in the pilot study area, and of these only Reese Air Force Base has extensive files. A sample of the extensive files on Reese Air Force Base at TNRCC Central Records was reviewed. Results of six aquifer tests, several high-quality stratigraphic cross sections, and data on large contaminant plumes are the data extracted from these files.

Eight TNRCC landfill permit files in Lubbock County were reviewed. In Lubbock County these files contained meager water-level information. Geographic location of the monitoring wells would require site visits or registered air photographs because the landfills lack mapped geographic reference locations.

## DOCUMENTATION AND METADATA

### Accuracy of Site Locations

The accuracy of digitally corrected GPS data collected using BEG protocols was quantified by collecting data at a first-order NGS reference point (Angle and others, 1996). Real-time differential

Table 5. Values for chemical-dependent constants for benzene (B), toluene (T), ethyl benzene (E), and xylene (X).

Parameter	B	T	E	X
$D_i$ [cm <sup>2</sup> -s <sup>-1</sup> ]	0.0933	0.0838	0.0748	0.074
$K_{as}$ [g-cm <sup>3</sup> ]	$1.38 \times 10^{-1}$	$4.35 \times 10^{-2}$	$1.20 \times 10^{-2}$	$6.01 \times 10^{-2}$
$K_{oc}$ (L-kg <sup>-1</sup> )	83	300	1,100	240
0.25*LEL action level (ppm)	3,200	3,000	2,500	2,500
H@20°C (atm-m <sup>3</sup> -mol <sup>-1</sup> )	$5.59 \times 10^{-3}$	$6.37 \times 10^{-3}$	$6.43 \times 10^{-3}$	$7.04 \times 10^{-3}$
$S$ [mg-l <sup>-1</sup> ]	1,750	535	152	198
$S_{Fo}$ (mg-kg <sup>-1</sup> -d <sup>-1</sup> )	$2.90 \times 10^{-2}$	-	-	-
$S_{Fi}$ (mg-kg <sup>-1</sup> -d <sup>-1</sup> )	$2.91 \times 10^{-2}$	-	-	-
$R_{fD}$ (mg-kg <sup>-1</sup> -d <sup>-1</sup> )	-	0.2	0.1	2.0
$R_{fC}$ (mg-kg <sup>-1</sup> -d <sup>-1</sup> )	-	0.4	1.0	-
$V_{Fi}$ (m <sup>3</sup> -kg <sup>-1</sup> )	20,885.87	39,650.59	79,890.87	35,828.56
$V_{Fr}$ (m <sup>3</sup> -kg <sup>-1</sup> )	22,903.45	43,480.84	87,608.33	39,289.60
$C_w$ (mg-l <sup>-1</sup> )	$5.0 \times 10^{-3}$	1	$7.0 \times 10^{-1}$	10

correction and averaging 180 data points reproduced better than  $\pm 1$  m accuracy, and individual points produced better than  $\pm 2$  m accuracy. Metadata were included in the digital files.

Comparison of the well locations from the GPS survey with the other site location methods showed two types of error. The site-map method used for most of the sites in the study showed an average error of 450 ft. Locations generated by the site-map method are systematically east of the correct positions as measured by GPS, indicating that the TNRIS Lubbock County base map is misregistered by that amount. The average error for the Geocode locations is 200 ft. This error appears to be the result of the random distribution of the monitoring wells with respect to their nominal street addresses. Errors in interpolation of the street addresses, especially in warehouse districts, is also a probable cause of error.

### Metadata

The only original data generated for this project were the representative GPS locations. For all other data used in this pilot study, most errors were inherited from the data sources and were poorly constrained. FGDC compliant metadata for the coverages generated for this pilot project are provided as digital read-me files. Quality-control documentation for other coverages obtained in digital form from various sources are included as available.

Quality control information on data extracted from reports is documented within the reports; no effort was made to capture laboratory duplicate information. Only obvious errors—for example, mislabeled units—could be corrected; therefore, we have retained all the reported data and dealt with errors statistically, as described in the methods and analyses.

## DEMONSTRATION OF THE USES OF GIS FOR LPST EVALUATION

Several analyses of the data compiled for this study are presented to demonstrate the uses of the HGIS. We selected a variety of types of analysis that are not equally applicable to the pilot area

but have potential to be useful in various Texas geologic and hydrologic settings. In each analysis we show spatial and statistical results.

### Site Location and Land Use

For this example, the digital locations created for each of the sites in the study are used to show relationships among sites and to extract additional information about the setting from available digital data. Figure 1 shows LPST sites in Lubbock County. Figure 2 is the same digital image enlarged (zoom-in on screen) to show monitoring-well locations on a street base map of the city of Lubbock. GIS could be used onscreen by contractors, regulators, or other users to compare a site of interest with other, adjacent sites. The existence of multiple sites in an area is significant to both site characterization and liability issues where plumes extend offsite and merge. The site listing in the pilot project is a representative sample and does not contain a comprehensive listing of all the sites in the study area.

As an example of cultural information, we have used the EPA (1979) land use/land coverage data base. These data, although out of date, are downloaded from the Internet at no cost. A wide variety of other census or cultural data could be used in a similar manner to systematically characterize sites for quantitative analysis and to complement the walking-tour information collected during site characterization. For this example, we used ArcView to create an overlay of LPST sites on the Land Use/Land Cover polygon theme (fig. 3). The ArcView "Select by Theme" function was used to quantify and identify the LPST sites that fell into each land use category. A histogram of the land use/land coverage classification of LPST sites sampled was prepared using Kaleidograph (fig. 4). More than half of the 150 LPST sites sampled lie in areas classified as commercial; residential and other land uses make up the remainder.

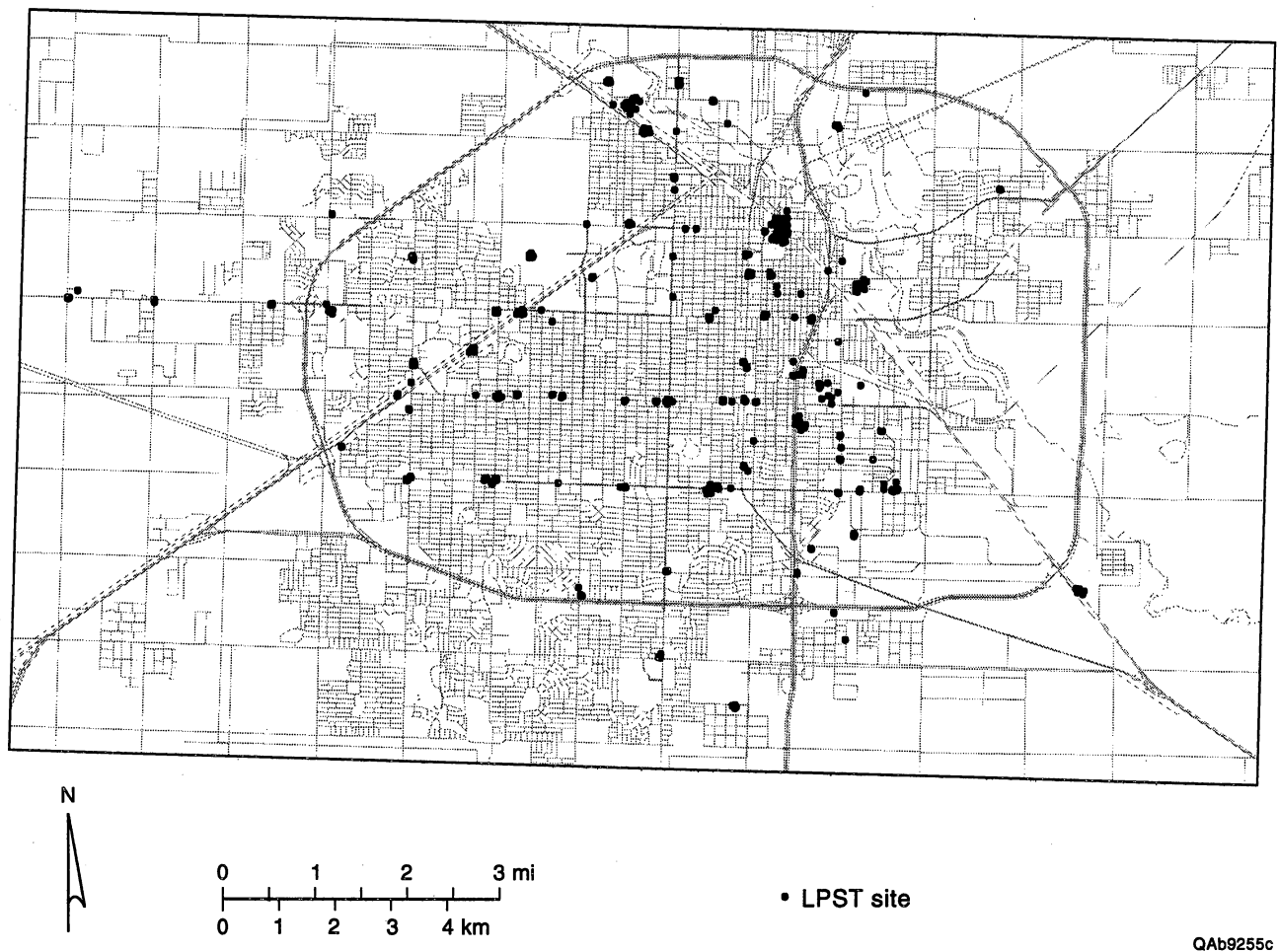


Figure 2. ArcView point coverage showing LPST-site monitoring wells in the city of Lubbock, compiled for this study. Road base map is the same as in figure 1 and the area is clipped to the two 7.5-minute USGS base maps, Lubbock East and Lubbock West. At this scale, individual monitoring wells can be seen as clusters and street patterns recognized. This is the area shown in city of Lubbock maps.

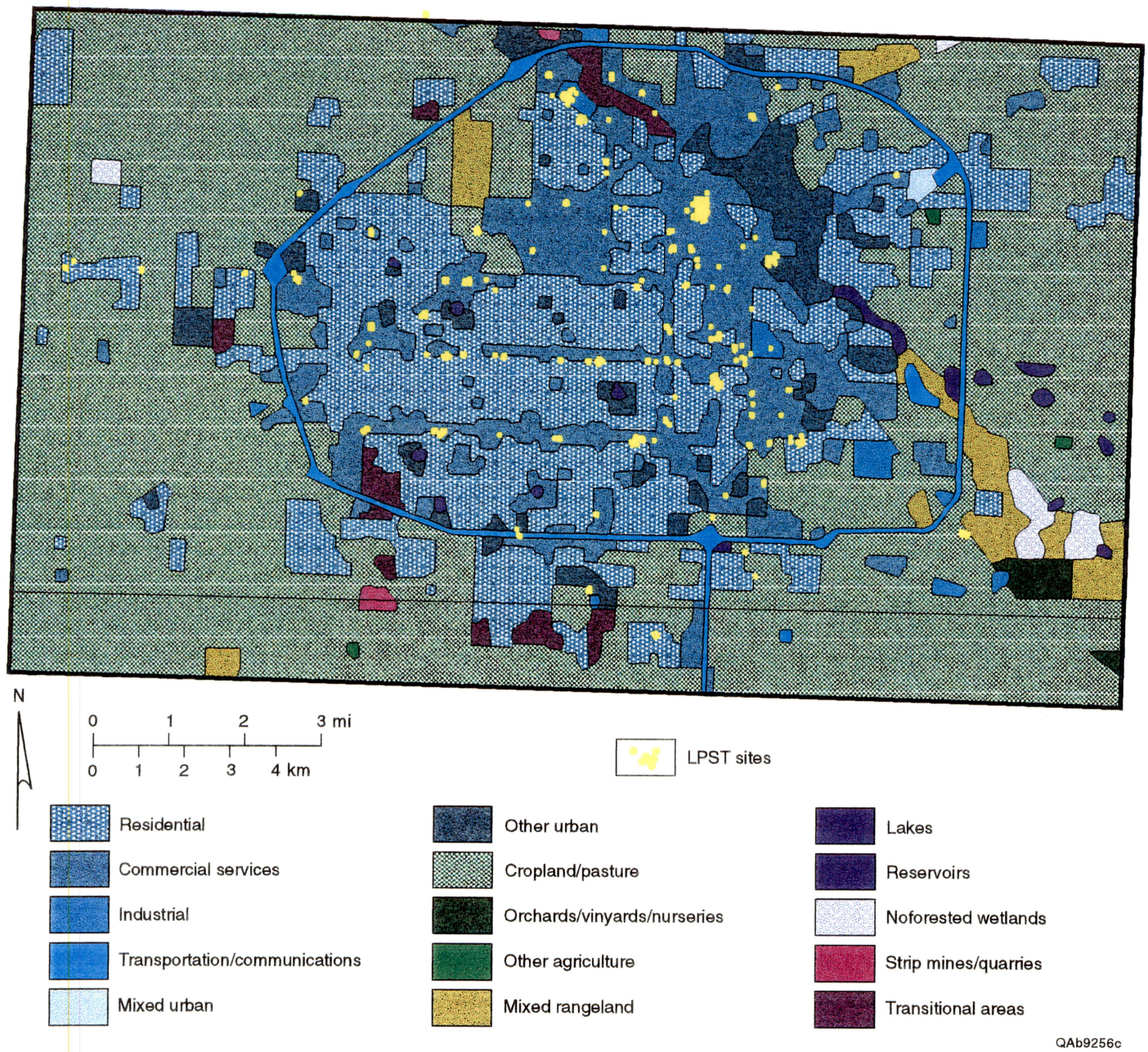


Figure 3. ArcView polygons showing land use/land cover and point coverage of LPST-site locations compiled for this study.

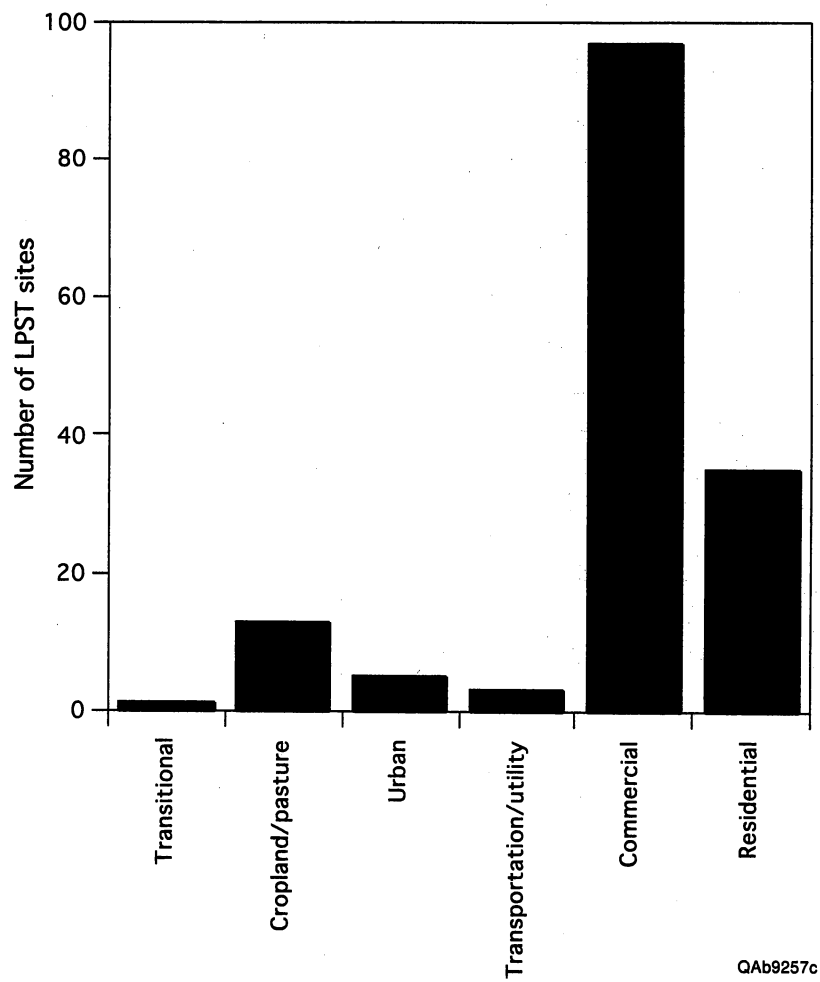


Figure 4. Histogram of land use/land cover classification at LPST sites in the Lubbock East and West 7.5-minute quadrangles inventoried for this study.

## Ambient Heterogeneity in the Unsaturated Zone and Aquifer

The unsaturated zone and top of the aquifer were described in each LPST site from borehole samples. Soil-boring logs were typically described in the field. At a small number of sites, the soil-boring logs were interpreted to create cross sections. Laboratory analyses of geotechnical soil parameters for an uncontaminated sample (organic carbon, moisture content, soil bulk density, effective porosity, and intrinsic permeability) were reported for 28 percent of the sites examined for this project. These geotechnical laboratory analyses were used for TNRCC Plan A and Plan B site-specific risk calculations described in the risk-assessment example below. Slug tests or pump tests were performed to calculate hydraulic conductivity and other hydrologic variables at fewer than 5 percent of the sites.

Comparisons of the sample intervals with the geologic cross sections show that the small number of samples tested was inadequate to describe the probable range of values at the site. We conducted an experiment in making additional cross sections from soil-boring logs. Although it was difficult to correlate and interpret boring logs created by different individuals, the same amount of heterogeneity as in the contractor-constructed cross sections was apparent over most of the city of Lubbock. Depths at which samples were collected ranged from 4 to 79 ft (fig. 5). The most typical sample depth was just above the saturated zone. The units sampled included the Pleistocene Blackwater Draw Formation and the top of the Miocene-Pliocene Ogallala Formation.

Heterogeneity in the Blackwater Draw Formation is the result of variation in depositional texture and calcic soil development (Holliday, 1989; Caran 1991; Hovorka, 1995). Fine-grained sediment and minor carbonate accumulation are typical of playa lakes; silty or loamy sand with abundant carbonate is typical of upland settings.

In the study area, the top of the saturated zone corresponds approximately to the top of the Ogallala Formation. LPST-monitoring wells generally penetrate only to the top of the saturated zone. Heterogeneity in this interval is probably mostly the result of variations in the properties of the well-developed pedogenic carbonate (Caprock) at the top of the formation, whereas the

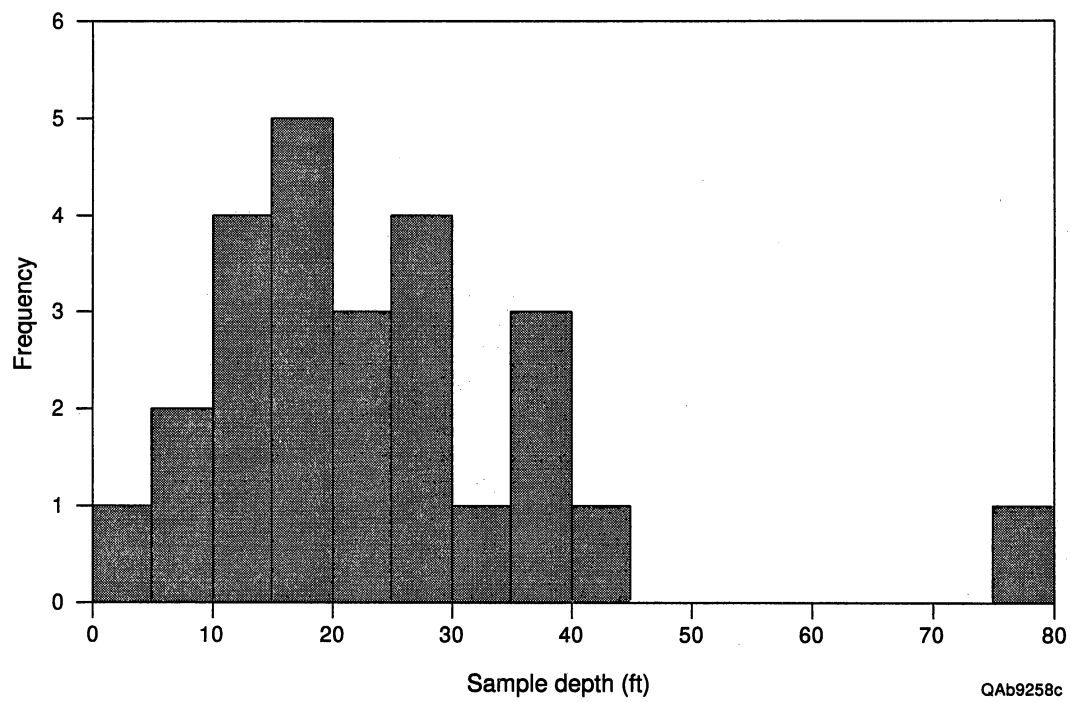


Figure 5. Depths at which geotechnical soil samples were collected.

properties of the entire Ogallala aquifer more likely are influenced by the thickness of fluvial gravels in the lower part of the aquifer (Foster, 1952; Reeves, 1976; Seni, 1980).

#### Demonstration of the uses of SSURGO soils data for characterizing heterogeneity

The purpose of this demonstration is to show how ArcView can be used to prepare statistical distributions of geotechnical parameters for geologic and hydrologically significant land areas. The use of these statistical distributions for risk assessment is shown in a later demonstration. In the Lubbock pilot study area, LPST sites are predominantly in upland areas because they make up most of the land surface, and because the contrasting clay and clay-loam soil types are found in playa basins, which flood during heavy rainfall events and therefore have been typically developed as city parks or lakes. LPST sites are also not found in the narrow belt of alluvial sediments along Yellowhouse Canyon and Blackwater Draw. Therefore, we did not expect to see more than one geologically significant population of geotechnical parameters in the study area. However, as an example of how ArcView can be used as a tool to identify separable populations of geotechnical parameters, we have used the SSURGO data base to test this assumption.

We downloaded the U.S. Department of Agriculture Natural Resources Conservation Service (NRCS) SSURGO data base for Lubbock County from the NRCS Web page. We used the Lubbock East and Lubbock West coverages to create a polygon coverage of soils in the area with most abundant LPST sites (fig. 6). Each polygon in the coverage has a soil code and soil association name. We superimposed the LPST borehole locations onto this coverage, extracted subsets of LPST sites for each polygon, and matched the geotechnical properties to the soil association. The distribution and cross-plot relationships of geotechnical properties were then examined for each soil group (figs. 7 and 8).

In the pilot area, all but one of the borehole locations lie within the Acuff, Estacado, Amarillo, and unclassified urban soil complexes developed on loamy upland parent materials (figs. 6 and 7). Histograms and cross plots of parameter values of the upland soils show a large scatter (figs. 7 and 8), reflecting the small-scale vertical and lateral variability seen in cross section and described

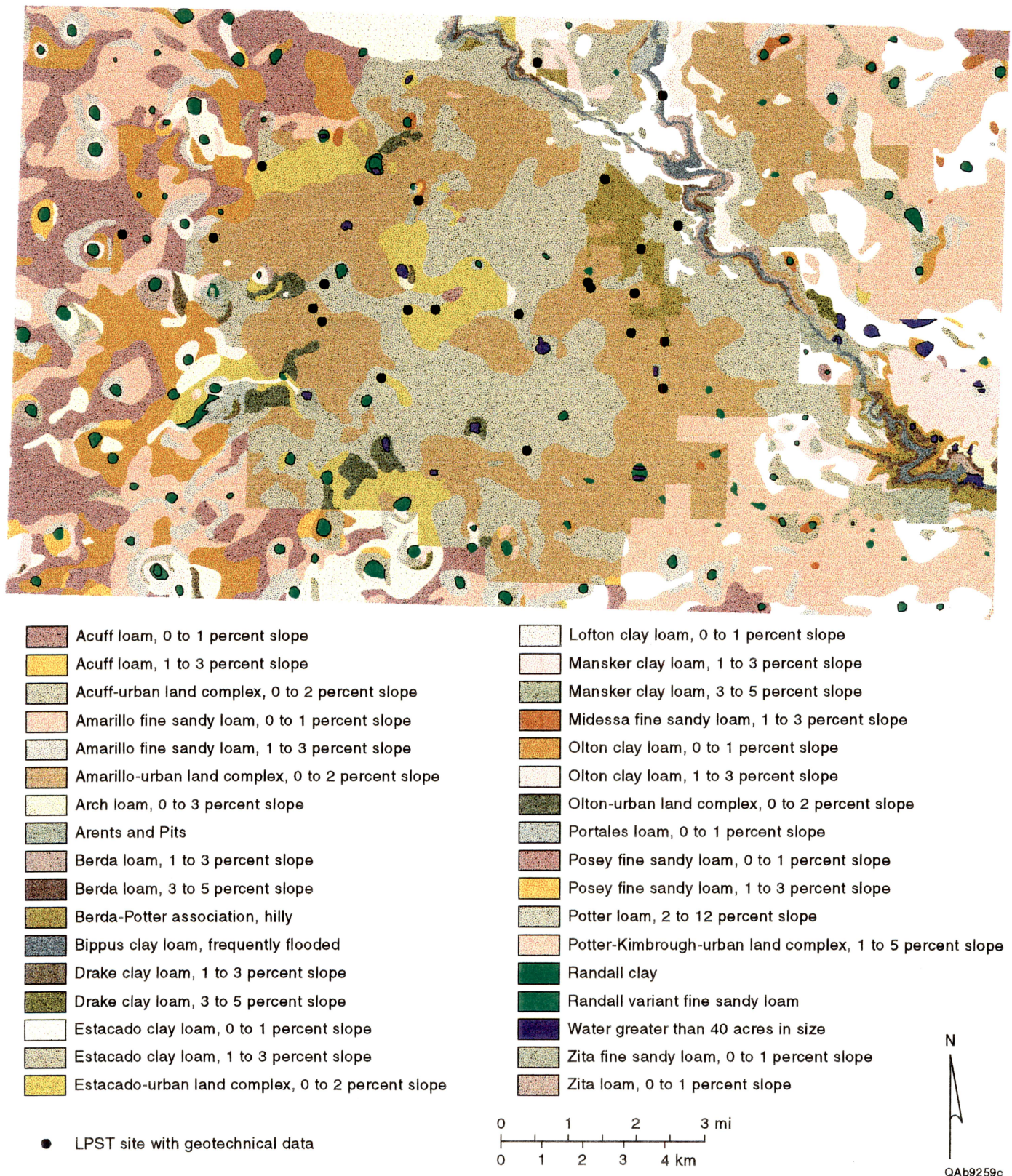
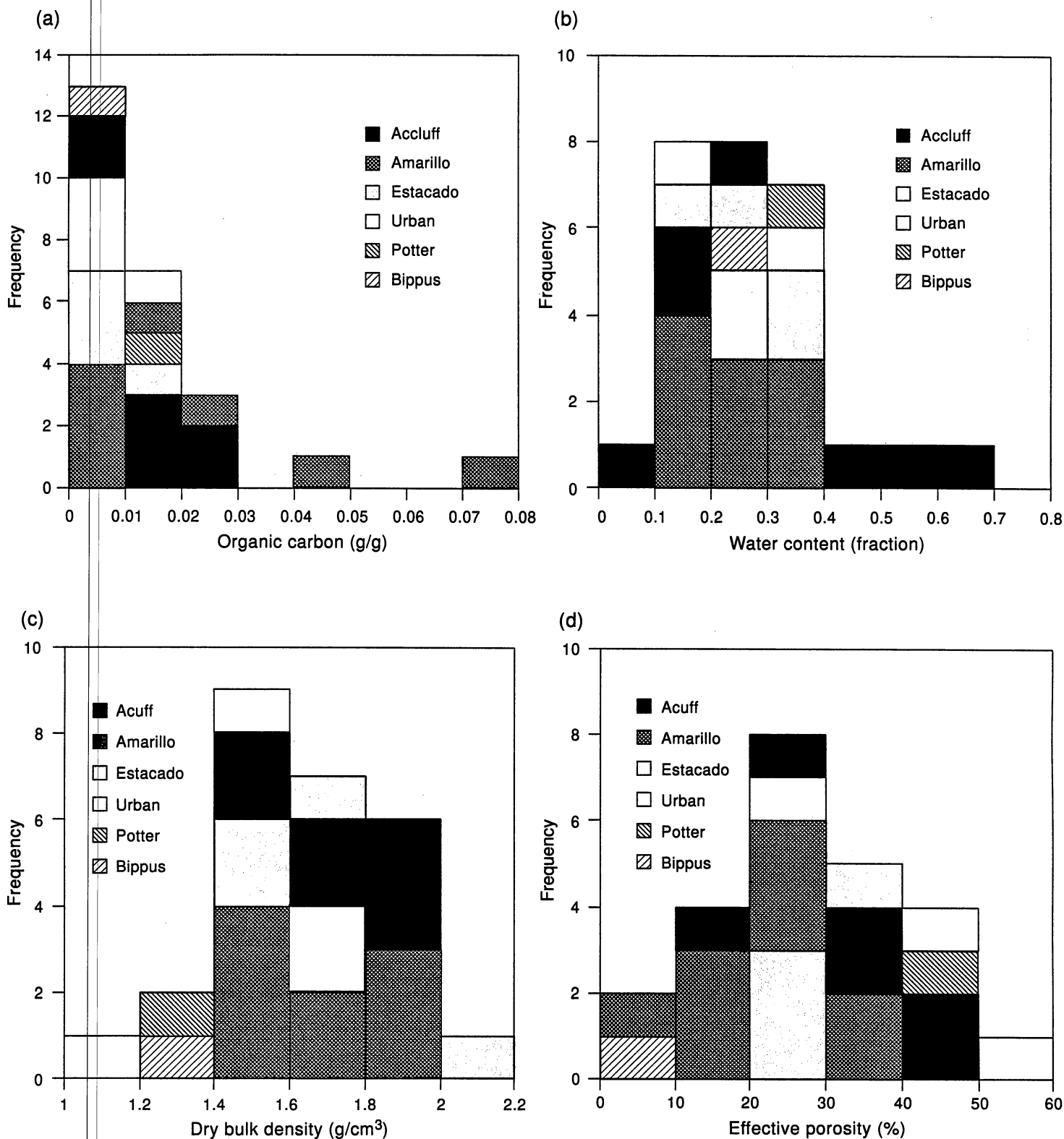
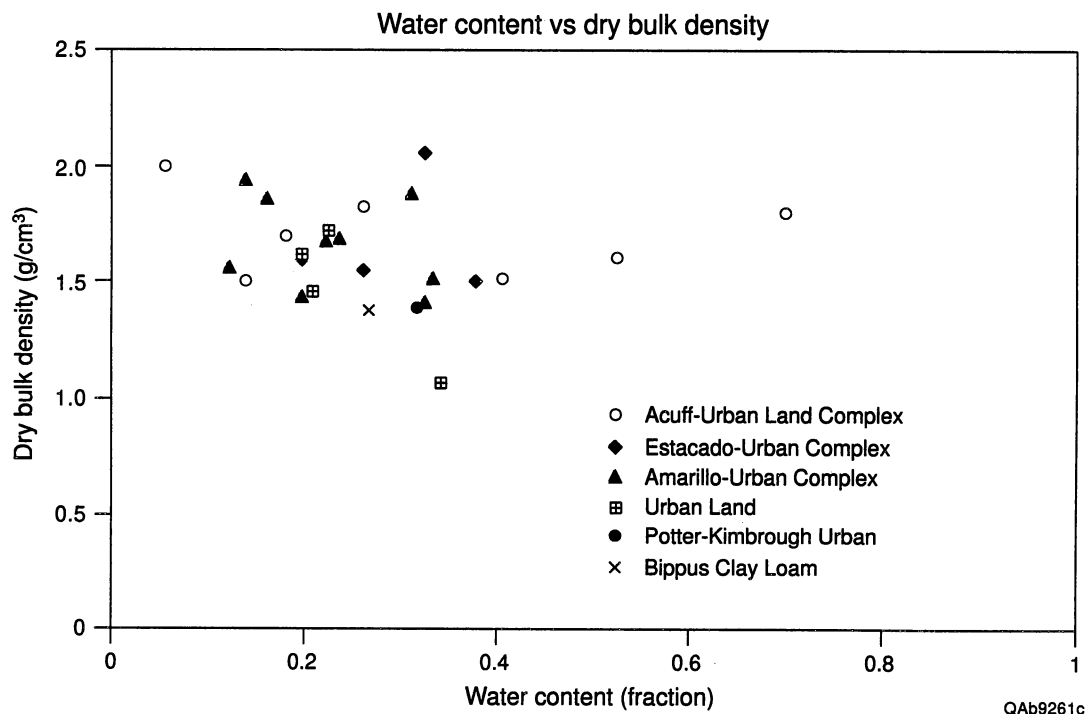
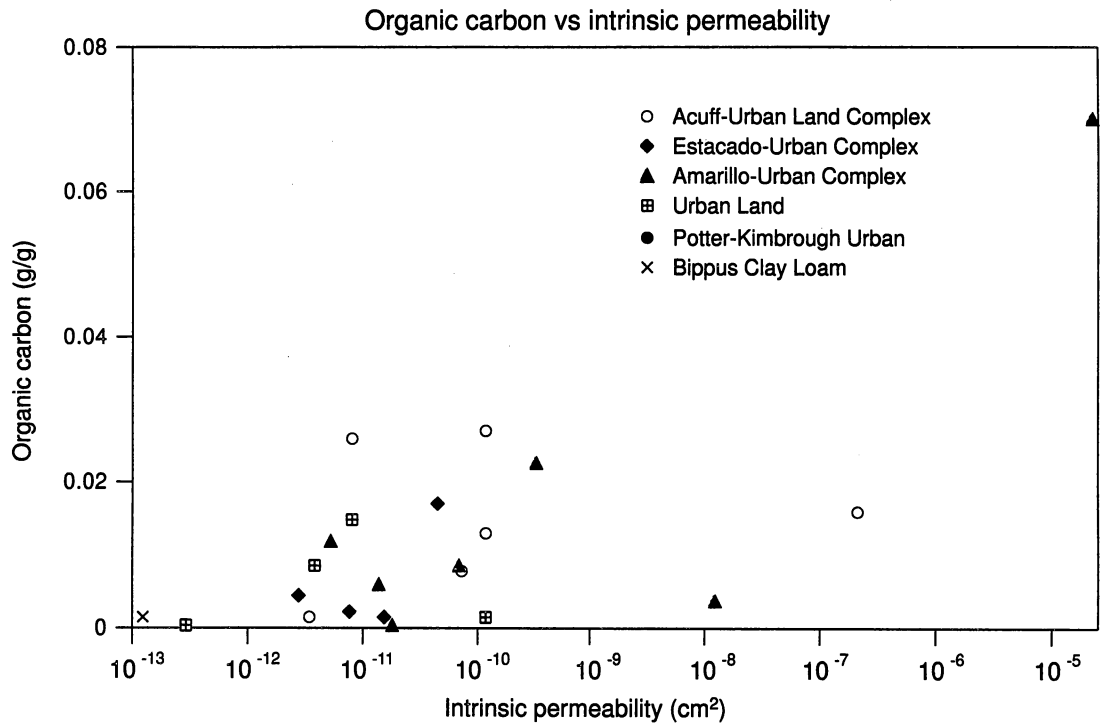


Figure 6. ArcView polygons showing soil classification and point coverage of LPST boreholes at which geotechnical data were collected. The soil mapping in the Lubbock East and Lubbock West quadrangles is derived from Natural Resources Conservation Service (NRCS) SSURGO data.



QAb9260c

Figure 7. Histograms of representative geotechnical data sorted by soil group.



QAb9261c

Figure 8. Cross plots of geotechnical soil parameters.

for the Blackwater Draw Formation in the literature (Holliday, 1989; Hovorka, 1995). As expected, no systematic variation in the geotechnical parameters among the soil types is identified, probably because the soils have similar parent materials and similar soil evolution.

In other geographic settings, different soil complexes form in contrasting geologic settings. For example, soils and subsoils on young and old alluvial materials and on several different types of bedrock have strong contrasts in composition and vertical permeability structure. We predict that the range of geotechnical parameters in these settings would be different for different parent material beneath the soil and for different depths of soil development. In such cases, distinctive statistical descriptions could be developed for each soil complex or group of complexes for which a separable population of geotechnical parameters is found. In the following example, we demonstrate the use of the single distribution of geotechnical parameters from the upland environment in Lubbock County for risk assessment.

#### Heterogeneity in the aquifer

Using GIS to merge several data sources to describe the heterogeneity of the aquifer parallels the example given above for using soils data to describe the vadose zone. Using ARCEDIT, we digitized and cleaned part of the regional facies map of the Ogallala Formation showing percent sand and gravel, then plotted from the aquifer data table the location of LPST monitoring wells at which aquifer tests were conducted (fig. 9). The percent and thickness of these coarse-grained materials should correlate positively with hydraulic conductivity. However, in the pilot area, this technique was not successful, because, even though the hydraulic conductivities varied over six orders of magnitude (fig. 10), all aquifer test sites fell within the 40 to 60 percent sand and gravel region.

Three factors interfered with the application of this method in the pilot area: (1) The Ogallala geologic facies map was prepared at a regional scale and lacked detail. (2) The LPST sites were concentrated in a single urban area that coincided with a single aquifer facies. (3) The aquifer tests were conducted in the very top of the aquifer or possibly in the overlying Blackwater Draw

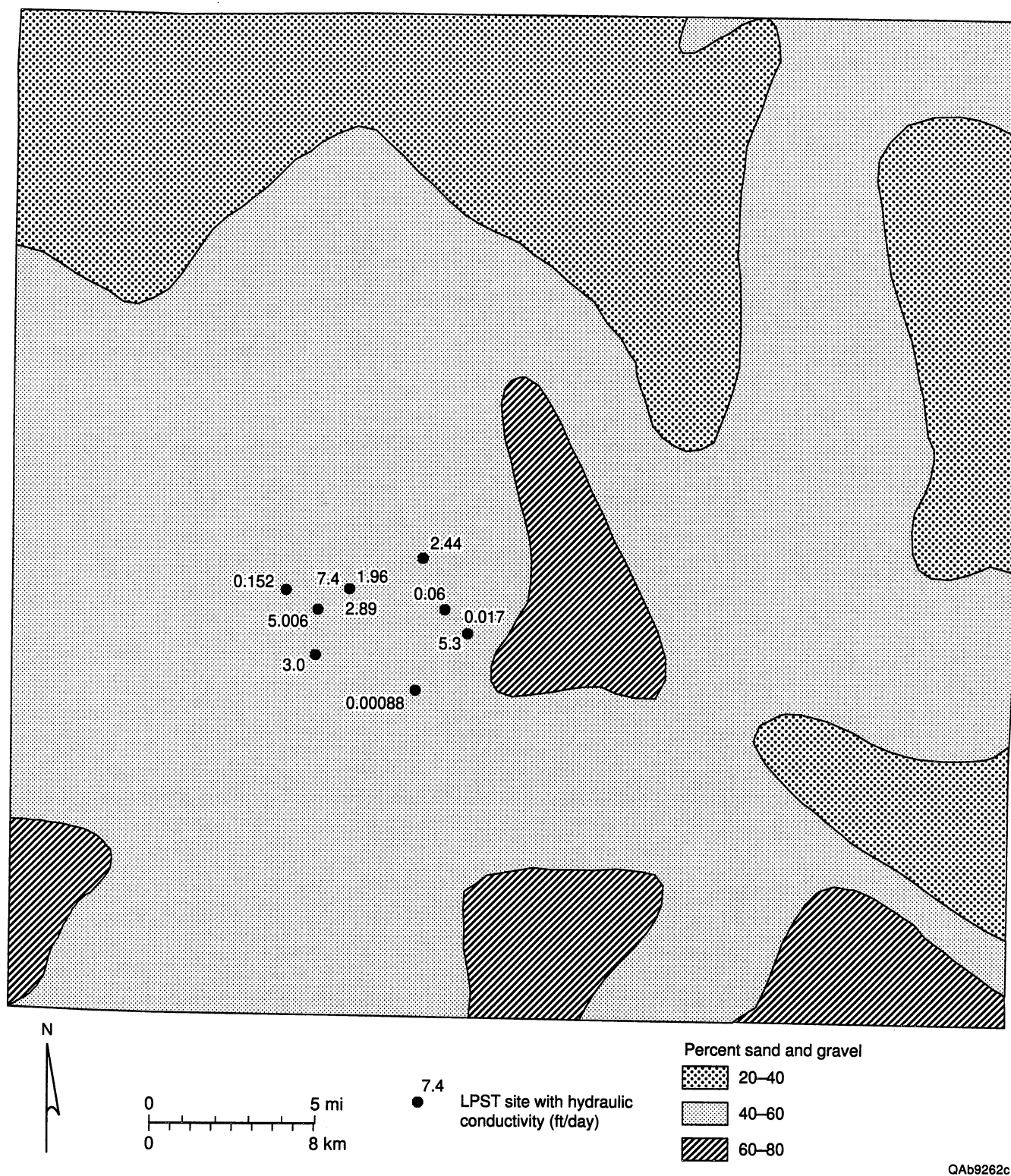


Figure 9. Hydraulic conductivity measured at LPST sites and percent sand and gravel in the aquifer (Seni, 1980) in Lubbock County.

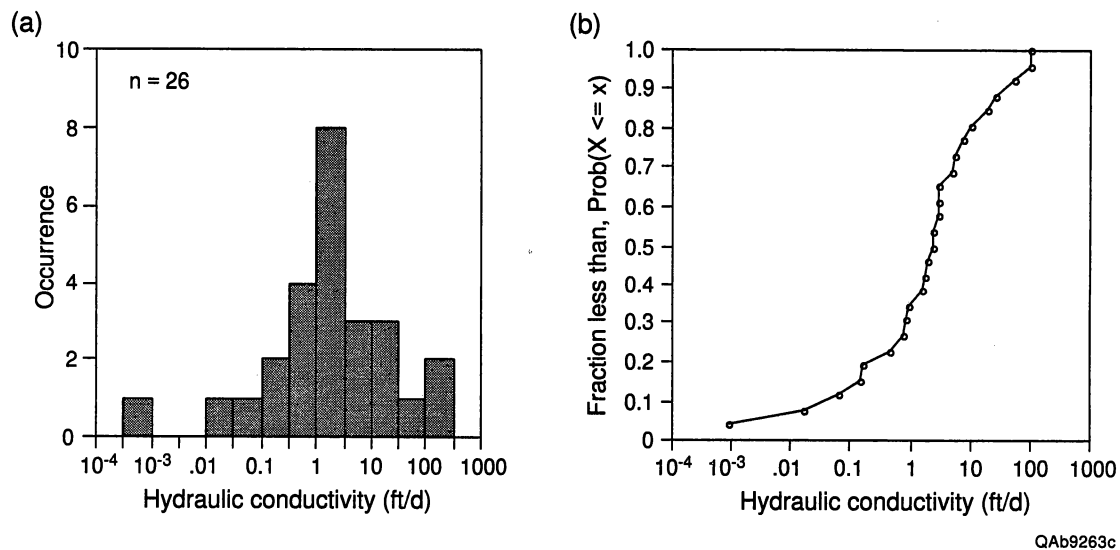


Figure 10. Histogram (a) and cumulative distribution function (b) for hydraulic conductivity. These data include six tests from the I&H waste site at Reese Air Force Base.

Formation, whereas the mapped sand and gravel intervals were concentrated toward the base. The range of hydraulic conductivities measured at contaminated sites was low relative to the contoured average hydraulic conductivities (67 to 134 ft/day) reported from Lubbock County by Knowles and others (1984). Aquifer tests probably included wells that penetrate the sand and gravel parts of the aquifer. The LPST monitoring-well tests indicated that hydraulic conductivity was generally lower near the top of the aquifer. This analysis develops a distribution of hydraulic conductivity in the top part of the aquifer to be used for risk assessment.

Identification of significant subdivisions of the aquifer might work better in other areas where the geologic units are thinner and mapped in more detail, and where pedogenic carbonate is a less significant factor. Larger urban areas where LPST sites are more widely distributed might have a better chance of showing systematic spatial variability in aquifer properties that can be correlated with aquifer geology.

GIS can assist during site characterization and risk assessment in several other ways. Scanning or digitizing and projecting in GIS makes existing maps available at a common scale that can be used to overlay on a detailed street map or site-location plot (fig. 11). GIS can be used on-screen to zoom in to or query for sites of interest to reduce labor in extracting data from existing maps. Examples of the types of data available for extraction and input on Plan A worksheets in the pilot study areas are aquifer thickness, digitized from Wyatt and others (1992), and total dissolved solids (TDS) from Chen and others (1988).

GIS can be used to analyze existing data to create new maps that are needed for input into worksheets. Techniques include plotting and contouring point data, generating continuous surfaces from existing contoured maps, and creating new mapped parameters from surfaces. For example, in ArcView, we created continuous cell-based grids from digitized maps of aquifer altitude and base elevation (Wyatt and others, 1992). Using Grid software, we were able to use these generated surfaces with the USGS 1:24,000 DEMs to calculate new grids for aquifer thickness and vadose-zone thickness (fig. 12). In the pilot study area, the High Plains Groundwater Conservation District has printed a comprehensive atlas of plates showing aquifer properties (Wyatt and others,

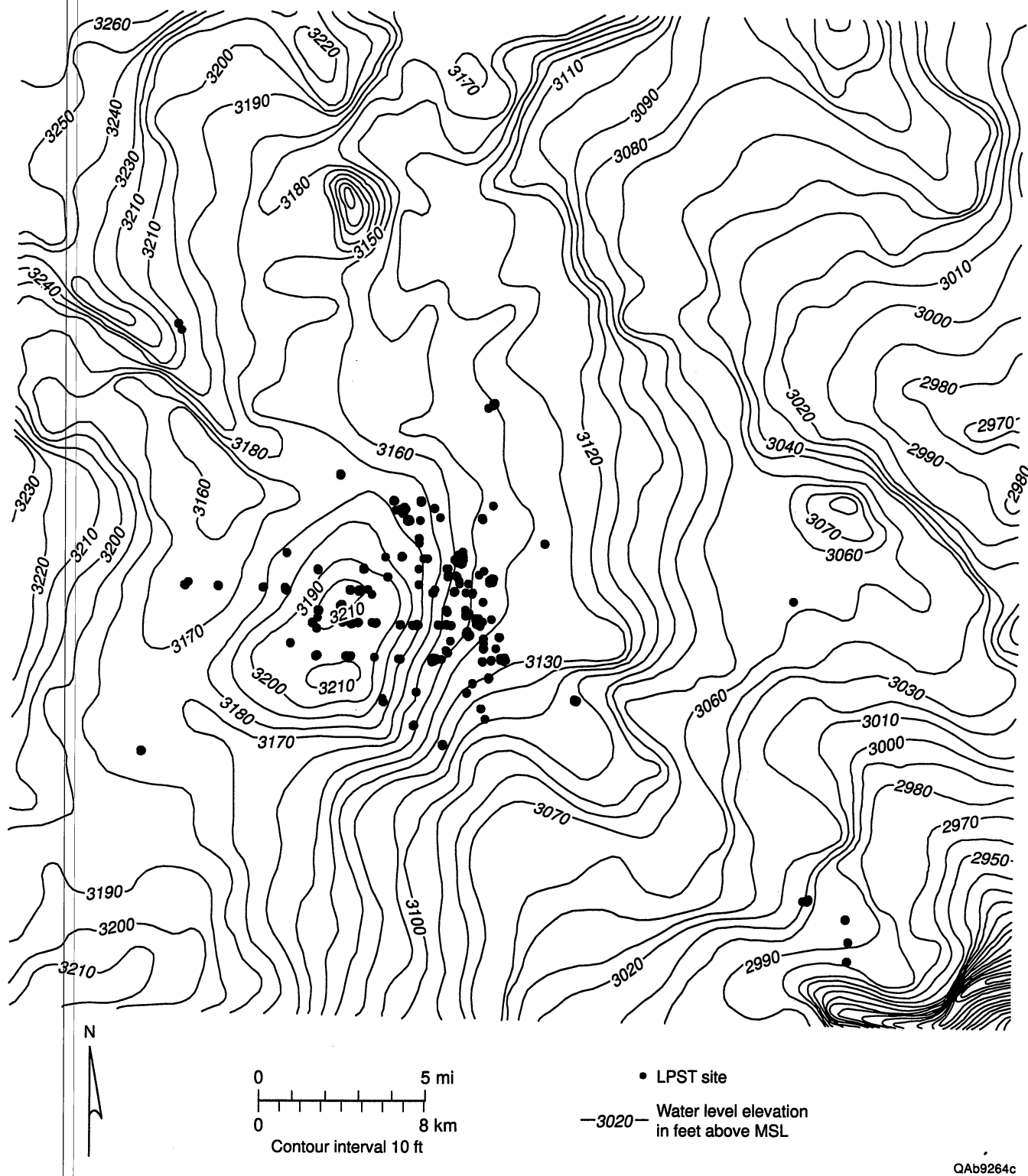


Figure 11. Digitized ArcView projection of the regional 1990 water elevation in feet above sea level (Wyatt and others, 1992) and LPST-monitoring wells.

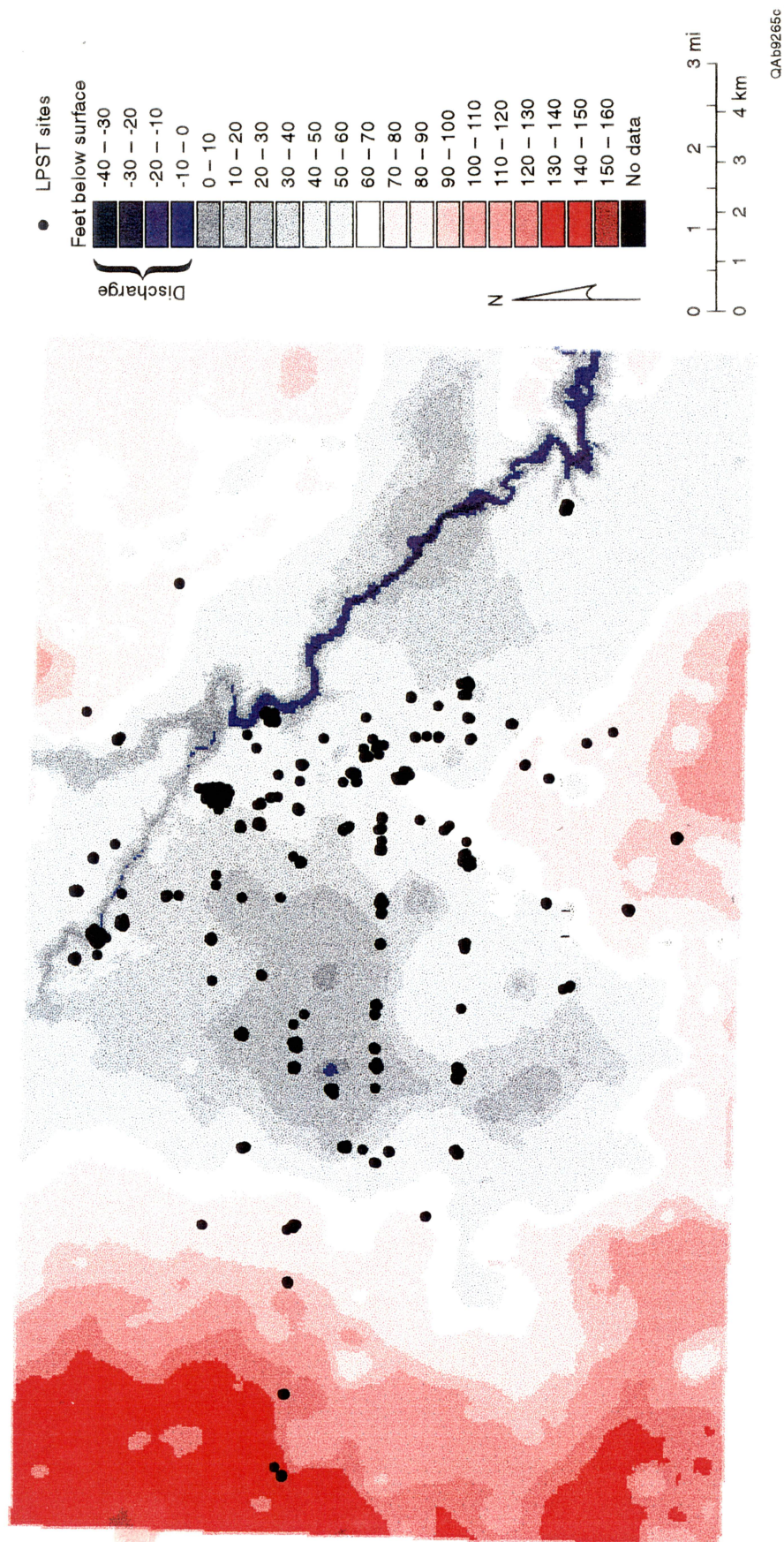


Figure 12. Arc GRID calculation of depth to water in the city of Lubbock calculated from USGS DEMs and 1990 water level (Wyatt and others, 1992). Locations of LPST sites compiled for this study are posted.

1992). In other areas where these data are lacking, GIS could be used to map properties such as TDS or aquifer thickness to provide input data for risk assessment.

Maps of specific yield, permeability, net sand and gravel, tritium concentration, and precipitation are also available and could be used to extend and interpret site-specific data by methods similar to the examples.

#### Hydraulic-head gradient

Hydraulic-head gradient is a key parameter controlling contaminant-plume dynamics. GIS is a tool that can be used to integrate a variety of data to look at regional and local gradient, fluctuations in water level (sweep), and variations in gradient through time. In this pilot project, experiments identified several steps required to integrate hydraulic-head gradient data.

GIS could be used to screen for erroneous water level or surface elevations and to correct the datum. Many sites provide elevation relative to an arbitrary on-site datum. For these sites, elevation is extracted from DEMs using ArcView Spatial Analyst, and water-level elevations relative to sea level are calculated in Excel before input of the aquifer data table into ARC/INFO. This approach is reasonably accurate on a between-site scale in an area of low topographic relief such as Lubbock County; however, it is not very satisfactory in areas where sites are closely spaced, because inaccuracies in elevation may be large relative to the hydraulic-head gradient between sites.

Simple contouring algorithms in ArcView had difficulty handling the irregular data distribution generated from LPST-well spacing, and edge effects and other contouring errors were unacceptable. Local drawdown due to pumping at remediation wells on site were also difficult to interpret with simple contouring.

We recommend use of a co-kriging contouring algorithm where one digital, well-constrained water-level map is used to control the contouring of other approximately parallel water-level surfaces for each time slice. As a test of how this technique would work, we hand-contoured LPST-site water levels for the first quarter of 1993 using ArcTools in ARCEDIT. A 1990 water-level map (Wyatt and others, 1992) was used for guidance where LPST-monitoring wells were not

present or not measured during this period. On-screen contouring used the capabilities of GIS to integrate data at several scales. Countywide 1990 water levels, citywide LPST sites, and site-specific hydraulic-head gradients were integrated into one best-fit map of water level in the first quarter of 1993 by contouring the 93-1 LPST water levels on-screen with the 1990 water levels as a background coverage (fig. 13).

### Risk-Based Corrective Action (RBCA) Calculations for Soil Vapor and Ground Water

Risk assessments are performed to identify target remediation concentrations upon which a site can be closed. These concentrations are generally higher than minimum concentration levels (MCLs) because they consider the residual risks to human health from the remaining contaminants. Texas uses two types of risk calculation: one that uses conservative assumed values of soil and aquifer properties to determine target concentrations (Plan A), though site-specific data can be used, and another that uses soil and aquifer properties measured at the site to determine target concentrations (Plan B).

Site-specific measurements of soil and aquifer properties used for Plan A and B risk assessments generally rely on a single sample. Because geologic material is heterogeneous, a single measurement might not indicate soil and hydrogeologic conditions at the site and could not be used to estimate the variability. If the sample selected for laboratory analysis happens to be from an interval with low permeability and high organic-carbon content, target concentrations may be set too high and the site may be closed while there is still risk to human health. On the other hand, target concentrations might be set too low and the site expensively remediated and monitored when there is little risk to human health.

We use the site data extracted from the GIS data base to address vadose-zone and aquifer heterogeneity beneath LPST sites when conducting risk assessments. For this example, measurements from each of the LPST sites in Lubbock County are used collectively as a realization of probable heterogeneity at any one site. Our spatial analysis described above, supplemented by

literature review and cross-section examination, shows that all but one of the soil borings lie within a single hydrogeologic environment with lateral and vertical variation at less than site scale.

Because the sites are heterogeneous, target concentrations determined from site measurements vary depending on the measurement values used. To quantify the range of target concentrations that might be expected from a thorough characterization of the site, we randomly sampled the statistical distributions of the collective measurements (Monte Carlo sampling) and determined the resulting distributions of Plan A and B target concentrations.

#### Statistical description of site-specific measurements

Several site-specific measurements are used for Plan A and Plan B risk assessments:

- Hydraulic conductivity
- Organic-carbon content
- Moisture content
- Dry-soil bulk density
- Hydraulic-head gradient

We compiled each of these parameters from site files to define statistical distributions for the population. Figures 10, 14, 15, 16, and 17 show the range, most probable values, and cumulative distribution function for each parameter. Some values are suspect: for example, moisture contents greater than 50 percent are highly unlikely (fig. 15). Experience suggests that some of the higher values of hydraulic-head gradient are likely to be wrong (fig. 17). We retained all the reported values for analysis, however, to show what effect they would have on risk calculations.

#### *Plan A equations*

Several Plan A risk calculations can be made using site-specific data for soil bulk density, moisture content, and fraction of organic carbon (TNRCC, 1994). These calculations determine target air and soil concentrations. Target ground-water concentrations are independent of site-

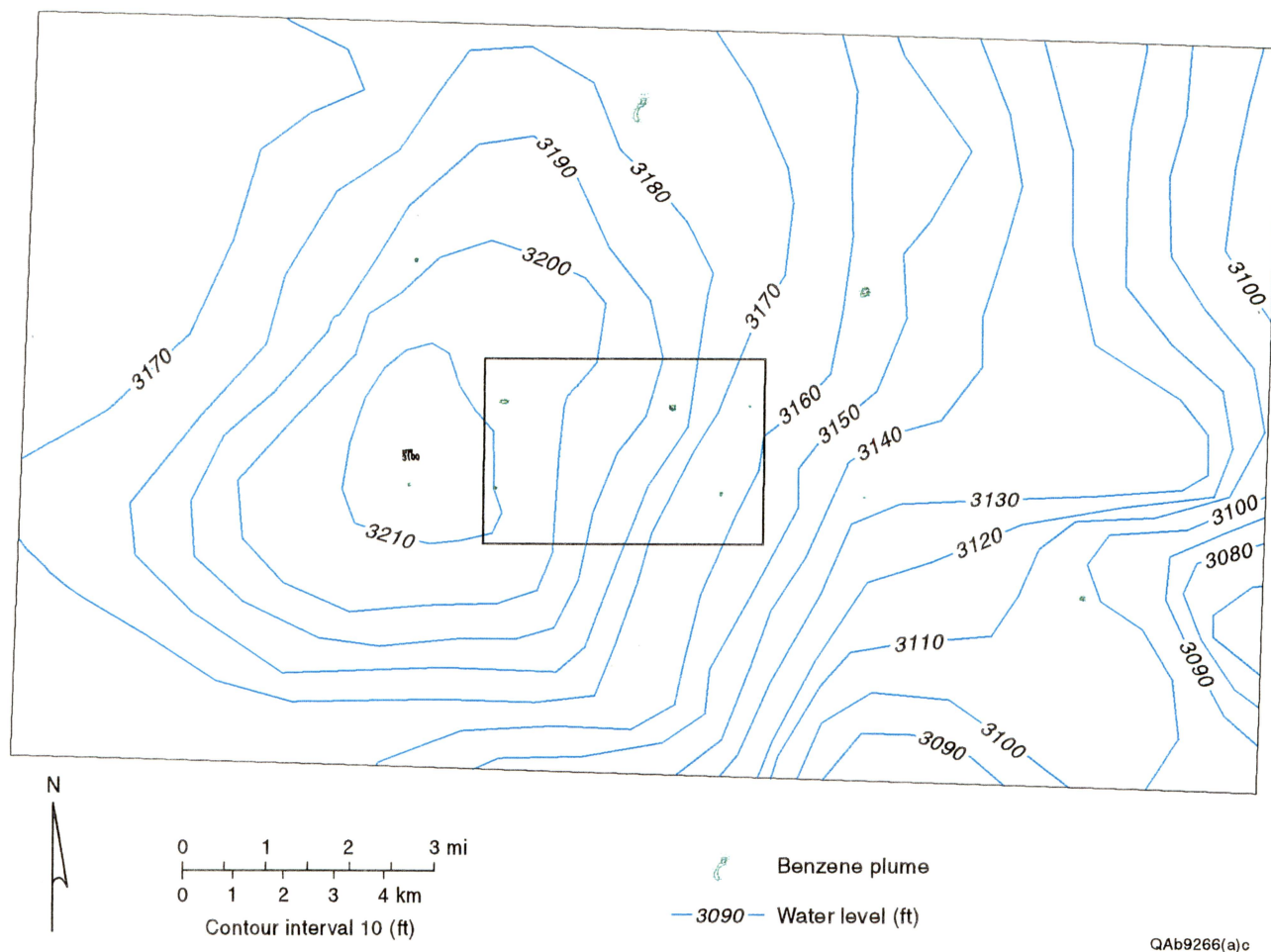
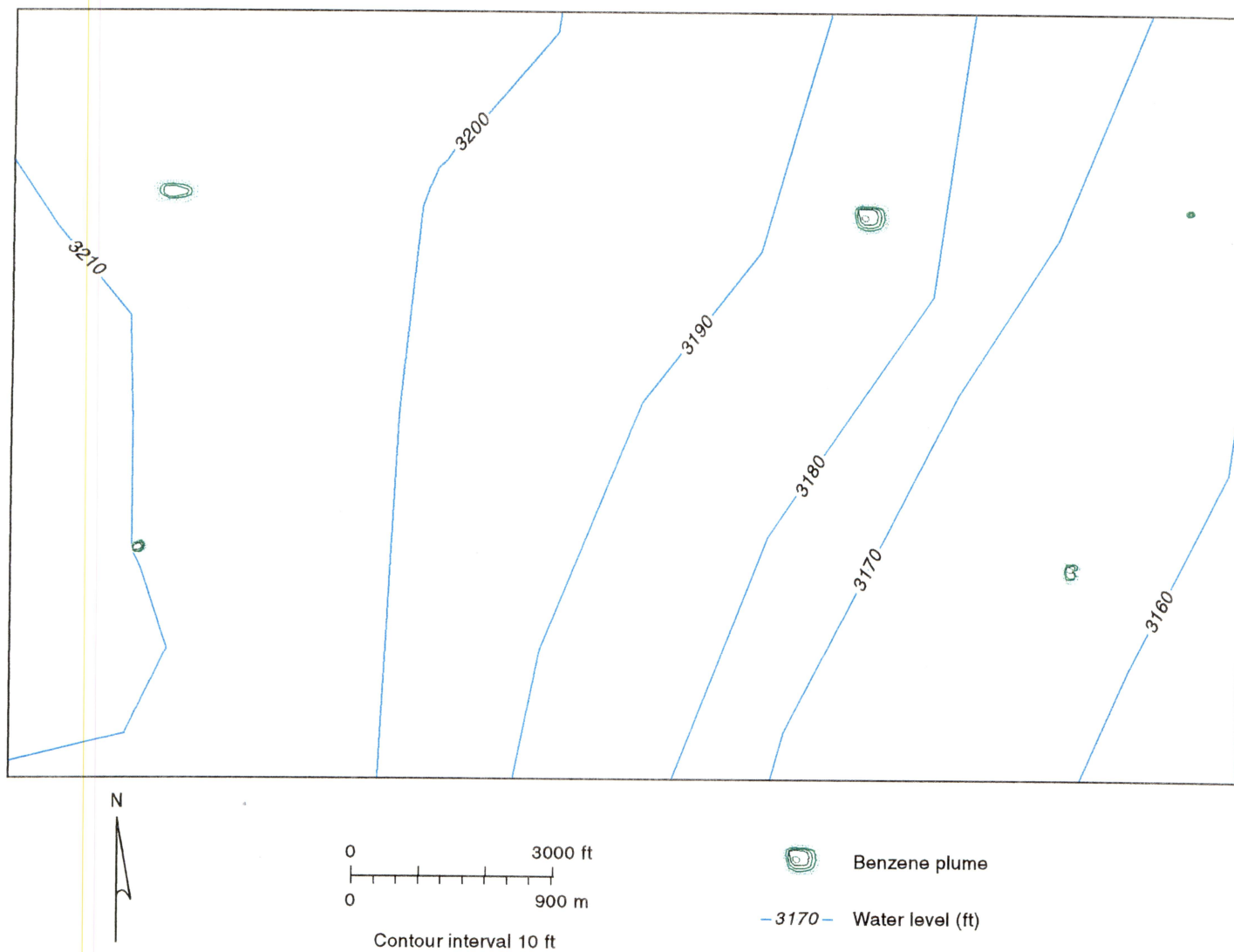


Figure 13. Hand-contoured interpretation of water level in the first quarter of 1993 and the extent of representative benzene plumes during the same period in the city of Lubbock (a). An enlarged view is shown in b. This demonstration used Arc tools in ARCEDIT with a back cover of the 1990 water levels.



QAb9266(b)c

Figure 13. (cont.)

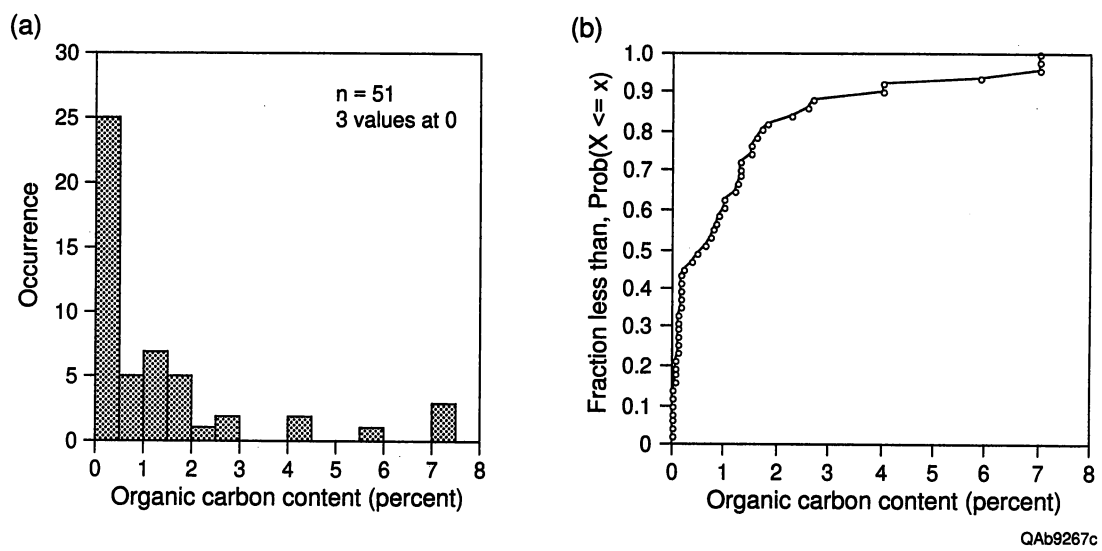


Figure 14. Histogram (a) and cumulative distribution function (b) for organic-carbon content.

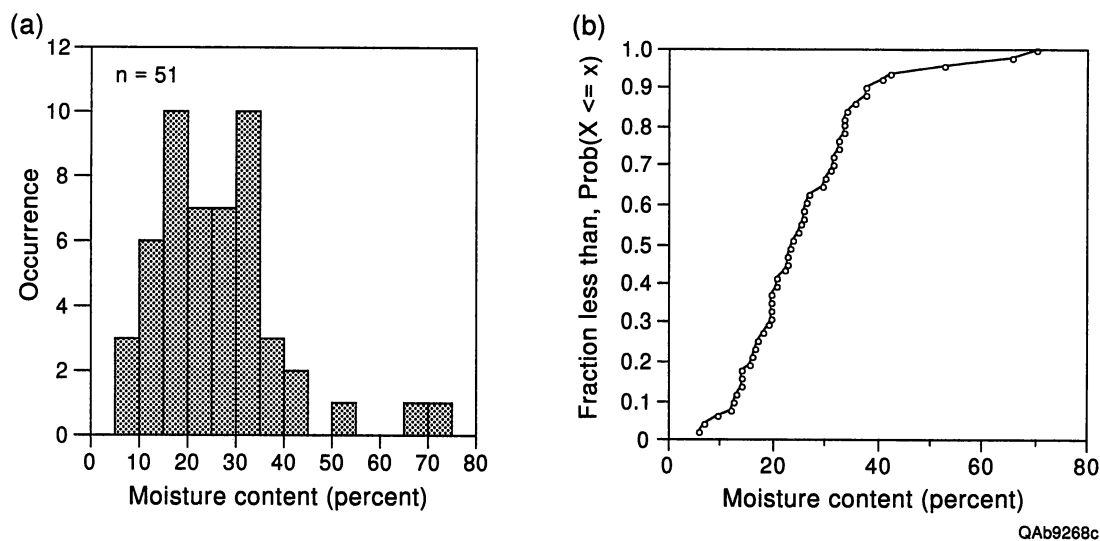


Figure 15. Histogram (a) and cumulative distribution function (b) for moisture content.

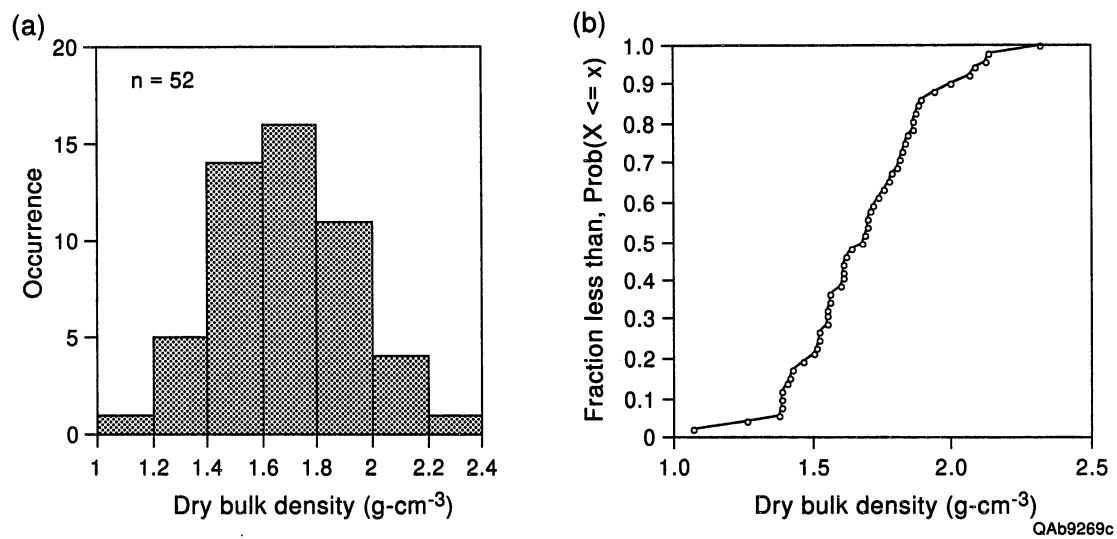


Figure 16. Histogram (a) and cumulative distribution function (b) for dry bulk density.

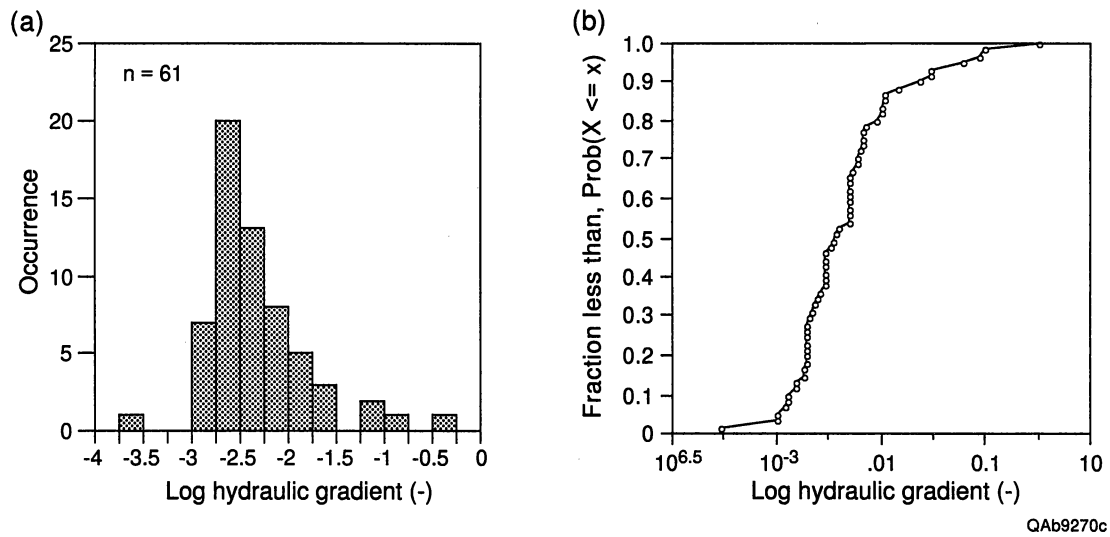


Figure 17. Histogram (a) and cumulative distribution function (b) for hydraulic-head gradient.

specific data for Plan A risk assessment. Many of these equations are unit dependent and therefore require the specific units listed with the default values shown in the tables.

### *Target Air Concentrations*

Target air concentrations cannot exceed 25 percent of the lower explosive limit (LEL) for the volatile constituent of concern. The expected concentration in the air,  $C_a$ , for a soil contaminated by a volatile is determined by

$$C_a = \frac{C_T H' \beta}{K_d \beta + \phi_w + \phi_a H'} \quad (2)$$

where  $C_T$  is the bulk soil concentration,  $H'$  is the dimensionless Henry's Law constant,  $\beta$  is the dry-soil bulk density,  $K_d$  is the soil-water partition coefficient,  $\phi_w$  is the water content, and  $\phi_a$  is the air-filled soil porosity.  $K_d$  is defined from

$$K_d = K_{oc} f_{oc} \quad (3)$$

where  $K_{oc}$  is the organic-carbon partition coefficient and  $f_{oc}$  is soil organic-carbon fraction. Air-filled soil porosity is found from

$$\phi_a = \phi - \phi_w \quad (4)$$

where  $\phi$  is defined by

$$\phi = 1 - \frac{\beta}{P_b} \quad (5)$$

where  $P_b$  is the particle density. The dimensionless Henry's Law constant is defined by

$$H' = H \times 41.57 \quad (6)$$

where  $H$  is the Henry's Law constant.

$H$  and  $K_{oc}$  are constants dependent on the chemical of concern (table 5).  $P_b$  is not generally measured at sites and is assumed to be 2.65 kg L<sup>-1</sup> (table 6).  $\beta$ ,  $\phi_w$ , and  $f_{oc}$  can be specified using site-specific measurements and also have default values when measured values are not available (table 6).  $C_T$  is determined from site-specific measurements of soil concentrations (TNRCC, 1994).

Table 6. Default values for soil properties.

**Default soil parameters:**

$f_{oc}$	=	0.2 percent
$\beta$	=	1.8 kg L <sup>-1</sup>
$P_b$	=	2.65 kg L <sup>-1</sup>
$\phi_w$	=	10 percent

### Target Soil Concentrations

The target soil concentrations for residential ingestion of soil are independent of site-specific data. However, the target soil concentrations for residential and worker ingestion and inhalation of volatiles and particulates are dependent on site-specific soil data.

The target soil concentration for residential ingestion and inhalation of volatiles and particulates with carcinogenic effects is

$$C_s = \frac{T_R \times 5,110}{\left[ (0.00798 \times S_{Fo}) + \left( S_{Fi} \times \left( \left( \frac{450}{V_{Fr}} \right) + (1.54 \times 10^{-8}) \right) \right) \right]} \quad (7)$$

where  $T_R$  is the target excess individual lifetime cancer risk ( $10^{-6}$  for benzene),  $S_{Fo}$  is oral cancer slope factor,  $S_{Fi}$  is the inhalation cancer slope factor, and  $V_{Fr}$  is the residential soil-to-air volatilization factor. The target soil concentration for worker ingestion and inhalation of volatiles and particulates with carcinogenic effects is

$$C_s = \frac{T_R \times 286.2}{\left[ (0.00005 \times S_{Fo}) + \left( S_{Fi} \times \left( \left( \frac{20}{V_{Fi}} \right) + (6.9 \times 10^{-10}) \right) \right) \right]} \quad (8)$$

where  $V_{Fi}$  is the industrial soil-to-air volatilization factor. There are also equations to determine target soil concentration for residential and worker ingestion and inhalation of volatiles and particulates with noncarcinogenic effects (TNRCC, 1994).

The parameters,  $T_R$ ,  $S_{Fo}$ ,  $S_{Fi}$ ,  $V_{Fr}$ , and  $V_{Fi}$ , required to determine these target soil concentrations can be determined using default values (table 5). However, the volatilization factors,  $V_{Fr}$  and  $V_{Fi}$ , can also be calculated using site-specific soil data:

$$V_F = \frac{(L_s \times V \times D_H)}{A} \times \frac{\sqrt{(3.14 \times \alpha \times T)}}{(2 \times D_{ei} \times \phi \times K_{as} \times 10^{-3})} \quad (9)$$

where

$$\alpha = \frac{D_{ei} \times \phi}{\phi + \frac{P_b(1 - \phi)}{K_{as}}} \quad (10)$$

$L_s$  is length of the contaminated area,  $V$  is the wind speed in the mixing zone,  $D_H$  is the diffusion height,  $A$  is the area of contamination,  $T$  is the exposure interval,  $D_{ei}$  is the effective diffusivity, and  $K_{as}$  is the soil/air partition coefficient, defined as

$$K_{as} = \frac{4.1H}{K_d} \quad (11)$$

Default values are available for the input parameters  $L_s$ ,  $V$ ,  $D_H$ ,  $A$ ,  $T$ ,  $D_{ei}$ , and  $K_{as}$  (table 7).

However, site-specific values can be used to calculate  $D_{ei}$  and  $K_{as}$  (table 7 and eqn. 11, respectively).

Equations 7 and 8 are valid only if soil contaminant concentrations are at or below saturation. Above saturation, the adsorptive limits of the solid phase and the solubility limits of the soil moisture are exceeded and liquid-phase contaminant is present. The bulk soil concentration,  $C_{sat}$  [mg kg<sup>-1</sup> dry weight], that coincides with the saturation limit of the soil is

$$C_{sat} = \frac{S(\beta K_d + \phi_w + \phi_a H')}{\beta} \quad (12)$$

where  $S$  is the pure-component solubility (table 7).

For sites where ground water is less than 15 ft below the land surface (fig. 12), a target soil concentration protective of ground water must be calculated (TNRCC, 1995). This target soil concentration,  $C_s$ , is

$$C_s = \frac{D_L C_w (\beta K_d + \phi_w + \phi_a H')}{\beta} \quad (13)$$

where  $D_L$  is the leachate concentration dilution factor and  $C_w$  is the Category I ground-water target concentration.  $D_L$  is assumed to be 100, and  $C_w$  is defined by default values for the chemical of concern (table 6).

### *Plan B equations*

Plan B risk assessments evaluate current and potential human health risks and the short- and long-term fate of contaminants, and typically require more thorough site assessment and regulatory review. Plan B risk assessments differ from Plan A risk assessments in that site-specific

Table 7. Default values for calculating volatilization factors.

Parameter	Units	Default value
$L_s$	[m]	21
$V$	[m-s <sup>-1</sup> ]	2.25
$D_H$	[m]	2
$A$	[cm <sup>2</sup> ]	1,500,000
$D_{ei}$	[cm <sup>2</sup> -s <sup>-1</sup> ]	$D_i \times \phi^{0.33}$
$\phi$	[-]	0.35
$P_b$	[g-cm <sup>-3</sup> ]	2.65
$T$	[s]	$9.5 \times 10^8$ (residential) $7.9 \times 10^8$ (industrial)

- $D_i$  is the molecular diffusivity [cm<sup>2</sup>-s<sup>-1</sup>], listed in table 1
- note that the equation for  $D_{ei}$  is different than in TNRCC (1994), which is in error (Chet Clarke, personal communication, 1997)

hydrogeologic properties are used with contaminant transport modeling to determine target concentrations. Contaminant transport depends strongly on the velocity,  $v$ , of ground water that is defined as

$$v = \frac{-Ki}{\phi} \quad (14)$$

where  $K$  is the hydraulic conductivity of the formations and  $i$  is the hydraulic gradient.

### *Analysis*

Our approach was to use Monte Carlo sampling to pick site-specific values and then calculate risk parameters that depend on those values. Monte Carlo sampling involves randomly choosing a value based on the distribution. For example, if we wanted to choose a value of organic-carbon fraction (fig. 14a) using Monte Carlo sampling, we would choose a random number between 0 and 1 and then find the corresponding organic-carbon fraction value for that random number from the cumulative distribution function (fig. 14b). When this is done many times, the original distribution can be reproduced. Therefore, when an equation depends on several-site specific parameters, each of the cumulative distribution functions for the parameters can be randomly sampled and substituted into the equation, resulting in a calculated value. When this is done hundreds of times, a distribution of calculated results shows the range of expected values. This analysis assumes that the input variables are independent of each other as tested by cross plotting (fig. 8).

After we used cross plots to verify that the site-specific measurements were independent of each other (fig. 8), we wrote a FORTRAN program to read in and randomly sample cumulative distribution functions and calculate risk-assessment parameters. Resulting distributions of risk-assessment parameters consisted of 1,000 random samples. We determined distributions for (1) the ratio of expected air concentration to bulk soil concentration (eqn. 2, divided by  $C_T$ ), (2) target soil concentrations for residential and worker ingestion and inhalation of volatiles and particulates (eqns. 7 and 8), (3) the bulk soil concentration that coincides with the saturation limit of the soil (eqn. 12), (4) the target soil concentration protective of ground water for depths to water

of less than 15 ft (eqn. 13), and (5) ground-water velocity (eqn. 14). We then plotted the distributions and compared them with the values determined using default input parameters.

The resulting distributions show the range of target concentrations and ground-water velocities that might be encountered if the heterogeneity of soil and aquifer parameters at the site were fully characterized. The ratio of expected air concentration to bulk soil concentration ranges from just over zero to about 3 (fig. 18). The default value is 0.93. Therefore, if a single sample is randomly collected at the site and the results of the analysis are used to determine the ratio of expected air concentration to bulk soil concentration, 75 percent of the time the ratio determined from site-specific data will be under the default value and 25 percent of the time it will be over (fig. 18b).

Target soil concentrations for residential ingestion and inhalation of volatiles and particulates ranges from just over zero to about 16 mg kg<sup>-1</sup> (fig. 19). About 45 percent of the target concentrations determined using site-specific data are less than the default value (fig. 19b). Target soil concentrations for worker ingestion and inhalation of volatiles and particulates ranges from just over zero to about 55 mg kg<sup>-1</sup>, with about 45 percent of the target concentrations determined using site-specific data less than the default value (fig. 20).

The bulk soil concentration that coincides with the saturation limit of the soil ranges from just over zero to about 11,000 mg kg<sup>-1</sup> (fig. 21). About 20 percent of the bulk soil concentrations determined from site-specific data are less than the default value. Target soil concentration protective of ground water for depths to water of less than 15 ft ranges from just over zero to 3,000 mg kg<sup>-1</sup> (fig. 22). About 90 percent of target soil concentrations determined from site-specific data are greater than the default concentration. Potential ground-water velocities range from  $3 \times 10^{-6}$  to about 10 ft d<sup>-1</sup> (fig. 23).

Each of the cumulative distribution functions for risk-assessment calculations shows a wide range of results dependent on the input values used. If our assumption is valid that the collective site measurements represent actual heterogeneity beneath a site, then the cumulative distribution functions represent the actual range of values that might be expected if the site were fully

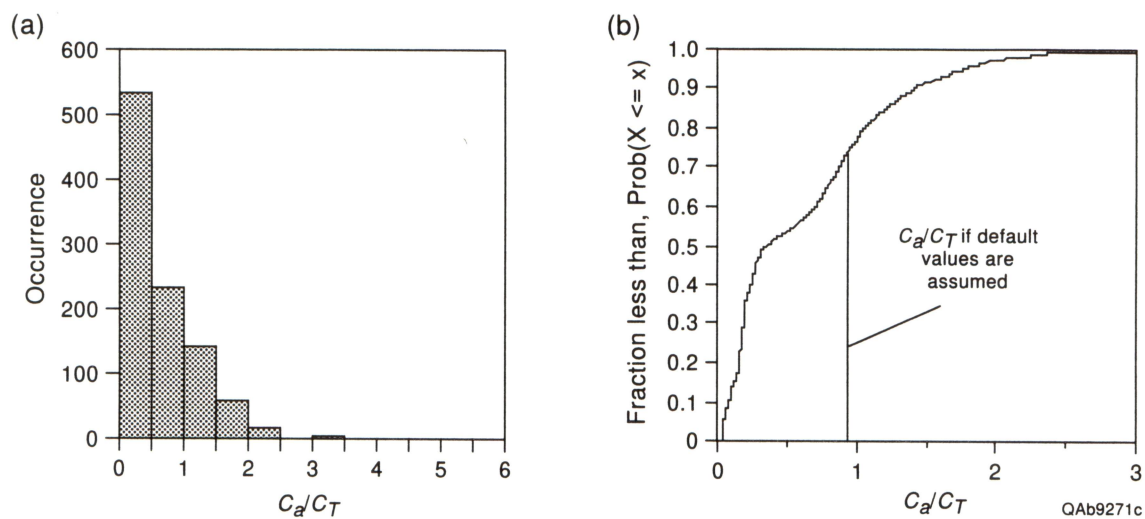


Figure 18. Histogram (a) and cumulative distribution function (b) for the ratio of expected air concentration to bulk soil concentration.

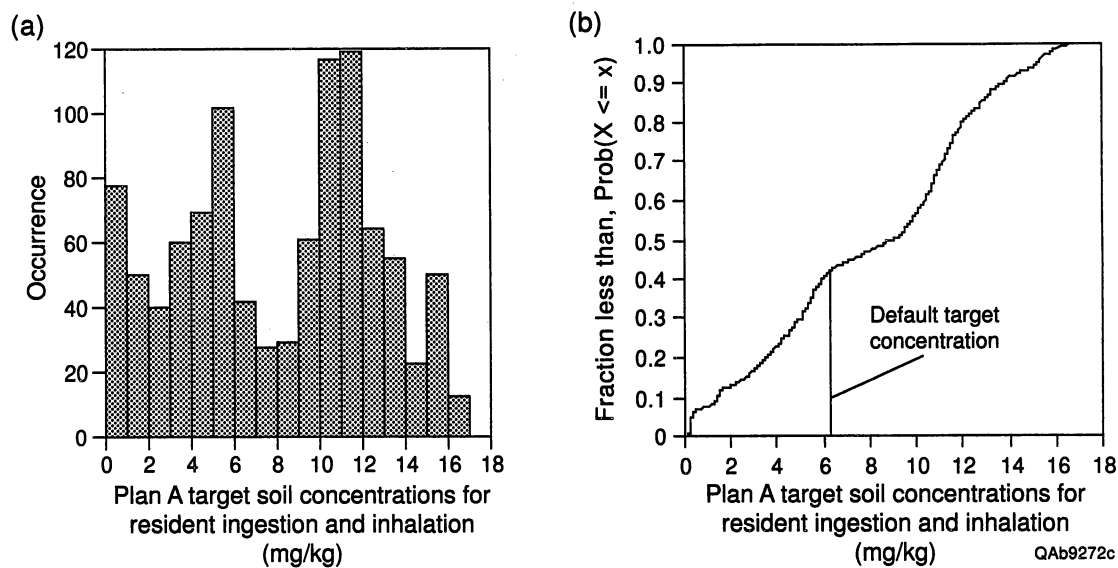


Figure 19. Histogram (a) and cumulative distribution function (b) for the target soil concentration for residential ingestion and inhalation of volatiles and particulates.

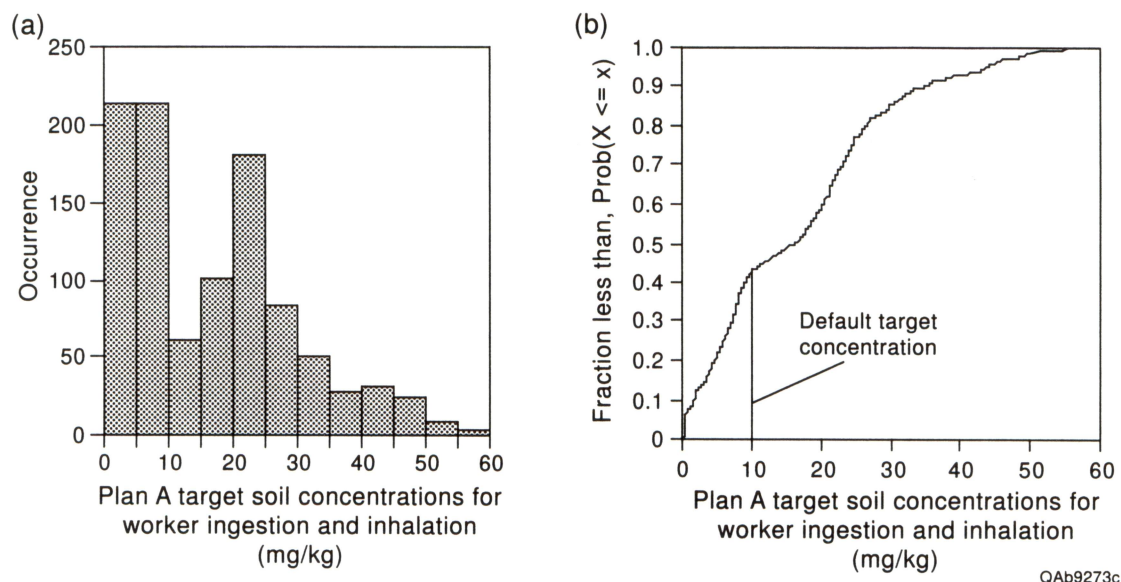


Figure 20. Histogram (a) and cumulative distribution function (b) for the target soil concentration for worker ingestion and inhalation of volatiles and particulates.

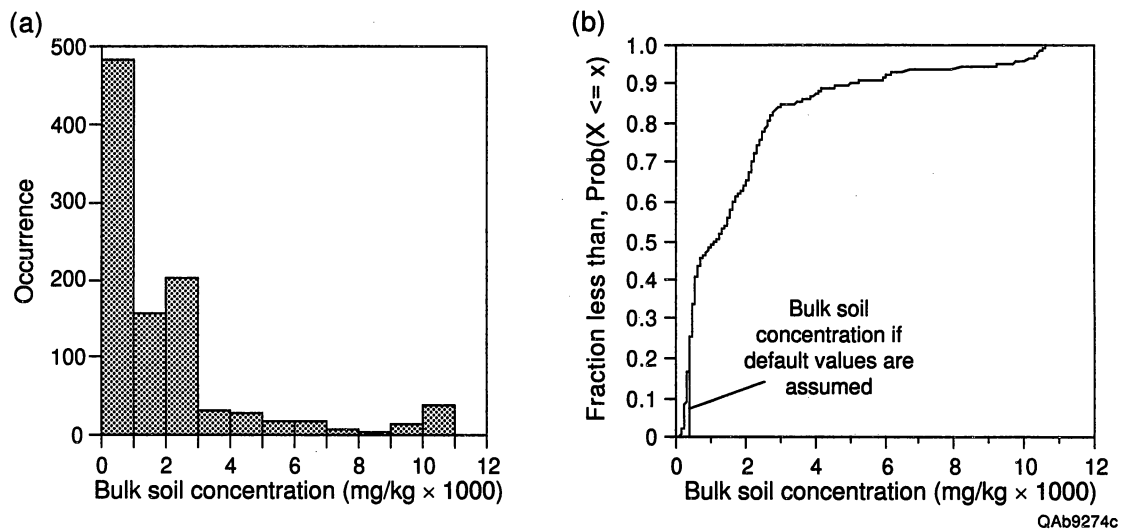


Figure 21. Histogram (a) and cumulative distribution function (b) for the bulk soil concentration that coincides with the saturation limit of the soil.

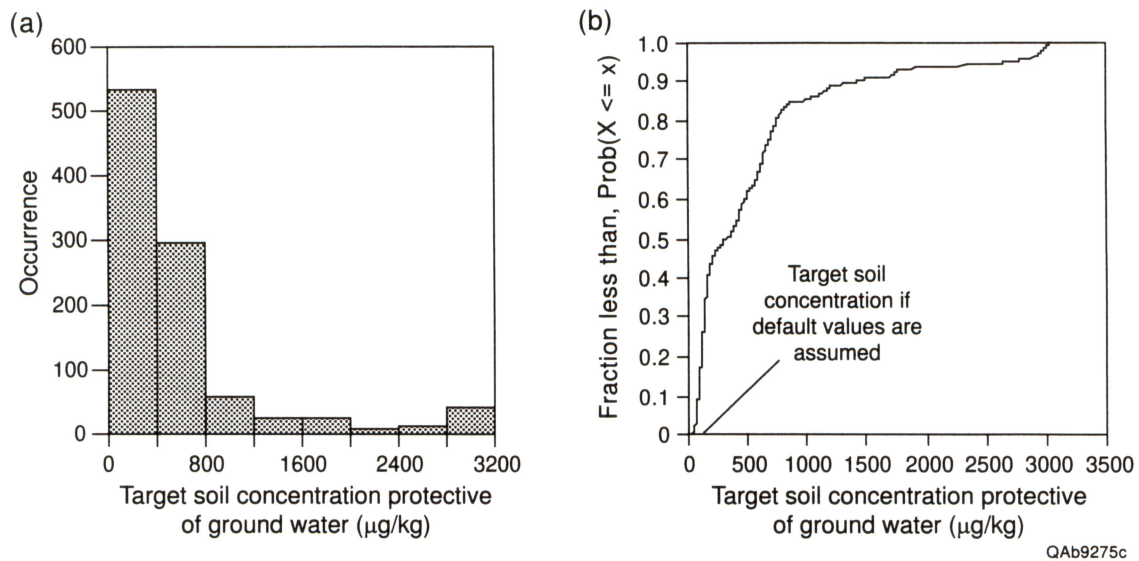


Figure 22. Histogram (a) and cumulative distribution function (b) for the target soil concentration protective of ground water for depths to water of less than 15 ft.

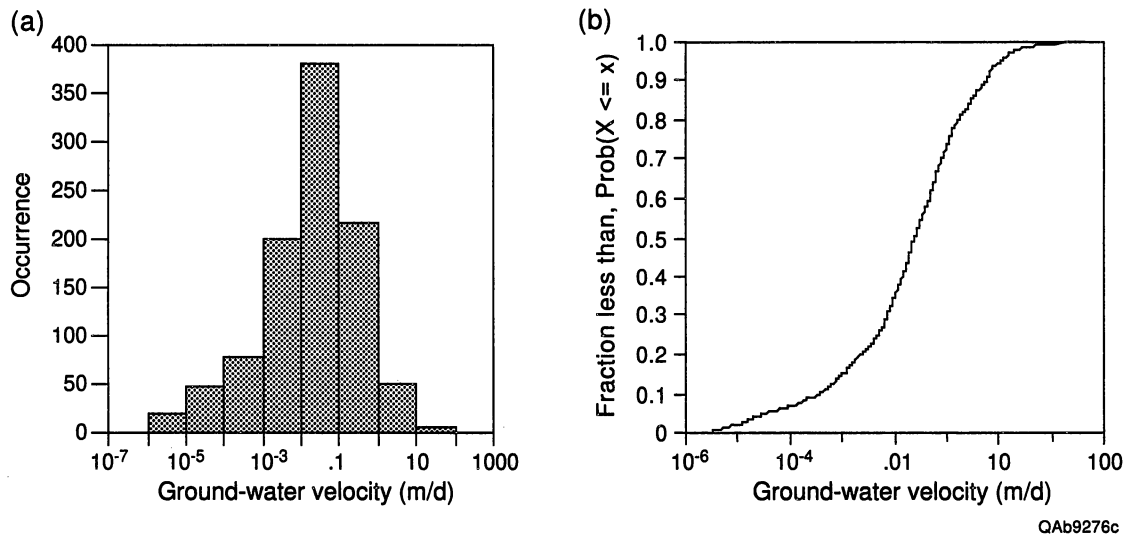


Figure 23. Histogram (a) and cumulative distribution function (b) for ground-water velocity.

characterized. Because site-specific values usually depend on a single measurement, all of the variability is missed and may result in a target concentration that is greater or smaller than appropriate.

At a minimum, this analysis allows a regulator a quantitative view of the true range in soil and aquifers parameters that may exist at a site and how these values impact risk calculations. Using these charts, a reviewer can determine how representative the site-specific values are of the impacted formation. This type of analysis might also be used to fine-tune the current default risk factors where defaults are assigned according to the specific formation and the heterogeneity encountered. For example, the default target concentration for residential ingestion and inhalation of volatiles and particulates might be set at the median value (probability = 0.5) of 9 for the Ogallala aquifer but at a different value, defined by statistical analysis, for the Gulf Coast aquifer. In areas that have an adequate statistical description of needed parameters derived from the HGIS, new sites may be able to focus soil and aquifer data collection, thus reducing site-specific data required. Similar distributions can be generated for other hydrogeologic settings by compiling available data for each aquifer or organizing a field collection program to sample site and formation heterogeneity.

### Contaminant-Plume Evolution

GIS offers a powerful tool for integrating all the variables needed to assess plume size, location, and evolution. This is especially true in areas of closely spaced sites where plumes may have multiple sources active over different time frames. GIS can show contaminant concentrations at adjacent sites without reference to surface features, such as roads and property boundaries, and assist in identifying areas where plumes from several sites may have merged.

GIS allows site-specific data to be integrated and better interpreted in regional context. For example, hydraulic-head gradient can be difficult to assess at sites with a nonoptimum monitoring-well distribution or a small number of wells. In Lubbock, seasonal variations in water level and local hydraulic-head gradient result from rapid recharge through modified playas (Chen and others,

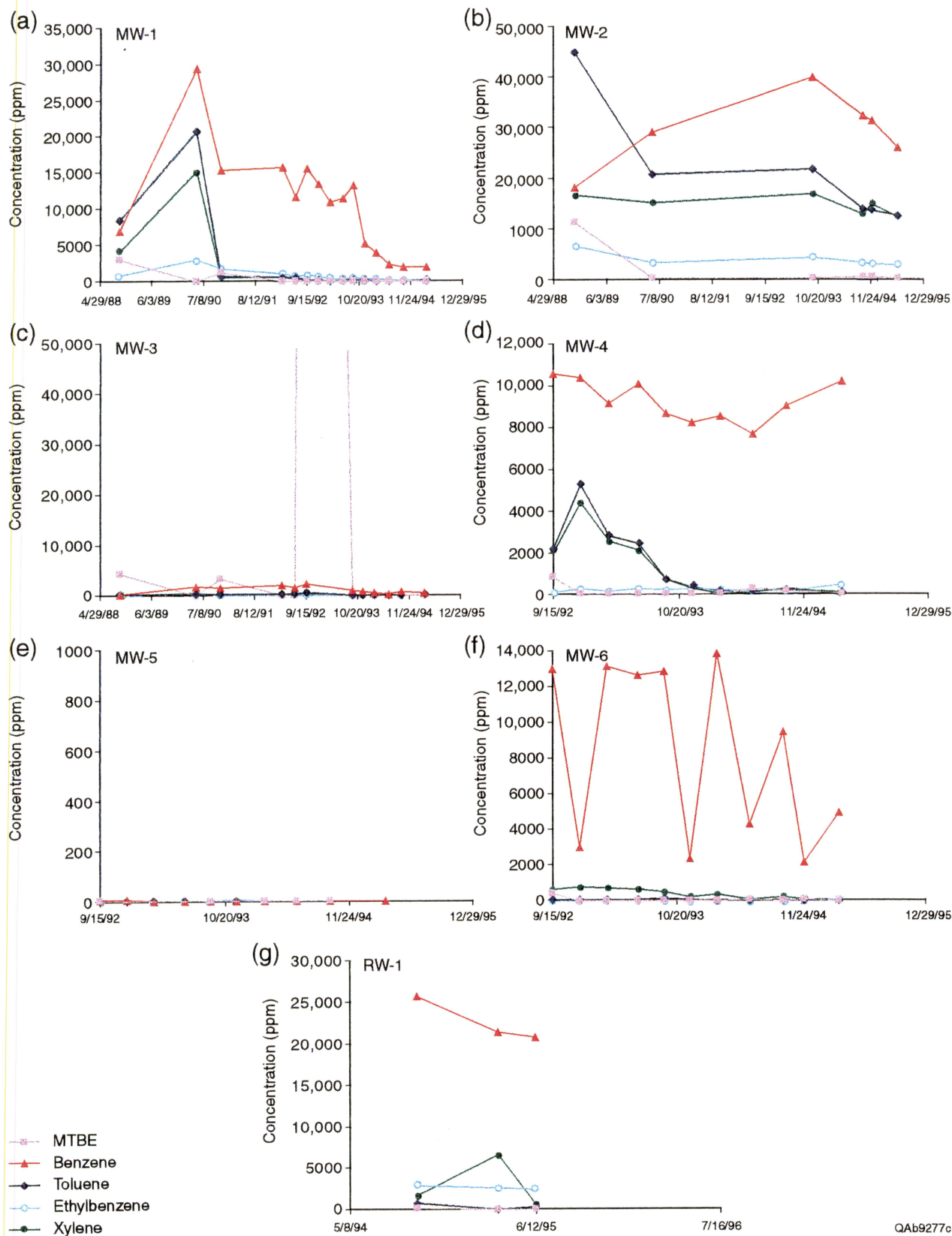
1988; Rainwater and Thompson, 1994). GIS can facilitate interpretation by allowing us to post ground-water elevations through time at numerous sites in the area of interest. Site data can be supplemented by regional data to control contouring away from areas with dense LPST data. This can improve assumptions made about the long-term average hydraulic-head gradient as well short-term fluctuations.

Because ground-water contaminant data are shown in a spatial context, a variety of plume-contouring options are available in GIS. For this pilot project, we hand-contoured a sample of benzene concentrations in representative plumes (fig. 13) to demonstrate that the data could be interpreted. More rigorous numerical analysis is required to extract quantitative data (Rice and others, 1995; Mace and others, 1997). GIS can be used to map plume evolution through time by posting the contaminant concentrations for each time slice. A combination of temporal slices of contaminant concentration co-kriged with moving average concentrations might be used to remedy irregularities in the data. An example of a common artifact is change in plume size when additional monitoring wells are drilled. The spatial analytic components of ESRI software could then be used to quantify changes in plume areas and contaminant concentrations through time.

The data base allows display of time series of contaminant data (benzene, toluene, ethyl benzene, xylene, TPH, and MTBE) during plume evolution and cleanup efforts (figs. 24, 25). These spatially distributed concentration data could be used for various calculations such as aquifer-specific half-lives for natural attenuation of each contaminant.

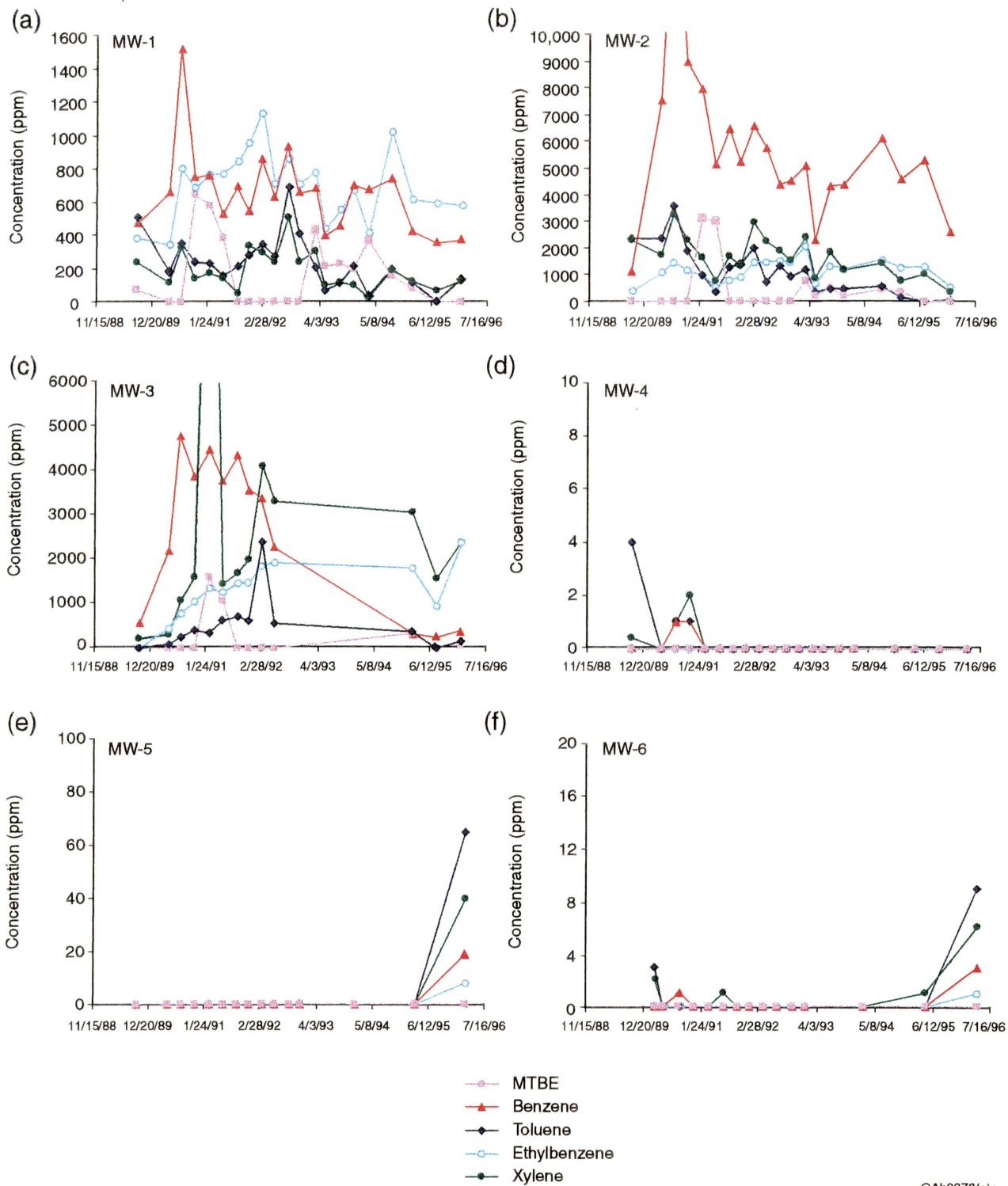
### Potential Receptors

One key element in assessing health risks is the spatial relationships between the contaminated sites and potential receptors. For ground-water contamination, this is translated into number of wells down-gradient from the contaminated site. For this project, a set of water-supply wells for which digital locations were available from TNRCC Water Utilities Division were selected to demonstrate the techniques. Spatial relationships between LPST sites and water wells were quantified using the Spatial Analyst extension of ArcView. A grid of distances from LPST sites



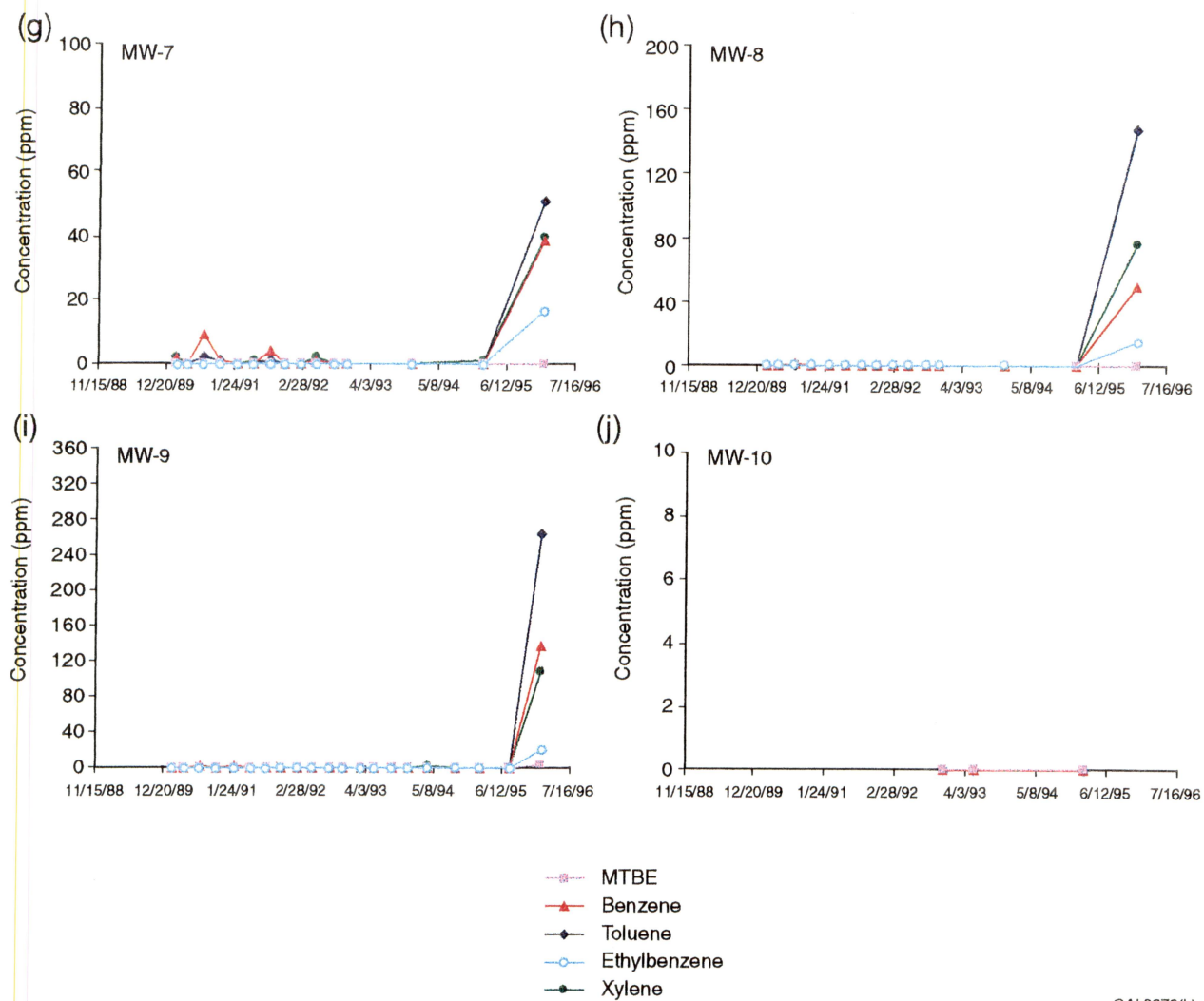
QAb9277c

Figure 24. Time series of contaminant concentrations at LPST site 91944 through its history.



QAb9278(a)c

Figure 25. Time series of contaminant concentrations at LPST site 93526 through its history.



QA b9278(b) c

Figure 25. (cont.)

was generated using the Distance From option in Spatial Analyst (fig. 26). By contouring and creating polygons from this grid, it was possible to extract subsets of water wells that fell within specified distances of LPST sites (fig. 27). The number of water wells within a specified radius of LPST sites can also be determined using the ArcView Select by Theme option, which will automatically intersect two or more themes to generate a desired subset.

In other geographic areas where alluvial settings are urbanized and surface drainage is more abundant, surface-water contamination is a common concern. GIS is an appropriate tool for quantifying relationships between sites, surface water, and surface-water discharge points. This use is not demonstrated in the Lubbock pilot area because surface drainage is poorly integrated; however, the same approaches can be used as for the water-well receptors.

## COSTS

As an aid to estimating the cost of creating an HGIS similar to this pilot in other areas, the labor used for each task of the pilot project is listed in table 8. This evaluation does not include time investment of the pilot project for literature and Internet research of the area; GIS design, testing, and troubleshooting; data review; example analysis; and report preparation specific to the pilot project. In a follow-up production project, time in the development phase would be saved and more time would be expended to extract results. In a production situation, we estimate data entry would use 75 percent of the time, location calculation about 4 percent, and building the GIS data base the remainder. However, any new analyses in GIS can be labor intensive, as it was during this pilot project, where about a third of the time was spent on GIS.

GPS data costs calculated for the El Paso area (Angle and others, 1996) provide guidance of the possible costs of this technology. In this study, 1,116 locations collected in 100 days cost \$76 each, including labor, field expenses, equipment, and subscription to GPS digital correction support services.

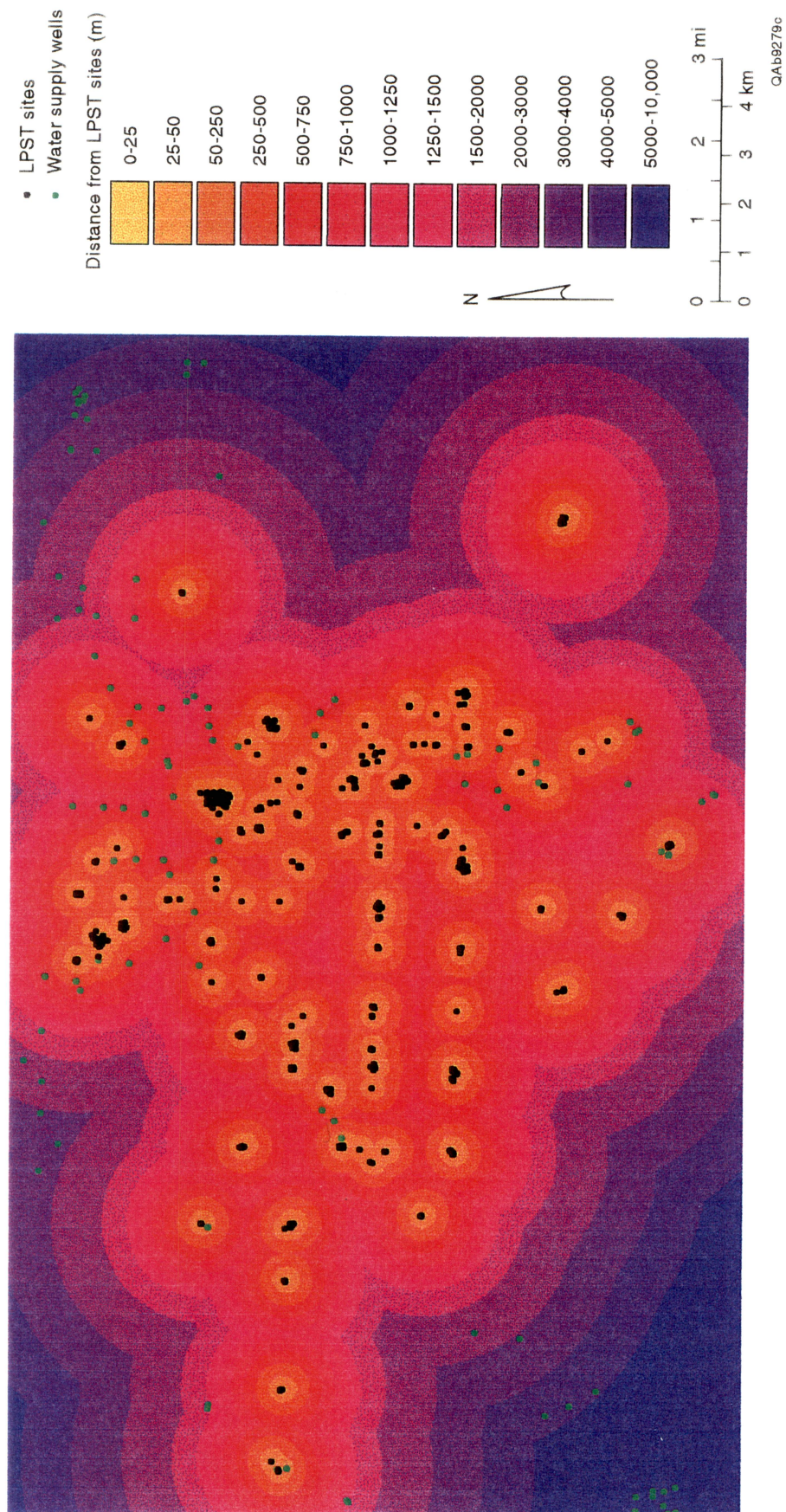


Figure 26. ArcView grid used to calculate distances from LPST sites inventoried in this pilot project to water wells. Water-well locations were downloaded from TNRRCC web page.

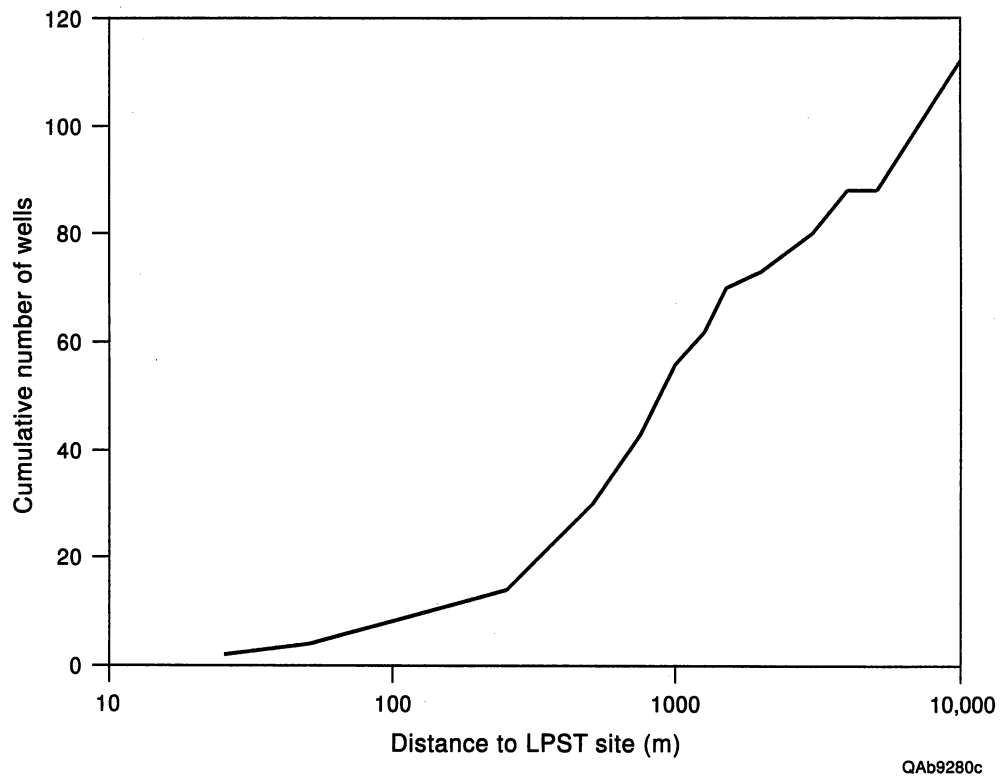


Figure 27. Cumulative number of water wells within measured distance of LPST sites.

Table 8. Labor costs for each task of pilot project.

Task	Number of units	Average manhours per unit
Extract data from TNRCC LPST files	217	1.2
Enter data from TNRCC files	217	1.6
Download and prepare GIS coverages from preexisting digital data sets	11	7.3
Digitize existing paper maps for study area	3	13.3
Create and check x-y locations for sites	207	0.6
Create metadata	5	6.7
Import spreadsheet data to GIS	7	1.4
GPS	35	0.6

## EVALUATION AND RECOMMENDATIONS

### Spatial Data

One key problem in creating a GIS data base is generating precise and accurate geographic coordinates for spatial features, in this case monitoring-well and soil-boring locations, and integrating them with other spatial data, such as mapped aquifer and soil parameters.

Surveying monitoring-well locations in the field using GPS technology produces the highest-quality location data. However, its usefulness is limited because of the relatively high cost of field work and because many monitoring wells are inaccessible because of site closure. Precise GPS field locations allow additional surveying during return visits.

The GPS survey determined that, in the pilot study, registration of the digital base map was the most significant source of error. Purchase of a better-registered street map from a vendor of digital data could reduce this error and is recommended for further GIS studies.

Geocode locations for wells were within 450 ft of the located monitoring well, with a average error of 200 ft. This inaccuracy reflects possible errors in interpolation of street addresses in industrial areas and random location of the monitoring wells with respect to the street address. Geocode site locations combined with scanned and scaled monitoring-well locations could speed generation of spatial well locations by about a third. However, Geocode locations might create undesirable errors in monitoring-well locations at closely spaced adjacent sites. Geocode locations are therefore recommended for rapid overviews at a multisite scale but not for analysis at a site-specific scale (plume geometry, for example).

The impact of map registration error on the analyses shown here is minor because spatial relationships among wells on sites, as well as between different sites, are maintained. Unquantifiable errors are probably present in other regional or countywide spatial data to which the monitoring-well locations are compared because of base-map quality during map compilation, interpolation between data, and errors introduced during digitizing. Therefore, at a very fine

(within-site) scale, spatial interpretation of most superposed map data is inappropriate, even with best-quality locations. The accuracy of the data created for this study is adequate to (1) classify sites according to attribute (such as soil type) extracted from superposed maps, (2) develop statistical descriptions for groups of related sites, (3) integrate information on plume geometry and hydraulic-head gradient between several sites, and (4) examine plume evolution and hydraulic-head gradient at a site scale.

### Data Entry

Data entry consumes a significant amount of time, more than one-third of the effort in this study. The data extraction from plan A was routine, and extraction of data from formatted digital files would be feasible. GIS can effectively use machine contouring or statistical evaluation to rapidly screen data for errors as well as significance.

### Impact of Heterogeneity on RBCA Assessments

GIS analysis offers potential for improved performance of assessment and risk-reduction strategies while containing costs. Heterogeneous natural environments are prohibitively costly to assess accurately because many boreholes, samples, and wells are required to adequately sample the materials and fluids present. GIS offers the potential of grouping sites according to geologic and hydrologic setting to place the results of limited sampling at a site in the context of previous experience with other sites in the setting. We envision a process where GIS is used to extract data and classify sites according to soil group, aquifer, and aquifer geology. Statistical characteristics of previously analyzed geotechnical soil parameters from the same soil group could be used to generate default input parameters in risk-assessment equations. A more conservative approach would be to plot a limited number of site-specific measurements against the typical distribution to check the validity of the assumptions. Similarly, aquifer properties such as hydraulic conductivity

that are collected only at a small percentage of sites could be derived from a population of these measurements compiled from tests in the same aquifer or in the same geologic facies of the aquifer.

The Lubbock pilot area was not ideal for demonstrating uses of site-specific data in a geologic and hydrologic context to better quantify heterogeneity, because most of the LPST sites in the county are located in one hydrogeologic setting. However, we used this data set to demonstrate the procedure. Further analysis of data from other areas is needed to determine the real impact of geologic and hydrologic heterogeneity on risk assessment.

In this study, the available data were used to generate distributions of soil and aquifer properties. Site-specific soil and aquifer measurements are more accurately defined by distributions rather than single values, owing to the natural heterogeneity at the site. Monte Carlo sampling was used to quantify the effects of heterogeneity in the soil and aquifer and to calculate values for target concentrations and ground-water velocities. These values can differ appreciably from default values. We recommend using an empirical cumulative frequency distribution to (1) determine whether site-specific soil and aquifer measurements are representative of the probable range of variability at the site, (2) determine whether site-specific soil and aquifer measurements will therefore result in meaningful risk calculations, and (3) fine-tune default values of target concentration levels for specific formations.

### Plume Analysis

GIS offers a powerful tool to assess plume size, location, and evolution. This is potentially most needed in areas of complex hydrology and for multisource plumes. A rigorous analysis of the plume data is beyond the scope of this project. Ideally, plume interpretation should (1) identify the average local hydraulic-head gradient and the amount of and variation in hydraulic-head gradient, (2) use rigorous numerical analysis to fit the plume to the measured values (Rice and others, 1995; Mace and others, 1997), and (3) track systematic evolution through time. For example, changes in interpreted plume size due to contaminant concentration measurement at additional monitoring wells should be corrected. Additional experimentation is required to determine whether contouring

algorithms within ESRI Spatial Analyst software can be customized to manage the highly irregular distribution of values and apply appropriate contouring algorithms. If ESRI software is not suitable, a method can be developed where the data base could be created in Arc software, then exported to create contoured surfaces suitable for creating Arc coverages or grids.

## CONCLUSIONS

The results of this study are

- creation of a test-case HGIS data base composed of the coverages and data tables listed in table 2 and containing the data listed in table 3,
- metadata and quality-control information, and
- several experiments in the application of the HGIS to site-characterization and risk-assessment problems.

The GIS format facilitates retrieval and analysis of data, including spatial information. Data tables facilitate statistical analysis of the data. The HGIS provides a tool to integrate experience and results of LPST-site remediation from multiple sites and place it in a regional hydrogeologic context.

## ACKNOWLEDGMENTS

This research was funded by Texas Natural Resource Conservation Commission (TNRCC) Contract 72-000000-38. We acknowledge the contribution of TNRCC staff from a variety of divisions, as well as staff from other State agencies, for making information available to us. Alan Dutton was program director for this project. We thank our BEG colleagues Tom Tremblay for advice in use of GIS, John Gamble for collecting the GPS data, and Ted Angle for overseeing GPS collection and downloading to GIS. Susan Palachek, an Austin AISD Anderson High School intern, assisted with interpretation of soil-boring logs and background data.

## REFERENCES

- Angle, E. S., White, Gerry, Salazar, Deborah, and Raney, J. A., 1996, GPS/GIS for regulated facilities in the EPA Region 6 international border area: The University of Texas at Austin, Bureau of Economic Geology final contract report, 16 p.
- Bonham-Carter, Graeme, 1994, Geographic information systems for geoscientists: modeling with GIS: Pergamon Press, Toronto, 398 p.
- Caran, S. C., 1991, Cenozoic stratigraphy, Southern Great Plains and adjacent areas, *in* Morrison, R. B., ed., Quaternary nonglacial geology, conterminous U.S.: Geological Society of America, The Geology of North America, v. K-2, plate 5.
- Chen, Yu-Chi, Rainwater, K. A., Ramsey, R. H., and Dvoracek, M. J., 1988, Economic potential for development of increasing groundwater storage beneath a High Plains municipality: Water Resources Center, Texas Tech University, 99 p.
- Eifler, G. K. Jr., Frye, J. C., Leonard, A. B., Hentz, T. F., and Barnes, V. E., 1967 (revised 1993), Lubbock Sheet: The University of Texas at Austin, Bureau of Economic Geology, Geologic Atlas of Texas, 1:250,000.
- Foster, J. A., 1952, The Pliocene series of southeast Lubbock County, Texas: Texas Technical College, Master's thesis, 133 p.
- Holliday, V. T. 1989, The Blackwater Draw Formation (Quaternary): a 1.4-plus- m.y. record of eolian sedimentation and soil formation on the Southern High Plains: Geological Society of America Bulletin, v. 101, p. 1598–1607.
- Hovorka, S. D., 1995, Quaternary stratigraphy of playa lakes on the Southern High Plains — a case study from the Amarillo area, Texas: The University of Texas at Austin, Bureau of Economic Geology Report of Investigations No. 236, 52 p.

- Knowles, T. R., Nordstrom, P., and Klemt, W. B., 1984, Evaluating the ground-water resources of the High Plains of Texas: Texas Department of Water Resources Report 288, v. 1, 119 p.
- Mace, R. E., Fisher, R. S., Welch, D. M., and Parra, S. P., 1997, Extent, mass, and duration of hydrocarbon plumes from leaking petroleum storage tank sites in Texas: The University of Texas at Austin, Bureau of Economic Geology Geological Circular, 97-1, 52 p.
- Nativ, Ronit, 1988, Hydrogeology and hydrochemistry of the Ogallala aquifer, Southern High Plains, Texas Panhandle and eastern New Mexico: The University of Texas at Austin, Bureau of Economic Geology Report of Investigations No. 177, 64 p.
- Rainwater, Ken, and Thompson, D. B., 1994, Playa lake influence on ground-water mounding in Lubbock, Texas: *in* Urban, L. V., and Wyatt, A. W., Proceedings of the Playa Basin Symposium, Texas Tech University, p. 113–118.
- Reeves, C. C., 1976, Quaternary stratigraphy and geologic history of the Southern High Plains, Texas and New Mexico, *in* Mahaney, W. C., ed., Quaternary stratigraphy of North America: Stroudsburg, Pennsylvania, Dowden, Hutchinson and Ross, p. 213–234.
- Rice, D. W., Grose, R. D., Michaelson, J. C., Dooher, B. P., MacQueen, D. H., Cullen, S. J., Kastenbergh, W. E., Everett, L. G., and Marino, M. A., 1995, California leaking underground fuel tank (LUFT) historical case analyses: Environmental Protection Department, Environmental Restoration Division, Lawrence Berkeley Laboratories, UCRL-AR-122207, variously paginated.
- Seni, S. J., 1980, Sand-body geometry and depositional systems, Ogallala Formation, Texas: The University of Texas at Austin, Bureau of Economic Geology Report of Investigations No. 105, 36 p.

TNRCC, 1994, Risk-based corrective action for leaking storage tank sites: Texas Natural Resource Conservation Commission, RG-36, 78 p.

TNRCC, 1995, Petroleum storage tank division assessment report form: Texas Natural Resource Conservation Commission, TNRCC-0562, 26 p.

TNRCC, 1996, Joint groundwater monitoring and contamination report 1995: Texas Natural Resource Conservation Commission, June 1996, SFR-36, variously paginated.

Wyatt, A. W., Smith, D. D., and McReynolds, Don, 1992, Hydrogeologic atlas for Lubbock County, Texas; High Plains Underground Water Conservation District No. 1, Lubbock, Texas, 8 p., 4 maps.