

United States  
Environmental Protection  
Agency

Office of Air Quality  
Planning and Standards  
Research Triangle Park, NC 27711

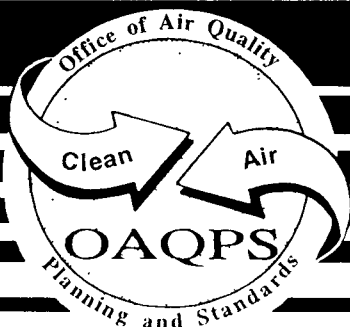
EPA-454/R-99-021 ✓  
July 1999

Air



# DISPERSION MODELING OF TOXIC POLLUTANTS IN URBAN AREAS

## GUIDANCE, METHODOLOGY AND APPLICATIONS



**AIR DISPERSION MODELING OF TOXIC POLLUTANTS  
IN URBAN AREAS**

**GUIDANCE, METHODOLOGY AND EXAMPLE APPLICATIONS**

Emissions, Monitoring and Analysis Division (MD-14)  
Office of Air Quality Planning and Standards  
U S. Environmental Protection Agency  
Research Triangle Park, NC 27711

July 1999

**U.S. Environmental Protection Agency  
Region 5, Library (PL-12J)  
77 West Jackson Boulevard, 12th Floor  
Chicago, IL 60604-3590**

## NOTICE

This report has been reviewed by the office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, and has been approved for publication. Any mention of trade names or commercial products is not intended to constitute endorsement or recommendation.

U.S. Environmental Protection Agency  
Office of Air Quality Planning and Standards  
Research Triangle Park, North Carolina 27711  
EPA-600/4-91-010

## TABLE OF CONTENTS

EXECUTIVE SUMMARY .....	vii
1. GUIDANCE FOR URBAN AIR TOXICS ANALYSES .....	1
1.1 INTRODUCTION .....	1
1.2 MODELING METHODOLOGY .....	2
1.2.1 Model Features .....	2
1.2.2 Model Options .....	2
1.2.3 Averaging Period .....	3
1.2.4 Receptors .....	3
1.2.5 Terrain .....	4
1.2.6 Meteorological Data .....	4
1.2.7 Chemistry .....	5
1.2.8 Background Concentrations .....	6
1.2.9 Model Evaluation .....	6
1.2.10 Study Limitations .....	7
1.3 SOURCE DEFINITIONS .....	8
1.3.1 Modeling Domain .....	8
1.3.2 Emission Inventory Definitions .....	8
1.3.3 Source Characterization for ISCST3 .....	9
1.3.4 Spatial and Temporal Distribution and Characterization of Area and Mobile Source Emissions .....	11
1.3.5 Default Source Parameters .....	12
1.3.6 Source Parameters for Deposition Calculations .....	14
1.3.7 Pollutants .....	14
1.3.8 Source Grouping .....	14
1.3.9 Quality Assurance .....	14
1.4 MODEL OUTPUT FOR ANALYSIS .....	15
2. CASE STUDY MODELING METHODOLOGY .....	16
2.1 INTRODUCTION .....	16
2.2 MODELING METHODOLOGY .....	17
2.2.1 Model Selection .....	17
2.2.2 Modeling Options .....	18
2.2.3 Receptor Locations .....	19
2.2.3.1 <i>Receptor Sampling Strategy</i> .....	19
2.2.3.2 <i>Treatment of Terrain Influences</i> .....	19
2.2.4 Meteorological Data .....	20
2.2.4.1 <i>Selection of Surface and Upper Air Stations</i> .....	20
2.2.4.2 <i>Meteorological Parameters for Deposition Calculations</i> .....	20
2.2.4.3 <i>Meteorological Preprocessing</i> .....	21



2.2	MODELING METHODOLOGY (continued)	
2.2.5	Emission Characteristics	22
2.2.5.1	<i>Determination of Background Concentrations</i>	22
2.2.5.2	<i>Default Source Parameter Values</i>	22
2.2.5.3	<i>Area and Mobile Source Modeling</i>	23
2.2.5.4	<i>Source Parameters for Dry Deposition Calculations</i>	23
2.2.5.5	<i>Source Grouping</i>	25
2.2.6	Model Evaluation	26
2.3	OVERVIEW OF EMISSION INVENTORIES	27
2.3.1	Phoenix, Arizona Inventory	27
2.3.2	Houston, Texas Inventory	28
2.4	APPROACH USED TO ESTIMATE ATMOSPHERIC SECONDARY FORMALDEHYDE PRODUCTION	31
2.4.1	Introduction	31
2.4.2	Simulation Specific Input	32
2.4.3	Analysis of Results for Use with ISCST3	33
2.4.4	Results	35
2.4.5	Conclusions	35
2.5	OVERVIEW OF MODELING RESULTS	40
2.5.1	Phoenix, Arizona Modeling Results	40
2.5.2	Houston, Texas Modeling Results	40
2.6	SUMMARY AND CONCLUSIONS	42
3.	REFERENCES	43
APPENDIX A	CASE STUDY FOR PHOENIX, ARIZONA	A-1
APPENDIX B	CASE STUDY FOR HOUSTON, TEXAS	B-1
APPENDIX C	PROPOSED METHODS FOR SELECTING RECEPTOR SAMPLES FOR THE APPLICATION OF THE ISCST3 DISPERSION MODEL TO URBAN AREAS	C-1
APPENDIX D	DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA	D-1
APPENDIX E	PARAMETERS RELATING TO THE FATES OF SELECT ATMOSPHERIC POLLUTANTS	E-1
APPENDIX F	ADDITIONAL INFORMATION ON HOW TO USE OZIPR	F-1

## LIST OF TABLES

Table 1.2-1 Sources of Meteorological Data .....	4
Table 2.3-1 Toxic Air Pollutant Emissions for Phoenix, Arizona, Based on Year 1993 .....	30
Table 2.3-2 Toxic Air Pollutant Emissions for Houston, Texas, Based on Year 1993 .....	30
Table 2.5-1 Highest Annual Average Concentrations from All Sources Combined for Phoenix, Arizona, Based on 5 Modeled Years 1987 - 1991 .....	41
Table 2.5-2 Highest Annual Average Concentrations from All Sources Combined for Houston, Texas, Based on 5 Modeled Years 1987 - 1991 .....	41

## LIST OF FIGURES

Figure 2.4-1: Formaldehyde concentrations for prototypical summer day in Houston Texas, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT_FORM (total formaldehyde). . . . .	36
Figure 2.4-2: Formaldehyde concentrations for prototypical autumn day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT_FORM (total formaldehyde). . . . .	37
Figure 2.4-3: Formaldehyde concentrations for prototypical winter day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT_FORM (total formaldehyde). . . . .	38
Figure 2.4-4: Formaldehyde concentrations for prototypical spring day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT_FORM (total formaldehyde). . . . .	39

## EXECUTIVE SUMMARY

The 1990 Clean Air Act Amendments (CAAA) Section 112(k) requires EPA to reduce urban air toxics. EPA is devising a broad strategy for reducing risks posed by air toxics from all sources in urban areas, which is discussed in the Draft Integrated Urban Air Toxics Strategy (U.S. EPA, 1999b). In order to fully understand the air toxics problem in urban areas, it is necessary to know the concentrations of air toxics to which people are exposed; however, air monitoring data are scarce and limited. Another means for understanding the air toxics problem is to estimate hazardous air pollutant (HAP) air concentrations through the use of dispersion models, relying on emissions measurements or estimates. Because urban areas can vary greatly in terms of air toxics, sources, meteorology, and the legal enforcement options provided by State and local programs to address air toxics, State and local measures will be needed to reduce urban air toxics risks. This document was prepared to provide:

- guidance for modeling urban area impacts of air toxics
- a demonstration of a methodology for modeling air toxics for use in city-specific analyses
- two example applications of city-specific air toxics modeling applications.

The first two sections of the document present generalized guidance and the overview of two modeling applications. Section 1 presents guidance for conducting air toxics dispersion modeling for an urban area. Section 2 presents highlights from two case studies of applying the modeling guidance in urban areas: Phoenix, Arizona and Houston, Texas. In these examples the impacts of five potential cancer causing air toxics were examined: benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium (chromium VI). Appendix A and B present more detail on the Phoenix, Arizona and Houston, Texas case studies, respectively. These appendices also include detailed documentation of the emission inventories used in the modeling analyses, as well as, detailed summaries of the results of the modeling analyses. The approach for preparation of a mobile source toxic emission inventory used for Houston is currently being updated from the approach used here. Therefore, persons conducting subsequent analyses should contact OAQPS for the current emission inventory preparation guidance.

The results of the modeling analyses show some significant differences between the two cities that were studied. For Phoenix, the mobile sources were clearly the dominant source of emissions for four of the five pollutants, the exception being hexavalent chromium, which occurred mostly from cooling towers and other major sources. The mobile source emissions also exhibited the strongest temporal variations, reflecting the diurnal patterns in road traffic, as well as, some influences of meteorology on emission estimates. These patterns in the emission inventory for Phoenix are also evident in the modeling results for that city. Since the majority of emissions were from mobile sources, they were distributed over the entire domain, with some spatial variability evident based on surrogate factors such as population. As a result, there was little evidence in the modeling results of localized "hot spots" (sharp gradients in concentrations over a relatively short distance) within the Phoenix domain.

While the Houston inventory showed emissions from mobile sources that were comparable in magnitude to Phoenix, major source (also called point source) emissions from Houston were significantly higher than major source emissions for Phoenix. Benzene emissions from major sources were almost 100 times higher for Houston than for Phoenix, and

1,3-Butadiene emissions were about 700 times higher. As a result of this, the modeling analysis for Houston does exhibit some significant "hot spots" associated with some of the larger sources of emissions for certain pollutants.

These applications illustrate a methodology that may be applied to similar urban-wide analyses of point and area sources of air toxics. The use of a plume model (a modified version of the Industrial Source Complex Short Term 3 - ISCST3) has certain advantages over puff and grid models in terms of less stringent input data requirements, and plume models require less computational resources for long term exposure analyses. Further insight into the applicability of such models on this scale of analysis may be gained in future studies by comparing modeled concentrations to monitored concentrations, a task begun in the study for Houston.

# **1. GUIDANCE FOR URBAN AIR TOXICS ANALYSES**

## **1.1 INTRODUCTION**

The 1990 Clean Air Act Amendments (CAAA) Section 112(k) requires EPA to reduce urban air toxics. EPA is devising a broad strategy for reducing risks posed by air toxics from all sources in urban areas, which is discussed in the Draft Integrated Urban Air Toxics Strategy (U.S. EPA, 1999b). In order to fully understand the air toxics problem in urban areas, it is necessary to know the concentrations of air toxics to which people are exposed; however, air monitoring data are scarce and limited. Another means for understanding the air toxics problem is to estimate hazardous air pollutant (HAP) air concentrations through the use of dispersion models, relying on emissions measurements or estimates. Because urban areas can vary greatly in terms of air toxics, sources, meteorology, and the legal enforcement options provided by State and local programs to address air toxics, State and local measures will be needed to reduce urban air toxics risks. This document was prepared to provide:

- guidance for modeling urban area impacts of air toxics
- a demonstration of a methodology for modeling air toxics for use in city-specific analyses
- two example applications of city-specific air toxics modeling applications.

Section 1 provides guidance and recommendations on specific issues for urban-wide analyses of air toxics. Urban areas contain major sources and numerous smaller, area sources. As a result modeling analyses for large numbers of air toxics sources poses special challenges. Since most modelers are more familiar with modeling applications for a single facility, this section should help provide guidance to transition from single facility applications to more complex urban-wide applications.

Section 2 provides an overview of two applications of the Industrial Source Complex Short Term 3 (ISCST3) model to urban-wide studies. ISCST3 was applied to the Phoenix, Arizona and Houston, Texas urban areas. Section 2 contains information that is general to both applications, while later sections present more detailed information and results of the analyses. Appendix A covers the Phoenix, Arizona study, while Appendix B covers the Houston, Texas study.

The guidance section begins with Section 1.2 covering the modeling methodology; Section 1.3 discusses the emissions inventory; Section 1.4 discusses modeling output/analysis.

## 1.2 MODELING METHODOLOGY

The recommended plume dispersion model for use in estimating urban-wide concentrations of toxic air pollutants is the ISCST3 model. Justification for selecting ISCST3 is provided in Section 2.2.1, where the needs for the case studies are discussed. This section describes some of the most important details needed to apply ISCST3 for these types of applications. Other sources of detailed guidance are listed below.

For general information on air quality modeling, consult Appendix W to CFR Part 51-Guideline on Air Quality Models.

For information on how to use the ISCST3 model, consult the ISC3 model user's guide to ISC Model (U.S. EPA, 1995c).

For information on how to preprocess the meteorological data for input in ISCST3, consult the PCRAMMET user's guide (U.S. EPA, 1996b) and the MPRM user's guide (U.S. EPA, 1996a).

All of the items listed above can be obtained from EPA's SCRAM web site at <http://www.epa.gov/scram001>.

Information on the "Integrated Urban Air Toxics Strategy" developed under the authority of Section 112(k) and 112(c) of the Clean Air Act is obtained from EPA's web site at <http://www.epa.gov/ttnuatw1>.

### 1.2.1 Model Features

Key features of the ISCST3 dispersion model that are useful for urban-wide air toxics applications include:

- handles multiple point, area, and mobile sources
- incorporates building downwash effects
- includes an urban dispersion option
- contains considerable flexibility for specifying receptor locations and for grouping of source impacts
- includes algorithms to treat the effects of elevated and/or complex terrain
- treats the effects of deposition of gaseous and particulate emissions
- includes an option to vary emissions by season and hour-of-day
- includes an option to treat atmospheric transformations by exponential decay.

### 1.2.2 Model Options

The regulatory default model option in ISCST3 should be selected for urban-wide applications. More information about the default option parameters can be found in the ISCST3 user's guide (U.S. EPA, 1995c).

The option to vary the emissions by season and hour-of-day should be selected, unless the objectives of the application or the form of the emissions inventory data dictate otherwise.

For best estimates, the use of the wet and dry deposition and plume depletion options should also be selected. Note that the selection of the plume depletion option will significantly increase model run time. To utilize the wet and dry deposition option, the model requires additional data for the meteorological and chemical parameters. For meteorological input, the user should consult the meteorological data preprocessors' user's guides (see above). Deposition estimates are very useful in multi-pathway exposure assessments. Information on chemical parameters may be obtained from the technical literature.

The urban modeling option should normally be selected. However, this guidance assumes that the sources are in an urban area and thus the urban option for ISC3 should be selected. To determine if the modeling domain is urban, apply the criteria in Section 8.2.8 of the Guideline on Air Quality Models. If the result of this analysis shows that the area to be modeled is rural, this guidance should not be used, and any toxics modeling should be done in close consultation with the EPA Regional Office.

The exponential decay option should be selected when half-life values are available for the air toxics under consideration.

### 1.2.3 Averaging Period

The ISCST3 model computes an hourly concentration for each receptor. Other averaging periods, e.g., 3-hour, daily, seasonal and annual can also be aggregated (U.S. EPA, 1995c). The averaging period selected is based on the intended use. For chronic (long-term) exposure studies, annual average air concentrations are generally needed. Some exposure studies require seasonal average air concentration estimates. Shorter term ambient concentrations are usually needed for determining acute exposure.

### 1.2.4 Receptors

A receptor is any location where ambient concentration estimates are calculated. Receptors are usually placed in "ambient air" which is outside of inaccessible plant property. The ISCST3 model requires the coordinates of the specified receptors. Receptor locations should be selected based on a case-by-case determination with expert judgement on the needs of the study. Often, receptors are selected at coordinates provided in the census data (census block, census block groups or census tracts). The procedure in Appendix C, Section C.7, Step-by-Step Guidelines for Using Sampling Method C, is recommended. Analysts should supplement Step 3 of those guidelines by selecting "automatic" (i.e., arbitrary) receptor points (termed "certainty points" in Appendix C) where, for example, there are likely to be very high emission gradients or other important source/receptor relationships that otherwise might lead to extreme concentration outliers affecting the variance of the sample mean. Method C is appropriate when there is no requirement to: (1) specifically estimate the maximum concentration, (2) estimate concentrations in ambient air exclusively, (3) estimating concentrations at locations over water, or (4) estimating concentrations on building roofs or within wakes or cavities.

In the two example studies below, receptor selections were based on the input needs of the Hazardous Air Pollutant Exposure Model (HAPEM4) (U.S. EPA, 1999a). The receptor points were defined as the centroids of census block groups (see Appendix C). Census data and



urban land use information can be used to identify locations (potential receptors) where individuals live, work, attend school, and spend time in recreation.

For many studies, the number of receptors selected will be very large, and since model run time is proportional to the number of receptors, unreasonable model run times can occur. For instance, the decision to place a receptor at each census block centroid will result in many more receptors than if they were placed at the centroid of the census block group. Following the guidelines in Appendix C should minimize such problems.

#### 1.2.5 Terrain

Terrain elevation at each source and receptor is required input for ISCST3. Digitized terrain elevation data are available from U.S. Geological Survey (USGS) maps. Source (stack) elevation is usually provided in the inventory. For many urban areas, terrain can be assumed to be flat and source and receptor elevations set to zero. Where the urban area is in mountainous terrain, terrain effects are important. First, the impact of individual plumes on elevated terrain results in higher air concentration (through placing the receptor at the correct higher air concentration, vertical location within the plume and estimating the impaction of the plume upon intervening terrain). Second, wind channeling due to terrain can cause higher air concentrations. The ISCST3 model does not address wind channeling effects other than if these effects are captured by the available meteorological data. If the urban area contains complex terrain features that are expected to significantly affect the modeled concentrations, a dispersion model that handles such situations should be selected from those listed in the Guideline on Air Quality Models (40CFR51).

#### 1.2.6 Meteorological Data

The ISCST3 model requires meteorological data consisting of hourly surface and upper air observations. Surface and upper air meteorological data files may be obtained from the National Climatic Data Center (NCDC) through their web site at <http://www.ncdc.noaa.gov>. Alternatively, several CD's are available which contain national and international meteorological data. These CD's are SAMSON, HUSWO, INSWO, and the Radiosonde Data of North America, Table 1.2-1 presents the years covered by each CD.

**Table 1.2-1 Sources of Meteorological Data**

Source - CD	Type	Years Covered	Reference
SAMSON	Surface	1961-1990	NCDC, 1993
HUSWO	Surface	1990-1995	NCDC, 1997a
INSWO	Surface	1982-1997	NCDC, 1998
Radiosonde Data of N.A.	Upper Air	1946-1996	NCDC, 1997b

If the INSWO CD is used and wet deposition estimates are required, an additional file is necessary. The TD-3240 precipitation data file can be obtained from NCDC through their web

site. A cross reference is also available from NCDC that is useful in matching TD-3240 station identifications to National Weather Service (NWS) station numbers.

In selecting surface and upper-air stations, consult with your State/Regional meteorologist for the most applicable stations for your area. The closest stations may not be the most representative due to the influences of terrain or water bodies.

Meteorological data must be preprocessed before use in ISCST3. Mixing heights can be computed using surface (SAMSON or HUSWO) and upper air data (Radiosonde Data of North America) via the mixing height program provided on the SCRAM web site. PCRAMMET and MPRM preprocessors use surface and mixing height data as input to create ISCST3 input meteorological files. PCRAMMET was developed for use with NWS data, while MPRM is used primarily for processing on-site meteorological data. In urban areas, on-site meteorological data are not often available.

The meteorological data preprocessor MPRM should be used to prepare the input files necessary for applying the gas dry deposition algorithm in ISCST3. Values for additional parameters needed in applying the gas dry deposition algorithms for the two case study cities are presented in Sections 2.2.4.2 and 2.2.5.4. MPRM can also be used for setting up a meteorological data file for ISCST3 to use in estimating both particle dry deposition, and gas and particle wet deposition. PCRAMMET does not contain the algorithms for setting up a file to support gas dry deposition, although it does prepare a meteorological data file for use in estimating particle dry and wet deposition and gas wet deposition. Finally, PCRAMMET can accept data directly from the SAMSON and HUSWO CD's, while MPRM has not been updated to read the HUSWO data. Neither preprocessor has yet been updated to read in the INSWO data.

The MPRM and PCRAMMET meteorological data preprocessors can occasionally produce very low mixing heights (less than 10 meters) based on the twice-daily values from the mixing height data file and the interpolation scheme used to provide hourly values of mixing height. Anomalously low mixing heights in the NCDC data files may be associated with a mid-day cold frontal passage. While the occurrence of very low mixing heights is more likely for the rural mixing heights than for the urban mixing heights, due to differences in the interpolation routines, low mixing heights may occur for both rural and urban conditions. The application of a very low mixing height with a near-surface level area source can produce anomalously large air concentrations due to the treatment of limited mixing effects in the ISCST3 model; expert judgement is needed to determine the minimum mixing height for a given urban area. For the two example studies, a minimum value of 100 meters was applied to the hourly mixing heights produced by MPRM to avoid this anomaly from influencing the results. For urban areas, building heights will limit the lower mixing heights and the 100-meter value was considered the upper limit to the minimum value for the depth of the well-mixed boundary layer in a large urban area (Sutton, 1953).

### 1.2.7 Chemistry

The ISCST3 model provides concentration estimates due to primary emissions and has a limited capability to consider atmospheric transformations by exponential decay (half-life). Some pollutants (e.g., formaldehyde, acetaldehyde, and acrolein) are also formed in the atmosphere due

to reactions among other pollutants (i.e., formed by secondary reactions). Thus, in addition to estimating concentrations due to primary emissions, an estimate of concentrations based on secondary reactions is usually needed and should be added to the ISCST3 output. EPA's OZIPR screening model (Gery and Crous, 1991) may be used to estimate the secondary transformation of pollutants. Appendix F contains additional detailed guidance on the use of OZIPR, and Section 2.4 contains details about the application of OZIPR to estimate secondary formaldehyde formation in Houston, TX.

#### 1.2.8 Background Concentrations

Background air quality includes pollutant concentrations due to natural sources, nearby sources other than those under consideration, and unidentified sources. For typical exposure assessments, background concentrations should be added to the modeled concentrations to provide total ambient air concentrations for estimating exposure. Air quality data from a HAP monitoring network in the vicinity of the analysis area are often used to establish background concentrations. Also, background concentrations of some air toxics may be found in the literature.

In the absence of measured or other reported values, the following approach for estimating background concentrations should be used. An expanded point source inventory should be obtained for an area surrounding each city from the National Toxic Inventory (NTI). The domain for this expanded point source inventory should extend approximately 50 km beyond the domain of the inventory being explicitly modeled in the analysis. The NTI point sources should be modeled to estimate background concentrations within the modeling domain as a function of wind direction. The modeled background concentration should be based on an average concentration computed from a coarse grid, about every 5 kilometers, across the modeling domain. These direction-specific background concentrations should then be added to the modeled concentrations by a post-processor that also reads the meteorological data to obtain the appropriate wind direction.

#### 1.2.9 Model Evaluation

Monitoring data can be used to check the validity of the modeled concentration estimates or determine background concentrations. Ideally, the monitoring and modeling data should span the same time period. Air toxics monitoring data are available from EPA's Aerometric Information Retrieval System (AIRS) web site at <http://www.epa.gov/airs>. In most instances, ambient data are collected at a frequency of one in six days. A variety of statistical tests can be used to compare modeled with observed estimates. How the model estimates compare to annual average monitored data is useful for determining the suitability for inputting the estimates into HAPEM4 (U.S. EPA, 1999a). For comparisons in urban areas, there are many uncertainties in all facets of the comparison effort. Thus a factor of two agreement between modeled and observed values is considered to be very good. Appendix B contains a discussion of the comparison of modeled and monitored air concentrations for Houston, TX.

#### 1.2.10 Study Limitations

As part of the conclusions in an urban-wide air toxics modeling study report, the limitations of the modeling effort should be clearly stated. The important limitations of the ISCST3 model are provided in the User's Guide (U.S. EPA, 1995c). Limitations due to data availability and other factors should also be described.

## 1.3 SOURCE DEFINITIONS

### 1.3.1 Modeling Domain

The urban area domain should be selected based on case-by-case determination with expert judgement. The urban area domain can include a city center or multiple counties. It should be carefully defined because the larger the modeling domain, the greater the number of sources and receptors to be considered and thus the greater the required computational resources. Guidance in the draft Integrated Urban Air Toxics Strategy (U.S. EPA, 1999a) should be consulted.

### 1.3.2 Emission Inventory Definitions

In many applications the initial step of the urban-wide air modeling process will be the assembly of the emissions inventory. Depending upon the end use of the analysis, this inventory may include:

- point sources - releases that can be attributed to individual stacks or release points,
- area sources - releases that cannot be attributed to individual stacks or release points, and
- mobile sources - releases attributed to engine emissions, both on and off-road.

Information about the most recent National Toxics Inventory (NTI) and documentation are available from EPA's web site at [ftp://www.epa.gov/pub/EmisInventory/nti\\_96](ftp://www.epa.gov/pub/EmisInventory/nti_96).

#### Major/Area Source Emissions

According to Title I, Section 112(a) of the CAA, a "major source" is any stationary source (including all emission points and units located within a contiguous area and under common control) of air pollution that has the potential to emit, considering controls, 10 tons or more per year of any HAP or 25 tons or more per year of any combination of HAPs. An "area source" is any stationary source of HAPs that does not qualify as a major source. Area sources are also defined as emission sources that are too small and numerous to inventory individually. Area and mobile source emissions are not attributed to a specific location. Instead, they are calculated as county-wide aggregated emissions. For example, all dry cleaners' emissions are summed to a single number that aggregates the emissions of all facilities in a county, instead of by individual dry cleaner facilities.

#### Major (Point) Source Emissions

For the purposes of the inventory used for air toxics analysis, major sources should be considered as point sources, meaning sources for which a location is known. This clarification is meant to distinguish point sources from area or mobile sources.

#### Mobile Source Emissions

Typically, mobile source emissions are split into on-road and off-road components. On-road mobile sources are those vehicles certified for highway use and to applicable emission

standards. They include cars, trucks (light-duty, such as pick-ups, sports utility vehicles, minivans; and heavy-duty, typically in the form of a semi tractor-trailer rig), buses, and motorcycles. They may be fueled using gasoline, highway diesel fuel, or alternative fuel (e.g., CNG, LPG, electricity). They do not include off-road equipment that is occasionally on highways in order to move to the work location, such as most types of construction equipment and agricultural equipment.

Off-road is a term that covers a diverse collection of engines, equipment, and vehicles within the mobile source realm. Also referred to as "non-road" or "off-highway," the off-road category includes recreational equipment, airport service equipment, industrial/commercial equipment, agricultural equipment, construction equipment, oil and mining equipment, lawn and garden equipment, logging equipment, and recreational and commercial marine vessels. Though dealt with separately in the Clean Air Act, locomotives and aircraft can also be considered categories of off-road engines.

### 1.3.3 Source Characterization for ISCST3

Generating the source inventory for modeling is intertwined with the creation of the pollutant inventory. Each emissions source and the constituents each source emits must be specifically identified. For ISCST3 dispersion model, each source will need to be classified as a point, area, volume, or line source. Building the source inventory usually begins with mapping the locations of emission sources, receptors and the study domain.

The ISCST3 model can accommodate a large number of sources and receptors, however, an optimum configuration is needed in order to minimize computer resources. Because source inputs vary with the type of source modeled, an important first step in creating the inventory is to identify each source of emissions as a point, area, volume, or line source. With the source types established, the appropriate model inputs can be determined. The following subsections describe the various source types and associated inputs for modeling.

#### Point Source Characterization

Point sources involve the release of emissions from a well-defined stack or vent, at a known temperature and flow rate. Consequently, characterizing point sources for modeling is fairly straightforward. The basic model inputs for any point source are: stack height above ground level; inside diameter at stack exit; gas velocity or flow rate at stack exit; gas temperature at stack exit; building dimensions, and emission rate. The location of the source will also need to be defined in terms of the model receptor grid used.

#### Area Source Characterization

Area sources are sources of air toxic pollutants that are emitted at or near ground level (e.g., landfills, waste lagoons, evaporation and settling ponds, etc.). The sizes of these sources can range from a few square meters in the case of settling ponds, to a few square kilometers or larger in the case of landfills. Emissions from area sources are assumed to be of neutral buoyancy. Therefore, plume phenomena such as downwash and impaction on elevated terrain features are not considered relevant for modeling area sources. The emission rate for area

sources is in units of mass per unit time per unit area [e.g., g/(s-meters squared[m<sup>2</sup>))]. It is an emission flux rather than an emission rate. As an example, assume the pollutant emission rate from a small lagoon is 150 g/s. The dimensions of the lagoon are 10 m by 20 m (total area is 200 m<sup>2</sup>). If this source were modeled as a single, square area source, then the modeled emission flux would be 0.75 g/s-m<sup>2</sup> (150 g/s ÷ 200 m<sup>2</sup>).

For dispersion modeling, the important parameters used to characterize area sources are location, geometry, and relative height. If the area source is not at ground level, a height for the source may be entered (for example, a non-zero value would typically be entered for the height of a land fill). If the release height of the source is greater than approximately 10 m, it should probably be modeled as a volume source.

The 1996 emissions inventory provides a county-wide emissions rate for area sources. Thus, the actual location of area sources is not available. In this situation, county-wide emissions must be distributed to locations in the county (see Section 1.3.4).

#### Volume Source Characterization

There are two basic types of volume sources: surface-based or ground-level sources that may also be modeled as area sources, and elevated sources. The effective emission height of a surface-based volume source, such as a surface rail line, is usually set equal to zero. An example of an elevated volume source is an elevated conveyor with an effective emission height set equal to the height of the conveyor. A source may be defined as a volume source for modeling when its emissions can be considered to occur over a certain area and within a certain depth of space. At refineries, fugitive exhaust from on-site structures such as tanks, or a treatment facility may be modeled as a volume source. A roadway over which contaminated soil is hauled may also be modeled as a series of volume sources. As with area sources, emissions from volume sources are assumed to be of neutral buoyancy.

The important parameters used to characterize volume sources for dispersion modeling are location and initial lateral and vertical dimensions. The particular model user's guide will have instructions on defining the initial lateral and vertical dimensions of the source. The length of the side of the volume source will need to be known, as will the vertical height of the source, and whether it is on or adjacent to a structure or building. Generally, the north-south and east-west dimensions of each volume source must be the same. For refined modeling, the location is simply expressed by a single east-west (X) and north-south (Y) coordinate.

Available emissions inventories, e.g., the 1996 NTI, do not contain sufficient information to use the volume source feature in the ISCST3 model.

#### Line Source Characterization

Line sources are typically used to represent roadways. Basic model inputs are the overall source length, width, and height. Emissions may be entered in units of grams per meter per second.

Line sources may also simply be modeled as a series of area or volume sources. In the case of a long and narrow line source, it may be impractical to divide the source into  $N$  volume sources, where  $N$  is given by the length of the line source divided by its width. Dividing the length of the line source by its width effectively splits the line source into a string of squares (for example, if the length of the line source was 100 m, and the width was 5 m, then the line source could be split into twenty, adjacent square volume sources). An approximate representation of the line source can be obtained by placing a smaller number of volume sources at equal intervals along the line source (for example, for the line source of length 100 m and width 5 m, a total of 10 square volume sources separated from one another by 5 m could be defined). With this option, the spacing between individual volume sources should not be greater than twice the width of the line source. A larger spacing can be used, however, if the ratio of the minimum source-receptor distance and the spacing between individual volume sources is greater than about 3.

Typically mobile sources are modeled as line sources. Using the county-wide mobile sources emissions from the 1996 NTI, mobile source emissions can be modeled as area sources by distributing them within the county (see Section 1.3.4). A better, but more resource intensive, approach is to allocate on-road mobile sources to actual roadways using Geographical Information System (GIS) software. These emissions can be modeled as line sources. Emissions from off-road mobile sources can be modeled as area sources.

#### 1.3.4 Spatial and Temporal Distribution and Characterization of Area and Mobile Source Emissions

Area and mobile source emissions are usually provided on a county-wide basis. These emissions must be allocated correctly to smaller areas contained in the modeling domain for use by ISCST3. Analysts should divide the urban area into two-by-two kilometer grids and apportion county wide emissions to these grids. Choose the allocation carefully since it will affect the accuracy/reliability of the concentration estimates. For example, assign emissions from large area sources such as landfills and airports to their actual locations in these grids rather than averaging them over the entire domain. Note that resources often preclude assigning emissions from sources with numerous locations, such as dry cleaners, gasoline stations, etc., to specific locations. For these sources, apportion county level emissions to the 2 kilometer grid cells within each county using surrogate distribution data, such as residential population, land use, or any other parameter whose distribution is known for the 2 kilometer grid cells. County wide emissions should be allocated to the grids based on the proportion of each cell's surrogate value of the total county surrogate. Tools such as GIS can be utilized to assign the county wide emissions from these sources to each grid cell.

For on-road mobile sources, GIS techniques can be used to estimate the relative length of major highways (roadway miles) in each county and grid cell. The ratio of roadway length in the grid cell to the roadway length in the county can be applied to the emissions rate. On-road mobile emissions are then assigned to grid cells in proportional amounts. Tables and maps should be produced to ensure that the allocation is performed properly. In gridding emissions in this manner, a pattern of highly variable emissions density in the modeling domain is established. The more comprehensive this emission allocation effort becomes, the sharper the gradient in the modeled concentrations.



Due to the time consuming nature of allocating emissions, EPA is developing a pre-processor Emissions Modeling System for Hazardous Air Pollutants (EMS HAP), to assist in this effort. Model code and a user's guide will be made available on the SCRAM website. EMS HAP will provide recommended spatial and temporal surrogates for area and mobile source emissions.

The area source algorithm in the ISCST3 model treats emissions as being uniformly distributed over each area source grid cell and allows for concentration estimates to be made within the area source itself. The area and mobile sources should be modeled using a nominal release height above ground, and an initial vertical dispersion value (sigma-z) to account for the fact that the area and mobile source emissions have some initial release height and initial depth, rather than being emitted passively from the ground surface. Expert judgement on the value used for this initial depth is needed to reflect local conditions (obstructions, proximity of receptors to roadway, etc.). For the example studies, the nominal release height was 2 meters and the initial vertical dispersion value (sigma-z) was 1 meter.

The ISCST3 model has the capability to address the temporal variations in emissions if the sources do not operate at the same rate for every hour of the year. For point sources, information on operating schedule or data from continuous emissions monitors can be input into the ISCST3 model. For area and mobile sources, temporal profiles are recommended, see for example (U.S. EPA, 1995b).

### 1.3.5 Default Source Parameters

Besides the emission rate, the parameters needed to model emissions from point sources include source location coordinates, physical release height, stack diameter, and stack exit velocity and temperature. Since most modeling analyses include a large number of sources over a relatively large area, it is inevitable that there will be gaps in the data for some of the sources. It is necessary to determine values for all the missing source characteristics, substitute them, and document the substitutions before the sources can be modeled. For national scale applications, using default source parameters is the only economical option.

The AIRS data base should be the primary source for identifying substitutions for missing source locations and missing stack parameters. For point sources with missing data that are not included in the AIRS data base, values for these missing data fields must be substituted before modeling those sources. An additional source for default values of stack height, stack diameter, exit velocity or flow rate, and exit temperature is the Ozone Transport Assessment Group (OTAG) defaults (OTAG, 1998), which are based on averages calculated by Source Classification Code (SCC). A list of the default parameters obtained from OTAG is provided in Appendix D. For any point sources which still have missing parameters and for which default parameters can not be identified from the OTAG data, the following conservative values are recommended for use in air toxics modeling analyses:

Stack height	10 meters
Stack diameter	1 meter
Exit temperature	ambient
Exit velocity	1 meter/second

If the actual location of the point source is missing, and it can not be identified from the AIRS data base, then the source should be assigned coordinates based on other available data (e.g., comparisons with other inventories, consultation with EPA Regional Office, etc.).

Since inventories available for analysis may not contain building dimension information, and since building downwash influences can significantly increase concentrations for receptors located close to the point source, the following approach may be used to set default values of building height and building width in the ISCST3 model. Default building dimensions of  $H_b = 0.625 * H_s$  and  $H_w = 2 * H_b$  (where  $H_b$  is building height,  $H_w$  is building width and  $H_s$  is stack height) may be used for stack heights of less than or equal to 65 meters, with a minimum building height of 3.05 meters, representative of a one-story structure. The value used for  $H_b$  places the stack height just above the Schulman-Scire criterion, except for stack heights that are less than about 4.6 meters, which is 1.5 times the minimum building height of 3.05 meters. The application of the Schulman-Scire downwash algorithm is therefore limited to the shorter stacks for which it is more likely to be applicable. The use of the Huber-Snyder downwash algorithm for stacks that are taller than 4.6 meters also avoids the potential for unrealistically increasing predicted impacts for these stacks based on relatively arbitrary building information, which could occur if the Schulman-Scire algorithm were to be applied to those stacks. For stack heights of greater than 65 meters, assume no building downwash occurs, since stacks of that height are likely to satisfy good engineering practice (GEP) stack height requirements to avoid building downwash influences.

Point source inventories, such as those available from AIRS, typically include either a stack height and stack gas exit parameters (temperature and flow rate), or a plume height for a given source. Sources with plume heights rather than stack heights and exit parameters can be considered non-buoyant releases (e.g., from isolated vents), where the release height is equal to the stated plume height. Therefore, for those point sources that include only plume height information and no stack parameters, the plume height should be taken as the release height, and other stack parameters should be set to zero to give no plume rise. Since these sources are likely to be from building vents or similar emission points, a building height equal to the stack height should be assumed, with building width equal to twice the building height. This automatically triggers the ISCST3 building downwash algorithm, and is a conservative approach.

#### 1.3.6 Source Parameters for Deposition Calculations

The ISCST3 model is capable of estimating wet and dry deposition rates of both gases and particles. While calculating deposition, the model also calculates the depletion of the deposited fraction from the plume during transport, resulting in a less conservative, more precise estimate of air concentrations. Calculating wet deposition requires additional meteorological data relating to precipitation and scavenging coefficients (U.S. EPA, 1995c). Values for the required deposition input parameters to ISCST3 should be obtained from the literature. The values used in the studies for Houston, Texas, and Phoenix, Arizona, below can be considered appropriate for those air toxic pollutants.

### 1.3.7 Pollutants

The ISCST3 model is run for one pollutant at a time. The number of pollutants should be carefully defined to minimize resources. Section 112 of the Clean Air Act lists 188 hazardous air pollutants (HAPs). The draft Integrated Air Toxics Strategy has identified 33 HAPs that are of primary concern in urban areas (U.S. EPA, 1999b).

### 1.3.8 Source Grouping

In some emission inventories, large industrial sources (e.g., paper mills, refineries, etc.) are grouped together so that hundreds of individual release points are assumed to exit from a few groups of stacks. If an inventory does not contain the individual location and release parameters, the analyst has little choice but to model the source as a group. Ambient concentrations from such facilities should be viewed with caution, especially at nearby receptors. If one assumes that ground level releases (e.g., leaks from pumps, seals or compressors, spilled liquids that form a puddle and then evaporate, lagoons, etc.) exit through an elevated stack, ground level concentrations will be underestimated.

From a post-analysis viewpoint, by grouping similar sources (e.g., mobile emissions), the analyst can more easily look at the impact of different source types. These groupings can be further subdivided into on-road and off-road mobile source groupings. ISC provides methods for grouping sources for these purposes.

### 1.3.9 Quality Assurance

Point source emissions must include all source parameters needed for input in the ISCST3 model: stack coordinates and source release parameters (stack height, temperature, exit velocity and diameter). Default values should be substituted for missing parameters using the guidance above.

Point source locations should be verified using GIS. Questionable locations should be identified. Large emission sources (e.g., greater than 10 tons/yr) should be verified where possible. A useful source of data for stack location is the AIRS database and, in particular, the AIRS Facility Subsystem (AIRS/AFS) which is a computer-based repository of information about airborne pollution. General information about the AIRS database is available at the Internet web site: <http://www.epa.gov/airs/airs2.html>. Although some hazardous air pollutant data are included, AIRS/AFS primarily houses data for criteria pollutants submitted by the States. The facility information includes data on emissions, process, control, stack, location, etc.

#### 1.4 MODEL OUTPUT FOR ANALYSIS

The ISCST3 model output includes concentration or deposition estimates for the various averaging periods. Annual average concentrations are used in an exposure model like HAPEM4 to estimate inhalation exposure as individuals move among different microenvironments such as from their homes to their work or school throughout the day. Model output can be further divided to show the impacts from the different types of sources, e. g., major, area, and onroad and off-road mobile. Estimates of deposition are used in multi media models.

The data analysis requirements for an air toxics modeling application are likely to be extensive. Large numbers of sources need to be modeled at a large number of receptor points that are representative of the exposure regimes/microenvironments and populations found in an urban area. As noted above, annual averages and the contributions of various sources/groupings are required, as a minimum.

## **2. CASE STUDY MODELING METHODOLOGY**

### **2.1 INTRODUCTION**

Air quality simulation models have a long history of use by the EPA in providing pollutant concentrations for use in specifying emission limits and assessing control strategies. The Guideline on Air Quality Models (40CFR51) was established to promote consistency in the use of models within the air management process. In an urban air toxics study, modeled concentrations are compared to health/exposure bench mark levels. The use of existing modeling tools for an urban air toxics study poses special challenges due to the large geographical scale in urban areas, the large number and variety of sources to be modeled and the variety of pollutants to be considered.

In this portion of the report the modeling methodology outlined in Section 1 is applied to two urban areas: Phoenix, Arizona and Houston, Texas. The pollutants modeled are benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium. This section addresses the selection of air quality model(s), describes the modeling options and modeling domains, selection of receptor locations, selection of meteorological data, data analysis requirements and important aspects of the emissions inventory. Also presented is the application of a scheme for estimating secondary formaldehyde formation and an overview of the model results.

This section of the report presents a discussion of the modeling methodology, while Appendix A and B present the case studies following this methodology for Phoenix, Arizona and Houston, Texas, respectively. The appendices include descriptions of the emission inventories used in the case studies, as well as, more detailed results of the air dispersion modeling analyses.

## 2.2 MODELING METHODOLOGY

### 2.2.1 Model Selection

The extent to which a specific model is suitable for the evaluation of source impact depends on several factors that include the meteorological and topographic complexities of the area; the level of detail and accuracy of the data base (i.e., emissions inventory), and the resources available.

There are a number of design criteria which need to be satisfied in order yield an acceptable modeling study of toxic pollutants. For the air dispersion model, for example, these include:

1. readily available/public domain/endorsed by EPA
2. represents state-of-modeling practice
3. applicable to urban areas and irregular terrain
4. capable of handling point, area and mobile sources
5. capable of accounting for dry deposition of pollutants
6. capable of treating atmospheric chemical transformations - pollutant chemistry
7. capable of accounting for pollutant emissions that vary by season and hour-of-day
8. ability to group source types for assessing impact
9. capable of providing annual average concentration estimates (as well as shorter time averages)
10. computationally efficient
11. good performance - estimated vs. observed concentrations.

The Gaussian plume model is a widely used technique for estimating the impacts of nonreactive pollutants because of its good performance against field measurements, and because it is computationally efficient relative to other types of models, such as grid and puff models.

The ISCST3 dispersion model includes the capability of handling multiple point, area, and mobile sources, incorporates building downwash effects, includes an urban dispersion option, and also contains considerable flexibility for specifying receptor locations and for grouping of source impacts. The ISCST3 model also includes algorithms to treat the effects of elevated and/or complex terrain, and the effects of dry and wet deposition of gaseous and particulate emissions. The ISCST3 model includes an option to vary emissions by season and hour-of-day, which was useful in meeting one of the design criteria for this modeling analysis, since the available emissions inventories reflect variations in emission rates by season and hour-of-day as inputs. This temporal resolution has also been selected for the model outputs based on the needs of a typical long term exposure assessment.

The model used for this study was the EPA Industrial Source Complex Short Term (ISCST3) dispersion model. The ISCST3 model is a steady-state Gaussian plume model which can be used to assess pollutant impacts from a wide variety of sources such as multiple point, area and mobile sources. This model was selected for this application to demonstrate what can be done with off the shelf modeling tools. For this modeling application annual and seasonal average concentrations, by hour-of-day, were generated. This selection for the temporal

resolution of the modeling results was based on the type of data that might be needed for use in a typical long term human exposure assessment. The ISCST3 model is applicable to receptors within about 50 km from the source and does not directly simulate the effects of pollutant chemistry (i.e., chemical transformation and reactivity).

Another approach considered for this study was the use of a photochemical model, such as the EPA Urban Airshed Model (UAM) (see Guideline on Air Quality Models). However, this approach had the following drawbacks: a) four of the five pollutants listed in this study were not treated specifically within UAM in the simulation of photochemical reactions (the exception is formaldehyde); b) UAM was designed for use during the summer months, hence the photolysis rates affecting some reactions leading to conversion of benzene, 1,3-butadiene, and POM for winter are not well characterized; c) UAM could not currently account for the effects of particle deposition, which were needed to treat hexavalent chromium; and d) there was little experience running UAM for an entire year (UAM normally is used for one to three day episodic periods). UAM also could not adequately account for pollutant source apportionment (required for developing control strategies).

For benzene, 1,3-butadiene and POM, the use of the simpler ISCST3 model was justified since UAM did not contain photochemical reactions for these pollutants. For hexavalent chromium, UAM did not handle particulate deposition, while ISCST3 does include a particle deposition algorithm. While UAM did contain photochemical reactions for formaldehyde, due to the dissimilarity between the two model's input and output, the time and effort required to use both the ISCST3 and the UAM models for this study was judged to be prohibitive. Also, considerable time and resources would have been needed to obtain annual estimates from UAM. Therefore, except for episodic concentrations of formaldehyde and for determining contributions from secondary formation of formaldehyde, ISCST3 was the preferred model.

Since ISCST3 did not address effects of secondary transformation for pollutants such as formaldehyde, a screening level photochemical model (OZIPR) was used to estimate the magnitude of secondary formaldehyde formations. A simplified approach to estimate secondary formaldehyde production for the ISCST3 model is described in Section 2.4.

### 2.2.2 Modeling Options

The regulatory default option of ISCST3 was selected for all modeling runs performed. This option specifies that the following will be used: final plume rise, buoyancy-induced dispersion, stack-tip downwash, calms processing routine, default wind profile exponents and default vertical potential temperature gradients. The modeling option to consider the influence of elevated and complex terrain was also selected for use for the Phoenix analysis, while flat terrain was assumed for the Houston analysis. The basis for these selections for treatment of terrain is discussed in more detail in Section 2.2.3.2.

Since the majority of the modeling domains for each city consists of urban land use categories, the urban modeling option was selected for all modeling runs, following the guidance contained in Section 8.2.8 of the Guideline on Air Quality Models (40CFR51).

Some of the toxic pollutants considered in this analysis are photochemically reactive, and the exponential decay option in the ISCST3 was utilized. The half-life values contained in Appendix A and B were used.

## 2.2.3 Receptor Locations

### 2.2.3.1 *Receptor Sampling Strategy*

The ISCST3 model incorporates numerous options for specification of receptor locations, including options for defining grids of receptors and/or discrete receptor locations in a cartesian and/or a polar coordinate system. These options provide the user with considerable flexibility in defining receptor locations for a particular application. Since the purpose of this analysis was the application of dispersion models to the assessment of human exposure to toxic pollutants on an urban scale, a study was undertaken to evaluate possible strategies for selecting receptor locations. The outcome of this study, which is documented in Appendix C, addresses the practical considerations of modeling a large number of sources spread over an urban area, as well as, the spatial resolution and sampling needs of a typical long term exposure assessment. The sampling strategy was relatively easy to implement, and significantly reduced the computing requirements of the analysis (by about a factor of five), while providing reasonable estimates of the mean and variance of the air concentrations (exposures) within the domain. The sampling strategy/procedure (Method C) in Section C.7 of Appendix C was applied to determine receptor points in both the Phoenix and Houston domains.

### 2.2.3.2 *Treatment of Terrain Influences*

The ISCST3 model may be run without terrain influences, i.e., flat terrain, or alternatively, the ISCST3 model will adjust the plume heights by the receptor elevation above or below stack base to account for the effects of elevated and complex terrain. The ISCST3 User's Guide (U.S. EPA, 1995c) contains information for handling terrain. The flat terrain option was used for the Houston analysis, while the elevated and complex terrain options in ISCST3 were used for the Phoenix analysis. The terrain within the Houston modeling domain is relatively flat with maximum height variations of about 50 feet. Given that a significant portion of the emissions for these pollutants is from area sources, and the ISCST3 model ignores terrain influences for area sources, the flat terrain assumption is considered adequate for Houston. While most of the Phoenix area averages about 1,100 to 1,300 feet above mean sea level (MSL), there are significant variations in terrain heights, up to about 2,700 feet MSL, within the modeling domain. Since the Phoenix modeling domain includes significant terrain features where high concentrations from stable plume impaction from elevated point sources may be of concern, terrain influences were addressed.

The terrain elevation for each receptor location modeled for Phoenix was determined from a file containing the 1-degree Digital Elevation Model (DEM) data, which provides terrain elevations from the U. S. Geological Survey (USGS) at a horizontal resolution of about 70-90 meters. The 1-degree DEM data can be downloaded for free from the USGS site on the World Wide Web at [http://edcwww.cr.usgs.gov/glis/hyper/guide/1\\_dgr\\_denfig/index1m.html](http://edcwww.cr.usgs.gov/glis/hyper/guide/1_dgr_denfig/index1m.html). The selection of 1-degree data for this analysis was based on the cost and availability factors and the fact that the resolution is considered more than adequate relative to the many other uncertainties



in the analysis. Stack base elevations for point sources in the Phoenix inventory were also determined from the 1-degree DEM data in a the same manner as receptor elevations.

#### 2.2.4 Meteorological Data

##### 2.2.4.1 *Selection of Surface and Upper Air Stations*

The ISCST3 model requires hourly surface observations of wind speed, wind direction, ambient temperature, and stability category, in addition to mixing heights derived from twice-daily upper air soundings as meteorological inputs. The mixing height data, processed by the National Climatic Data Center (NCDC), and the hourly surface data for major National Weather Service (NWS) stations are currently available for most cities for years up to 1991 from EPA's SCRAM Internet web site at <http://www.epa.gov/scram001>. The appropriate data for the five most recent years of readily available data were obtained from SCRAM for each city. The stations and years of data used for each city are summarized below (the numbers in parentheses are each station's identification number):

<u>City</u>	<u>Surface Station</u>	<u>Upper Air Station</u>	<u>Years</u>
Houston	Houston (12960)	Lake Charles, LA (03937)	1987-1991
Phoenix	Phoenix (23183)	Tucson, AZ (23160)	1987-1991

The selection of the surface stations was based on the only available first-order NWS station for each city. The selection of the upper air stations for mixing heights was based on the station considered to be the most representative for each city. For Phoenix, the Tucson upper air station is located about 100 miles away, while the next nearest upper air station is located at Albuquerque, NM, which is over 300 miles away. For Houston, the Lake Charles upper air station is located about 135 miles away, while the upper air station at Victoria, TX is about 120 miles away. However, the Victoria station was relocated to Corpus Christi, TX in January 1990, and 1989 upper air data is missing from SCRAM for both Victoria and Corpus Christi. Since Victoria is located about the same distance inland from the Gulf of Mexico as both Houston and Lake Charles, and the distance from Houston to Victoria is comparable to the distance from Houston to Lake Charles, both stations would be equally representative for use with the Houston surface data. However, given the fact that the Victoria station was moved to Corpus Christi resulting in a gap for 1989, and the fact that Corpus Christi is located nearer to the Gulf coast, Lake Charles was considered to be the better choice for use with Houston. This selection also corresponds with the recommendation of the Texas Natural Resource Conservation Commission (TNRCC) for modeling in Harris County, where Houston is located (TACB, 1992).

##### 2.2.4.2 *Meteorological Parameters for Deposition Calculations*

Several additional meteorological parameters are needed as inputs to the Meteorological Processor for Regulatory Models (MPRM) in order to implement the dry deposition algorithms in the ISCST3 model for particulate and gaseous emissions. Additional parameters related to wet deposition were not needed, since wet deposition was not included in the analysis. The additional dry deposition parameters are listed below:

Albedo  
Bowen Ratio  
Roughness Length (measurement site)  
Roughness Length (application site)  
Minimum Monin-Obukhov Length  
Surface Heat Flux (fraction of net)  
Anthropogenic Heat Flux  
Leaf Area Index

These parameters were estimated on a seasonal basis for this analysis, since many of the parameters will vary significantly by season, and the modeling analysis is designed to produce average concentrations by season. Based on a review of the guidance provided in Section 3.3 of the MPRM User's Guide (U.S. EPA, 1996a) for specifying these parameters, the following values were selected for this analysis:

<b><u>Phoenix:</u></b>	<b><u>Winter</u></b>	<b><u>Spring</u></b>	<b><u>Summer</u></b>	<b><u>Fall</u></b>
Albedo	0.20	0.14	0.16	0.18
Bowen Ratio	2.0	2.0	4.0	4.0
Roughness Length (measurement site) (m)	0.15	0.15	0.15	0.15
Roughness Length (application site) (m)	1.00	1.00	1.00	1.00
Minimum Monin-Obukhov Length (m)	50.0	50.0	50.0	50.0
Surface Heat Flux (fraction of net)	0.25	0.25	0.25	0.25
Anthropogenic Heat Flux (W/m <sup>2</sup> )	10.0	10.0	10.0	10.0
Leaf Area Index	0.5	0.5	0.5	0.5

<b><u>Houston:</u></b>	<b><u>Winter</u></b>	<b><u>Spring</u></b>	<b><u>Summer</u></b>	<b><u>Fall</u></b>
Albedo	0.20	0.14	0.16	0.18
Bowen Ratio	1.5	1.0	2.0	2.0
Roughness Length (measurement site) (m)	0.15	0.15	0.15	0.15
Roughness Length (application site) (m)	1.00	1.00	1.00	1.00
Minimum Monin-Obukhov Length (m)	50.0	50.0	50.0	50.0
Surface Heat Flux (fraction of net)	0.25	0.25	0.25	0.25
Anthropogenic Heat Flux (W/m <sup>2</sup> )	10.0	10.0	10.0	10.0
Leaf Area Index	1.0	1.0	1.0	1.0

The Bowen ratio values for Phoenix reflect dry conditions for an urban area, while the Bowen ratio values for Houston reflect average conditions for an urban area. Since the surface meteorological data used in the analysis are from major airports, it is assumed that the measurements are taken from well-sited instruments, away from major obstructions, with a nominal surface roughness length for the measurement site of 0.15 meters. For the application of the ISCST3 model in the selected urban areas, the roughness length at the application site was set at 1.0 meter.

#### 2.2.4.3 Meteorological Preprocessing

The MPRM program was used to preprocess the meteorological data for use with the ISCST3 model. Both the MPRM (U.S. EPA, 1996a) and PCRAMMET (U.S. EPA, 1996b)

meteorological preprocessors can be used to preprocess NWS surface and mixing height data for use with the ISCST3 model. However, PCRAMMET does not allow for specifying temporal (e.g., seasonal) or spatial variations of the surface parameters identified in the previous section. Also, PCRAMMET does not support the additional parameters needed to utilize the dry deposition algorithm for gaseous pollutants. These additional parameters are leaf area index (input by the user), and incoming solar radiation (calculated by MPRM). An estimate of minimum mixing depth for both study areas was determined based on guidance in Section 1.2.6.

#### 2.2.5 Emission Characteristics

The air toxics emission used for this analysis were extracted from the air toxics emissions inventory prepared by the State of Arizona and the State of Texas. These emissions inventories were already gridded (e.g., spatially and temporally allocated) and could not be changed. These aspects are not further discussed.

##### 2.2.5.1 *Determination of Background Concentrations*

As the guidance in Section 1.2.8 above indicates, background concentrations should be added to the modeled concentrations to provide total concentration/exposure. However, no background concentrations of the five modeled chemicals were available for Houston, nor were POM or hexavalent chromium for Phoenix. Resource limitations did not allow detailed analyses. For the simulation period, background concentrations of zero were assumed for all five pollutants in this analysis. This assumption creates uncertainties about the magnitude of the total estimated concentration.

##### 2.2.5.2 *Default Source Parameter Values*

Besides the emission rate, the parameters needed to model emissions from point sources include source location coordinates, physical release height, stack diameter, and stack exit velocity and temperature. Since the modeling analysis included a large number of sources over a relatively large area, it was inevitable that there were gaps in the data for some of the sources. It was therefore necessary to determine default values of the necessary source characteristics to be substituted for missing data before the sources could be modeled. This was a more significant problem for the Phoenix inventory, since the only information provided besides the emission rate was the grid cell containing the point source, and the facility and stack IDs. Source parameters were available for nearly all of the Houston point and major sources.

Since the Phoenix point source data were identified by facility and stack IDs, the AIRS data base was used as the primary source to identify substitutions for missing source locations and missing stack parameters. For point sources with missing data that were not included in the AIRS data base, default values of the missing data fields were substituted before modeling those sources. The default values of stack height, stack diameter, exit velocity or flow rate, and exit temperature were taken from the OTAG defaults, as described in Section 1.3.3. The procedures for compensating/substituting for missing or erroneous stack parameters are described in Chapter 3, Part 1, Section 3 of the OTAG Technical Supporting Document (OTAG, 1998). For any point sources which had missing parameters and for which default parameters could not be identified from the OTAG data, the values given in Section 1.3.3 above were used in the modeling analysis.

If the actual location of any point source had been missing, and it could not be identified from the AIRS data base, then the source would have been assigned coordinates within the grid cell on a random basis (see Section 1.3.5 above). None of the sources from the Phoenix or Houston inventories fell into this category. The guidance in Section 1.3.5 concerning building dimensions was followed.

#### *2.2.5.3 Area and Mobile Source Modeling*

The area source emissions that were provided as part of the emission inventory for each city were spatially allocated to rectangular grid cells of varying sizes, depending on the city. The grid size for Houston was 2 km x 2 km, and the grid size for Phoenix was 4 km x 4 km. The area source algorithm for the ISCST3 model was used to physically model these area source emissions, which is equivalent to modeling the emissions as being uniformly distributed over each area source grid cell. The ISCST3 area source algorithm also allows for air concentration estimates to be calculated within the area source itself. The guidance concerning values for initial dispersion, Section 1.3.4 above, was followed.

#### *2.2.5.4 Source Parameters for Dry Deposition Calculations*

ISCST3 is capable of estimating wet and dry deposition rates of both gases and particles. While calculating the deposition, the model also calculates the depletion of the deposited fraction from the plume, resulting in a less conservative, more precise, estimate of air concentrations. In this analysis only the dry deposition algorithms were selected. Chemical-specific scavenging coefficients were not available for the gaseous pollutants in order to estimate wet deposition. The neglect of wet deposition, which requires additional meteorological data related to precipitation, results in a conservative air concentration estimate. Dry deposition of particles was modeled for hexavalent chromium, with dry deposition of gases modeled for the other four pollutants.

In order to apply the gas dry deposition algorithm in the ISCST3 model to the gaseous pollutants of interest (benzene, 1,3-butadiene, formaldehyde and POM), several additional parameters (see ISC **Keywords**, U.S. EPA, 1995c) must be specified:

Molecular diffusivity in air (**Diff**) was obtained from Fletcher, et al., 1997.

The solubility enhancement factor (**Alphas**) (also referred to as Alpha\*) is used when applying the deposition algorithm over wet surfaces such as moisture on vegetation due to precipitation or water bodies. No major water surfaces were included in the respective modeling domains. Since a value for this aqueous phase dissociation is not available for the pollutants in this study, a value of 1.0 for SO<sub>2</sub>, suggested in the ISCST3 model User's Guide was used.

The reactivity parameter (**Reac**) is the scaling factor for "stickiness" of the pollutant and is pollutant specific. In the absence of observed data in this study, it was set to 10 (a moderate value).

The mesophyll resistance (**Rsubm**) can be set to zero for soluble compounds (e.g. formaldehyde, maximum water solubility = 550000 mg/L), while non-soluble compounds (e.g.

naphthalene, maximum water solubility = 30 mg/L) are set to a high value i.e. 100. Values for maximum water solubility may be found in chemical engineering handbooks and various publications, such as the Air/Superfund National Technical Guidance Study Series' air emission model (U.S. EPA, 1993).

Henry's Law coefficient (**Henry**) is used when applying the deposition algorithm over wet surfaces. It is a measure of the vapor/water partitioning of a compound. This dimensionless value is obtained by dividing the Henry's Law constant (values based on Fletcher, et al., 1997) by the gas constant, R, ( $8.314 \text{ Pa}\cdot\text{m}^3/\text{mol}\cdot\text{K}$ ) times the ambient temperature, T, (assumed to be 293 K).

#### Benzene:

Molecular diffusivity in air (Diff)	= 0.0912 $\text{cm}^2/\text{sec}$
Alphas (Alpha*)	= 1.0
Reactivity parameter (Reac)	= 10.0
Mesophyll resistance (Rsubm)	= 10.0
Henry's Law coefficient (Henry)	= 0.24
Maximum water solubility	= 1,780 mg/L
Henry's Law constant	= 543 $\text{Pa}\cdot\text{m}^3/\text{mol}$

#### 1,3-Butadiene:

Molecular diffusivity in air (Diff)	= 0.1158 $\text{cm}^2/\text{sec}$
Alphas (Alpha*)	= 1.0
Reactivity parameter (Reac)	= 10.0
Mesophyll resistance (Rsubm)	= 20.0
Henry's Law coefficient (Henry)	= 2.95
Maximum water solubility	= 735 mg/L
Henry's Law constant	= 7,180 $\text{Pa}\cdot\text{m}^3/\text{mol}$

#### Formaldehyde:

Molecular diffusivity in air (Diff)	= 0.1698 cm <sup>2</sup> /sec
Alphas (Alpha*)	= 1.0
Reactivity parameter (Reac)	= 10.0
Mesophyll resistance (Rsubm)	= 0.00
Henry's Law coefficient (Henry)	= 0.00
Maximum water solubility	= 550,000 mg/L
Henry's Law constant	= 0.032 Pa-m <sup>3</sup> /mol

POM can be defined in a number of ways, e.g., a group of 16-PAH's (polycyclic aromatic hydrocarbons). Since a precise definition of POM was not provided in the emissions inventory, parameter values used here are for a surrogate compound, naphthalene. Note, there is uncertainty about this approach. Values for benzo(a)pyrene, or another PAH, could have been used if available. These values, based on naphthalene were used for POM:

#### Naphthalene:

Diffusivity in air (Diff)	= 0.0590 cm <sup>2</sup> /sec
Alphas (Alpha*)	= 1.0
Reactivity parameter (Reac)	= 10.0
Mesophyll resistance (Rsubm)	= 100.00
Henry's Law coefficient (Henry)	= 0.02
Maximum water solubility	= 30 mg/L
Henry's Law constant	= 48.6 Pa-m <sup>3</sup> /mol

The fifth pollutant modeled, hexavalent chromium, is emitted in a particulate form. To apply the dry deposition algorithm for hexavalent chromium, a particle density and particle size distribution was input to the model for each source. Particle size distributions are provided in AP-42 for many source categories that typically emit particulate matter (U.S. EPA, 1995a). As an example, for chrome electroplating facilities (SCC 3-09-010-18), Section 12.20 of AP-42 provides the following particle size distribution for uncontrolled emissions: 6.9 percent for particles less than 0.5 $\mu$ m; 60.8 percent for particles between 0.5 and 2.4 $\mu$ m; 14.9 percent for particles between 2.4 and 8.0 $\mu$ m; and 17.4 percent for particles larger than 8.0 $\mu$ m. This distribution was used for chrome electroplating facilities, with a particle density of 1.0 gm/cm<sup>3</sup>. Similar distributions were determined for other major categories of hexavalent chromium emissions, based on review of the inventories.

#### *2.2.5.5 Source Grouping*

Source grouping was used to allow for tracking and comparison of impacts for various source types, such as major sources versus area and mobile source types. All source types were combined into three source groups: major/point sources, area sources, and mobile sources. There were a few small point sources for which no coordinates were available; these were placed in the area source group. Modeled concentrations were generated for all sources combined, as well as, for each of these groups of sources.

#### 2.2.6 Model Evaluation

As a preliminary, exploratory model evaluation analysis, available ambient monitoring data in the Houston area were obtained and compared with modeled estimates. No suitable ambient monitoring data were available for Phoenix. Appendix B contains a discussion of the comparison study. Resource permitting, further investigations are warranted.

## 2.3 OVERVIEW OF EMISSION INVENTORIES

This section provides an overview of the emission inventories used in the case studies. Section 2.3.1 discusses the Phoenix, Arizona inventory, and Section 2.3.2 discusses the Houston, Texas inventory. Each section describes the sources of the emissions data, the temporal and spatial resolution of the emissions data, summarizes the data processing that was performed to get the data into model ready format, and provides brief summaries of the total emissions by pollutant and source category. Each of the inventories is described in more detail in the respective case study appendix.

### 2.3.1 Phoenix, Arizona Inventory

The Phoenix inventory used in this study was extracted from an inventory that was developed as part of the Arizona HAP research program by the Arizona Department of Environmental Quality (ADEQ), and the Arizona Department of Health Services (ENSR, 1995). The original inventory included gridded emissions of 163 pollutant species on a 4 km-by-4 km grid resolution covering the regions around Phoenix, Tucson, Casa Grande, and Payson, Arizona. The gridded inventory incorporated all source categories, including major, area and mobile sources. Temporal variations of the emissions on a season by hour-of-day basis were also included in the inventory. Table 2.3-1 summarizes the emissions information. Detailed information about the Phoenix emission inventory can be found in Appendix A, Section 2.

The Phoenix inventory consisted of gridded emission estimates on a 4km-by-4km grid resolution. A total of 850 grid cells was used to cover the modeling domain. The data consisted of separate files by season for each pollutant, with each file containing seasonal average emission estimates by hour of day. The original data also included emission estimates by source classification code (SCC) for three-level categories. For the five pollutants modeled in this analysis, the original raw emissions data files consisted of about 66 Mb of data. The process of converting the Phoenix emissions data to model-ready format for the ISCST3 model involved the following steps for each pollutant, which were accomplished using several utility programs that were written in the Fortran programming language:

1. Defining a unique source ID for each grid cell based on the i- and j-cell values in the data files;
2. Extracting the hour-of-day emission estimates from the four seasonal files for a particular pollutant;
3. Summing the emissions by source category to the total emissions for the particular source (i.e., grid cell);
4. Converting the emissions data to the proper units for model input; and
5. Writing out the source locations, physical parameters, and season by hour-of-day emission rates in the proper format for model input.



These basic steps applied to both the major source inventory for Phoenix and the area source inventory. However, the raw data files for major sources included all pollutants within a single data file for each season, while the area source data were in separate data files for each pollutant. Minor adjustments to the utility programs to process the data were therefore needed for each type of source.

The physical source characteristics used to model the area and mobile source emissions are described in Section 2.2.5.3. Since the raw data files did not include physical source characteristics for major sources, additional processing was needed to prepare these data for input to the model. The general approach taken for this is described in Section 2.2.5.2. The approach was implemented by means of a Fortran utility program that read the raw data files, attempted to match the facility and source ID with data from AIRS, assigned source locations and physical stack parameters based on data in the AIRS data base, or applied default values if no other data were available.

### 2.3.2 Houston, Texas Inventory

The Houston inventory used in this study included multiple components. Major source emissions of the selected pollutants for the Houston area were contained in a spreadsheet. The area source emission inventory was taken from the inventory developed as part of the Houston Area Source Toxic Emissions (HASTE) project (Radian, 1995 a and b). It consisted of several components including total annual emission estimates by pollutant and source category for Harris County, gridded values of various activity factors, such as population, number of dry cleaners, etc., and an association between activity factors and source categories that could be used to spatially allocate the pollutant emissions across the domain. The activity factors, and therefore the area source emissions derived from them, were gridded with a 2km-by-2km grid resolution for the Harris County domain.

Table 2.3-2 summarizes the emissions information. Note that in Table 2.3-2, major and area source emissions are estimates for 1993, and mobile source emission estimates are for the base year 1990. It is assumed that mobile source emissions did not change appreciably from 1990 to 1993, since no fuel additives were introduced in Houston during that period. More detailed information about the emissions can be found in Appendix B, Section B.2.

As with the Phoenix inventory, several steps were involved in converting the emissions inventory data as provided to a model-ready format for the ISCST3 model. However, there were significant differences in the type and structure of data provided for each city that placed different requirements on this conversion process. As noted above, the Houston inventory consisted of three main components: major sources; non-mobile area sources; and mobile sources. Each of these components had somewhat different processing requirements. For major sources, the source locations and physical release parameters were provided for nearly all sources, and required simple reformatting to the input format required for the ISCST3 model after extracting the data from the spreadsheet file. The more difficult task in preparing the major sources was in separating the emissions by pollutant and applying toxic speciation factors to certain sources that were listed as emitting only kerosene, fuel oil, etc. These toxic fractions varied by source category, and matching the sources provided to the source categories had to be accomplished manually based on the information available.

The area source data provided from the HASTE project required several steps to prepare them for input to the ISCST3 model. The gridded activity factor data that were used to spatially allocate the toxic emissions across the Houston domain were provided in the form of RAM model (40CFR51) input files. A separate RAM model file was provided for each of the 17 activity factors used in the analysis. These activity factors were provided in absolute units per grid cell, such as number of people per grid cell, number of dry cleaners per grid cell, etc. The gridded activity factors were combined into a single file that contained the relative activity factor for each grid cell, and the grid cells were converted to source coordinates for area sources to be used in the model input files. In addition, some activity factors were combined for use in spatially allocating certain source categories. A utility program was written in Fortran to combine data from the gridded activity factors with total emissions by source category into model-ready inputs for ISCST3.

The mobile source inventory for Houston, provided by EPA's Office of Mobile Sources (OMS) consisted of gridded VOC emission estimates by season and hour-of-day for on-road and off-road mobile sources by four-level SCC. The raw data consisted of separate files by season for both on-road and off-road sources. The OMS also provided toxic fractions for benzene, 1,3-butadiene, formaldehyde, and POM to be applied to the VOC emissions. The process of converting the mobile source data to model-ready inputs was similar to that applied to the Phoenix area source data, with the additional step required of applying the toxic fractions by source category and pollutant. The on-road and off-road emissions were also combined to reduce the number of area sources (i.e., grid cells) that needed to be modeled.

**Table 2.3-1 Toxic Air Pollutant Emissions for Phoenix, Arizona  
Based on Year 1993**

Source Category		Pollutant Emissions (Mg/yr)				
		Benzene	1,3-Butadiene	Formaldehyde	POM <sup>(1)</sup>	CrVI <sup>(2)</sup>
Mobile Sources	On-Road	1,003.0	100.0	605.0	195.0	-- <sup>(3)</sup>
	Off-Road	727.0	165.0	392.0	34.4	-- <sup>(3)</sup>
	Total Mobile	1,730.0	265.0	997.0	229.4	-- <sup>(3)</sup>
Area Sources		79.5	3.5	20.4	29.2	0.082
Major Sources		12.8	0.4	7.1	-- <sup>(3)</sup>	0.028
TOTAL		1,822.3	268.9	1,024.5	258.6	0.11
<p>(1) POM = Polycyclic Organic Matter. Individual constituents of POM vary. For mobile source, POM is defined as sum of 16 chemicals. For area &amp; mobile sources, data provided by TNRCC do not define POM constituents. Data reported as POM by TNRCC are used here.</p> <p>(2) CrVI = Hexavalent Chromium</p> <p>(3) Not Applicable</p>						

**Table 2.3-2 Toxic Air Pollutant Emissions for Houston, Texas  
Based on Year 1993**

Source Category		Pollutant Emissions (Mg/yr)				
		Benzene	1,3-Butadiene	Formaldehyde	POM <sup>(1)</sup>	CrVI <sup>(2)</sup>
Mobile Sources	On-Road	971.8	159.2	439.5	0.13	-- <sup>(3)</sup>
	Off-Road	266.4	69.4	156.3	0.22	-- <sup>(3)</sup>
	Total Mobile	1,238.2	228.6	595.8	0.35	-- <sup>(3)</sup>
Area Sources		116.9	4.9	59.0	1.70	2.2
Major Sources		1,080.7	386.5	75.7	0.00	11.9
TOTAL		2,435.8	620.0	730.5	2.05	14.1
<p>(1) POM = Polycyclic Organic Matter. Individual constituents of POM vary. For mobile source, POM is defined as sum of 16 chemicals. For area &amp; mobile sources, data provided by TNRCC do not define POM constituents. Data reported as POM by TNRCC are used here.</p> <p>(2) CrVI = Hexavalent Chromium</p> <p>(3) Not Applicable</p>						

## 2.4 APPROACH USED TO ESTIMATE ATMOSPHERIC SECONDARY FORMALDEHYDE PRODUCTION

### 2.4.1 Introduction

Formaldehyde, one of the five frequently emitted pollutants in this analysis, is the only one of the five formed in significant amounts in the atmosphere. It is necessary to estimate the amount of formaldehyde formed and destroyed by subsequent reactions in the atmosphere. ISCST3 does not contain a chemical reaction module to handle the transformations of formaldehyde or other chemicals that occur after emission from a source. These transformations are often termed "secondary reactions" and the results "secondary production." A simplified approach to estimate annual average secondary formaldehyde production for the ISCST3 model is described below. The option of using a refined photochemical model was explored but rejected because of: 1) the expense of using a photochemical model; 2) the inability of current models to provide annual average estimates; and 3) data limitations.

An annual average estimate of secondary formation of formaldehyde was approximated using EPA's research oriented version of the Ozone Isopleth Plotting Package (OZIPR), an air quality screening model (Gery and Crouse, 1991) which can use very complex chemical mechanism, in this case the Carbon Bond 4 (CB-4) mechanism, to explore the photochemistry of a well mixed column of air traveling along a trajectory<sup>1</sup>. The column of air extends from the ground surface to the top of the mixed layer, the height varying throughout the day. Initial chemical species concentrations are changed via dilution by air aloft as the mixed layer height increases, by time dependent pollutant emissions along the trajectory path, and by the model photochemistry. A reaction mechanism for use in OZIPR has been adapted to distinguish between formaldehyde due to primary sources (emissions), and formaldehyde formed in situ (secondary production). Estimates based on OZIPR are added by a post processor to the initial ISCST3 model estimates on a seasonal basis and then summed for the annual average. This simplified approach provides a "ball park" estimate of secondary formaldehyde formation.

The complete process consists of the following:

1. Obtaining the input variables necessary to run OZIPR: meteorological data, chemical species background concentrations, and chemical species hourly emissions;
2. Running OZIPR;
3. Using the OZIPR results to estimate secondary formaldehyde concentrations for use with the ISCST3 model results; and
4. Applying the derived secondary values to the ISCST3 modeled concentrations based on primary formaldehyde emissions to approximate total formaldehyde concentrations.

---

<sup>1</sup>The OZIPR model may be obtained from EPA's Office of Research and Development, National Exposure Research Lab. in Research Triangle Park, North Carolina.

The analysis is specific to both a city and season. For Houston, Texas, a prototypical day was determined for each season: winter, spring, summer, and autumn. Seasonally averaged hourly meteorological data were used. For example, Houston's 8 am temperature values for each spring day were averaged to simulate the 8 am temperature value; 9 am, 10 am, and so on, averages were created for the meteorological data. The emissions data for each season were handled in a like fashion.

#### 2.4.2 Simulation Specific Input

Appendix B describes in detail how to run the OZIPR model for an application similar to the Houston one. It assumes a user has little familiarity with the model. To run OZIPR for determining total formaldehyde the following is needed:

1. Location (latitude, longitude);
2. Date (to be used in zenith angle calculation);
3. Meteorology for the period of the simulation;
4. Background concentrations of chemical species; and
5. Hourly emissions throughout the run.

Input is defined for a prototypical day for each season. The prototypical day is defined using seasonally averaged hourly meteorology, solar radiation, and pollutant emissions. The date selected (used in zenith angle calculations) for the prototypical day for a given season was the day that fell in the middle of the season (i.e. July 16 for summer). Seasons for this purpose are:

Winter:	December, January, February
Spring:	March, April, May
Summer:	June, July, August
Fall:	September, October, November

The time period selected for a simulation cannot exceed 24 hours. The initial time is selected so the simulation does not extend into more than one period of daylight. For example, a 24-hour run starts before sunrise and goes through one consecutive period of daylight.

Seasonally averaged meteorology is extracted from local or state records. It is possible to approximate the hourly values using long term climatological data sets that include morning and afternoon values for temperature, relative humidity, or mixing height; linearly interpolating between given values. Mixing heights can be entered at two points in time using the model's DILUTION option. Morning and afternoon mixing heights are from Holzworth (1972). Minimum and maximum temperatures and morning and afternoon relative humidities were obtained from a climatic atlas (NCDC, 1984).

Background concentrations values for NO, NO<sub>2</sub>, O<sub>3</sub>, and NO<sub>3</sub> for this run used using 6 am concentrations from a previous day's simulation as noted in Appendix F. However, these values

are lower than expected. For example, NO<sub>x</sub> was initialized to 0.04 ppb which then results in an initial O<sub>3</sub> value of 4.5 ppb instead of an expected value of 100 ppb. It is uncertain at this time how final model results would differ if different initialization values were used. The user should consult with a State modeling expert for advice in values to use in initialization.

### 2.4.3 Analysis of Results for Use with ISCST3

Results from the OZIPR giving both the primary and secondary formaldehyde for the location, and for a prototypical day for each season, are shown in Figures 2.4-1 through 2.4-4. The plots shown are for a 24-hour run from 6 am until 5 am the following morning, hours 6-29. Note the morning maximum of primary formaldehyde, FORM, around 10 am in all four seasonal simulations. The time of the maximum in formaldehyde emissions for these simulations is 4 pm in the afternoon, but it is the photochemistry that defines the shape of the curve and depletes the primary formaldehyde in daylight hours. As the primary emissions of formaldehyde are the same for all seasons, the differences in the FORM curves are due to meteorology and photochemical depletion, and are not large.

It should be noted that the secondary formaldehyde, FRMS, contributes a much greater share of the total formaldehyde. FRMS also exhibits greater seasonal differences, influenced by both seasonal emissions of precursors and by meteorology.

These results, presented in Figures 2.4-1 through 2.4-4, are used to define a seasonally dependent scaling factor which relates the amount of secondary to primary formaldehyde at 10 am. The values for 10 am are chosen, as this corresponds to a time of day when primary formaldehyde values are relatively high for the day. The 10 am seasonally dependent scaling factor (SF) is then the 10 am secondary formaldehyde concentration from OZIPR divided by the 10 am primary formaldehyde concentration from OZIPR:

$$SF = \frac{\text{OZIPR [10 am Secondary FORM]}}{\text{OZIPR [10 am Primary FORM]}}$$

This scaling factor is then used with the ISCST3 model results for the prediction of secondary formaldehyde from the primary formaldehyde in the ISCST3 model. (The scaling factor calculated for Houston, TX in summer was 6.74.)

$$\text{ISCST3[10 am Secondary FORM]} = SF * \text{ISCST3[10 am Primary FORM]}$$

Using a scaling factor for each hour is not recommended, as large errors can be introduced. The 10 am value was chosen for its minimal SF. For hours other than 10am, the primary formaldehyde simulated by OZIPR is lower, and any uncertainty in this value will be amplified in computing an hourly SF value. This method keeps in check amplification of high outlier values. Instead, the OZIPR model results for secondary formaldehyde are used to determine the amount of secondary formaldehyde relative to that at 10 am for each hour. The result is an hourly fraction value (FR), calculated for each hour, which can be used to determine the values of secondary formaldehyde with respect to the calculated 10 am value. For example, the 6 am fraction would be:

$$FR(6 \text{ am}) = \frac{OZIPR [6 \text{ am secondary FORM}]}{OZIPR [10 \text{ am secondary FORM}]}$$

Using the approximated value for ISCST3 10 am secondary formaldehyde, and fractions relating the 10 am secondary formaldehyde for all hours, the final calculations for approximated ISCST3 secondary formaldehyde can be made. For example, the 6 am value for ISCST3 secondary formaldehyde would be:

$$ISCST3 [6 \text{ am secondary FORM}] = ISCST3 [10 \text{ am secondary FORM}] * FR(6 \text{ am})$$

A data file was created for each season for use with ISCST3 which contains the scaling factors for each season, and the hourly fractional values for each season:

HOUSTON TEXAS SUMMER:

USE HOUR 10 DOMAIN AVERAGE PRIMARY FORMALDEHYDE MULTIPLY BY  
6.73699 TO GET 10 AM SECONDARY FORM = FRMS10AM

FRMS(10AM)=(10AM Primary FORM)\*(SF summer)

$$FRMS(HR)=FRMS(10AM)*FR(HR)$$

FRMS	600.000	= FRMS(10AM) *	1.44
FRMS	700.000	= FRMS(10AM) *	1.24
FRMS	800.000	= FRMS(10AM) *	1.10
FRMS	900.000	= FRMS(10AM) *	1.01
FRMS	1000.000	= FRMS(10AM) *	1.00
FRMS	1100.00	= FRMS(10AM) *	1.04
FRMS	1200.00	= FRMS(10AM) *	1.09
FRMS	1300.00	= FRMS(10AM) *	1.13
FRMS	1400.00	= FRMS(10AM) *	1.17
FRMS	1500.00	= FRMS(10AM) *	1.25
FRMS	1600.00	= FRMS(10AM) *	1.28
FRMS	1700.00	= FRMS(10AM) *	1.29
FRMS	1800.00	= FRMS(10AM) *	1.28
FRMS	1900.00	= FRMS(10AM) *	1.28
FRMS	2000.00	= FRMS(10AM) *	1.31
FRMS	2100.00	= FRMS(10AM) *	1.33
FRMS	2200.00	= FRMS(10AM) *	1.35
FRMS	2300.00	= FRMS(10AM) *	1.36
FRMS	2400.00	= FRMS(10AM) *	1.38
FRMS	2500.00	= FRMS(10AM) *	1.40
FRMS	2600.00	= FRMS(10AM) *	1.41
FRMS	2700.00	= FRMS(10AM) *	1.42
FRMS	2800.00	= FRMS(10AM) *	1.43
FRMS	2900.00	= FRMS(10AM) *	1.43

Total formaldehyde, secondary plus primary, was calculated as following:

$$\text{FRMT}(\text{HR}) = \text{FRMP}(\text{HR}) + \text{FRMS}(\text{HR})$$

#### 2.4.4 Results

Time series plots were generated for total formaldehyde, the sum of the primary and secondary formaldehyde. A comparison of the time series plots for the primary formaldehyde and the total formaldehyde in Figures 2.4-1 through 2.4-4 shows that the general diurnal pattern remains the same with an hourly increase of  $0.4 - 0.6 \mu\text{g}/\text{m}^3$  when the secondary is added. The isopleth maps in Appendix B also show a slight increase in the spatial distribution of the total formaldehyde levels within the study area. The maximum concentration increased to  $2.5 \mu\text{g}/\text{m}^3$  from  $2.1 \mu\text{g}/\text{m}^3$ .

#### 2.4.5 Conclusions

The OZIPR model has been used to approximate the secondary formation of formaldehyde in Houston, Texas by its relationship to the amount of primary formaldehyde using the CB-4 mechanisms. In all cases, the secondarily formed formaldehyde was the greater contributor to total formaldehyde. The resulting approximation for secondary formaldehyde was applied to ISCST3 model results uniformly throughout the Houston domain.

Because there is no direct chemical relationship between the primary and secondary formaldehyde, relating these species empirically can only be applied in a small area where the local chemistry can be well defined. Using OZIPR requires information on both biogenic and anthropogenic emissions, which might not readily be on hand. The model is not very flexible and it was not trivial to use for this application. For example, because OZIPR is a 24-hour model, running for additional time requires setting up a new input file based on the model output. This includes initializing of all species represented in the model. The OZIPR results were based on the conditions for the geographic area of Houston and will not necessarily apply to other areas.



### Formaldehyde Concentration

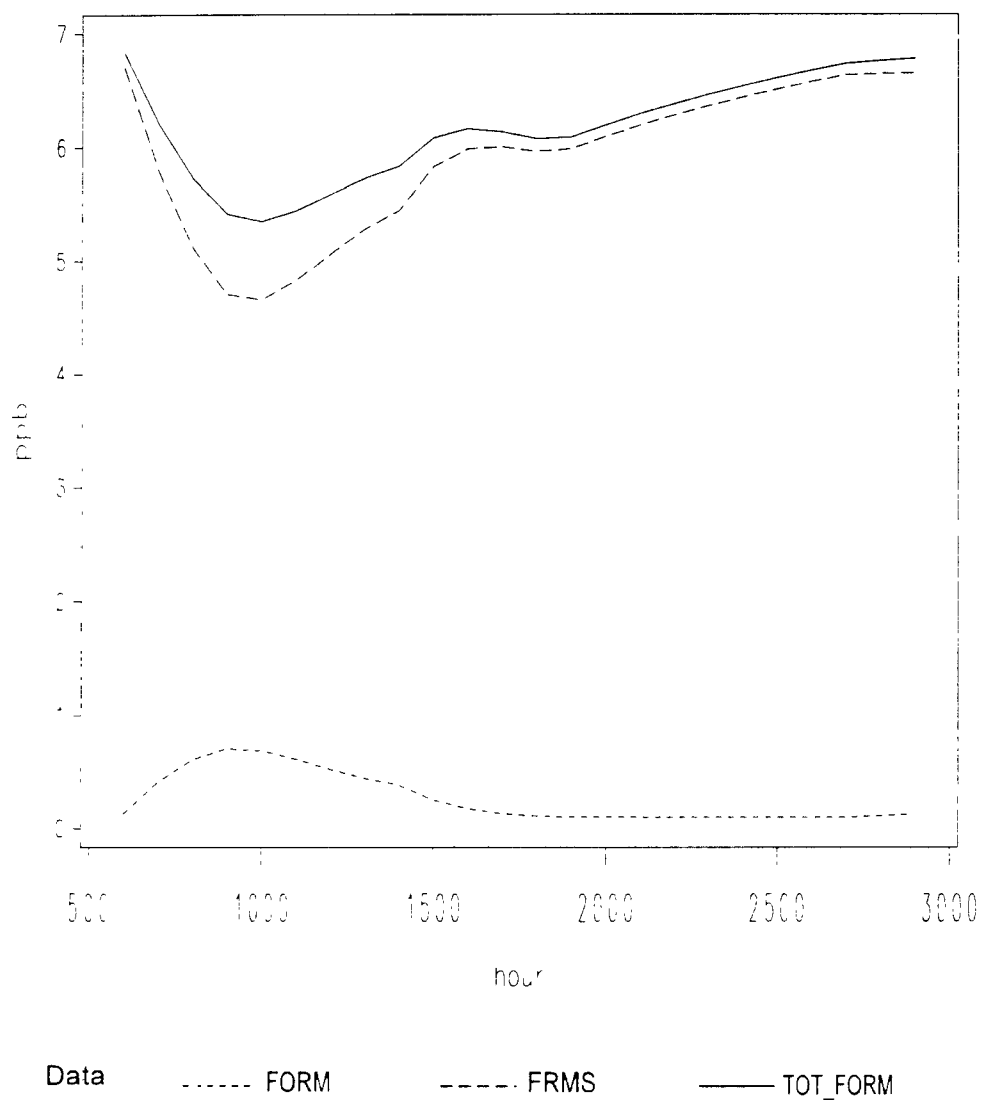


Figure 2.4-1: Formaldehyde concentrations for prototypical summer day in Houston Texas, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT\_FORM (total formaldehyde).

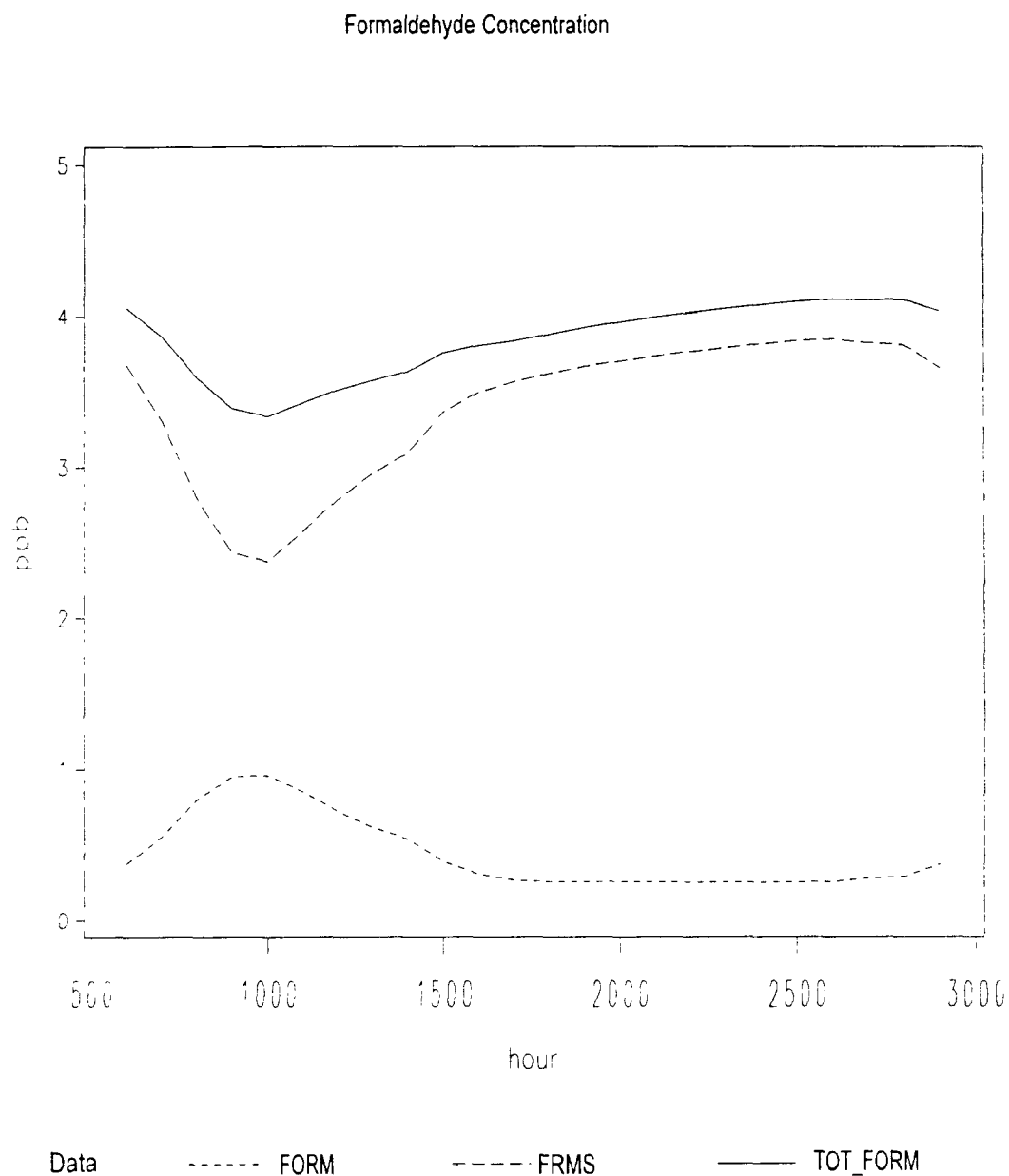


Figure 2.4-2: Formaldehyde concentrations for prototypical autumn day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT\_FORM (total formaldehyde).

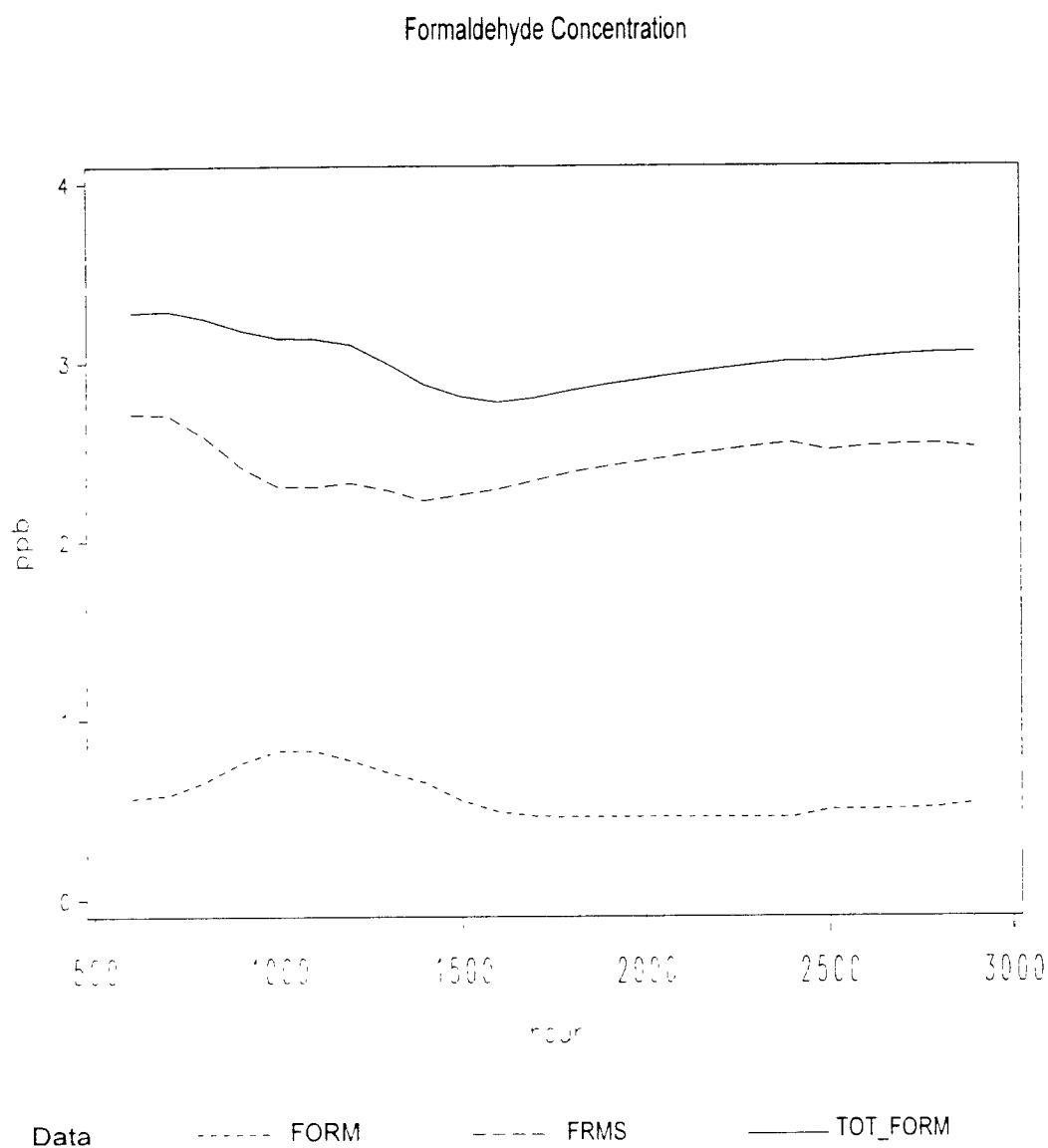


Figure 2.4-3: Formaldehyde concentrations for prototypical winter day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT\_FORM (total formaldehyde).

### Formaldehyde Concentration

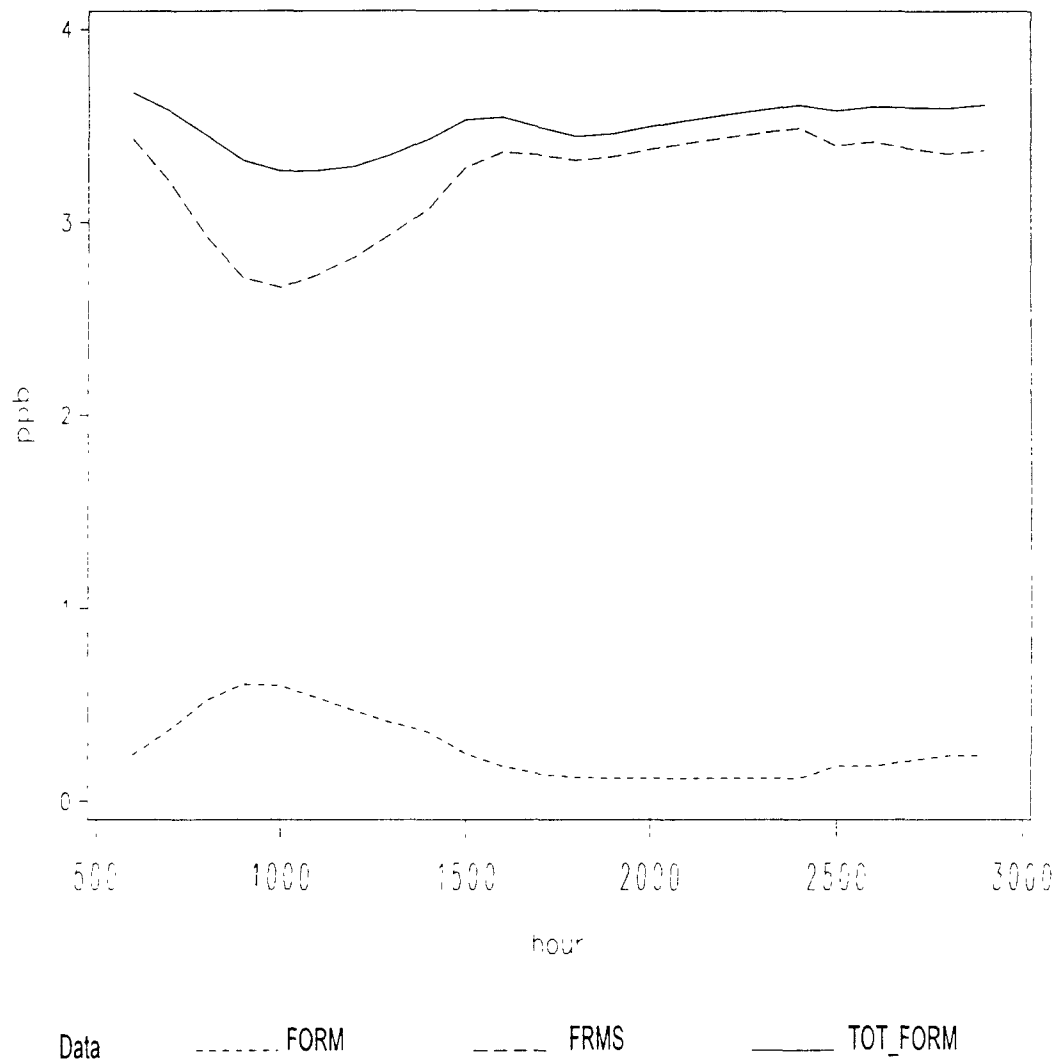


Figure 2.4-4: Formaldehyde concentrations for prototypical spring day in Houston, TX, at approximate steady state. OZIPR with hour of simulation vs ppb of: FORM (primary formaldehyde), FRMS (secondarily produced formaldehyde), and TOT\_FORM (total formaldehyde).

## 2.5 OVERVIEW OF MODELING RESULTS

This section provides an overview of the modeling results obtained from the two case studies for Phoenix, Arizona and Houston, Texas. A summary of the annual average concentrations and of the temporal concentration variations is provided. More detailed summaries of the modeling results are provided in the appendix for each city.

### 2.5.1 Phoenix, Arizona Modeling Results

Table 2.5-1 summarizes the highest annual-average impacts for each modeled pollutant. The impacts listed in this table are the combined impacts of all source categories (major, area and mobile) in the Phoenix area. The largest contributions to the total impacts for all pollutants, except hexavalent chromium, were due to the mobile sources. The area sources contributed significantly less than the mobile sources, and the major sources contributed the least. For hexavalent chromium, for which there were no mobile sources, the major and the area sources contributed equally to the highest total impacts. The contributions from the individual source categories, as well as further discussions on the modeling results, including isopleths of concentrations, can be found in Appendix A.

### 2.5.2 Houston, Texas Modeling Results

Table 2.5-2 summarizes the highest annual-average impacts for each modeled pollutant. The impacts listed in this table are the combined impacts of all source categories (major, area and mobile) in the Houston area. For benzene, 1,3-butadiene and POM, the largest contributions to the total impacts were due to the major sources, with the area and the mobile sources contributing only small fractions. For formaldehyde, the largest contributions were from the mobile sources, while the major and the area source contributions were significantly smaller. For hexavalent chromium, for which there were no mobile sources, the area sources account for almost all of the highest total concentration, while the major sources have an insignificant contribution. The contributions from the individual source categories, as well as further discussions on the modeling results, including isopleths of concentrations, can be found in Appendix B.

**Table 2.5-1 Highest Annual Average Concentrations from All Sources Combined for  
Phoenix, Arizona Based on 5 Modeled Years 1987 - 1991**

Pollutant	Highest Annual Average Concentration ( $\mu\text{g}/\text{m}^3$ )	Receptor Location <sup>(1)</sup> (X,Y) (meters)
Benzene	1.37	(396936, 707195)
1,3-Butadiene	0.17	(396936, 707195)
Formaldehyde	0.62	(393553, 707047)
POM <sup>(2)</sup>	0.18	(396936, 707195)
Chromium VI <sup>(3)</sup>	0.0001	(395455, 699171)
<p>(1) Receptor locations are in Universal Transverse Mercator (UTM) coordinates for Zone 15. While modeling, the first digit of the Y coordinate (North UTM) was removed.</p> <p>(2) POM = Polycyclic Organic Matter</p> <p>(3) Hexavalent Chromium</p>		

**Table 2.5-2 Highest Annual Average Concentrations from All Sources Combined for  
Houston, Texas Based on 5 Modeled Years 1987 - 1991**

Pollutant	Highest Annual Average Concentration ( $\mu\text{g}/\text{m}^3$ )	Receptor Location <sup>(1)</sup> (X,Y) (meters)
Benzene	10.41	(296660, 299970)
1,3-Butadiene	26.17	(281902, 287136)
Formaldehyde	2.13	(275162, 319329)
POM <sup>(2)</sup>	0.004	(305619, 292378)
Chromium VI <sup>(3)</sup>	0.11	(305619, 292378)
<p>(1) Receptor locations are in Universal Transverse Mercator (UTM) coordinates for Zone 15. While modeling, the first digit of the Y coordinate (North UTM) was removed.</p> <p>(2) POM = Polycyclic Organic Matter</p> <p>(3) Hexavalent Chromium</p>		

## 2.6 SUMMARY AND CONCLUSIONS

This section provides an overall summary of the study, including the major conclusions, based on the emissions data and the modeling results presented in the case study sections of this report.

The results of the modeling analyses show some significant differences between the two cities that were studied. For Phoenix, the mobile sources were clearly the dominant source of emissions for benzene, 1,3-butadiene, formaldehyde and POM; the mobile source emissions for these pollutants were an order of magnitude higher than area source emissions and two orders of magnitude higher than major source emissions. For hexavalent chromium, however, the emissions were due to cooling towers and other major sources. The mobile source emissions also exhibited the strongest temporal variations, reflecting the diurnal patterns in road traffic, as well as some influences of meteorology on emission estimates. These patterns in the emission inventory for Phoenix are also evident in the modeling results for that city. The mobile sources were clearly the largest contributors to the overall highest impacts. Also, since the majority of emissions were from mobile sources, they were distributed over the entire domain, with some spatial variability based on surrogate factors such as population. As a result, there was little evidence in the modeling results of localized "hot spots" (sharp gradients in concentrations over a relatively short distance) within the Phoenix domain.

While the Houston inventory showed emissions from mobile sources that were comparable to Phoenix, major source emissions from Houston were significantly higher than major source emissions for Phoenix. Benzene emissions from major sources were almost 100 times higher for Houston than for Phoenix, and 1,3-Butadiene emissions were about 700 times higher. Also, the major source emissions of benzene, 1,3-butadiene and POM were the same order of magnitude as the corresponding mobile source emissions. As a result of this, the major sources were found to be the largest contributors to the overall highest impacts in Houston for these pollutants and the modeling analysis for Houston does exhibit some significant "hot spots" associated with some of the larger sources of emissions.

These analyses illustrate a methodology that may be applied to similar urban-wide analyses of area sources. The use of a plume model has certain advantages over puff and grid models in terms of less stringent input data requirements, and plume models are also more easily applied to long term exposure analyses. Limited comparisons with modeled data were made where monitored data were available and are presented in Appendix B. Further insight into the applicability of such models on this scale of analysis may be gained in future studies involving more extensive comparisons of modeled concentrations and monitored values.

Although some of the factors that affected the pollutant impacts were identified as part of this study, there are other factors that were not fully analyzed. Therefore, it may be beneficial to conduct further studies. For example, a study of the affect of the temporal variations in emissions when coupled with temporal variations in meteorology could provide further insight into the observed concentration patterns.

### 3. REFERENCES

- 40CFR51. Guideline on Air Quality Models. Appendix W to 40CFR51.
- ENSR Consulting and Engineering, 1995. Arizona Hazardous Air Pollution Research Program Volumes 1 & 2. Arizona Department of Environmental Quality, Phoenix, Arizona.
- Fletcher, K., Strommen, M., and Kamens, R., 1997. Final Report on Parameters Relating to the Fates of Selected Atmospheric Pollutants. Department of Environmental Sciences and Engineering, UNC Chapel Hill.
- Gery, M. W. and Crouse, R. R., 1991. User's Guide for Executing OZIPR, U. S. Environmental Protection Agency, Research Triangle Park, NC.
- Holzworth, G. C., 1972. Mixing Heights, Wind speeds, and Potential for Urban Air Pollution Throughout the Contiguous United States, AP-101, U. S. Environmental Protection Agency, Research Triangle Park, NC.
- OTAG (Ozone Transport Assessment Group), 1998. Technical Supporting Document. The OTAG Final Report is available on the Internet at <http://www.epa.gov/ttn/rto/otag/finalrpt/>.
- NCDC (National Climatic Data Center), 1984. Comparative Climatic Data for the United States. Asheville, NC.
- NCDC (National Climatic Data Center), 1993. Solar and Meteorological Surface Observation Network, Version 1.0. Asheville, NC.
- NCDC (National Climatic Data Center), 1997a. Hourly United States Weather Observations. Asheville, NC.
- NCDC (National Climatic Data Center), 1997b. Radiosonde Data of North America. Asheville, NC.
- NCDC (National Climatic Data Center), 1998. International Surface Weather Observations. Asheville, NC.
- Radian Corp., 1995a. Development of the Houston Area Source Toxics Emissions (HASTE) Inventory. Prepared for Texas Natural Resources Conservation Commission.
- Radian Corp., 1995b. Air Quality Dispersion Modeling for the Houston Area Source Toxic Emissions (HASTE) Project. Prepared for Texas Natural Resources Conservation Commission.
- Sutton, O. G., 1953. Micrometeorology. McGraw-Hill, New York, NY



- TACB (Texas Air Control Board), 1992. Air Quality Modeling Guidelines. Texas Air Control Board, Austin, Texas.
- U.S. EPA (Environmental Protection Agency), 1993. Air/Superfund National Technical Guidance Study Series, Models for Estimating Air Emission Rates from Superfund Remedial Actions. Office of Air Quality Planning and Standards, Research Triangle Park, NC. EPA-451/R-93-001.
- U.S. EPA (Environmental Protection Agency), 1995a. Compilation of Air pollutant Emission Factors - Volume I: Stationary Point and Area Sources. AP-42. Office of Air Quality Planning and Standards, Research Triangle Park, NC.
- U.S. EPA (Environmental Protection Agency), 1995b. The Development and Improvement of Temporal Allocation Factor Files. Joint Emissions Inventory Oversight Group, Prepared for the Office of Air Quality Planning and Standards, prepared by Air and Energy Engineering Research Lab, Washington, DC. EPA-600/R-95-004.
- U.S. EPA (Environmental Protection Agency), 1995c. User's Guide for the Industrial Source Complex (ISC3) Dispersion Models. Office of Air Quality Planning and Standards, Research Triangle Park, NC. EPA-454/B-95-003b.
- U.S. EPA (Environmental Protection Agency), 1996a. Meteorological Processor for Regulatory Models (MPRM) User's Guide. Office of Air Quality Planning and Standards, Research Triangle Park, NC. EPA-454/B-96-002.
- U.S. EPA (Environmental Protection Agency), 1996b. PCRAMMET User's Guide. Office of Air Quality Planning and Standards, Research Triangle Park, NC. EPA-454/B-96-001.
- U.S. EPA (Environmental Protection Agency), 1999a. The Hazardous Air Pollution Exposure Model (HAPEM4 ) User's Guide, Draft, Office of Air Quality Planning and Standards, Research Triangle Park, NC.
- U.S. EPA (Environmental Protection Agency), 1999b. Integrated Urban Air Toxics Strategy to Comply with Section 112(d), 112(c)(3) and Section 202(I) of the Clean Air Act; Draft Notice. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

## **APPENDIX A**

### **AIR DISPERSION MODELING OF TOXIC POLLUTANTS IN URBAN AREAS**

#### **CASE STUDY FOR PHOENIX, ARIZONA**

## APPENDIX A - TABLE OF CONTENTS

A. CASE STUDY FOR PHOENIX, ARIZONA .....	A-1
A.1 INTRODUCTION .....	A-1
A.2 SUMMARY OF EMISSION INVENTORY FOR PHOENIX .....	A-2
A.3 OVERVIEW OF MODELING RESULTS .....	A-4
A.3.1 Annual Average Modeling Results .....	A-4
A.3.1.1 <i>Concentration Contributions by Source Category</i> .....	A-4
A.3.1.2 <i>Isopleths for Benzene</i> .....	A-5
A.3.1.3 <i>Isopleths for 1,3-Butadiene</i> .....	A-5
A.3.1.4 <i>Isopleths for Formaldehyde</i> .....	A-5
A.3.1.5 <i>Isopleths for Polycyclic Organic Matter (POM)</i> .....	A-6
A.3.1.6 <i>Isopleths for Hexavalent Chromium</i> .....	A-6
A.3.2 Modeling Results for Average Maximum Concentrations by Hour-of-Day .....	A-6
A.3.3 Maximum Seasonal Average Concentrations .....	A-7
A.3.4 Pollutant Decay .....	A-7
A.4 SUMMARY AND CONCLUSIONS .....	A-40
A.5 REFERENCES .....	A-41

## APPENDIX A - LIST OF TABLES

Table A.2-1 Toxic Air Pollutant Emissions for Phoenix, Arizona Based on Year 1993 . . . .	A-3
Table A.3-1 Highest Annual Average Concentrations from All Sources Combined for Phoenix, Arizona Based on 5 Modeled Years 1987 - 1991 . . . . .	A-7
Table A.3-2 Source Category Contributions to Total Annual Average Concentrations for Phoenix, Arizona . . . . .	A-8
Table A.3-3 Half-Life Decay Values . . . . .	A-9

## APPENDIX A - LIST OF FIGURES

Figure A.3-1. Modeled Receptor Location Urban Area Source Modeling, Phoenix, Arizona	A-10
Figure A.3-2. Population Density, Phoenix, Arizona	A-11
Figure A.3-3. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Benzene, All Sources (1987-1991)	A-12
Figure A.3-4. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Benzene, Major Sources (1987-1991)	A-13
Figure A.3-5. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Benzene, Area Sources (1987-1991)	A-14
Figure A.3-6. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Benzene, Mobile Sources (1987-1991)	A-15
Figure A.3-7. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
1,3-Butadiene, All Sources (1987-1991)	A-16
Figure A.3-8. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
1,3-Butadiene, Major Sources (1987-1991)	A-17
Figure A.3-9. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
1,3-Butadiene, Area Sources (1987-1991)	A-18
Figure A.3-10. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
1,3-Butadiene, Mobile Sources (1987-1991)	A-19
Figure A.3-11. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Formaldehyde, All Sources (1987-1991)	A-20
Figure A.3-12. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Formaldehyde, Major Sources (1987-1991)	A-21
Figure A.3-13. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Formaldehyde, Area Sources (1987-1991)	A-22
Figure A.3-14. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Formaldehyde, Mobile Sources (1987-1991)	A-23
Figure A.3-15. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
POM, All Sources (1987-1991)	A-24
Figure A.3-16. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
POM, Area Sources (1987-1991)	A-25
Figure A.3-17. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
POM, Mobile Sources (1987-1991)	A-26
Figure A.3-18. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Hexavalent Chromium, All Sources (1987-1991)	A-27
Figure A.3-19. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Hexavalent Chromium, Major Sources (1987-1991)	A-28
Figure A.3-20. Isopleths of Annual Average Concentrations, Phoenix, Arizona	
Hexavalent Chromium, Area Sources (1987-1991)	A-29
Figure A.3-21. Urban Area Source Modeling, Phoenix, Arizona, Benzene Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data	A-30
Figure A.3-22. Urban Area Source Modeling, Phoenix, Arizona, 1,3-Butadiene Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data	A-31

Figure A.3-23. Urban Area Source Modeling, Phoenix, Arizona, Formaldehyde Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-32
Figure A.3-24. Urban Area Source Modeling, Phoenix, Arizona, POM Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data ....	A-33
Figure A.3-25. Urban Area Source Modeling, Phoenix, Arizona, Hexavalent Chromium Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-34
Figure A.3-26. Urban Area Source Modeling, Phoenix, Arizona, Benzene Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-35
Figure A.3-27. Urban Area Source Modeling, Phoenix, Arizona, 1,3-Butadiene Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-36
Figure A.3-28. Urban Area Source Modeling, Phoenix, Arizona, Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-37
Figure A.3-29. Urban Area Source Modeling, Phoenix, Arizona, POM Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-38
Figure A.3-30. Urban Area Source Modeling, Phoenix, Arizona, Hexavalent Chromium Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	A-39

## **A. CASE STUDY FOR PHOENIX, ARIZONA**

### **A.1 INTRODUCTION**

This appendix documents the modeling methodology employed by the EPA in estimating ambient air concentrations of selected toxic pollutants for the city of Phoenix, Arizona. The pollutants modeled were benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium. The modeling study serves as an example of guidance on the application of dispersion models to the assessment of exposure to toxic pollutants on an urban scale.

## A.2 SUMMARY OF EMISSION INVENTORY FOR PHOENIX

Five pollutants were included in this modeling study; benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM) and hexavalent chromium. The emissions for these pollutants were extracted from an air toxics emission inventory developed by the Arizona Department of Environmental Quality and the Arizona Department of Health Services as part of the Arizona Hazardous Air Pollution Research Program (ENSR, 1995). The original inventory, which was based on the emissions for the year 1993, included gridded emissions of 163 species of pollutants on a 4km-by-4km grid resolution covering the regions around Phoenix, Tucson, Casa Grande, and Payson, Arizona. The gridded inventory incorporated all source categories, including point, area and mobile sources. Temporal variations of the emissions on a season by hour-of-day basis were also included in the inventory.

A subset of the original inventory was selected to cover the Phoenix area for this analysis. This subset consisted of a rectangular domain extending from 349 kilometers to 485 kilometers Easting and 3,663 kilometers to 3,763 kilometers Northing in Universal Transverse Mercator (UTM) coordinates in UTM Zone 12.

The combined total 1993 emissions of the five modeled pollutants was 3,374 megagrams (3,719 tons) in the Phoenix metropolitan area. The largest of these five pollutants was benzene (1,822 Mg or 2,008 tons), followed by formaldehyde (1,024 Mg or 1,129 tons), 1,3-butadiene (269 Mg or 296 tons), POM (258 Mg or 285 tons) and hexavalent chromium (0.11 Mg or 0.12 ton). A breakdown of emissions of each pollutant by source category (major, area and mobile) is presented in Table A.2-1.

As shown in Table A.2-1, the mobile sources account for the largest emissions for all pollutants except hexavalent chromium. Area sources (i.e., small stationary sources which are too numerous and diverse to be counted as individual point sources) contributed much less than mobile sources, and major sources contributed less than 1% of the total pollutant emissions, where applicable, except for hexavalent chromium (25%). Within the mobile sources, the light-duty gasoline vehicles and trucks and the gasoline-powered lawn and garden equipment were among the largest contributors.

There is uncertainty about the POM and hexavalent chromium emissions. POM may have been undercounted in the Phoenix study because emissions of POM from gasoline-powered off-road sources were not reported. The study reported emissions of hexavalent chromium from cooling towers (which are being phased-out), but did not report emissions from chrome-plating facilities or surface coating operations.



**Table A.2-1**

**Toxic Air Pollutant Emissions for Phoenix, Arizona  
Based on Year 1993**

Source Category		Pollutant Emissions (Mg/yr)				
		Benzene	1,3-Butadiene	Formaldehyde	POM <sup>(1)</sup>	CrVI <sup>(2)</sup>
Mobile Sources	On-Road	1,003	100	605	195	-- <sup>(3)</sup>
	Off-Road	727	165	392	34.4	-- <sup>(3)</sup>
	Total Mobile	1,730	265	997	229	-- <sup>(3)</sup>
Area Sources		79.5	3.5	20.4	29.2	0.082
Major Sources		12.8	0.4	7.1	-- <sup>(3)</sup>	0.028
TOTAL		1,822.4	269	1,024.5	258	0.11
<p>(1) POM = Polycyclic Organic Matter. Individual constituents of POM vary. For mobile source, POM is defined as sum of 16 chemicals. For area &amp; mobile sources, data provided by TNRCC do not define POM constituents. Data reported as POM by TNRCC are used here.</p> <p>(2) CrVI = Hexavalent Chromium</p> <p>(3) Not Applicable</p>						

### A.3 OVERVIEW OF MODELING RESULTS

This section presents and discusses the results of the modeling analysis conducted for the sources located in Phoenix, Arizona. As discussed earlier, the five pollutants that were included in this study are benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium.

A total of 927 sources of benzene were modeled using the ISCST3 model. This number consisted of 87 major sources and 840 area/mobile sources. For 1,3-butadiene, a total of 859 sources were modeled including 19 major sources and 840 area/mobile sources. For formaldehyde, a total of 888 sources were modeled including 48 major sources and 840 area/mobile sources. For POM, a total of 589 sources were modeled all of which were area/mobile sources. For hexavalent chromium, a total of 236 sources were modeled including 28 major sources and 208 area sources. There were no mobile sources for hexavalent chromium.

A total of 356 receptors were modeled. Figure A.3-1 shows the modeled receptor locations. The modeling was conducted using the five year period 1987 through year 1991 National Weather Service surface meteorological data from Phoenix Airport with mixing heights from Tucson, Arizona. Figure A.3-2 shows the population density for Phoenix.

Both the annual average concentrations as well as the seasonal average concentrations by hour of day were calculated. Results for both of these averaging periods, and a discussion on the contributions to the total annual average concentrations from each source category (major, area and mobile) are presented below for all pollutants modeled.

The effects of pollutant decay on predicted concentrations were included in this analysis for three modeled pollutants; 1,3-butadiene, formaldehyde, POM. Hexavalent chromium was not modeled with decay due to its particulate nature, and benzene was not modeled with decay due to its long half-life. The modeling results presented below for these pollutants are based on seasonally variable (cold vs warm) half-lives for each pollutant. A discussion on the effects of decay is also presented in this section.

It should be noted that low mixing heights, i.e., less than 100 meters, occurred in less than 2.5 percent of all the hours during the five year period that was modeled. Thus the need to adjust the hourly values up to 100 meters arose quite infrequently.

#### A.3.1 Annual Average Modeling Results

Table A.3-1 presents the highest annual average concentrations for each of the five pollutants. The corresponding receptor locations are also listed in the table. The listed concentrations represent the combined total concentrations of all modeled sources for each pollutant.

##### A.3.1.1 *Concentration Contributions by Source Category*

Table A.3-2 presents the contributions of each source category (major, area and mobile) to the highest annual concentrations listed in Table A.3-1. As can be seen from this table, the

mobile sources have the largest contribution to the total impacts. The next largest contribution is from the area sources. The major sources account for only a small fraction of the total impacts.

For comparison purposes, Table A.3-2 also presents the highest concentrations of all pollutants due to each of the source categories separately. These results show that, out to the three source categories, the mobile sources cause the highest concentrations, the area sources cause the next highest concentrations and the major sources have the smallest concentrations. However, given that the receptor placement for this study was based on population density, it is possible that there may not be a receptor close enough to a large source to “capture” its maximum impacts.

#### *A.3.1.2 Isopleths for Benzene*

Figures A.3-3 through A.3-6 show the isopleths of concentration for benzene.<sup>1</sup> Figure A.3-3 shows the isopleths for the highest total concentrations due to all sources. Figures A.3-4, A.3-5, and A.3-8 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. Similar to the results shown in Table A.3-2, the major source category is the largest contributor (accounting for two-thirds or greater contribution) to the total concentrations. By contrast, the mobile and area sources contribute only a small fraction (2% or less) in most of the modeling domain.

#### *A.3.1.3 Isopleths for 1,3-Butadiene*

Figures A.3-7 through A.3-10 show the isopleths of concentration for 1,3-butadiene. Figure A.3-7 shows the isopleths for the highest total concentrations due to all sources. Figures A.3-8, A.3-9, and A.3-10 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. Similar to the results shown in Table A.3-2, the mobile source category is the largest contributor to the total concentrations (accounting for two-thirds or greater contribution in the central part of the domain and almost 100% contribution near the edges of the domain). By contrast, the major sources contribute only a small fraction (accounting for 1% or less contribution) in most of the modeling domain.

#### *A.3.1.4 Isopleths for Formaldehyde*

Figures A.3-11 through A.3-14 show the isopleths of concentration for formaldehyde. Figure A.3-11 shows the isopleths for the highest total concentrations due to all sources. Figures A.3-12, A.3-13, and A.3-14 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. Again, the mobile source category is the largest contributor to the total concentrations (accounting for two-thirds or greater contribution). By contrast, the major sources contribute only a small fraction (accounting for 0.5% or less contribution) in most of the modeling domain.

---

<sup>1</sup>Isopleth contours should be viewed with caution because this shape is also dependent on the software package used.

#### A.3.1.5 *Isopleths for Polycyclic Organic Matter (POM)*

Figures A.3-15 through A.3-17 show the isopleths of concentration for POM. Figure A.3-15 shows the isopleths for the highest total concentrations due to all sources. Figures A.3-16 and A.3-17 show the isopleths for the highest concentrations for each source category, i.e., area and mobile, respectively. Note that there were no major sources for POM. Again, the mobile source category contributes a majority of total concentrations, accounting for approximately two-thirds or greater contribution. The area sources contribute approximately one-third or less to the total concentrations.

#### A.3.1.6 *Isopleths for Hexavalent Chromium*

Figures A.3-18 through A.3-20 show the isopleths of concentration for hexavalent chromium. Figure A.3-19 shows the isopleths for the highest total concentrations due to all sources. Figures A.3-19 and A.3-20 show the isopleths for the highest concentrations for each source category, i.e., major and area, respectively. Note that there were no mobile sources for hexavalent chromium. A comparison of Figures A.3-18 and A.3-20 shows that the major sources contribute the most to the total concentrations in most of the modeling domain. However, as was shown in Table A.3-2, both the major and the area source categories contribute equally to the highest total hexavalent chromium concentration.

#### A.3.2 Modeling Results for Average Maximum Concentrations by Hour-of-Day

The modeling results for the annual average maximum concentrations by hour-of-day from the entire modeling domain are presented in a series of figures. These values represent the maximum concentration of each of the four seasonal averages, averaged over the five year period. These figures show the temporal variation of the annual average concentrations. A figure was generated for each pollutant showing the concentrations due to all sources and for each source category (major, area and mobile).

Figures A.3-21 through A.3-25 shows the temporal variation curves for benzene, 1-3 butadiene, formaldehyde, POM, and hexavalent chromium, respectively.

With the exception of the figures for the major source category, all figures show a distinct peak during the morning hours (7-8 a.m.) and a distinct peak during the evening hours (7-8 p.m.). The peaks generally occur at the same time for all four seasons. There is also a slight increase in concentrations observed between the hours of 12 noon and 3 p.m. Although a detailed study would be necessary to determine the exact reason(s) for the patterns seen, the most likely reason is the temporal variation of emissions from mobile sources, as well as, related to meteorological conditions during these hours, as well as. For example, mobile source emissions are higher during the morning and evening rush hours, and will contribute to the morning and evening peaks. Since a majority of the emissions are from low-level releases, the peaks may also be related to an increase in concentrations due to more stable atmospheric conditions at night.

For the major sources, although there is a general trend that shows high concentrations during the morning and evening hours, as compared to the rest of the day, the high

concentrations consist of several peaks. Further analysis would be required to explain these occurrences in detail.

### A.3.3 Maximum Seasonal Average Concentrations

From Figures A.3-26 through A.3-30 showing seasonal averages, it can also be noted that, generally, the concentrations of all pollutants during winter and fall seasons are higher than spring and summer seasons. This may be associated with the seasonal variations in the operation of certain types of sources. For example, residential wood combustion occurs primarily during winter with practically no wood combustion during the summer. The only exceptions to this are the major sources which do not show any apparent seasonal trend.

Similar to the annual average concentrations, the total seasonal average concentrations are also dominated by mobile sources. For example, a comparison of Figure A.3-21 and A.3-22 shows that the mobile sources contribute approximately 75% to the observed evening peak. By contrast, major sources contribute only a small fraction to the total concentrations.

### A.3.4 Pollutant Decay

As noted above, pollutant decay was modeled for 1,3-butadiene, formaldehyde and POM. Pollutant decay was not modeled for benzene because of its long half life or for hexavalent chromium because of its particulate nature. Decay has been defined seasonally (cold versus warm). With the study area being located in the southern latitudes, winter has been designated the cold season with spring, summer and fall designated warm seasons. Appropriate half-life decay values, shown in Table A.3-3 have been assigned to pollutants using these criteria.

**TABLE A.3-1 Highest Annual Average Concentrations from All Sources Combined for Phoenix, Arizona Based on 5 Modeled Years 1987 - 1991**

<b>Pollutant</b>	<b>Highest Annual Average Concentration (ug/m<sup>3</sup>)</b>	<b>Receptor Location <sup>(1)</sup> (X,Y) (meters)</b>
Benzene	1.37	(396936, 707195)
1,3-Butadiene	0.17	(396936, 707195)
Formaldehyde	0.62	(393553, 707047)
POM <sup>(2)</sup>	0.18	(396936, 707195)
Chromium IV <sup>(3)</sup>	0.0001	(395455, 699171)
<p>(1) Receptor locations are in Universal Transverse Mercator (UTM) coordinates for Zone 15. While modeling, the first digit of the Y coordinate (North UTM) was removed.</p> <p>(2) POM = Polycyclic Organic Matter</p> <p>(3) CrVI = Hexavalent Chromium</p>		

**TABLE A.3-2 Source Category Contributions to Total Annual Average Concentrations  
for Phoenix, Arizona**

<b>Pollutant</b>	<b>Concentrations for All Sources combined (ug/m<sup>3</sup>)</b>				<b>Highest Concentrations for Individual Source Categories<sup>(2)</sup> (ug/m<sup>3</sup>)</b>		
	<b>Highest Total Concentrations</b>	<b>Contributions to the Total Concentrations <sup>(1)</sup></b>			<b>Major</b>	<b>Area</b>	<b>Mobile</b>
		<b>Major</b>	<b>Area</b>	<b>Mobile</b>			
Benzene	1.37	0.0046	0.031	1.34	0.13	0.033	1.34
1,3-Butadiene	0.17	0.000056	0.002	0.17	0.0047	0.0021	0.17
Formaldehyde	0.62	0.00041	0.017	0.604	0.0034	0.0178	0.604
POM <sup>(3)</sup>	0.18	N/A <sup>(4)</sup>	0.025	0.15	N/A <sup>(4)</sup>	0.026	0.15
Chromium VI <sup>(3)</sup>	0.00015	0.000074	0.000074	N/A <sup>(4)</sup>	0.000074	0.000114	N/A <sup>(4)</sup>
<p>(1) Represent the contributions from the individual source categories to the listed total concentrations for all sources combined.</p> <p>(2) Represent the highest concentrations due to the individual source categories. The locations of these concentrations may be different from the locations of the highest total concentrations for all sources combined.</p> <p>(3) POM = Polycyclic Organic Matter, Chromium VI = Hexavalent Chromium</p> <p>(4) N/A = Not Applicable</p>							

**Table A.3-3 Half-Life Decay Values**

<b>Pollutant</b>	<b>Half-Life (hours) Cold Season (Winter)</b>	<b>Half-Life (hours) Warm Seasons (Spring, Summer, Fall)</b>
Benzene <sup>(1)(2)</sup>	1560	144
1,3-Butadiene <sup>(1)</sup>	8	2
Formaldehyde <sup>(1)</sup>	6	2
POM <sup>(2)</sup>	6	1.2
<p>(1) From Appendix E, Table E.1</p> <p>(2) From Appendix E, Table E.3. using the shortest reported value</p> <p>(3) Due to long half-life value, pollutant decay was not used for this pollutant.</p>		

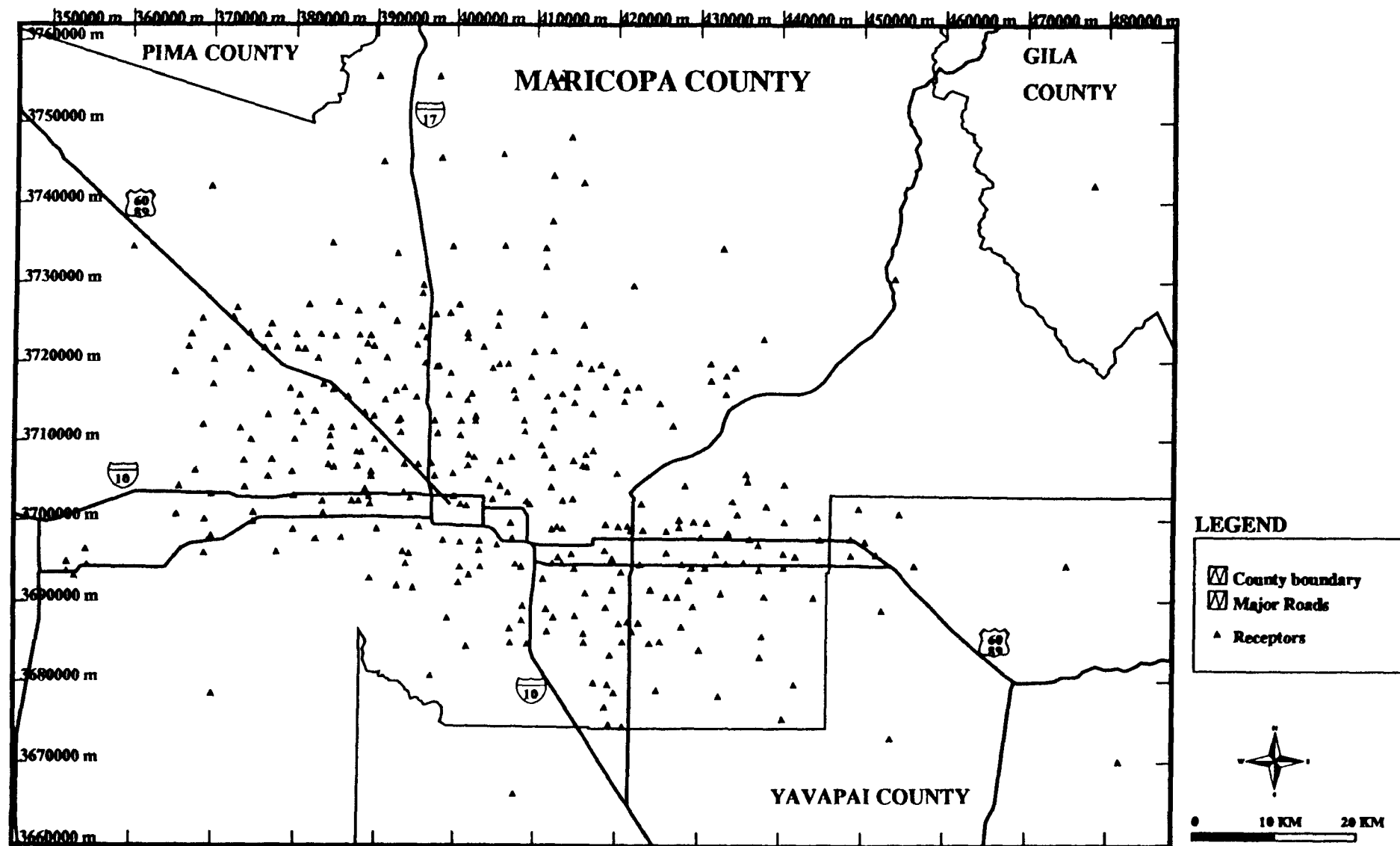


Figure A.3-1. Modeled Receptor Location Urban Area Source Modeling, Phoenix, Arizona



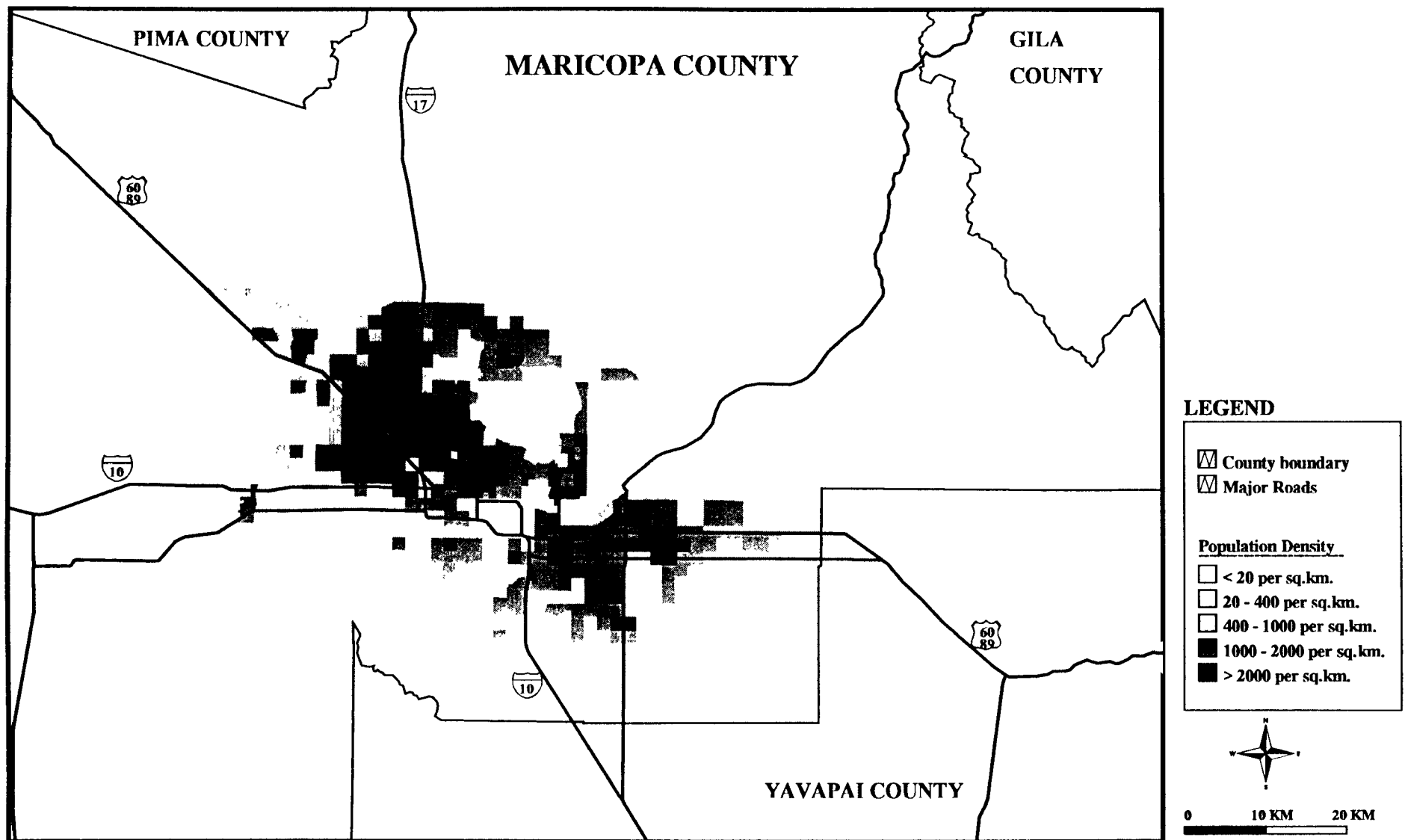


Figure A.3-2. Population Density, Phoenix, Arizona

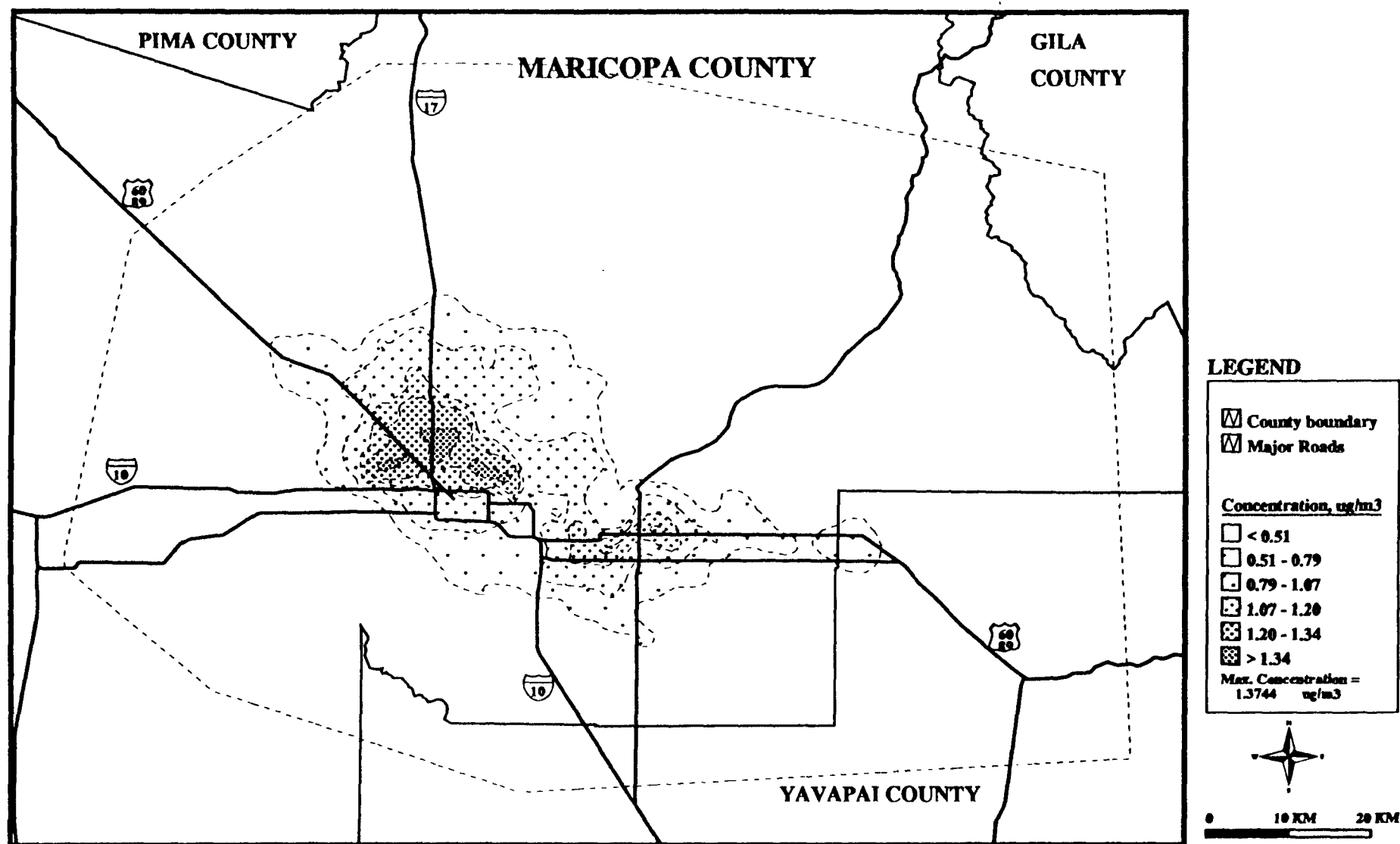
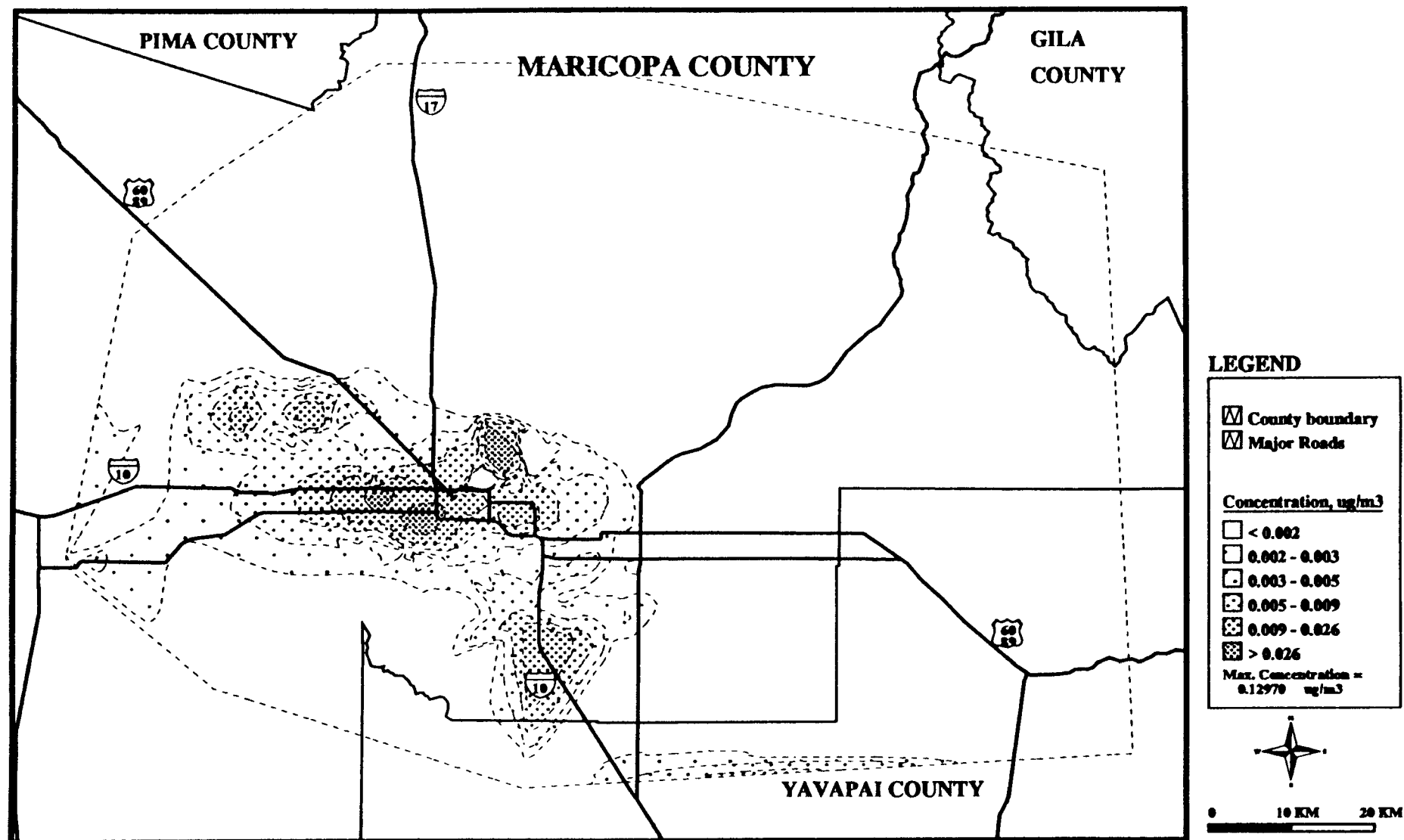


Figure A.3-3. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Benzene, All Sources (1987-1991)



**Figure A.3-4. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Benzene, Major Sources (1987-1991)**

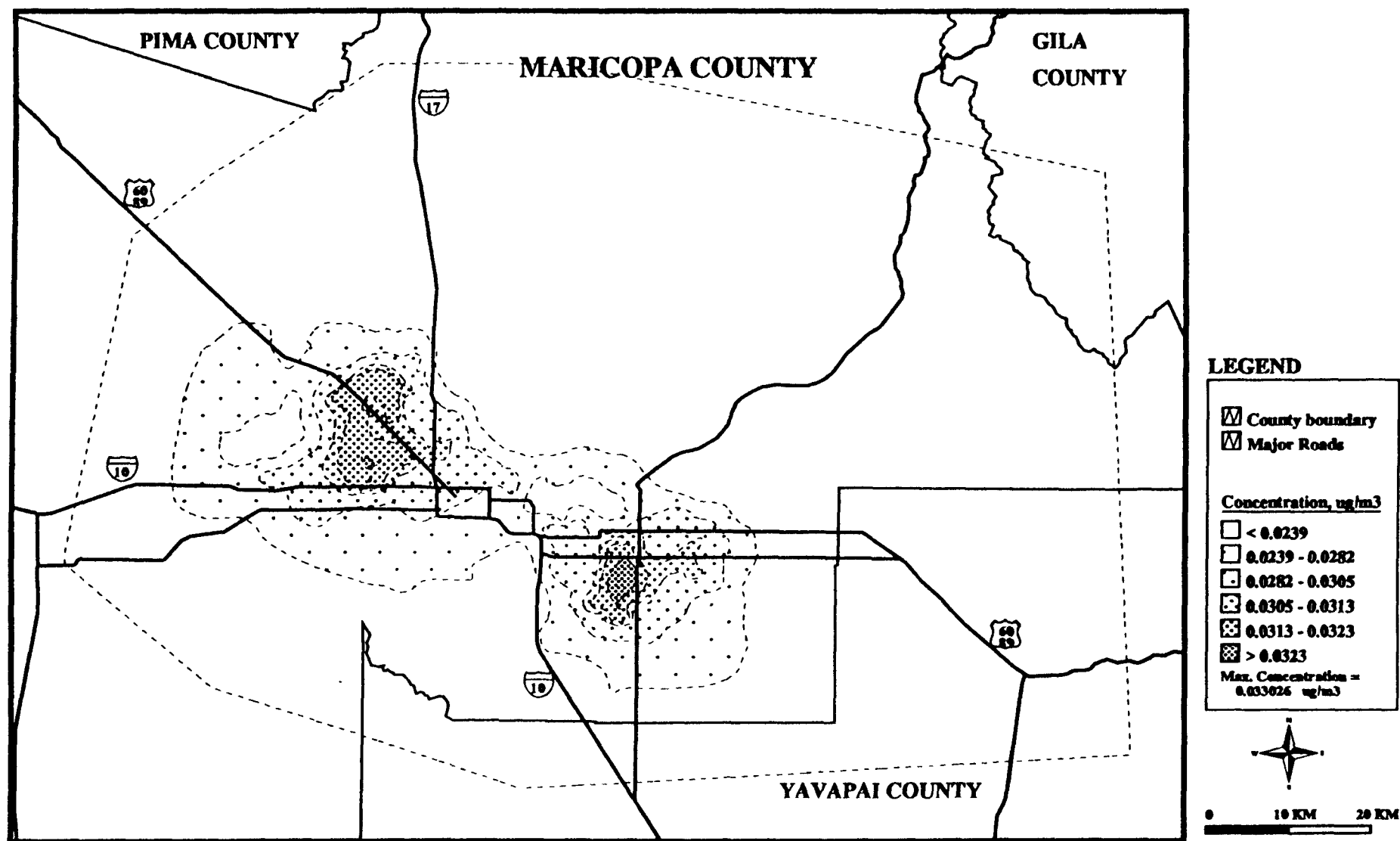


Figure A.3-5. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Benzene, Area Sources (1987-1991)

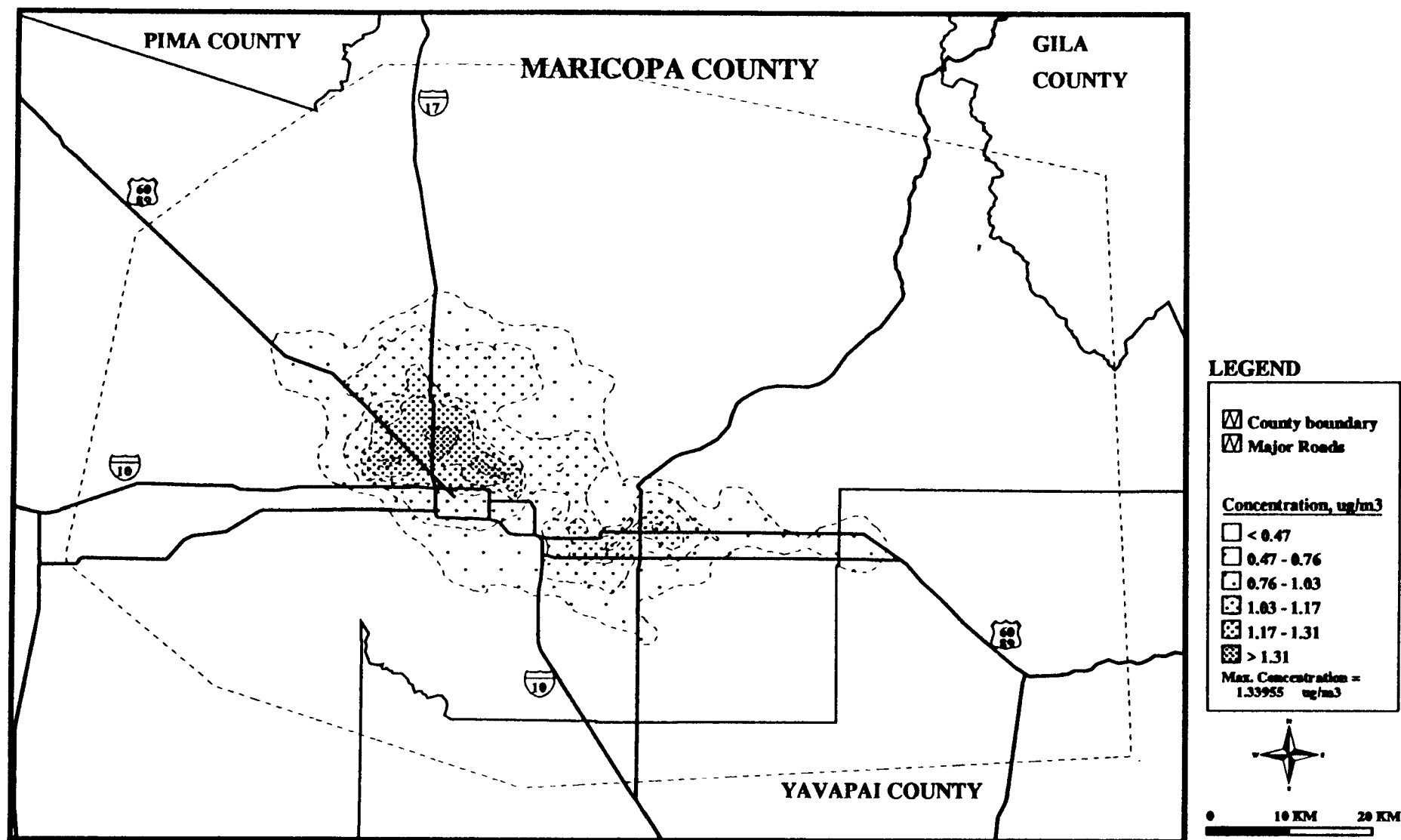


Figure A.3-6. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Benzene, Mobile Sources (1987-1991)

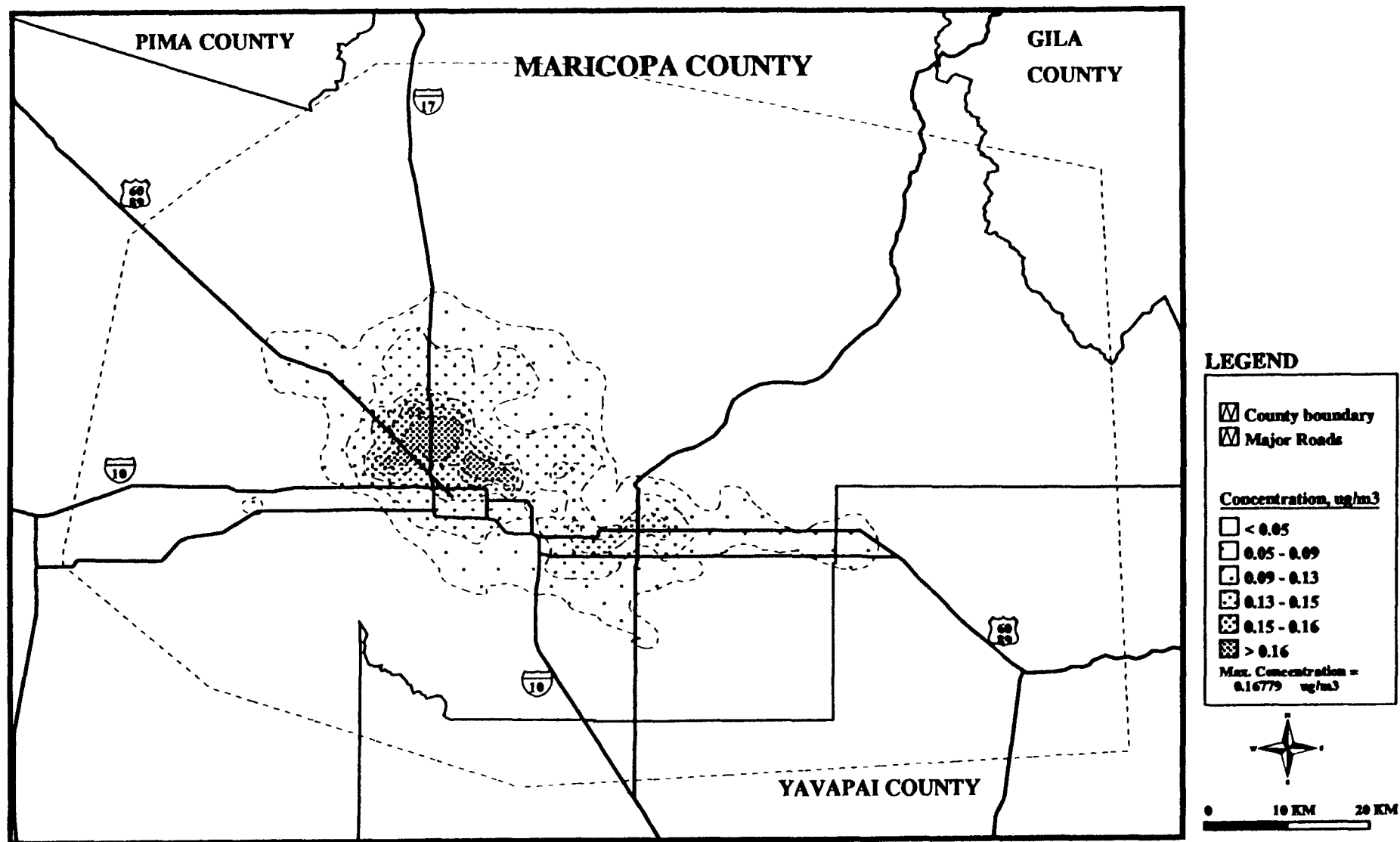


Figure A.3-7. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
1,3-Butadiene, All Sources (1987-1991)

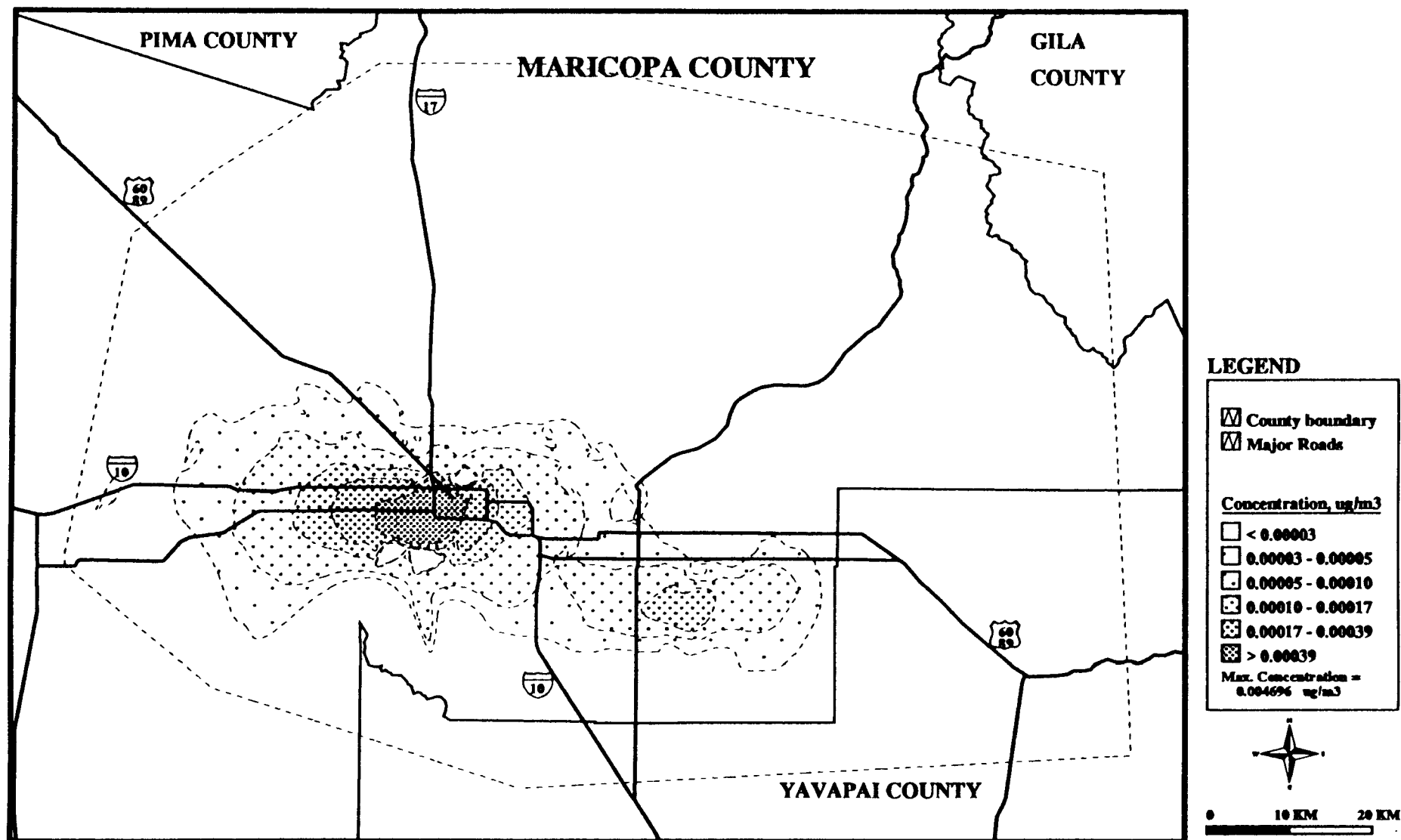


Figure A.3-8. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
1,3-Butadiene, Major Sources (1987-1991)

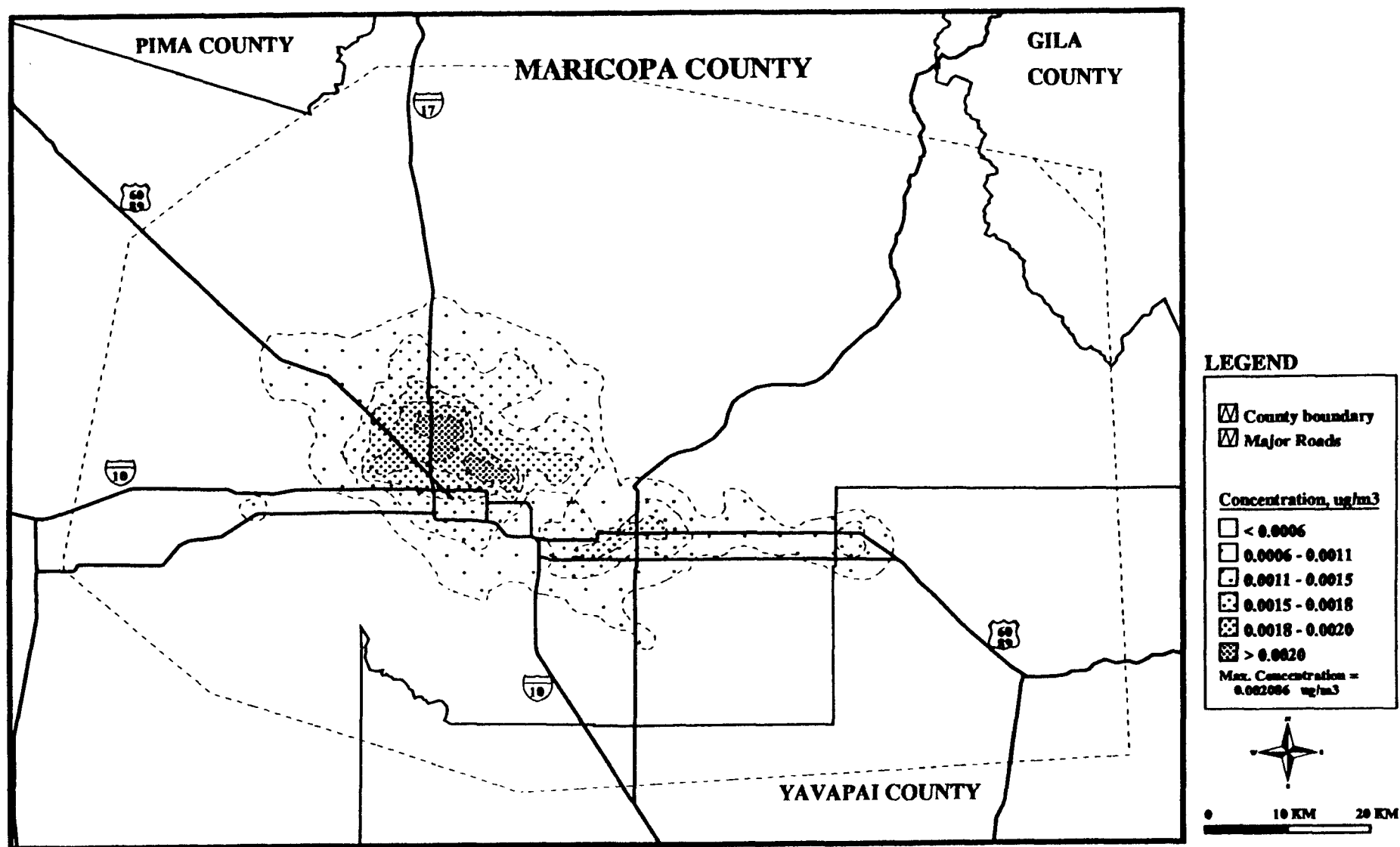


Figure A.3-9. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
1,3-Butadiene, Area Sources (1987-1991)



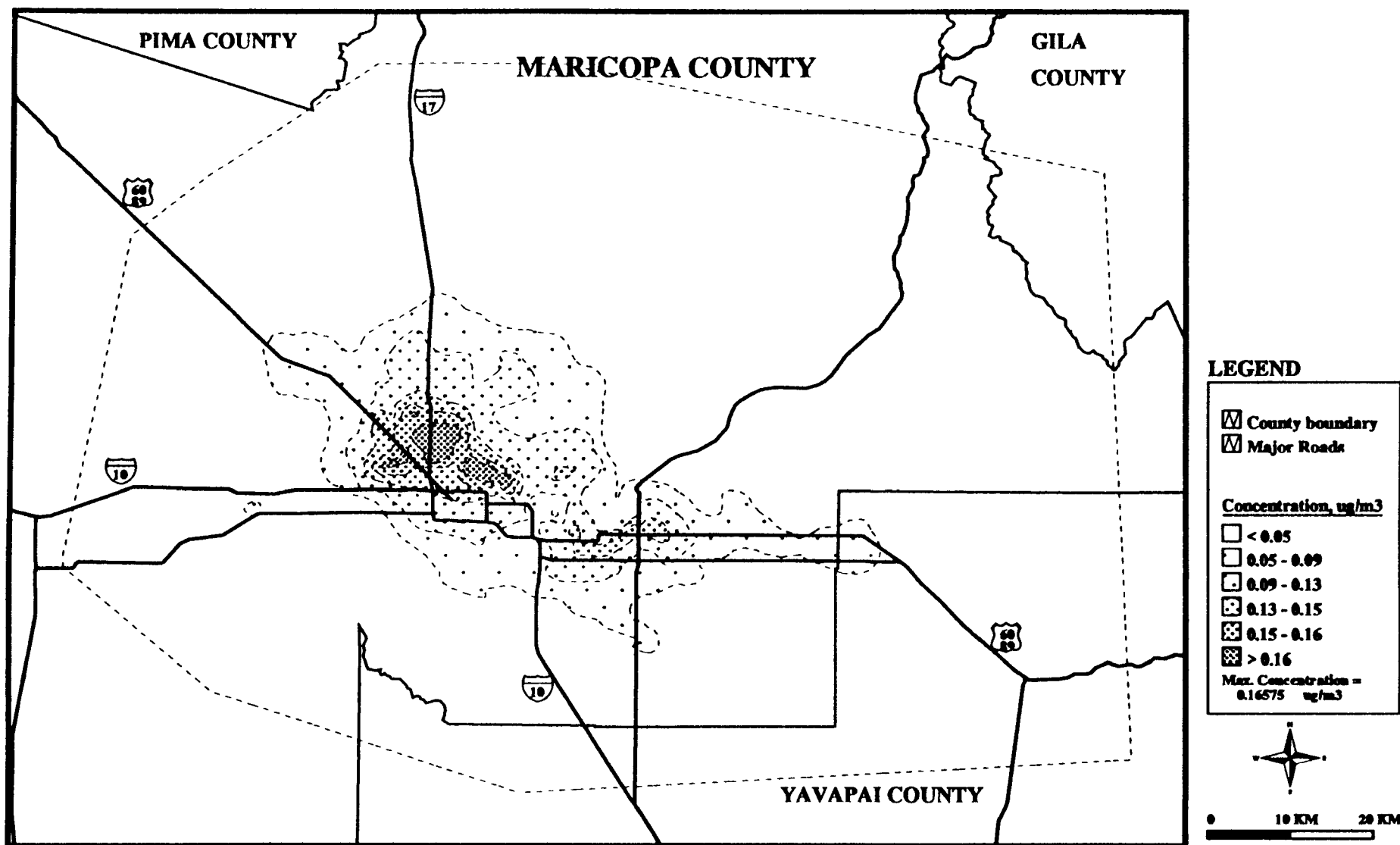


Figure A.3-10. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
1,3-Butadiene, Mobile Sources (1987-1991)

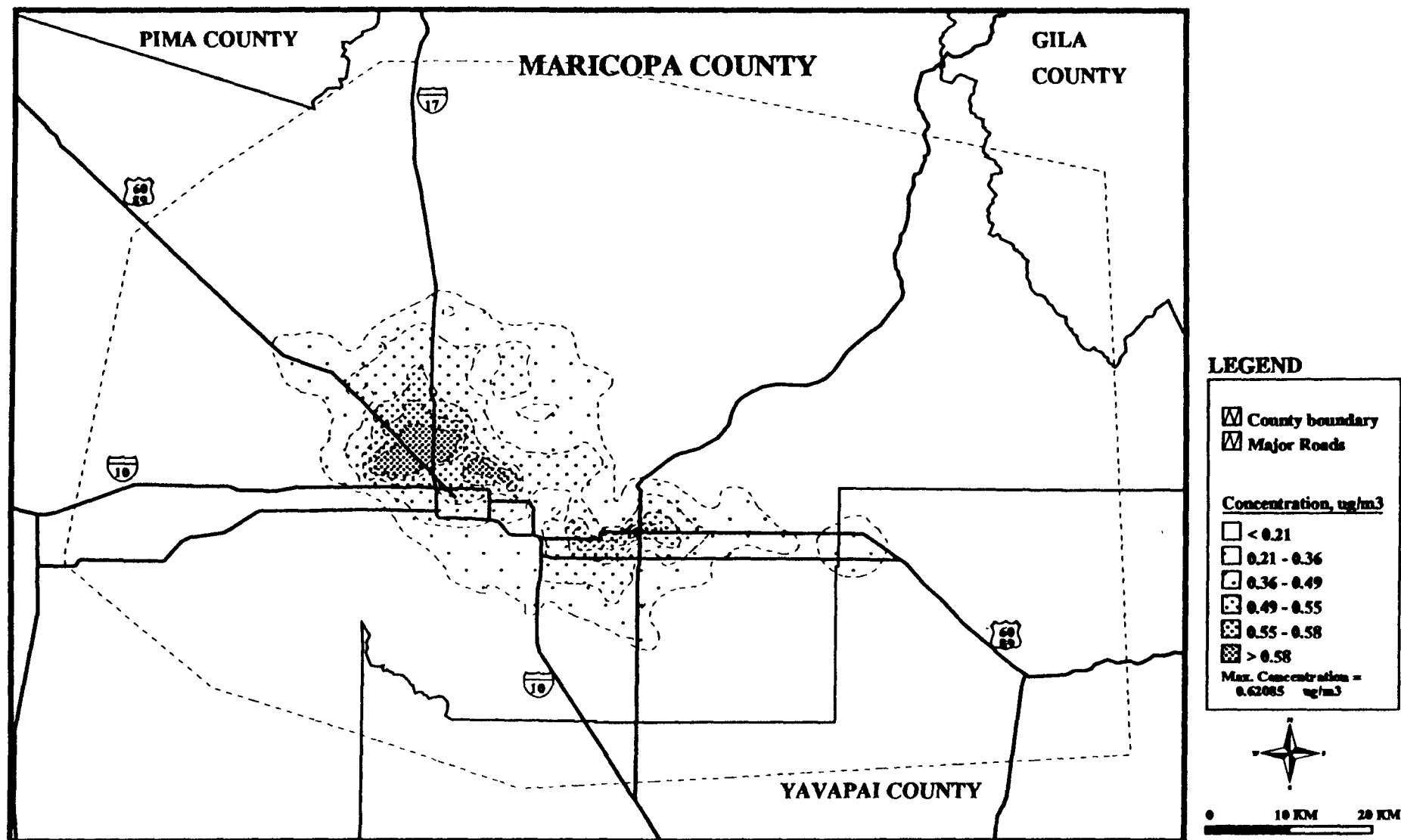
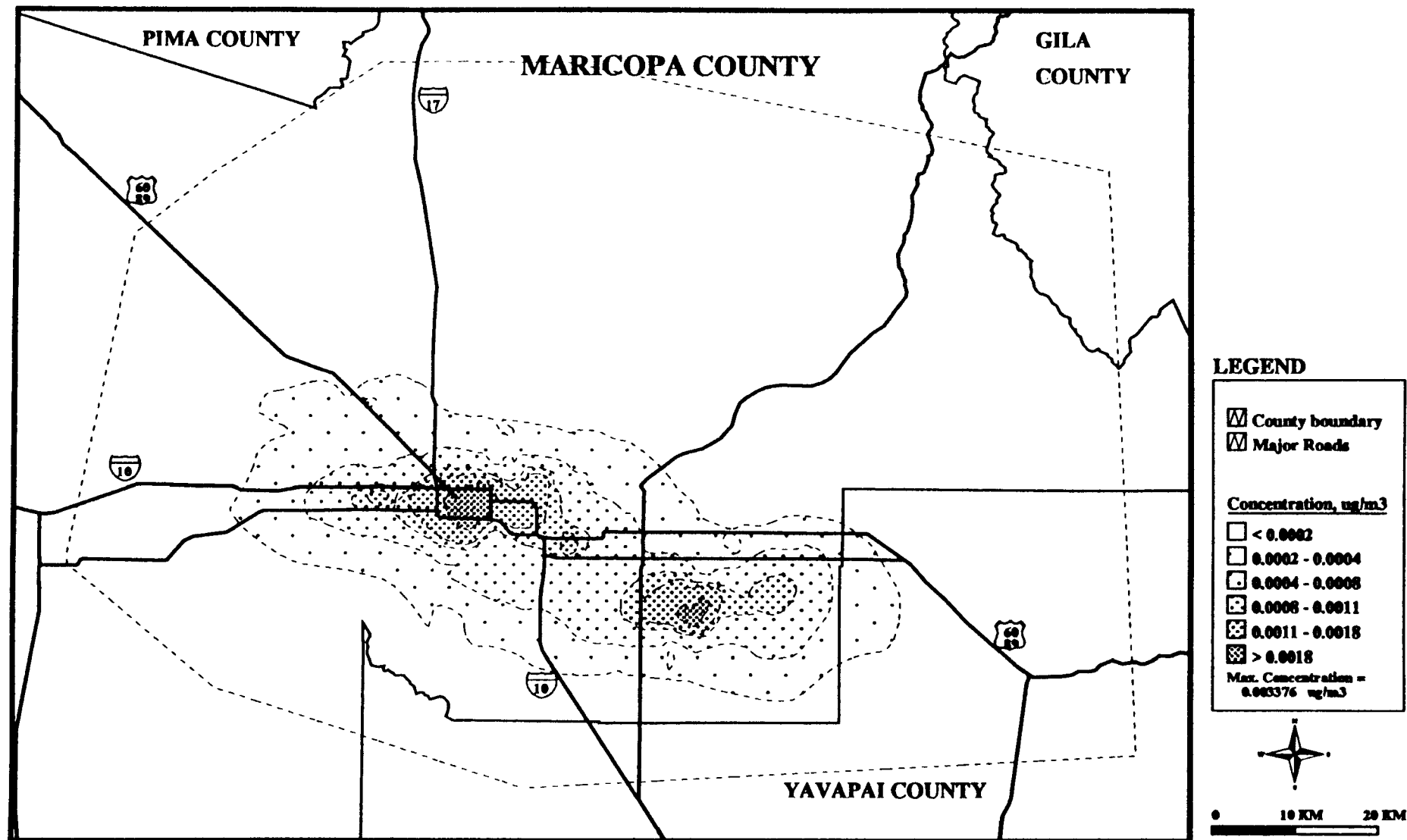


Figure A.3-11. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Formaldehyde, All Sources (1987-1991)



**Figure A.3-12. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Formaldehyde, Major Sources (1987-1991)**

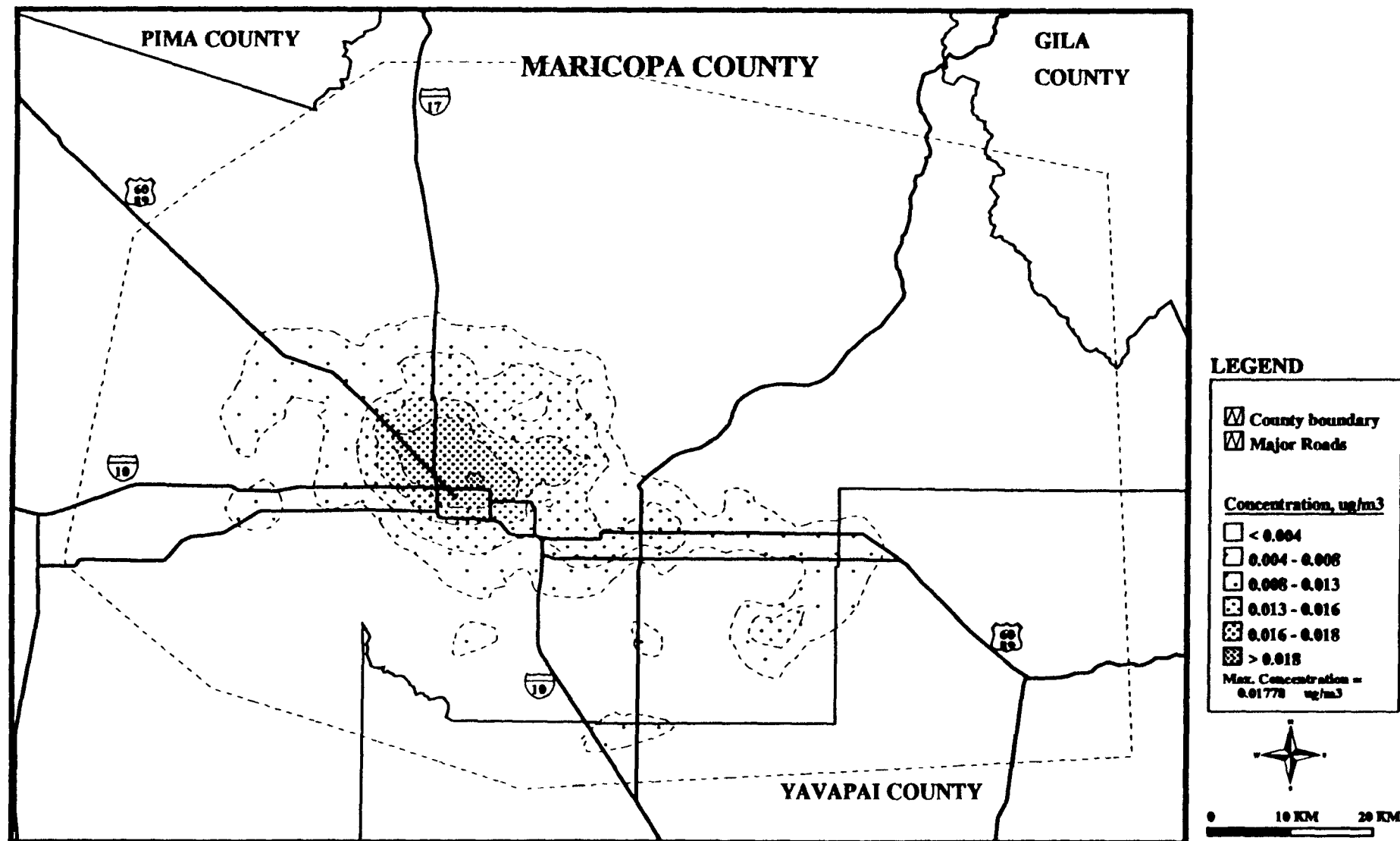


Figure A.3-13. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Formaldehyde, Area Sources (1987-1991)

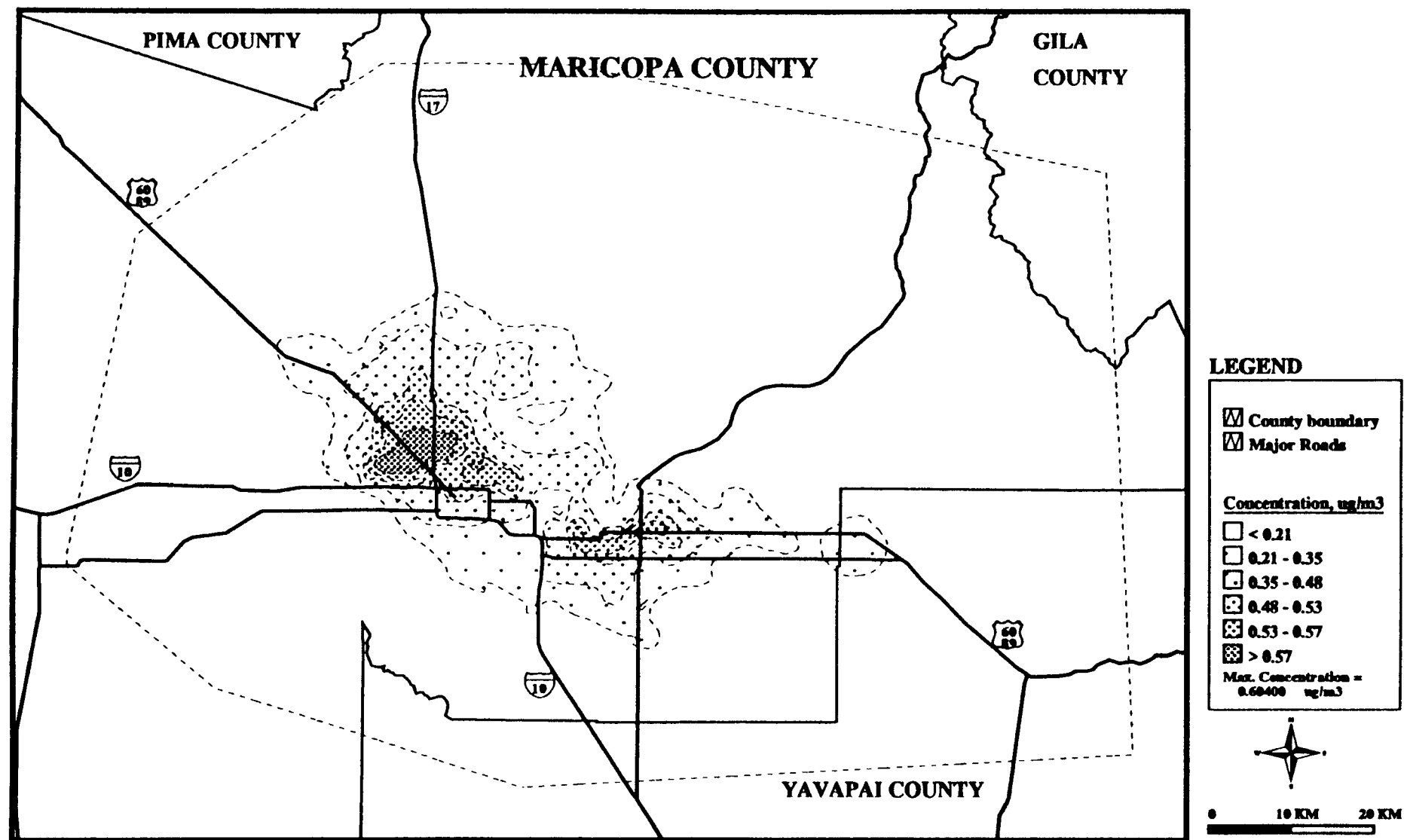


Figure A.3-14. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Formaldehyde, Mobile Sources (1987-1991)

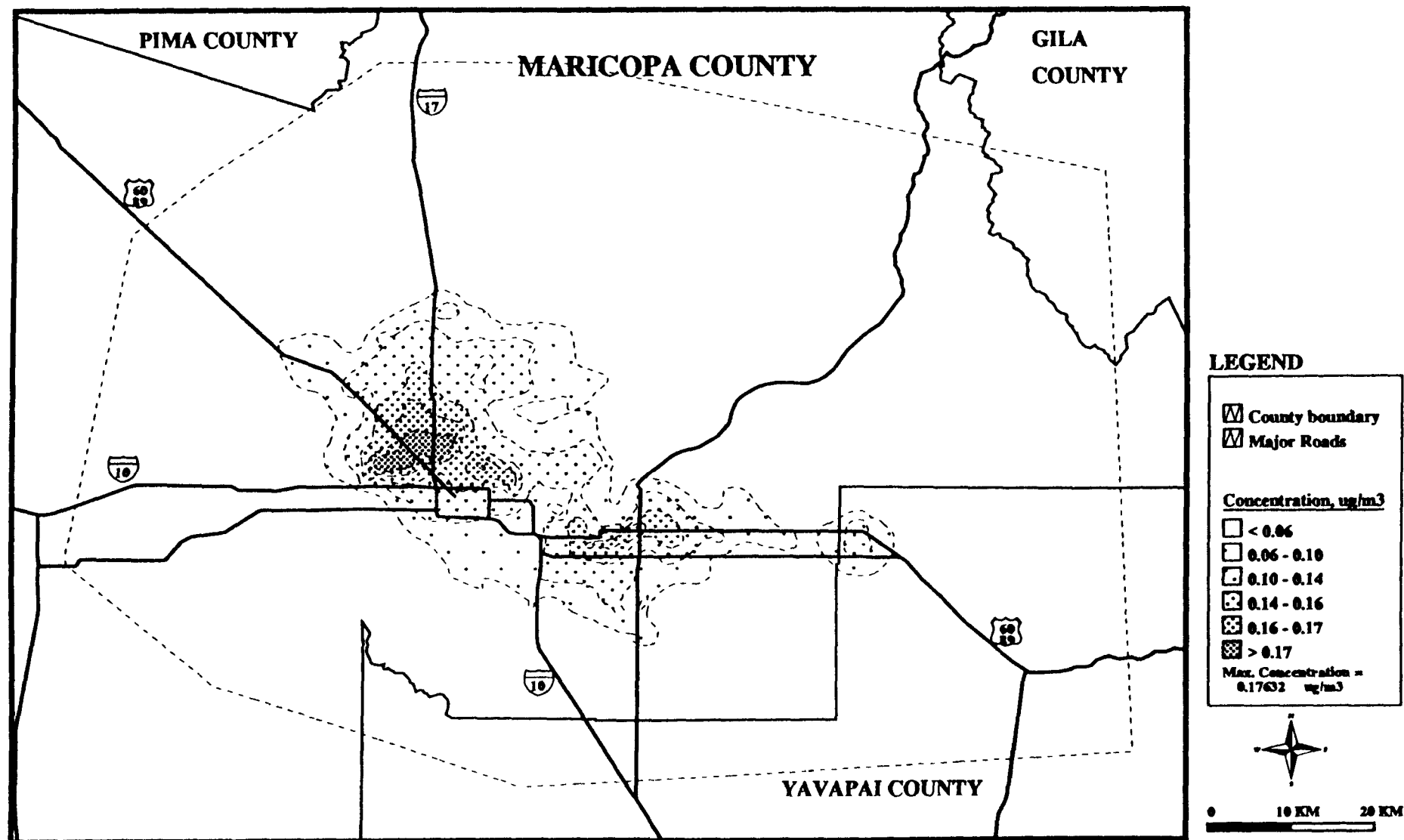


Figure A.3-15. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
POM, All Sources (1987-1991)

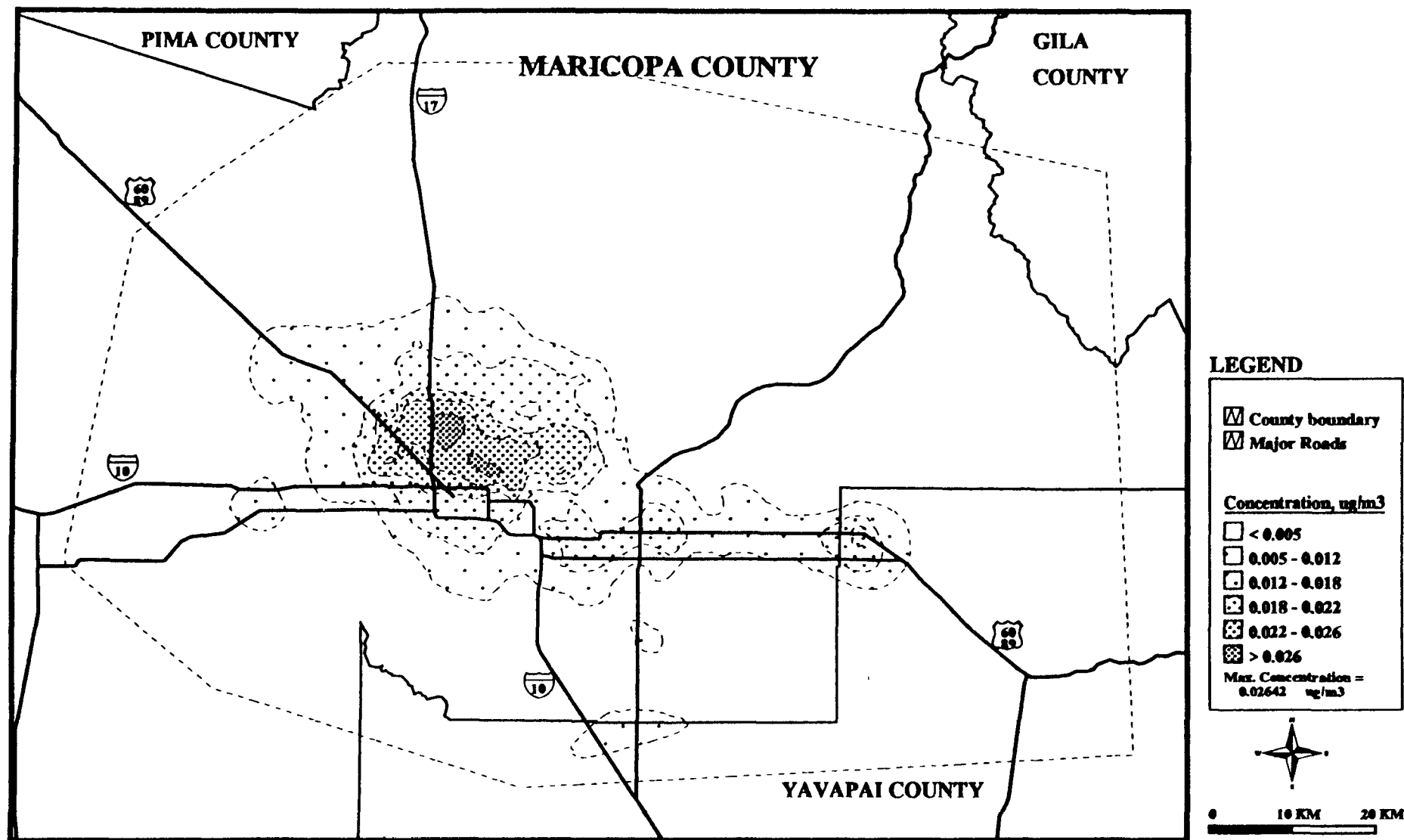


Figure A.3-16. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
POM, Area Sources (1987-1991)

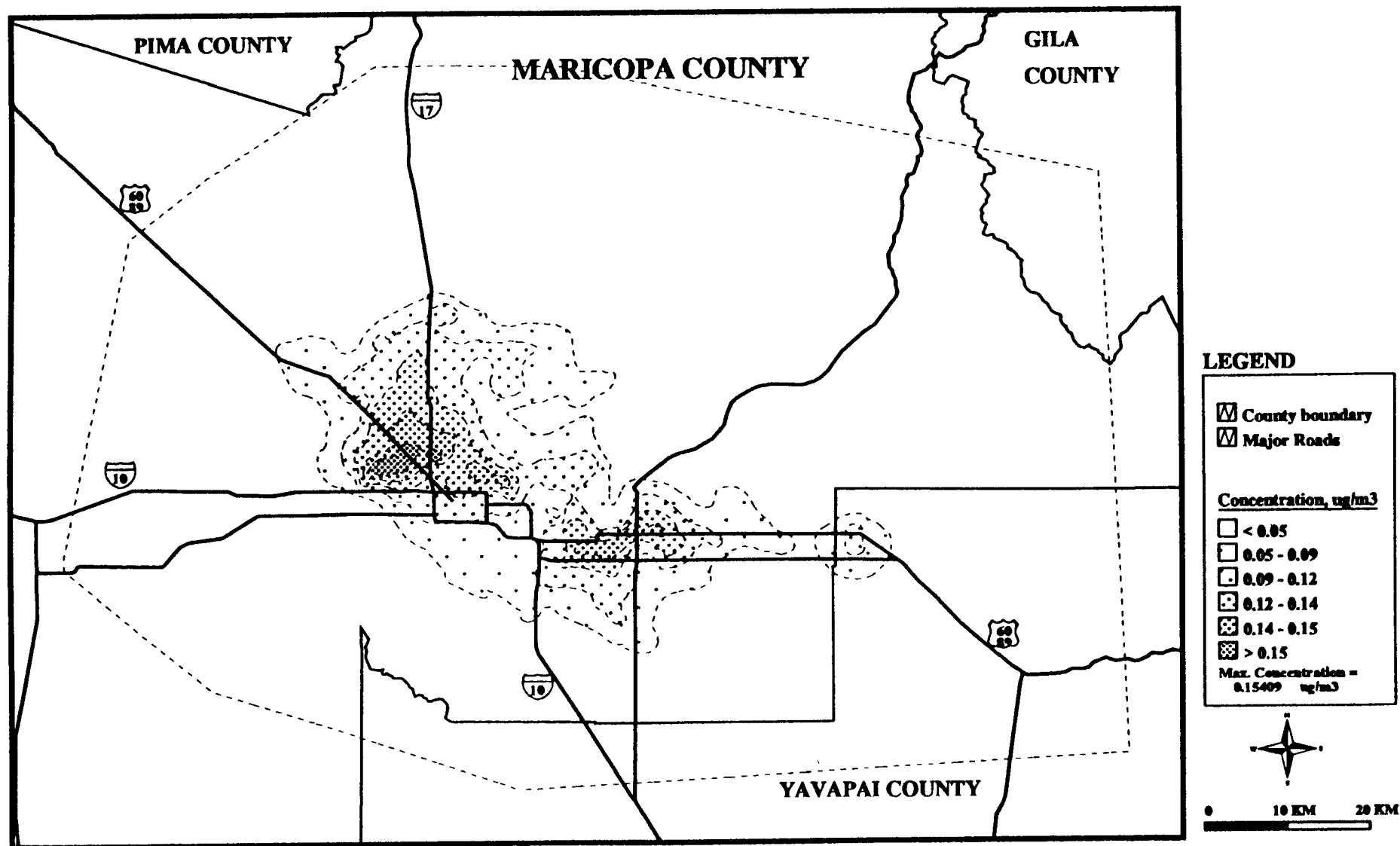


Figure A.3-17. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
POM, Mobile Sources (1987-1991)



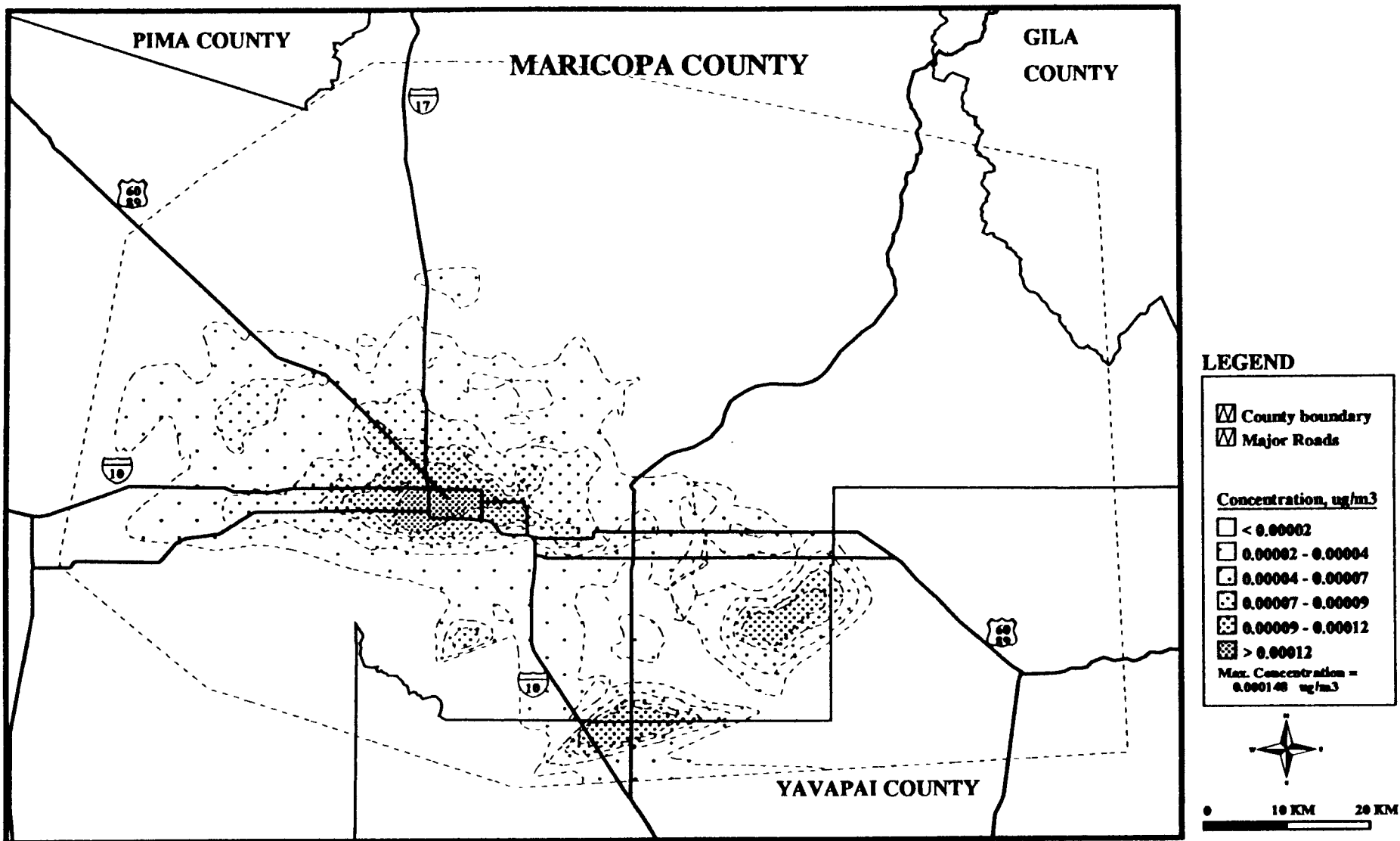


Figure A.3-18. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Hexavalent Chromium, All Sources (1987-1991)

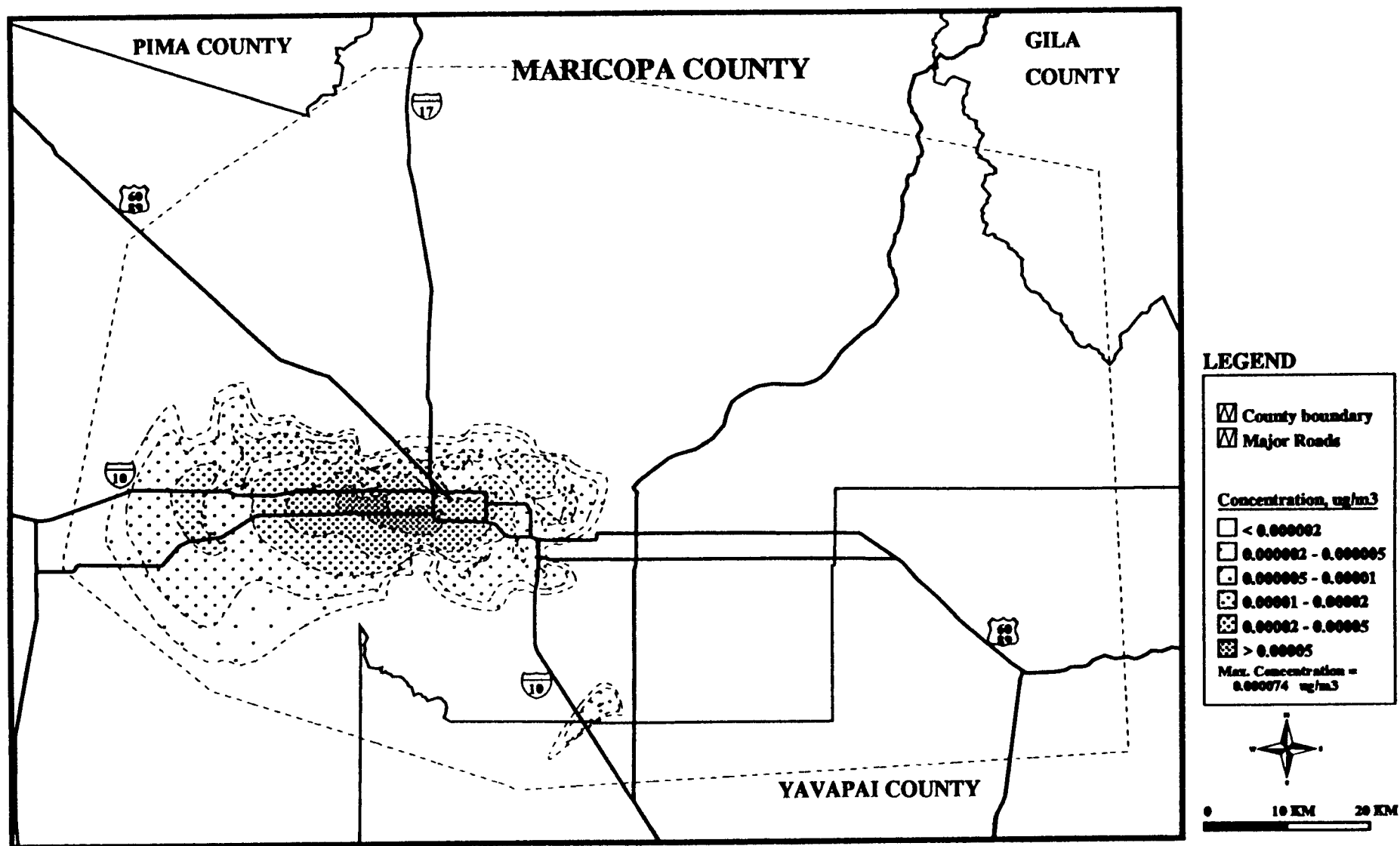
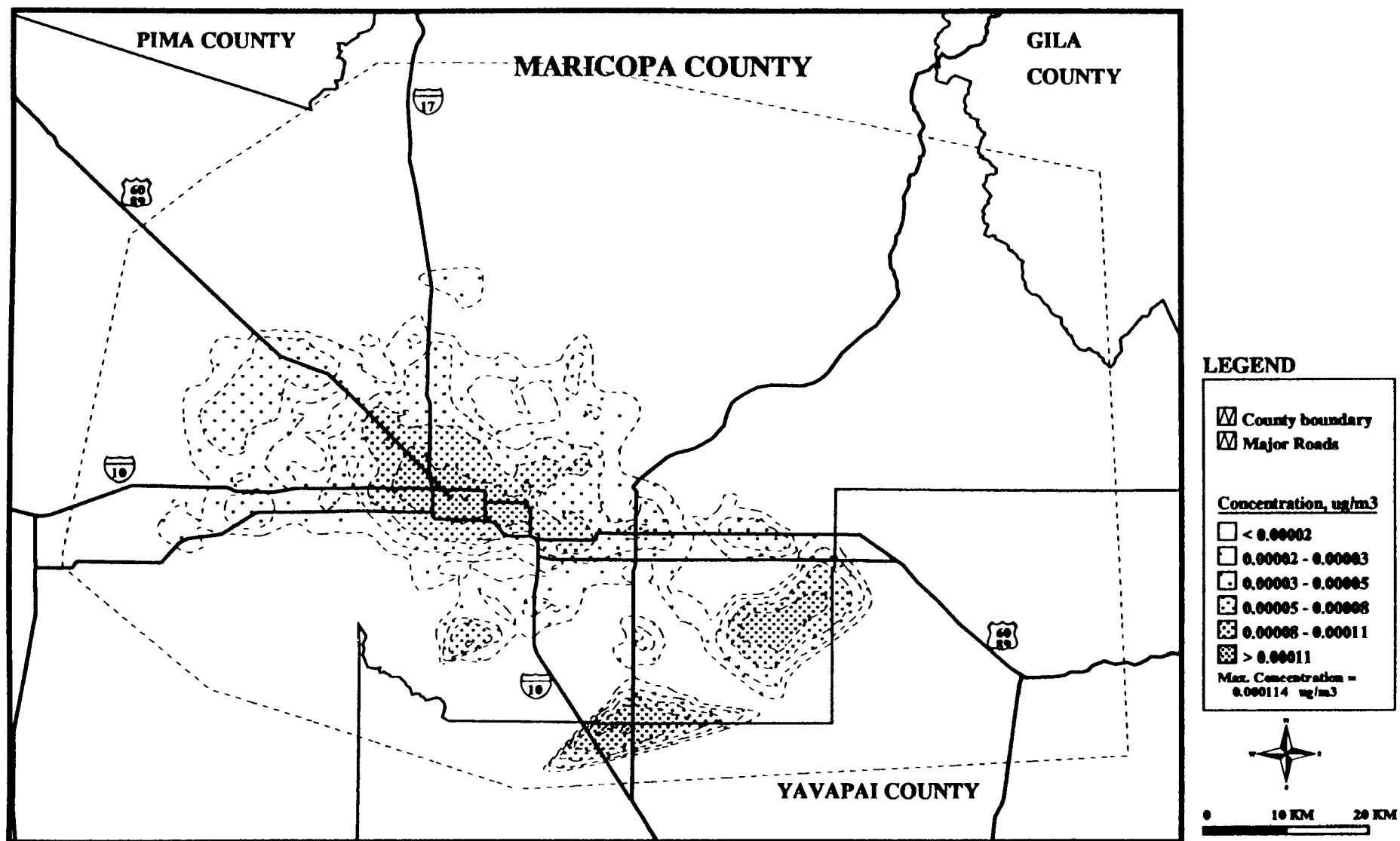


Figure A.3-19. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Hexavalent Chromium, Major Sources (1987-1991)



**Figure A.3-20. Isopleths of Annual Average Concentrations, Phoenix, Arizona  
Hexavalent Chromium, Area Sources (1987-1991)**

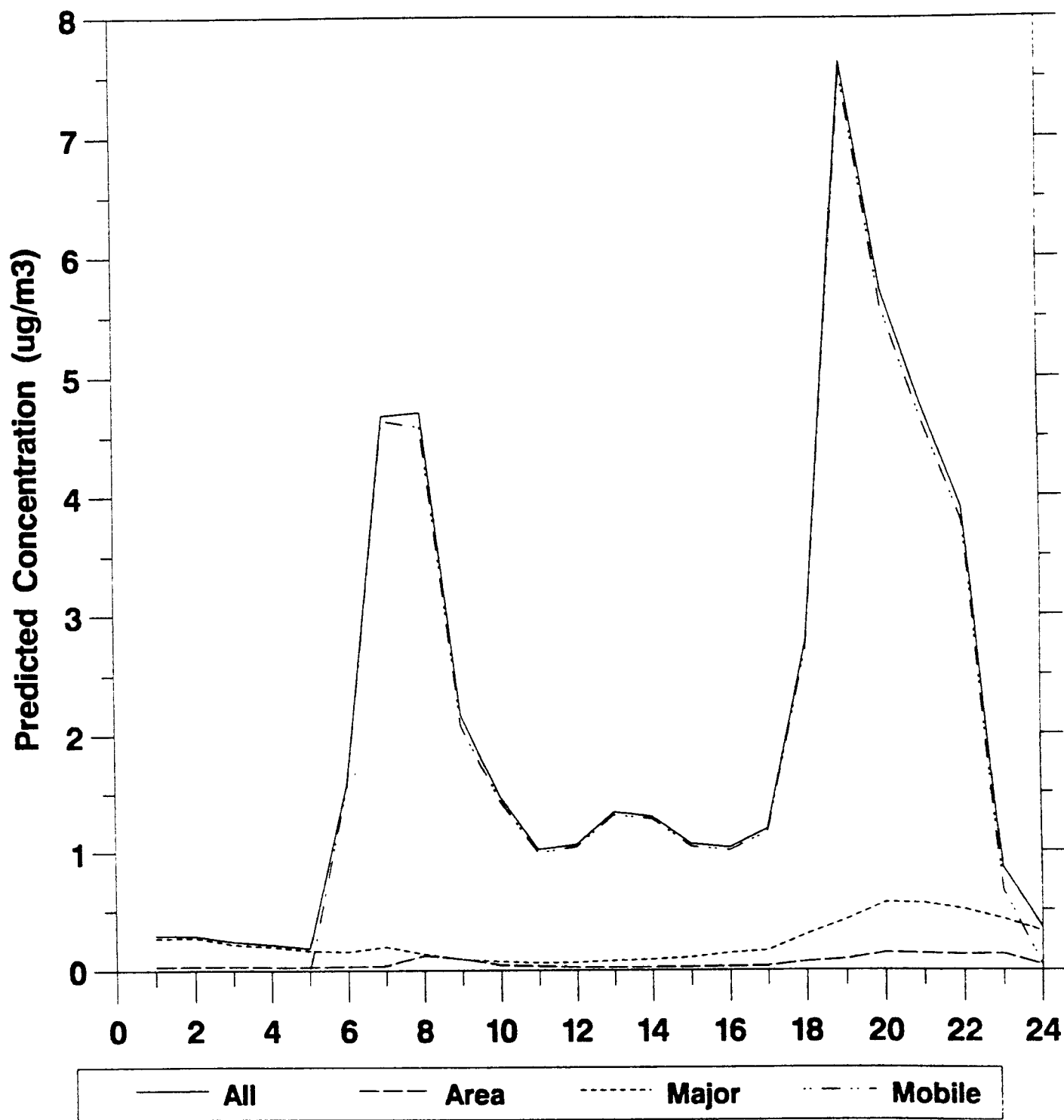
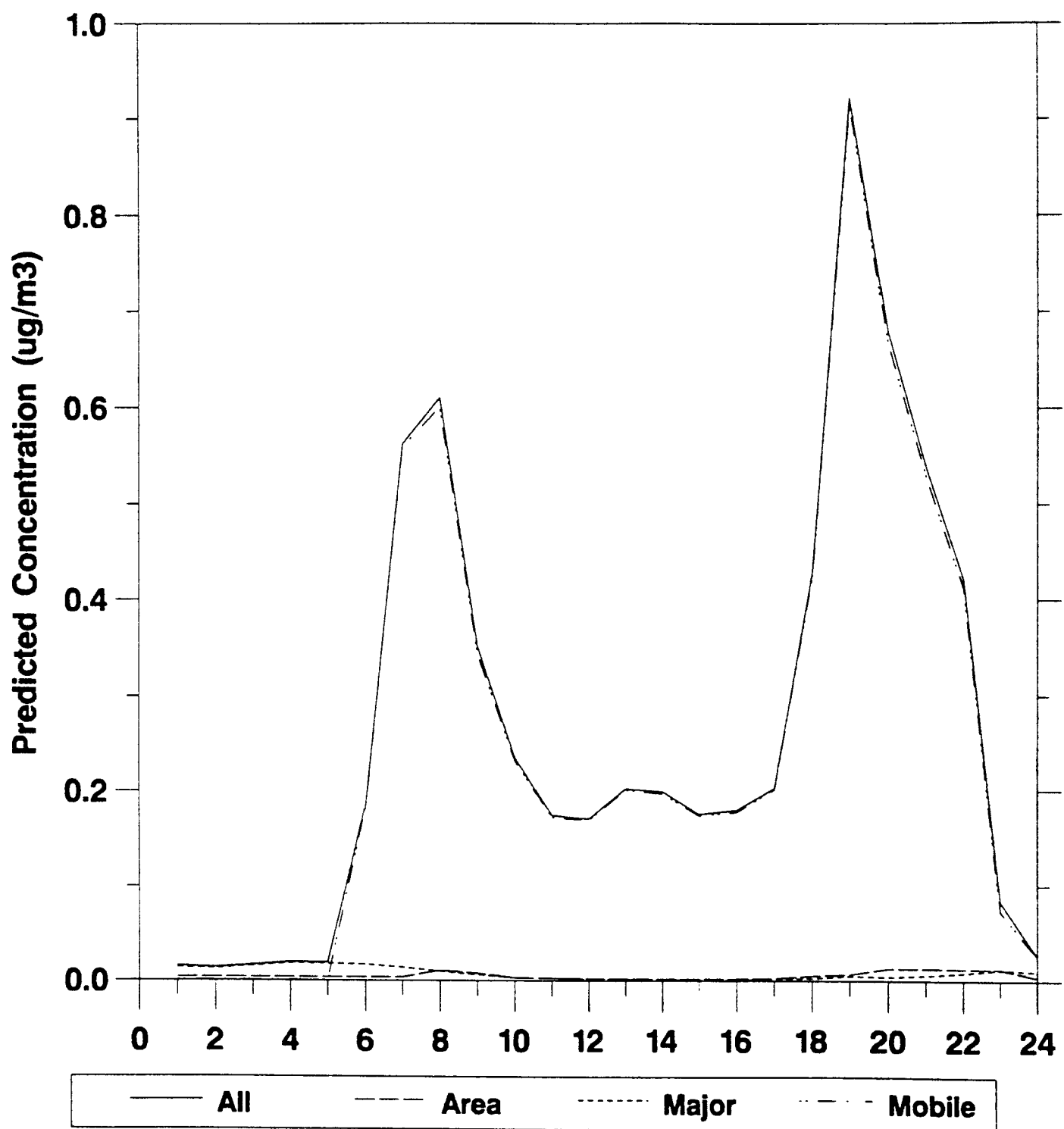
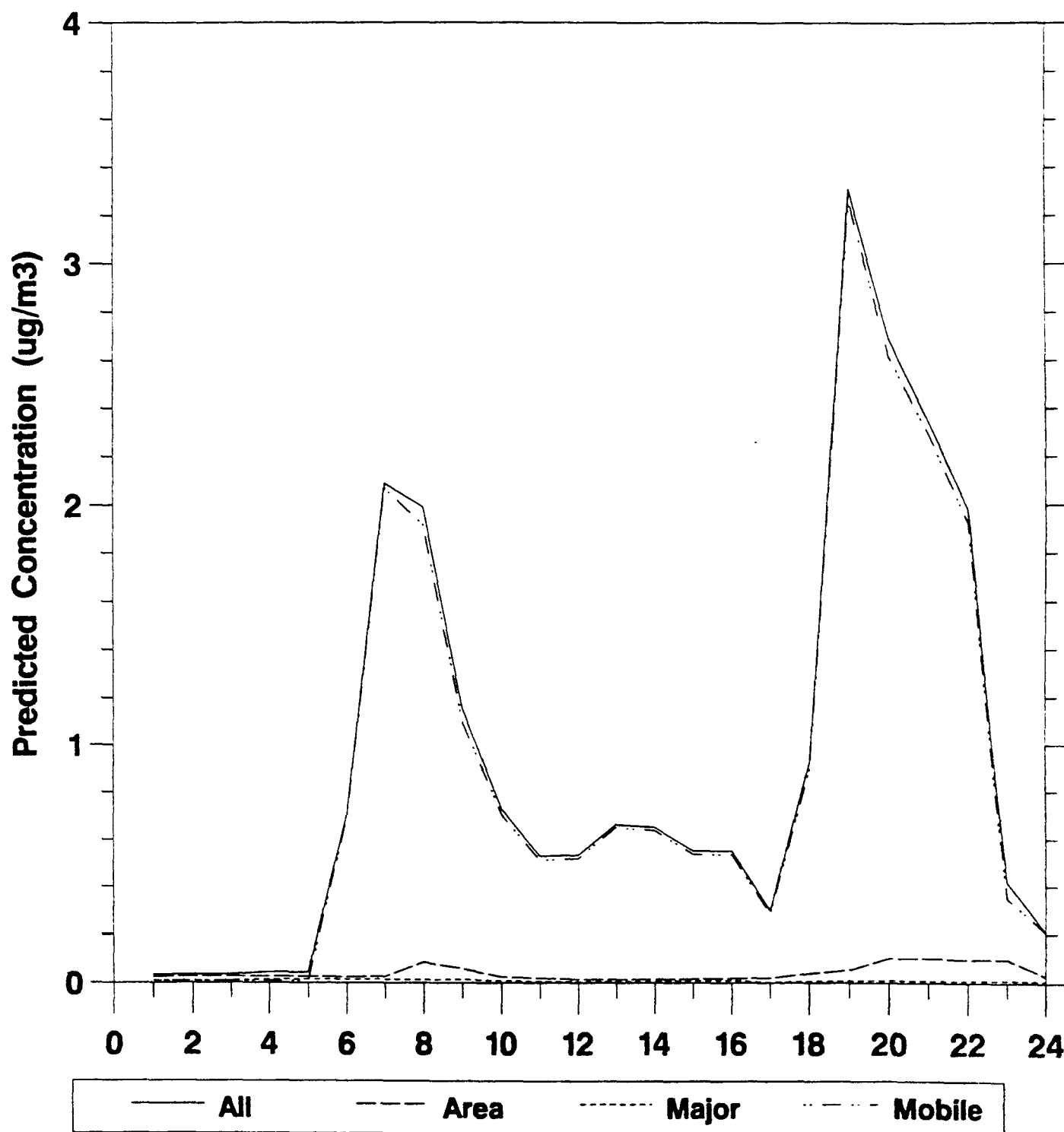


Figure A.3-21. Urban Area Source Modeling, Phoenix, Arizona  
Benzene Average Maximum Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data



**Figure A.3-22. Urban Area Source Modeling, Phoenix, Arizona  
1,3-Butadiene Average Maximum Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**

**Figure 3–40: Urban Area Source Modeling, Phoenix, Arizona, Formaldehyde  
Average Maximum Concentrations by Hour of Day  
All Sources, 1987–1991 Meteorological Data**



**Figure A.3-23. Urban Area Source Modeling, Phoenix, Arizona  
Formaldehyde Average Maximum Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**

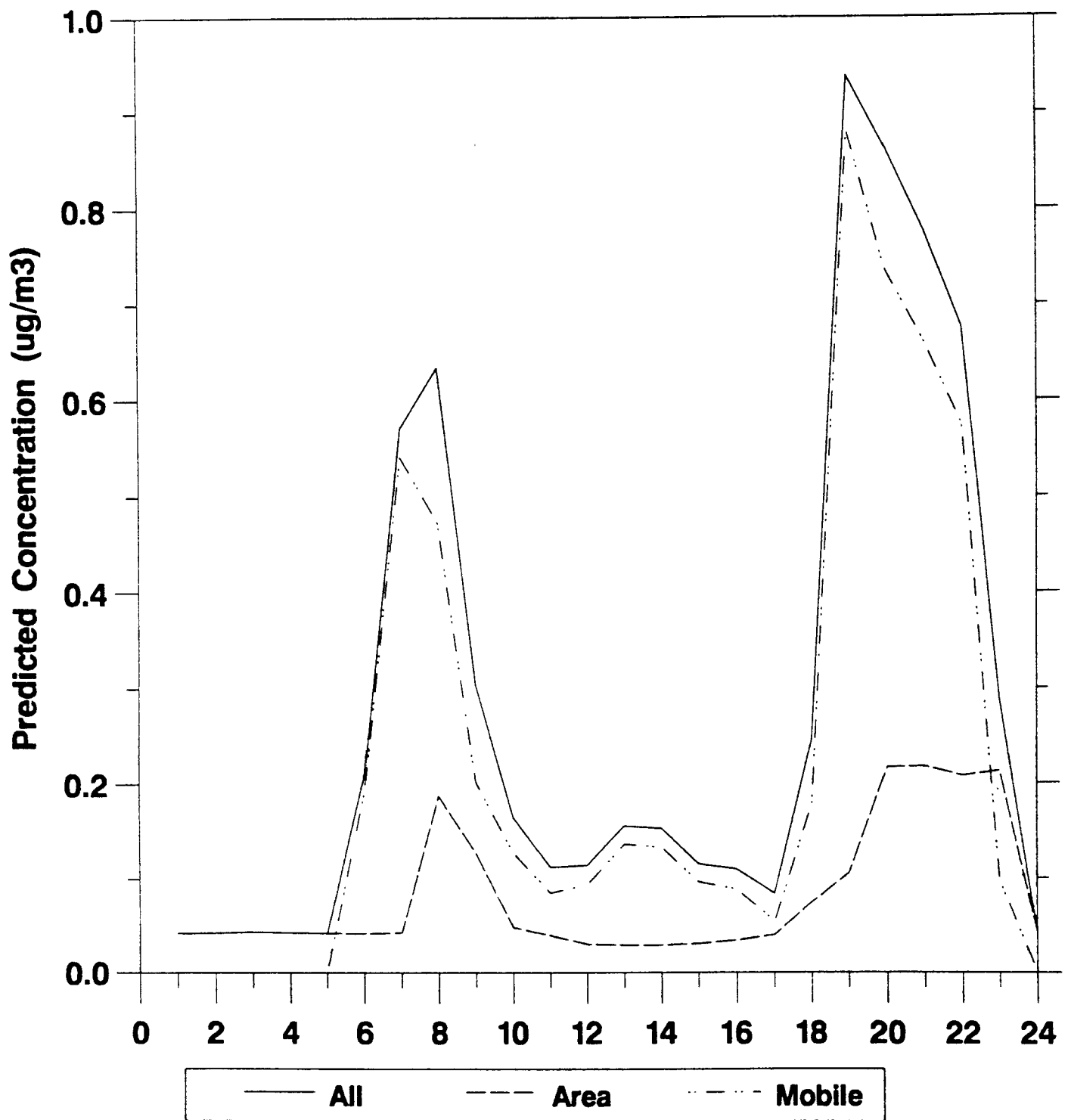


Figure A.3-24. Urban Area Source Modeling, Phoenix, Arizona  
 POM Average Maximum Concentrations by Hour of Day  
 All Sources, 1987-1991 Meteorological Data

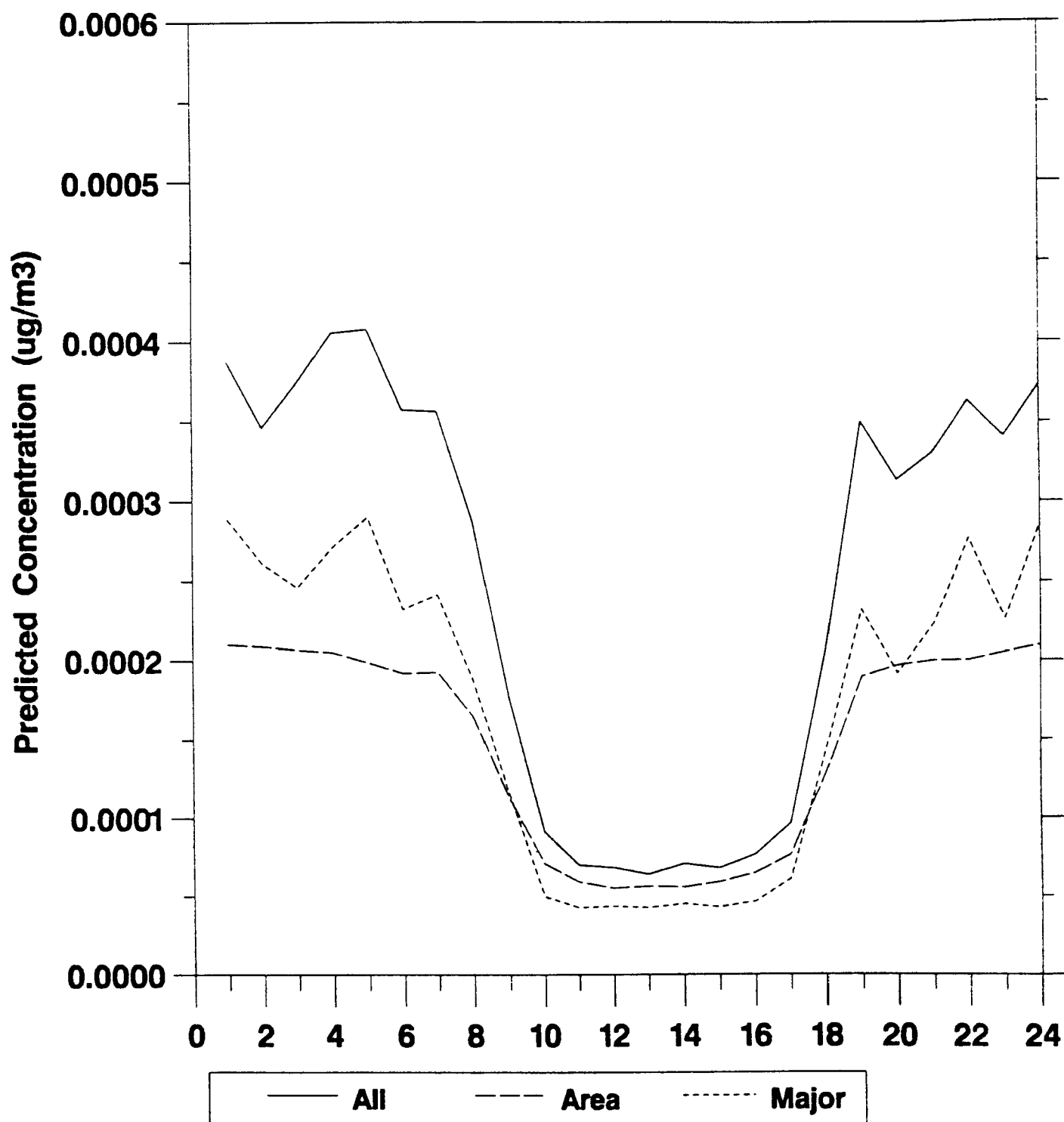
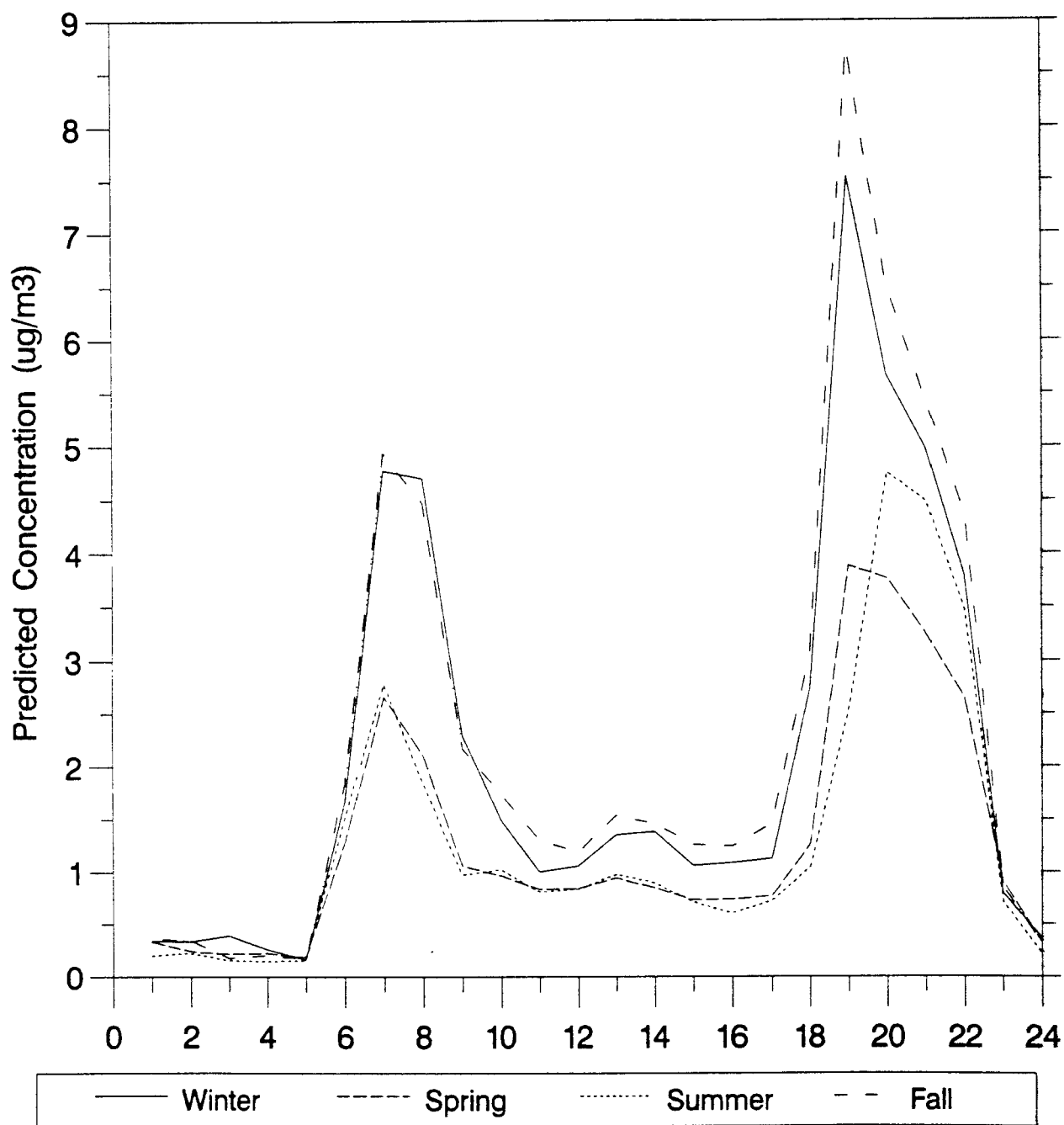
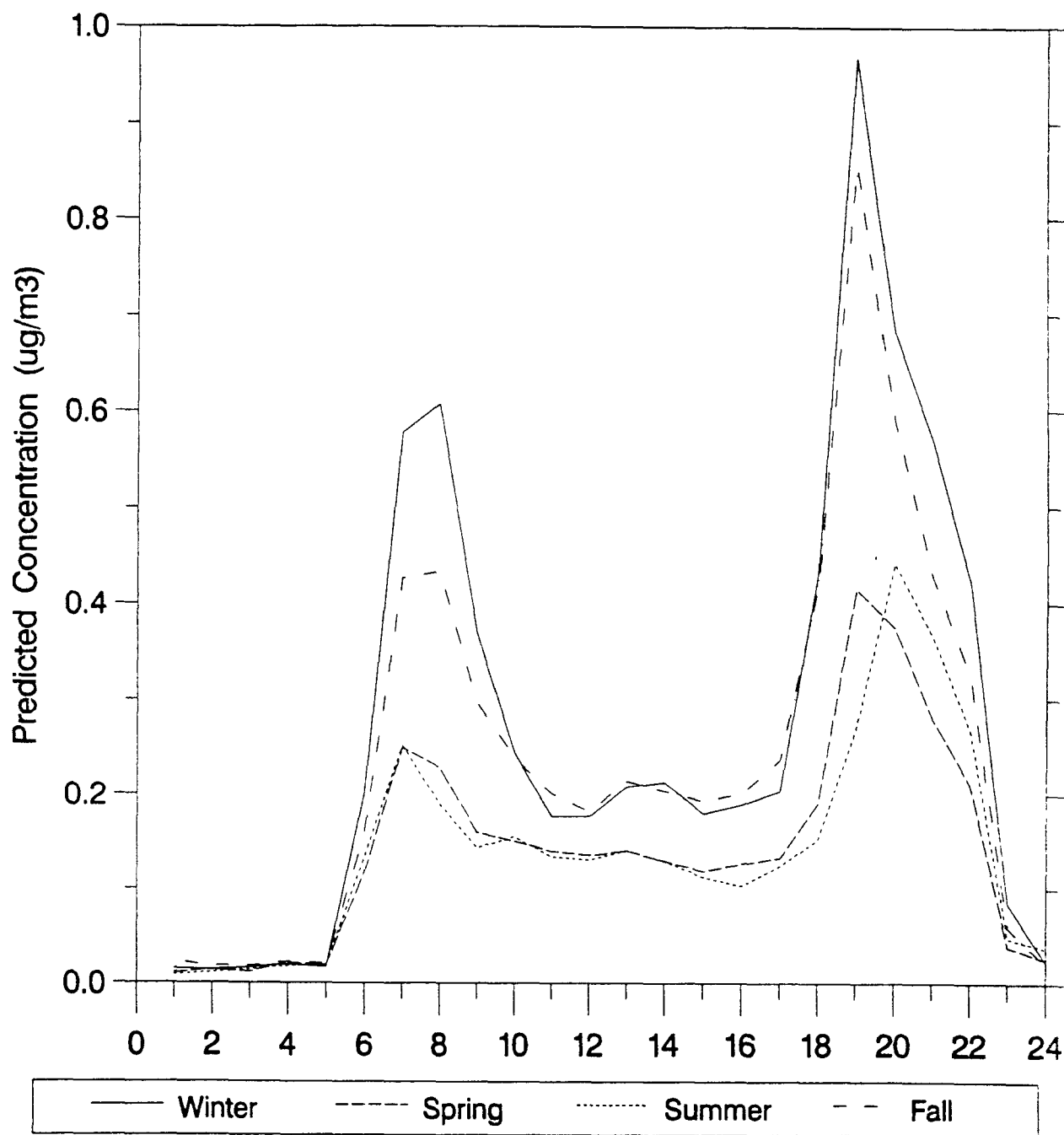


Figure A.3-25. Urban Area Source Modeling, Phoenix, Arizona  
Hexavalent Chromium Average Maximum Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data

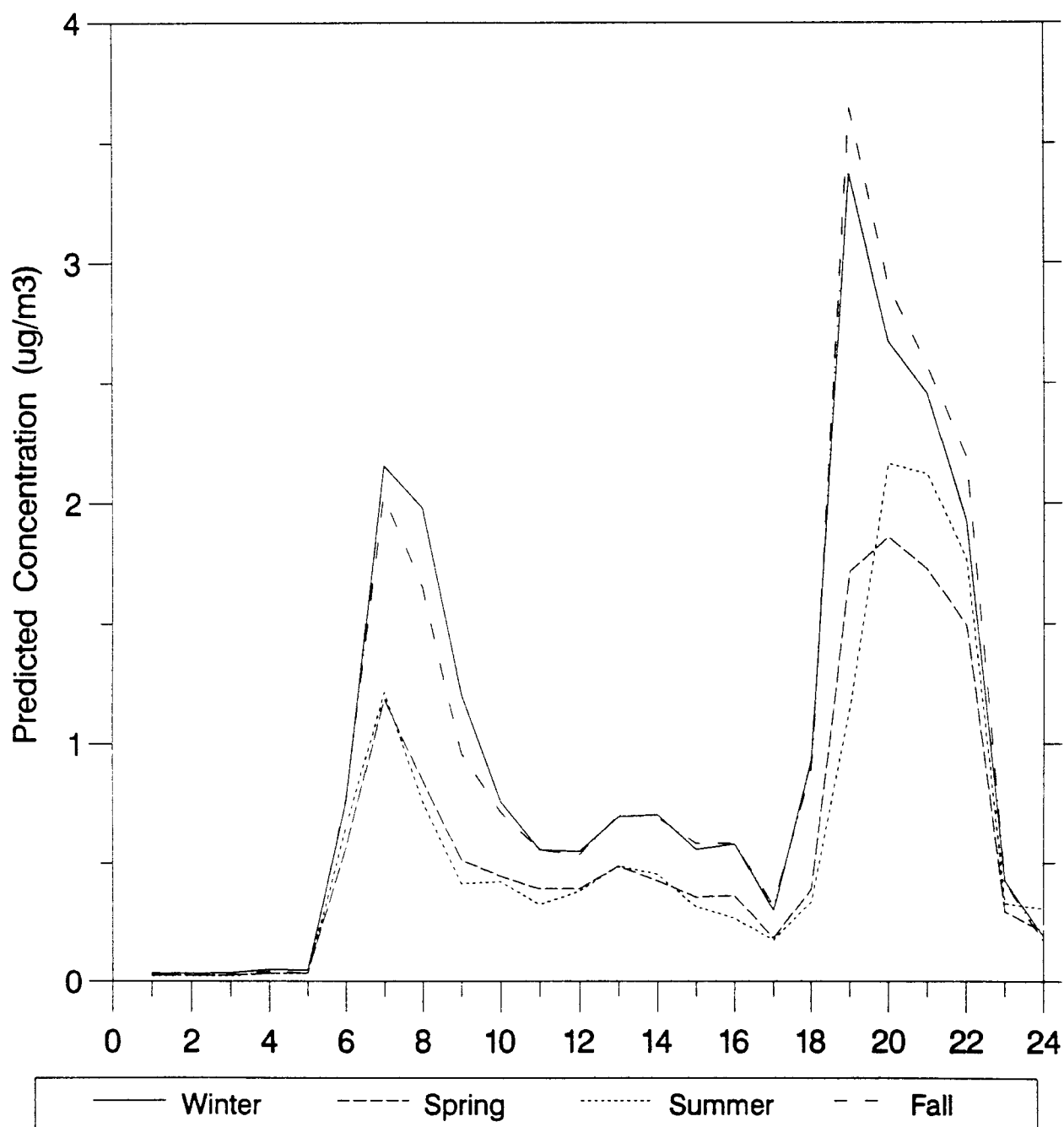




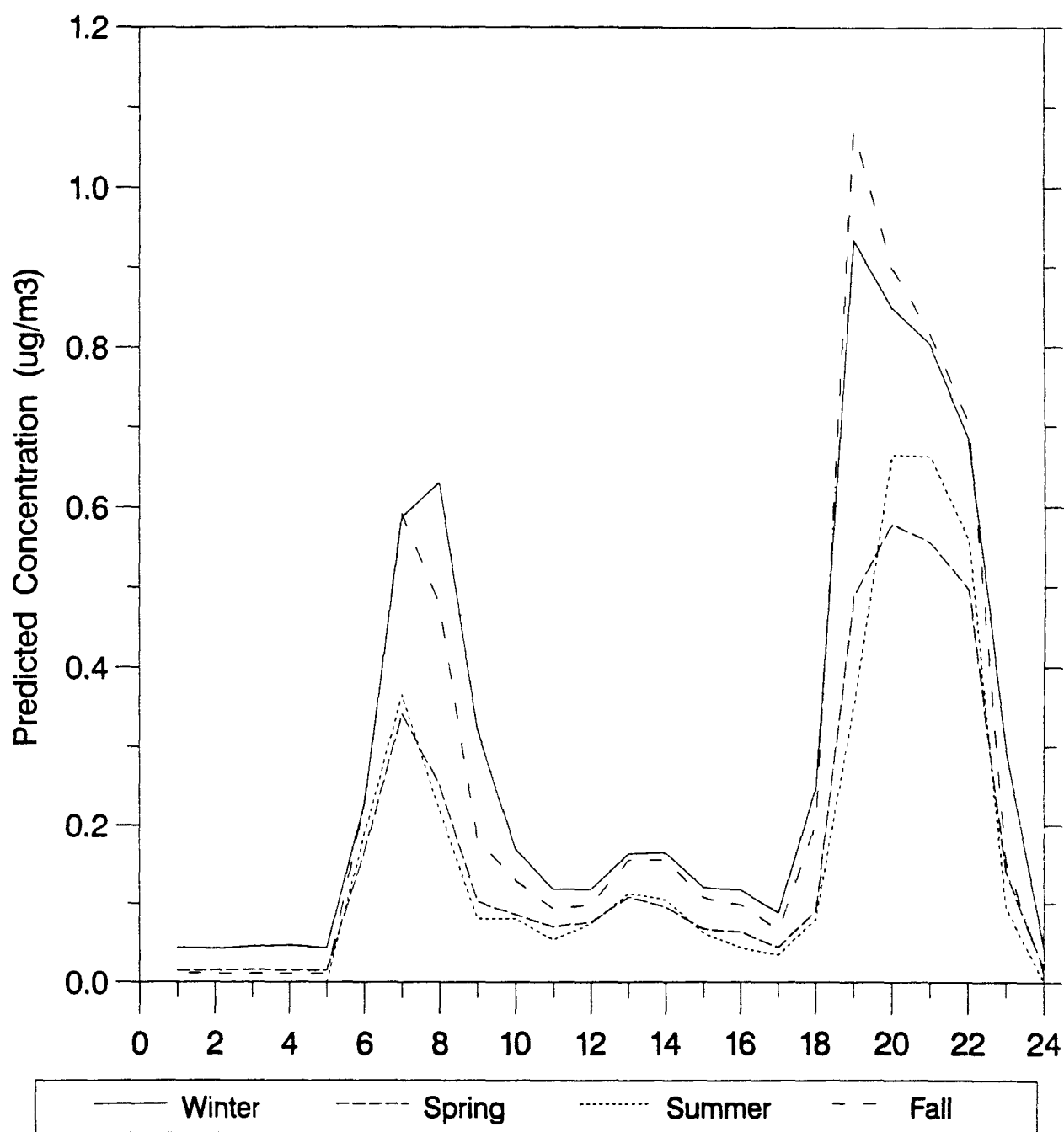
**Figure A.3-26. Urban Area Source Modeling, Phoenix, Arizona  
Benzene Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



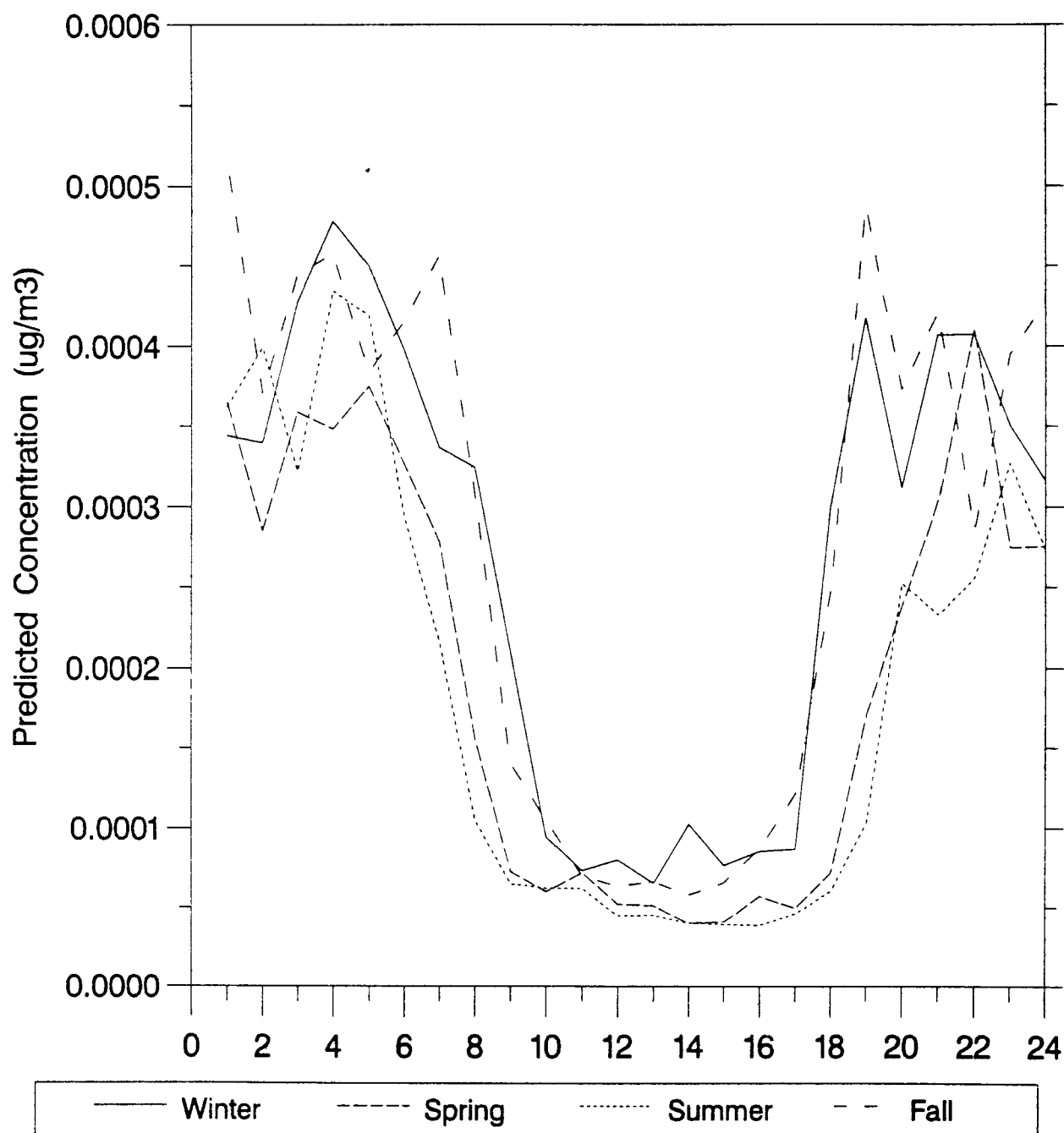
**Figure A.3-27. Urban Area Source Modeling, Phoenix, Arizona  
1,3-Butadiene Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



**Figure A.3-28. Urban Area Source Modeling, Phoenix, Arizona**  
**Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**



**Figure A.3-29. Urban Area Source Modeling, Phoenix, Arizona  
POM Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



**Figure A.3-30. Urban Area Source Modeling, Phoenix, Arizona  
Hexavalent Chromium Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**

#### A.4 SUMMARY AND CONCLUSIONS

This appendix presents the results of a study conducted in support of an air quality impact analysis for five toxic air pollutants emitted from major, area and mobile sources located in Phoenix, Arizona. In this study, the ambient concentrations attributable to these sources were estimated through the application of the ISCST3 dispersion model.

Both the annual average concentrations, as well as the seasonal average concentrations by hour of day, were estimated. The results of the modeling study show that a majority of the total air concentrations for Phoenix can be attributed to the mobile sources. The area sources were found to be the next largest contributors. The major sources were found to contribute insignificantly to the total concentrations.

A study of the hour-by-day variations of seasonal average concentrations showed that the concentrations are higher during the morning and evening hours, and that the concentrations are generally higher during winter and fall seasons as compared to spring and summer seasons.

In order to appropriately explain the observed patterns in concentrations, it is recommended that further detailed studies be conducted. The detailed studies should focus on an analysis of the temporal variations in emissions for various types of sources and their contributions to the predicted total concentrations. The studies should also take into consideration the effects of the meteorological conditions. It is also recommended that these studies be conducted using at least the same five years of meteorological data as used in this report to ensure continuity. Long term trends could be identified using additional years of data.

## A.5 REFERENCES

ENSR (Consulting and Engineering), 1995. Arizona Hazardous Air Pollution Research Program Volumes 1 & 2. Arizona Department of Environmental Quality, Phoenix, Arizona.

## **APPENDIX B**

### **AIR DISPERSION MODELING OF TOXIC POLLUTANTS IN URBAN AREAS**

#### **CASE STUDY FOR HOUSTON, TEXAS**



## APPENDIX B - TABLE OF CONTENTS

B. CASE STUDY FOR HOUSTON, TEXAS .....	B-1
B.1 INTRODUCTION .....	B-1
B.2 SUMMARY OF EMISSION INVENTORY FOR HOUSTON .....	B-2
B.3 OVERVIEW OF MODELING RESULTS .....	B-4
B.3.1 Annual Average Modeling Results .....	B-4
B.3.1.1 <i>Concentration Contributions by Source Category</i> .....	B-5
B.3.1.2 <i>Isopleths for Benzene</i> .....	B-5
B.3.1.3 <i>Isopleths for 1,3-Butadiene</i> .....	B-5
B.3.1.4 <i>Isopleths for Primary Formaldehyde</i> .....	B-5
B.3.1.5 <i>Isopleths for Polycyclic Organic Matter (POM)</i> .....	B-6
B.3.1.6 <i>Isopleths for Hexavalent Chromium</i> .....	B-6
B.3.1.7 <i>Isopleths for Total (Primary and Secondary) Formaldehyde</i> .....	B-6
B.3.2 Modeling Results for Average Maximum Concentrations by Hour-of-Day .....	B-6
B.3.3 Maximum Seasonal Average Concentrations .....	B-7
B.3.4 Pollutant Decay .....	B-7
B.4 PRELIMINARY ANALYSIS OF AIR QUALITY DATA AND MODELED ESTIMATES .....	B-43
B.5 SUMMARY AND CONCLUSIONS .....	B-49
B.6 REFERENCES .....	B-50

## **APPENDIX B - LIST OF TABLES**

Table B.2-1 Toxic Air Pollutant Emissions for Houston, Texas Based on Year 1993 . . . . .	B-3
Table B.3-1 Highest Annual Average Concentrations from All Sources Combined for Houston, Texas Based on 5 Modeled Years 1987 - 1991 . . . . .	B-7
Table B.3-2 Source Category Contributions to Total Annual Average Concentrations for Houston, Texas . . . . .	B-8
Table B.3-3 Half-Life Decay Values . . . . .	B-9
Table B.4-1 Statistics for Predicted and Observed Concentrations . . . . .	B-44
Table B.4-2 Model Prediction Statistics by Source Type . . . . .	B-45

## APPENDIX B - LIST OF FIGURES

Figure B.3-1. Modeled Receptor Locations Urban Area Source Modeling, Houston, Texas	B-10
Figure B.3-2. Population Density, Houston, Texas	B-11
Figure B.3-3. Isopleths of Annual Average Concentrations, Houston, Texas	
Benzene, All Sources (1987-1991)	B-12
Figure B.3-4. Isopleths of Annual Average Concentrations, Houston, Texas	
Benzene, Major Sources (1987-1991)	B-13
Figure B.3-5. Isopleths of Annual Average Concentrations, Houston, Texas	
Benzene, Area Sources (1987-1991)	B-14
Figure B.3-6. Isopleths of Annual Average Concentrations, Houston, Texas	
Benzene, Mobile Sources (1987-1991)	B-15
Figure B.3-7. Isopleths of Annual Average Concentrations, Houston, Texas	
1,3-Butadiene, All Sources (1987-1991)	B-16
Figure B.3-8. Isopleths of Annual Average Concentrations, Houston, Texas	
1,3-Butadiene, Major Sources (1987-1991)	B-17
Figure B.3-9. Isopleths of Annual Average Concentrations, Houston, Texas	
1,3-Butadiene, Area Sources (1987-1991)	B-18
Figure B.3-10. Isopleths of Annual Average Concentrations, Houston, Texas	
1,3-Butadiene, Mobile Sources (1987-1991)	B-19
Figure B.3-11. Isopleths of Annual Average Concentrations, Houston, Texas	
Primary Formaldehyde, All Sources (1987-1991)	B-20
Figure B.3-12. Isopleths of Annual Average Concentrations, Houston, Texas	
Primary Formaldehyde, Major Sources (1987-1991)	B-21
Figure B.3-13. Isopleths of Annual Average Concentrations, Houston, Texas	
Primary Formaldehyde, Area Sources (1987-1991)	B-22
Figure B.3-14. Isopleths of Annual Average Concentrations, Houston, Texas	
Primary Formaldehyde, Mobile Sources (1987-1991)	B-23
Figure B.3-15. Isopleths of Annual Average Concentrations, Houston, Texas	
POM, All Sources (1987-1991)	B-24
Figure B.3-16. Isopleths of Annual Average Concentrations, Houston, Texas	
POM, Area Sources (1987-1991)	B-25
Figure B.3-17. Isopleths of Annual Average Concentrations, Houston, Texas	
POM, Mobile Sources (1987-1991)	B-26
Figure B.3-18. Isopleths of Annual Average Concentrations, Houston, Texas	
Hexavalent Chromium, All Sources (1987-1991)	B-27
Figure B.3-19. Isopleths of Annual Average Concentrations, Houston, Texas	
Hexavalent Chromium, Major Sources (1987-1991)	B-28
Figure B.3-20. Isopleths of Annual Average Concentrations, Houston, Texas	
Hexavalent Chromium, Area Sources (1987-1991)	B-29
Figure B.3-21. Isopleths of Annual Average Concentrations, Houston, Texas	
Total Formaldehyde (Primary and Secondary), All Sources (1987-1991)	B-30

Figure B.3-23. Urban Area Source Modeling, Houston, Texas, 1,3-Butadiene Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-32
Figure B.3-24. Urban Area Source Modeling, Houston, Texas, Primary Formaldehyde Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-33
Figure B.3-25. Urban Area Source Modeling, Houston, Texas, POM Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-34
Figure B.3-26. Urban Area Source Modeling, Houston, Texas, Hexavalent Chromium Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-35
Figure B.3-27. Urban Area Source Modeling, Houston, Texas, Total Formaldehyde Average Maximum Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-36
Figure B.3-28. Urban Area Source Modeling, Houston, Texas, Benzene Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-37
Figure B.3-29. Urban Area Source Modeling, Houston, Texas, 1,3-Butadiene Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-38
Figure B.3-30. Urban Area Source Modeling, Houston, Texas, Primary Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-39
Figure B.3-31. Urban Area Source Modeling, Houston, Texas, POM Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-40
Figure B.3-32. Urban Area Source Modeling, Houston, Texas, Hexavalent Chromium Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-41
Figure B.3-33. Urban Area Source Modeling, Houston, Texas, Total Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day, All Sources, 1987-1991 Meteorological Data .....	B-42
Figure B.4-1. Box Plot of Houston Benzene 1987-1991 Modeled vs. 1993-1994 Monitored .....	B-46
Figure B.4-2. Scatter Plots of Monitored Benzene Values (ug/m3) .....	B-47
Figure B.4-3. Scatter Plots of Modeled Benzene Values (ug/m3) .....	B-48

## **B. CASE STUDY FOR HOUSTON, TEXAS**

### **B.1 INTRODUCTION**

The purpose of this appendix is to document the modeling methodology employed by EPA in estimating ambient air concentrations of selected toxic pollutants for Houston, Texas. The pollutants modeled were benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium. The modeling study serves as an example of guidance on the application of dispersion models to the assessment of exposure to toxic pollutants on an urban scale.

## B.2 SUMMARY OF EMISSION INVENTORY FOR HOUSTON

The Houston inventory used in this study included multiple components. Major source emissions of the selected pollutants for the Houston area were contained in a spreadsheet. The area source emission inventory was taken from the inventory developed as part of the Houston Area Source Toxic Emissions (HASTE) project (Radian 1995 a and b). It consisted of several components including total annual emission estimates by pollutant and source category for Harris County, gridded values of various activity factors, such as population, number of dry cleaners, etc., and an association between activity factors and source categories that could be used to spatially allocate the pollutant emissions across the domain. The activity factors, and therefore the area source emissions derived from them, were gridded with a 2km-by-2km grid resolution for the Harris County domain. The modeling domain extends from 214 kilometers to 316 kilometers Easting and from 3,266 kilometers to 3,342 kilometers Northing in Universal Transverse Mercator (UTM) coordinates in UTM Zone 15 (see Figure B.3-1).

Since the area source inventory developed under the HASTE project did not include mobile source emissions, this component was obtained separately from the U.S. EPA Office of Mobile Sources (OMS). The mobile source inventory provided by OMS included VOC emissions that were gridded with the same 2km-by-2km resolution as the HASTE area sources, and toxic fractions that could be applied to the VOC emissions to obtain emission estimates for the pollutant species of interest in this study. Since the VOC emission estimates included averages by season and hour-of-day, this temporal resolution was also applied to the toxic emissions from mobile sources. Since the benzene toxic fractions for some mobile source categories varied depending on whether the emissions were from exhaust or evaporative emissions, OMS also provided a data file containing exhaust versus evaporative fractions for VOC emissions by source category and county within the modeling domain<sup>1</sup>.

Emissions of benzene, 1,3-butadiene, formaldehyde, POM, and hexavalent chromium from the Houston modeling domain totaled 3,803 megagrams (Mg), or 4,192 tons for the base year of 1993. The largest of these five pollutants was benzene (2,436 Mg or 2,685 tons), followed by formaldehyde (731 Mg or 806 tons), 1,3-butadiene (620 Mg or 683 tons), POM (2.05 Mg or 2.26 tons) and hexavalent chromium (14.1 Mg or 15.5 ton). Table B.2-1 summarizes the information.

The HASTE report assumed no emission controls on electroplater chromium emissions. However, TNRCC indicated that these sources were using emission controls. Therefore, chromium emissions for the electroplater source category was reduced to 2.345 tons per year. For POM emissions, TNRCC suggested only using reported POM emissions and not the sum of 16 definition for POM.

---

<sup>1</sup>The approach for preparation of a toxics mobile source emission inventory is currently being updated from the approach used here. Guidance on the preparation of emission inventories for toxics air quality modeling is being prepared by OAQPS. For more information concerning this guidance contact the Info CHIEF help line at (919)541-5285.

In contrast to the Phoenix inventory described in Appendix A, Houston major sources accounted for the largest portion of emissions for 1,3-butadiene, and were a much more significant component of total emissions for formaldehyde. The mobile source emissions for Houston were slightly less than for Phoenix for benzene, 1,3-butadiene, and formaldehyde (ranging from about 15 to 40 percent lower), but were about 650 times lower than Phoenix for POM emissions. The much lower POM emissions for Houston can not readily be explained.

The Houston inventory shows almost 150 times as much hexavalent chromium emissions as reported for Phoenix, with a similar ratio of emissions from area sources versus major sources. The majority of chromium emissions for Houston are from metal plating facilities, with cooling towers also contributing a significant portion. Hexavalent chromium use in cooling towers is being phased out.

**Table B.2-1 Toxic Air Pollutant Emissions for Houston, Texas  
Based on Year 1993**

Source Category		Pollutant Emissions (Mg/yr)				
		Benzene	1,3-Butadiene	Formaldehyde	POM <sup>(1)</sup>	CrVI <sup>(2)</sup>
Mobile Sources	On-Road	971.8	159.2	439.5	0.13	-- <sup>(3)</sup>
	Off-Road	266.4	69.4	156.3	0.22	-- <sup>(3)</sup>
	Total Mobile	1,238.2	228.6	595.8	0.35	-- <sup>(3)</sup>
Area Sources		116.9	4.9	59.0	1.70	2.2
Major Sources		1,080.7	386.5	75.7	0.00	11.9
TOTAL		2435.8	620.0	730.5	2.05	14.1
<p>(1) POM = Polycyclic Organic Matter. Individual constituents of POM vary. For mobile source, POM is defined as sum of 16 chemicals. For area &amp; mobile sources, data provided by TNRCC do not define POM constituents. Data reported as POM by TNRCC are used here.</p> <p>(2) CrVI = Hexavalent Chromium</p> <p>(3) Not Applicable</p>						

## B.3 OVERVIEW OF MODELING RESULTS

This section presents and discusses the results of the modeling analysis conducted for the sources located in Houston, Texas. As discussed earlier, the five pollutants that were included in this study are benzene, 1,3-butadiene, formaldehyde, polycyclic organic matter (POM), and hexavalent chromium.

A total of 4609 sources of benzene were modeled using the ISCST3 model. This number consisted of 1432 major sources and 3177 area/mobile sources. For 1,3-butadiene, a total of 3372 sources were modeled including 195 major sources and 3177 area/mobile sources. For formaldehyde, a total of 3324 sources were modeled including 147 major sources and 3177 area/mobile sources. For POM, a total of 3177 sources were modeled all of which were area/mobile sources. For hexavalent chromium, a total of 1296 sources were modeled including 20 major sources and 1276 area sources. There were no mobile sources for hexavalent chromium.

A total of 573 receptors were modeled. Figure B.3-1 shows the modeled receptor locations. The modeling was conducted using the five year period 1987 through 1991 National Weather Service surface meteorological data from Houston Airport with mixing heights from Lake Charles, Louisiana. Figure B.3-2 shows the population density for Houston.

Both the annual average concentrations as well as the seasonal average concentrations by hour of day were calculated. Results for both of these averaging periods, and a discussion on the contributions to the total annual average concentrations from each source category (major, area and mobile) are presented below for all pollutants modeled.

The effects of pollutant decay on predicted concentrations were included in this analysis for three modeled pollutants; 1,3-butadiene, formaldehyde, POM. Hexavalent chromium was not modeled with decay due to its particulate nature, and benzene decay was not modeled due to the HAPs long half-life. The modeling results presented below for these pollutants are based on seasonally variable (cold vs warm) half lives for each pollutant. A discussion on the effects of decay is also presented in this section.

It should be noted that low mixing heights, i.e., less than 100 meters, occurred in less than 1.4 percent of all the hours during the five year period that was modeled. Thus the need to adjust the hourly values up to 100 meters arose quite infrequently.

### B.3.1 Annual Average Modeling Results

Table B.3-1 presents the highest annual average concentrations for each of the five pollutants. The corresponding receptor locations are also listed in the table. The listed concentrations represent the combined total concentrations of all modeled sources for each pollutant.



#### B.3.1.1 *Concentration Contributions by Source Category*

Table B.3-2 presents the contributions of each source category (major, area and mobile) to the highest annual concentrations listed in Table B.3-1. As can be seen from this table, the major sources have the largest contribution to the total impacts for benzene (98.8%), 1,3 butadiene (99.9%) and hexavalent chromium (100%). Mobile sources have the largest contribution to the total impact for formaldehyde (100%) and area sources have the largest contribution to the total impact for POM (100%).

For comparison purposes, Table B.3-2 also presents the highest concentrations of all pollutants due to each of the source categories separately. These concentration values follow the same trend seen in the contributions for all sources combined as stated above.

#### B.3.1.2 *Isopleths for Benzene*

Figures B.3-3 through B.3-6 show the isopleths of concentration for benzene.<sup>2</sup> Figure B.3-3 shows the isopleths for the highest total concentrations due to all sources. Figures B.3-4, B.3-5, and B.3-6 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. Similar to the results shown in Table B.3-2, the major source category is the largest contributor (accounting for 98.8%) to the total concentrations. By contrast, the mobile and area sources contribute only a small fraction (2% or less) in most of the modeling domain.

#### B.3.1.3 *Isopleths for 1,3-Butadiene*

Figures B.3-7 through B.3-10 show the isopleths of concentration for 1,3-butadiene. Figure 3-7 shows the isopleths for the highest total concentrations due to all sources. Figures B.3-8, B.3-9, and B.3-10 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. Similar to the results shown in Table B.3-2, the major source category is the largest contributor to the total concentrations (accounting for 99.9%). By contrast, the mobile and area sources contribute only a small fraction (accounting for 0.1% or less contribution) in most of the modeling domain.

#### B.3.1.4 *Isopleths for Primary Formaldehyde*

Figures B.3-11 through B.3-14 show the isopleths of concentration for primary formaldehyde. Figure B.3-11 shows the isopleths for the highest total concentrations due to all sources. Figures B.3-12, B.3-13, and B.3-14 show the isopleths for the highest concentrations for each source category, i.e., major, area and mobile, respectively. The mobile source category is the largest contributor to the total concentrations (accounting for greater than 99.9%). By contrast, the major and area sources contribute negligibly (accounting for 0.1% or less contribution) in most of the modeling domain.

---

<sup>2</sup>Isopleth contours should be viewed with caution because this shape is also dependent on the software package used.

#### B.3.1.5 *Isopleths for Polycyclic Organic Matter (POM)*

Figures B.3-15 through B.3-17 show the isopleths of concentration for POM. Figure B.3-15 shows the isopleths for the highest total concentrations due to all sources. Figures B.3-16 and B.3-17 show the isopleths for the highest concentrations for each source category, i.e., area and mobile, respectively. Note that there were no major sources for POM. The area source category contributes a majority of total concentrations (accounting for 99.8%) contribution. The mobile sources contribute 0.2% or less to the total concentrations.

#### B.3.1.6 *Isopleths for Hexavalent Chromium*

Figures B.3-18 through B.3-20 show the isopleths of concentration for hexavalent chromium. Figure B.3-18 shows the isopleths for the highest total concentrations due to all sources. Figures B.3-19 and B.3-20 show the isopleths for the highest concentrations for each source category, i.e., major and area, respectively. Note that there were no mobile sources for hexavalent chromium. A comparison of Figures B.3-19 and B.3-20 shows that the major sources contribute the most to the total concentrations in most of the modeling domain.

#### B.3.1.7 *Isopleths for Total (Primary and Secondary) Formaldehyde*

Figure B.3-21 shows the isopleths of concentration for total formaldehyde. Figure B.3-21 shows the isopleths for the highest total concentrations due to primary (all sources) plus secondary formaldehyde. A comparison of Figures B.3-11 and B.3-21 shows the isopleths to be the same, however, the concentrations have increased with the maximum concentration increasing by 16.6%.

### B.3.2 Modeling Results for Average Maximum Concentrations by Hour-of-Day

The modeling results for the annual average maximum concentrations by hour-of-day for the entire modeling domain are presented in a series of figures. These values represent the maximum concentration of each of the four seasonal averages, averaged over the five year period. These figures show the temporal variation of the annual average concentrations. For each pollutant, figures were generated for the total concentrations due to all sources and for each source category (major, area and mobile).

Figures B.3-22 through B.3-27 show the temporal variation curves for benzene, 1,3 butadiene, formaldehyde, POM, hexavalent chromium and total formaldehyde, respectively. All figures show a distinct diurnal patterns. Benzene and 1,3 butadiene and hexavalent chromium, whose main constituents are major sources, show maxima occurring during the night hours (6 p.m. - 8 a.m.) and a distinct peak during the early morning hours (1-6 a.m.). POM concentrations, are dominated by primarily mobile sources, and to a lesser extent area sources. The diurnal pattern follows the same pattern as that of benzene and 1,3 butadiene. For the major sources, there is a general trend that shows high concentrations during the night hours, as compared to the rest of the day. Again, further analysis would be required to explain these occurrences in detail.

Formaldehyde follows the same general pattern, however, there are distinct peaks at 7 a.m. and 6 p.m.. The major contributor to formaldehyde concentrations is mobile sources. These peaks may be attributed to increased motor vehicle activity during peak rush hour. The peaks generally occur at the same time for all four seasons. Although a detailed study would be necessary to determine the exact reason(s) for the patterns seen, the most likely reason is the temporal variation of emissions from mobile sources, as well as, related to meteorological conditions during these hours. For example, mobile source emissions are higher during the morning and evening rush hours, and will contribute to the morning and evening peaks. The peaks may also be related to an increase in concentrations due to more stable atmospheric conditions at night.

### B.3.3 Maximum Seasonal Average Concentrations

From Figures B.3-28 through B.3-33 showing seasonal averages, it can also be noted that, generally, the concentrations of all pollutants during winter and fall seasons are higher than spring and summer seasons. This may be associated with the seasonal variations in the operation of certain types of sources, or the fact that the half lives for all pollutants modeled are shorter during warm conditions versus cold conditions. For example, residential wood combustion occurs primarily during winter with practically no wood combustion during summer. The only exceptions to this are the major sources which do not show any apparent seasonal trend.

### B.3.4 Pollutant Decay

As noted above, pollutant decay was modeled for 1,3-butadiene, formaldehyde and POM. Pollutant decay was not modeled for hexavalent chromium because of its particulate nature, and for benzene due to its long half-life. Decay has been defined seasonally (cold versus warm). With the study area being located in the southern latitudes, winter has been designated the cold season with spring, summer and fall designated warm seasons. Appropriate half-life decay values have been assigned to pollutants using these criteria, as presented in Table B.3-3.

**Table B.3-1 Highest Annual Average Concentrations from All Sources Combined for Houston, Texas Based on 5 Modeled Years 1987 - 1991**

Pollutant	Highest Annual Average Concentration (ug/m <sup>3</sup> )	Receptor Location <sup>(1)</sup> (X,Y) (meters)
Benzene	10.41	(296660, 299970)
1,3-Butadiene	26.17	(281902, 287136)
Formaldehyde	2.13	(275162, 319329)
POM <sup>(2)</sup>	0.004	(305619, 292378)
Chromium VI <sup>(3)</sup>	0.11	(305619, 292378)
<p>(1) Receptor locations are in Universal Transverse Mercator (UTM) coordinates for Zone 15. While modeling, the first digit of the Y coordinate (North UTM) was removed.</p> <p>(2) POM = Polycyclic Organic Matter</p> <p>(3) Chromium VI = Hexavalent Chromium</p>		

**Table B.3-2 Source Category Contributions to Total Annual Average Concentrations  
for Houston, Texas**

Pollutant	Concentrations for All Sources combined (ug/m <sup>3</sup> )				Highest Concentrations for Individual Source Categories <sup>(2)</sup> (ug/m <sup>3</sup> )		
	Highest Total Concentrations	Contributions to the Total Concentrations <sup>(1)</sup>			Major	Area	Mobile
		Major	Area	Mobile			
Benzene	10.41	10.29	0.029	0.097	10.29	0.28	0.76
1,3-Butadiene	26.17	26.15	0.002	0.04	26.15	0.011	0.63
Formaldehyde	2.13	0.0029	0.0041	2.13	0.21	0.03	2.13
POM <sup>(3)</sup>	0.0038	N/A <sup>(4)</sup>	0.0038	0.00002	N/A <sup>(4)</sup>	0.0038	0.0023
Chromium VI <sup>(3)</sup>	0.11	0.11	0.00016	N/A <sup>(4)</sup>	0.11	0.017	N/A <sup>(4)</sup>
<p>(1) Represent the contributions from the individual source categories to the listed total concentrations for all sources combined.</p> <p>(2) Represent the highest concentrations due to the individual source categories. The locations of these concentrations may be different from the locations of the highest total concentrations for all sources combined.</p> <p>(3) POM = Polycyclic Organic Matter, Chromium VI = Hexavalent Chromium</p> <p>(4) N/A = Not Applicable</p>							

**Table B.3-3 Half-Life Decay Values**

<b>Pollutant</b>	<b>Half-Life (hours) Cold Season (Winter)</b>	<b>Half-Life (hours) Warm Seasons (Spring, Summer, Fall)</b>
Benzene <sup>(1)(3)</sup>	1560	144
1,3-Butadiene <sup>(1)</sup>	8	2
Formaldehyde <sup>(1)</sup>	6	2
POM <sup>(2)</sup>	6	1.2
(1) From Volume I, Appendix E, Table E.1 (2) From Volume I, Appendix E, Table E.3, using the shortest reported value (3) Due to long half-life value, pollutant decay was not used for this pollutant.		

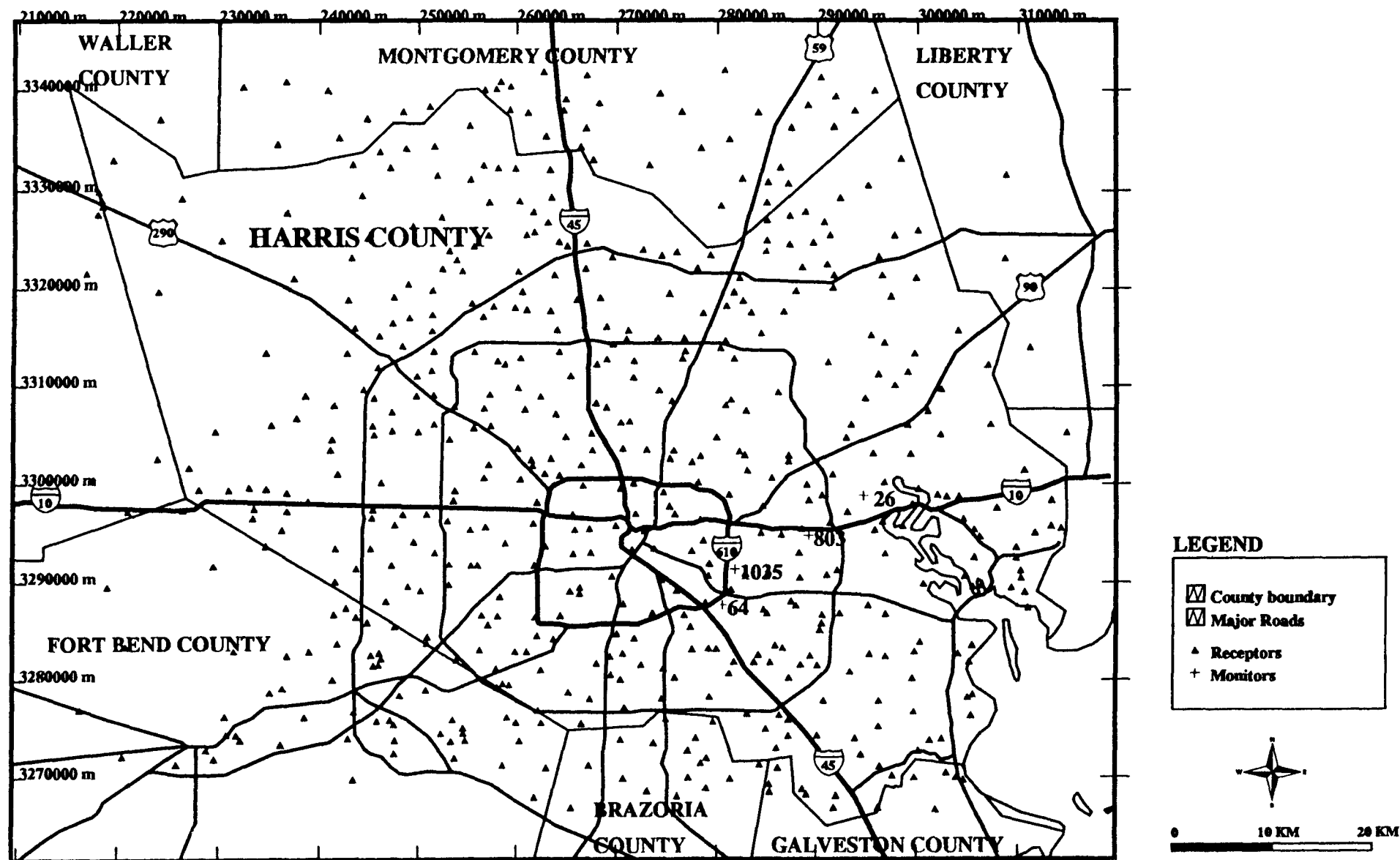


Figure B.3-1. Modeled Receptor Locations Urban Area Source Modeling, Houston, Texas

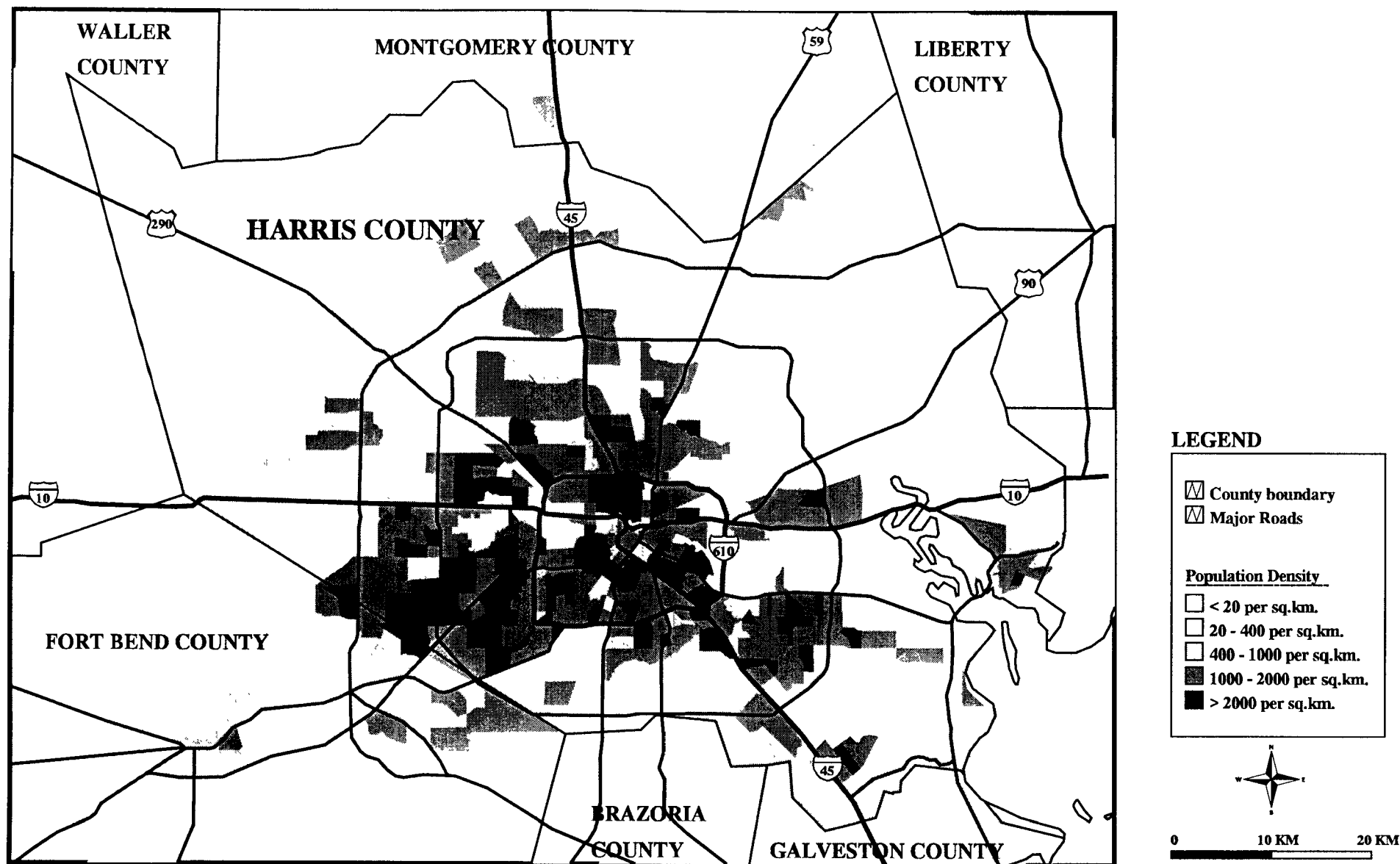


Figure B.3-2. Population Density, Houston, Texas

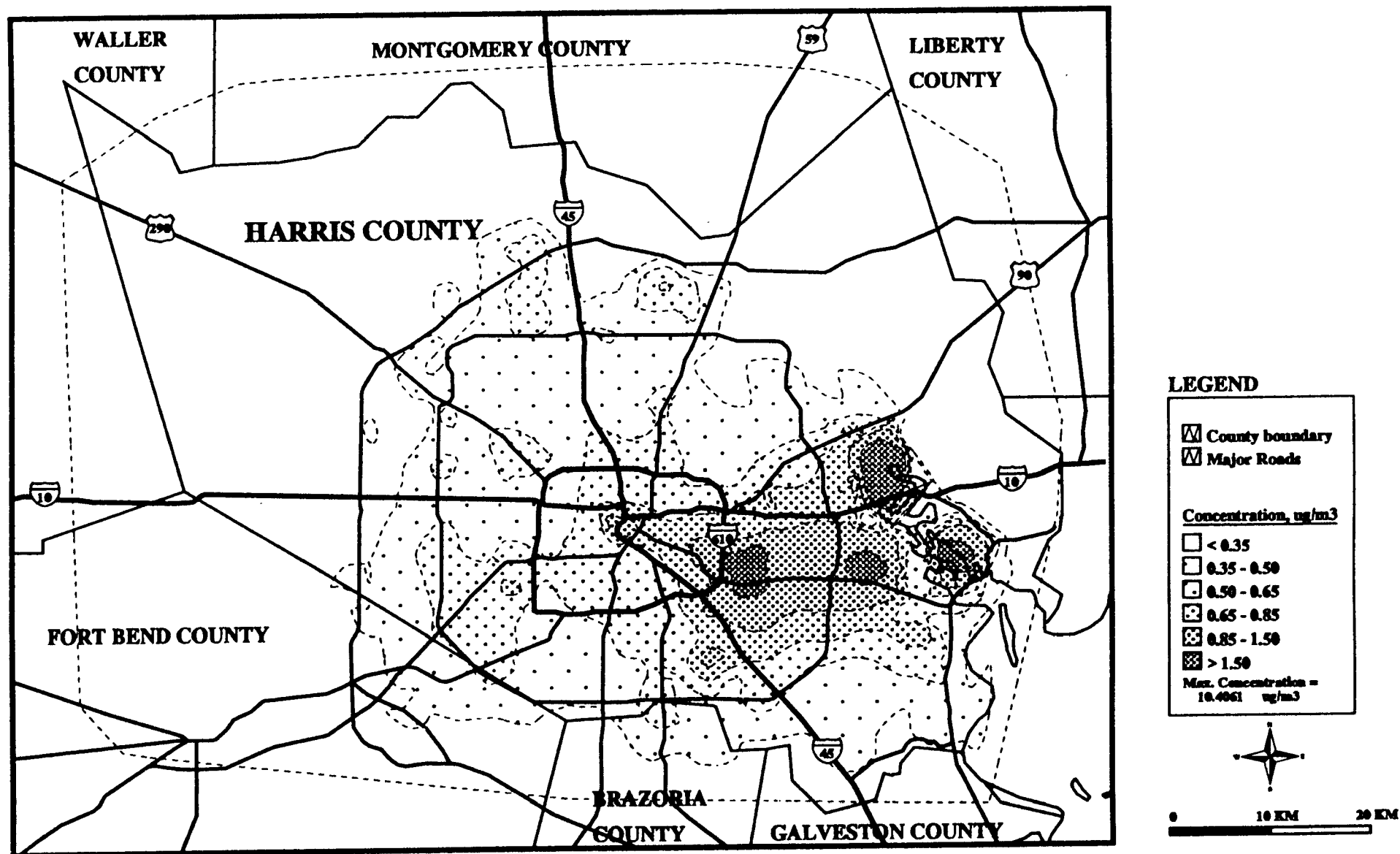


Figure B.3-3. Isopleths of Annual Average Concentrations, Houston, Texas  
Benzene, All Sources (1987-1991)



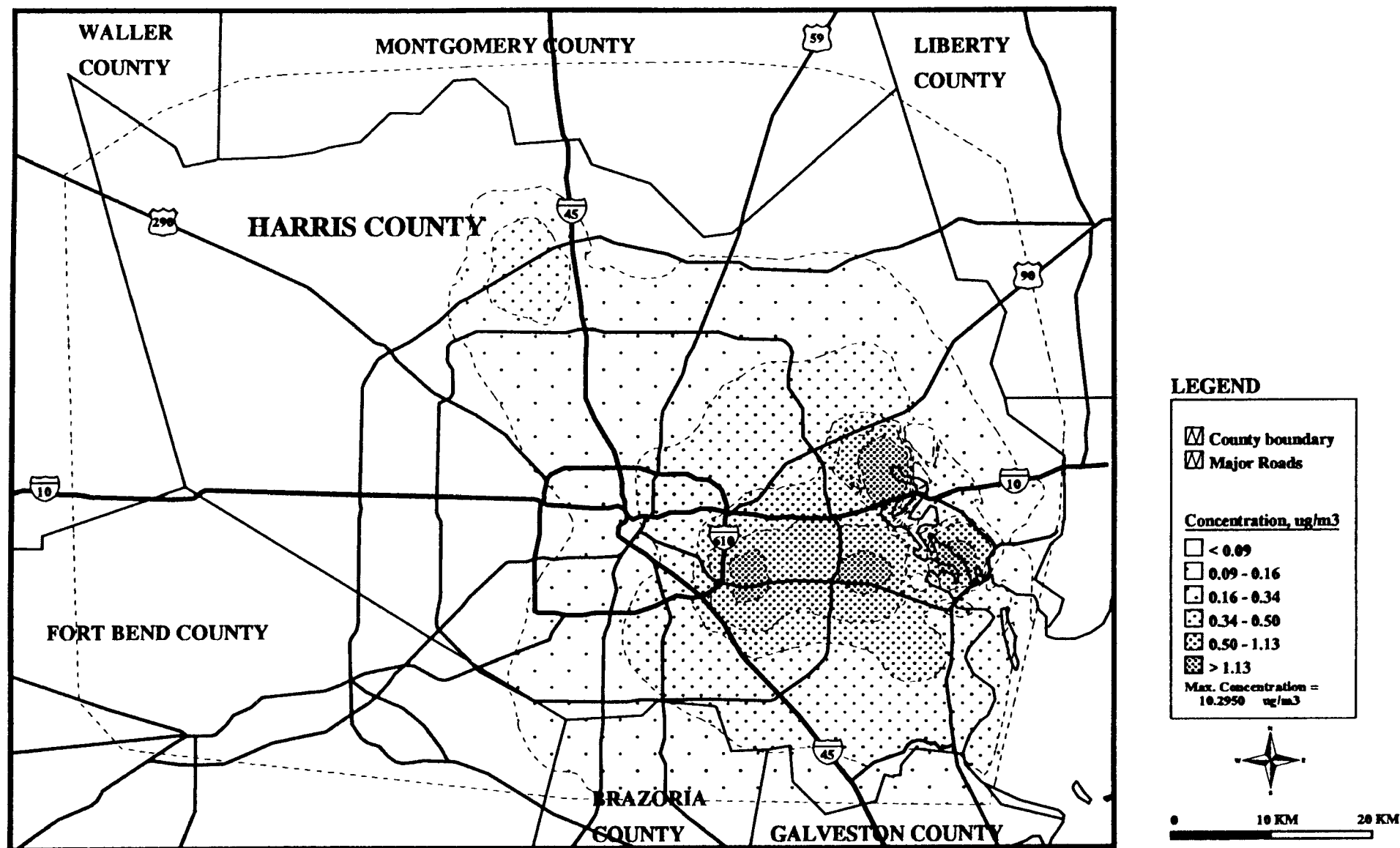


Figure B.3-4. Isopleths of Annual Average Concentrations, Houston, Texas  
Benzene, Major Sources (1987-1991)

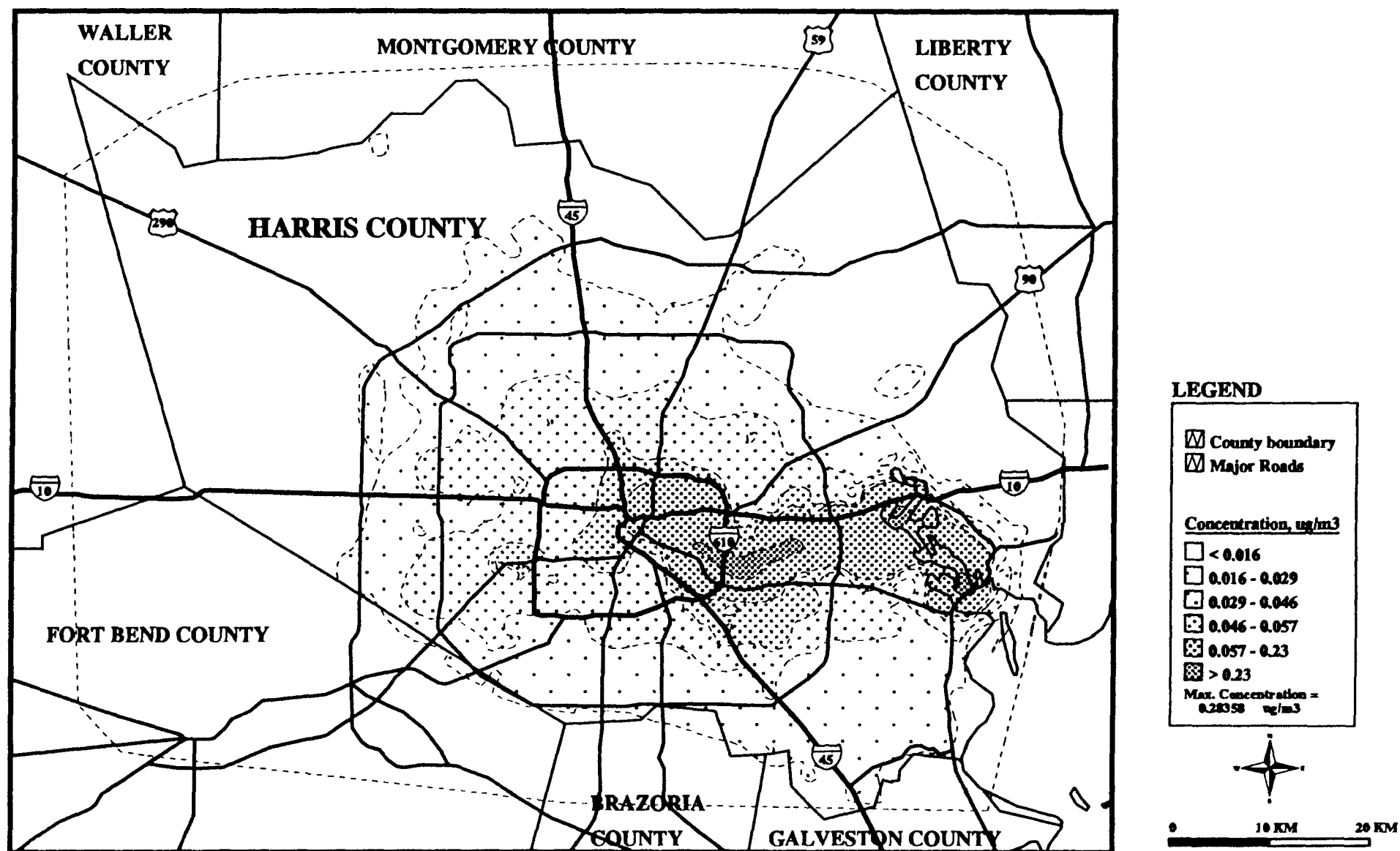


Figure B.3-5. Isopleths of Annual Average Concentrations, Houston, Texas  
Benzene, Area Sources (1987-1991)

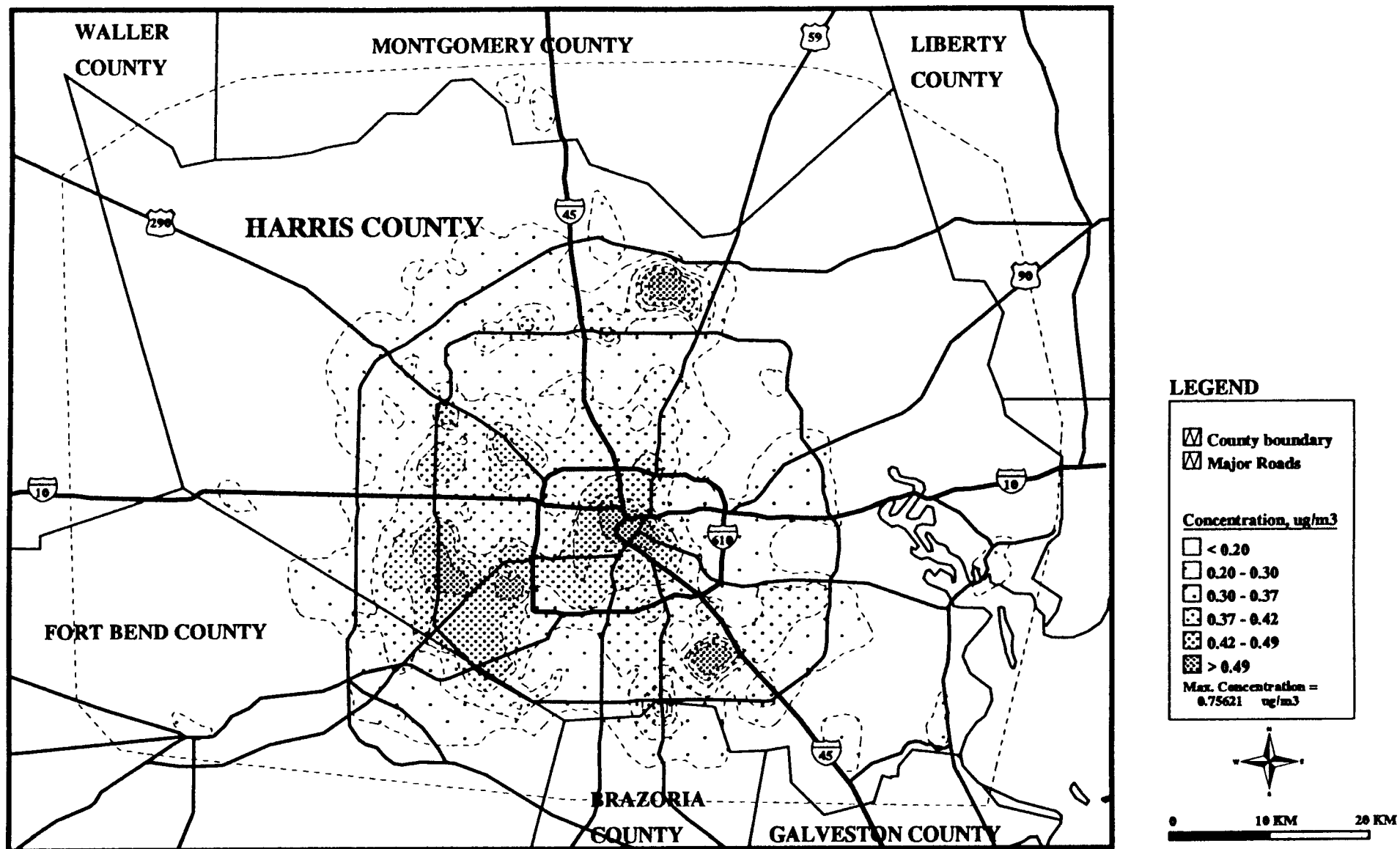


Figure B.3-6. Isopleths of Annual Average Concentrations, Houston, Texas  
Benzene, Mobile Sources (1987-1991)

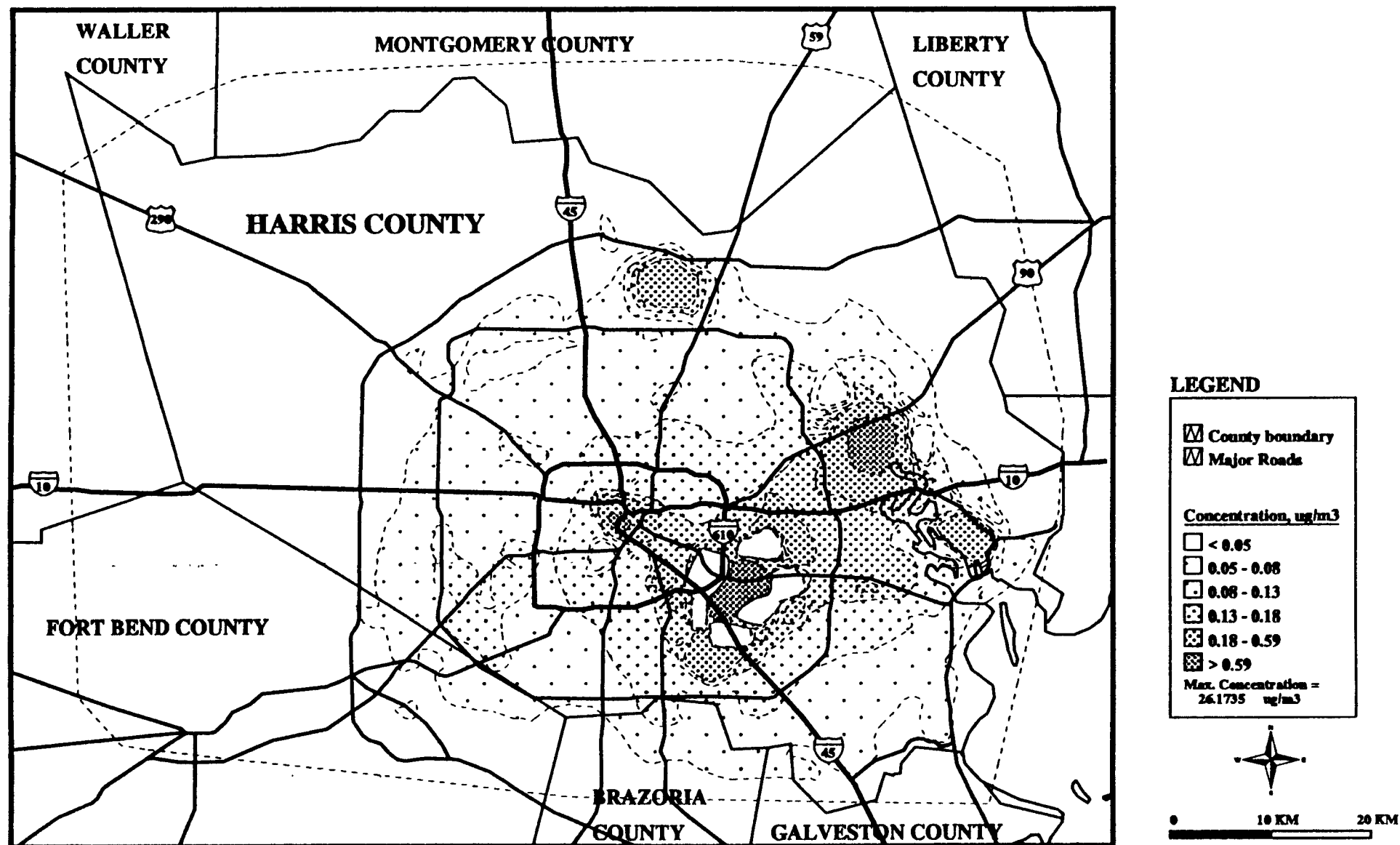


Figure B.3-7. Isopleths of Annual Average Concentrations, Houston, Texas  
1,3-Butadiene, All Sources (1987-1991)

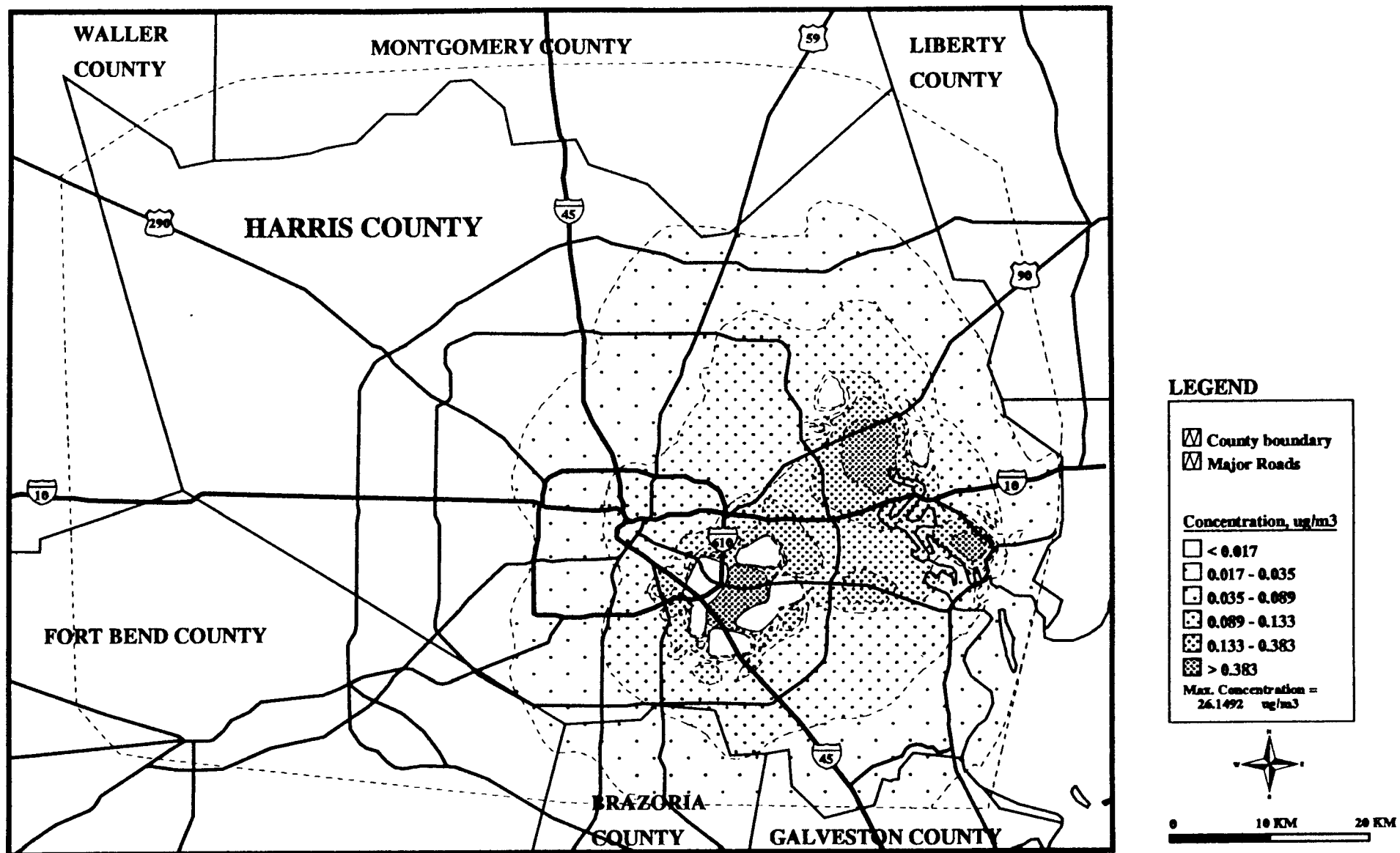


Figure B.3-8. Isopleths of Annual Average Concentrations, Houston, Texas  
1,3-Butadiene, Major Sources (1987-1991)

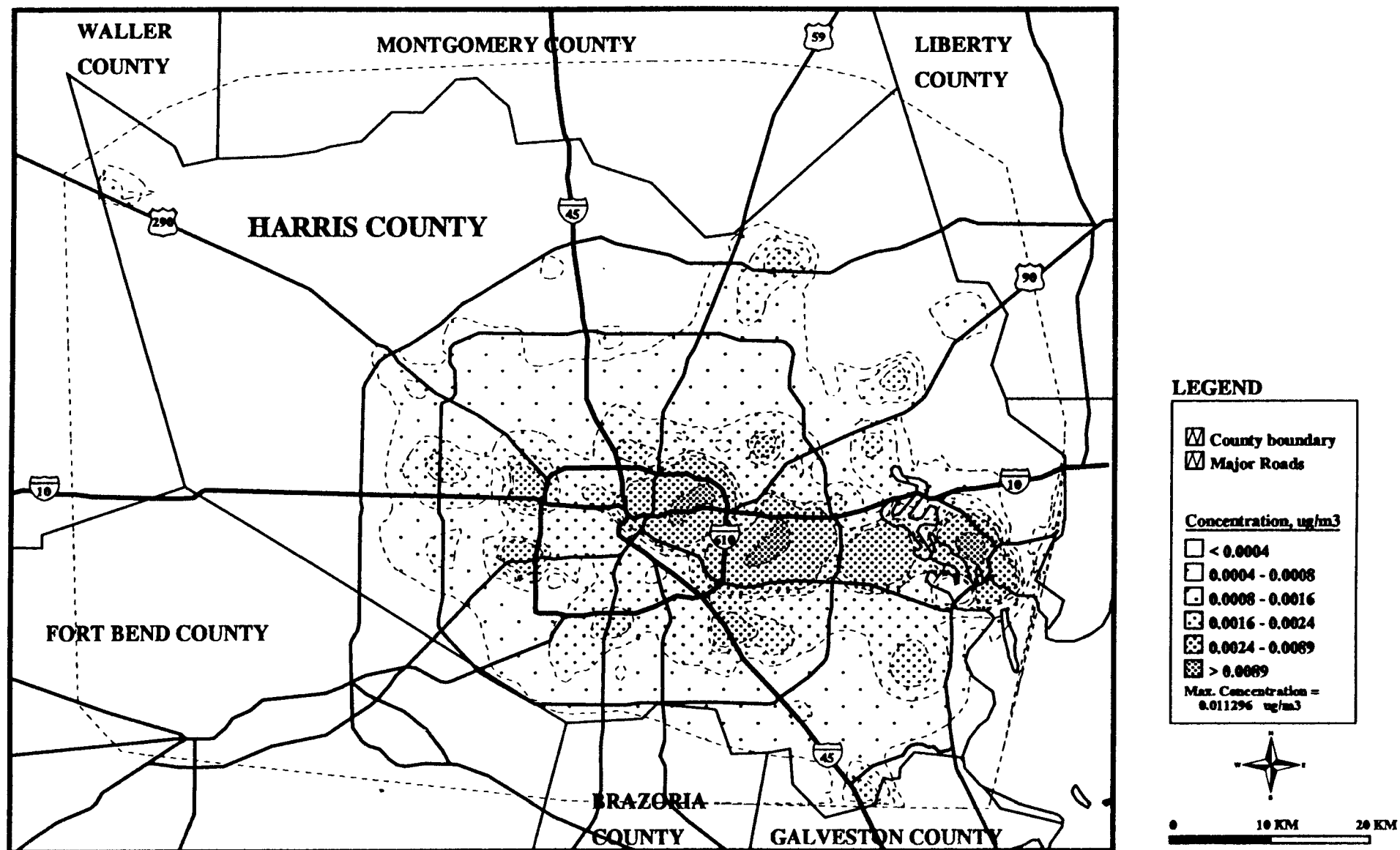


Figure B.3-9. Isopleths of Annual Average Concentrations, Houston, Texas  
1,3-Butadiene, Area Sources (1987-1991)

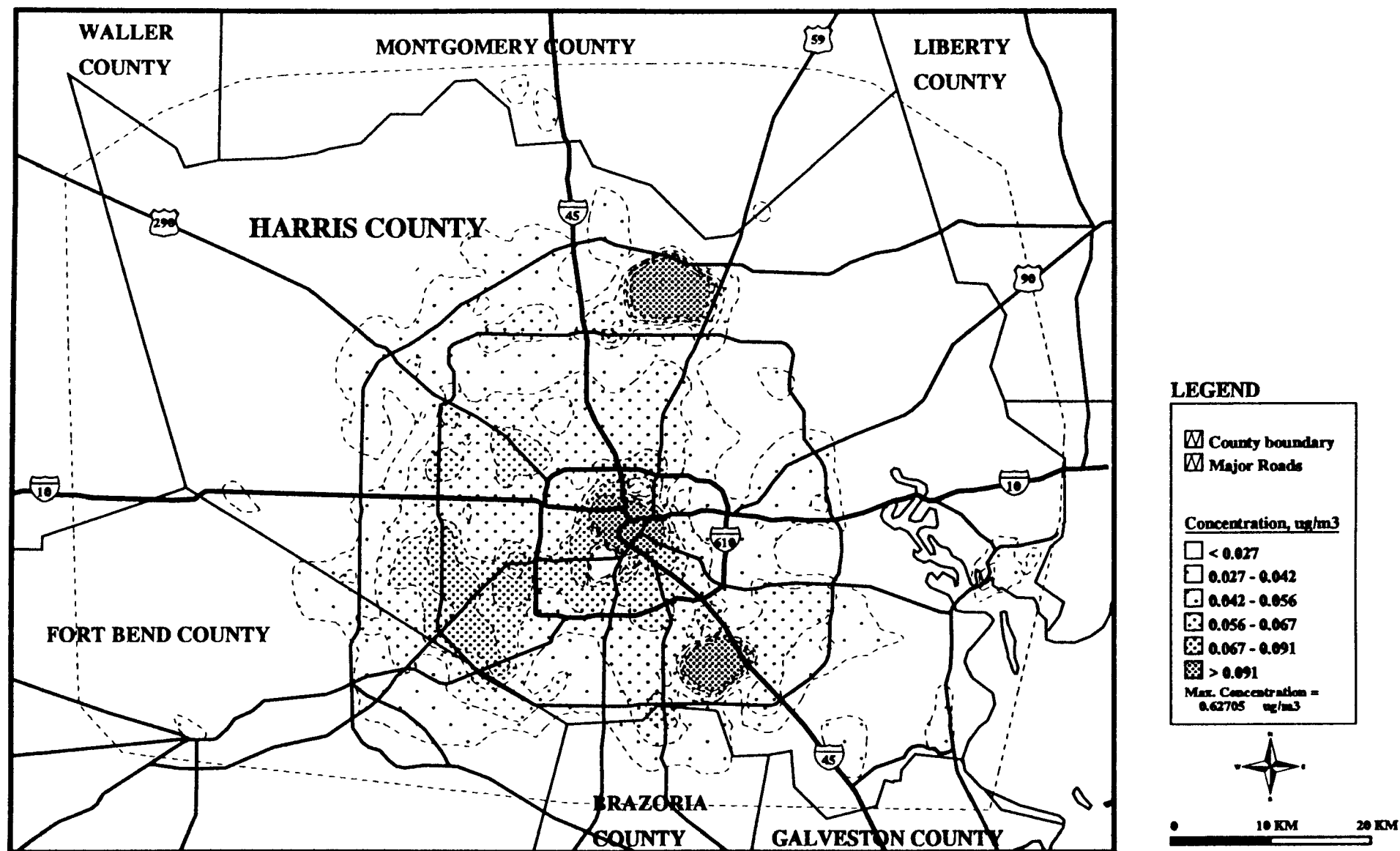


Figure B.3-10. Isopleths of Annual Average Concentrations, Houston, Texas  
1,3-Butadiene, Mobile Sources (1987-1991)

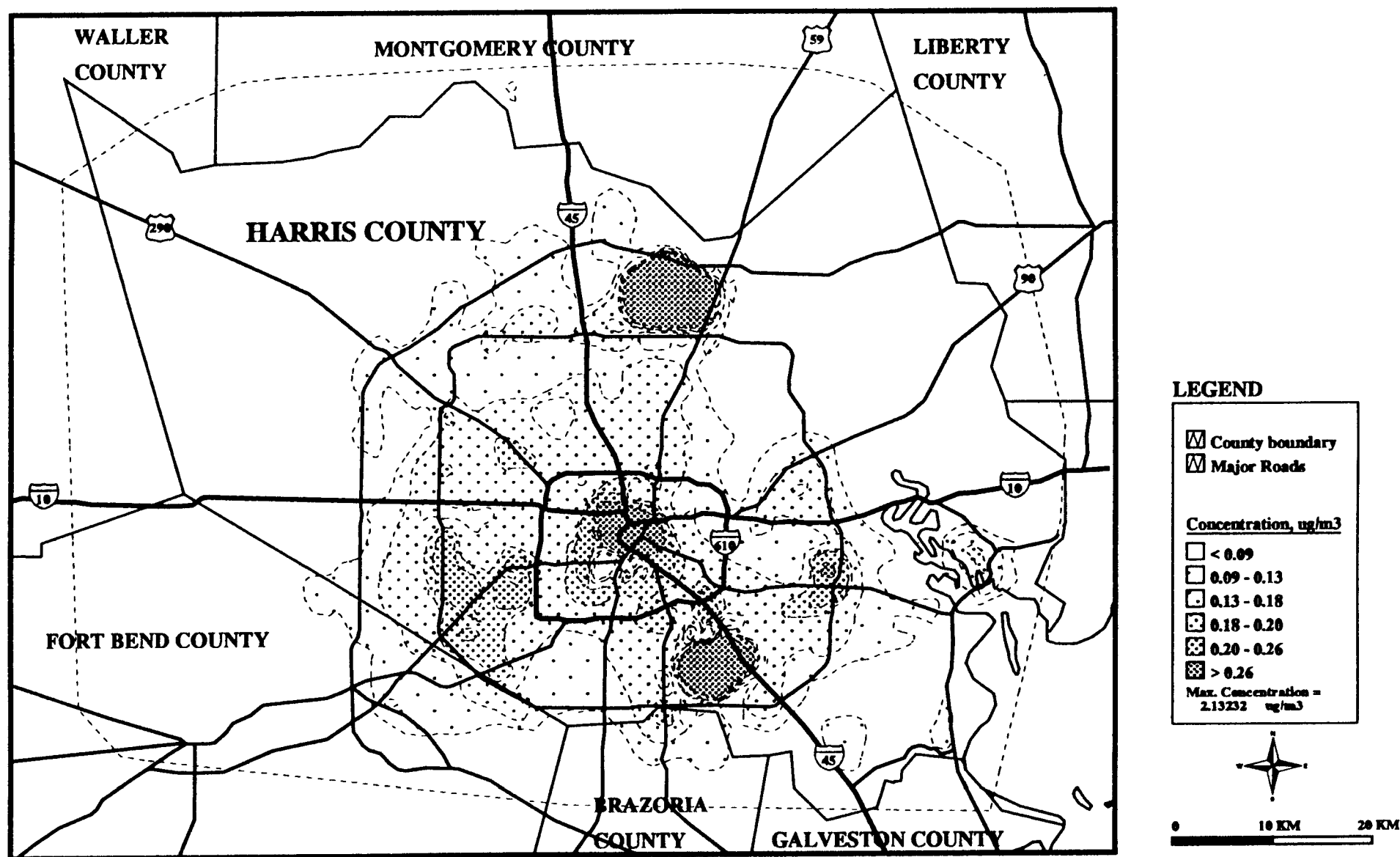


Figure B.3-11. Isopleths of Annual Average Concentrations, Houston, Texas  
Primary Formaldehyde, All Sources (1987-1991)



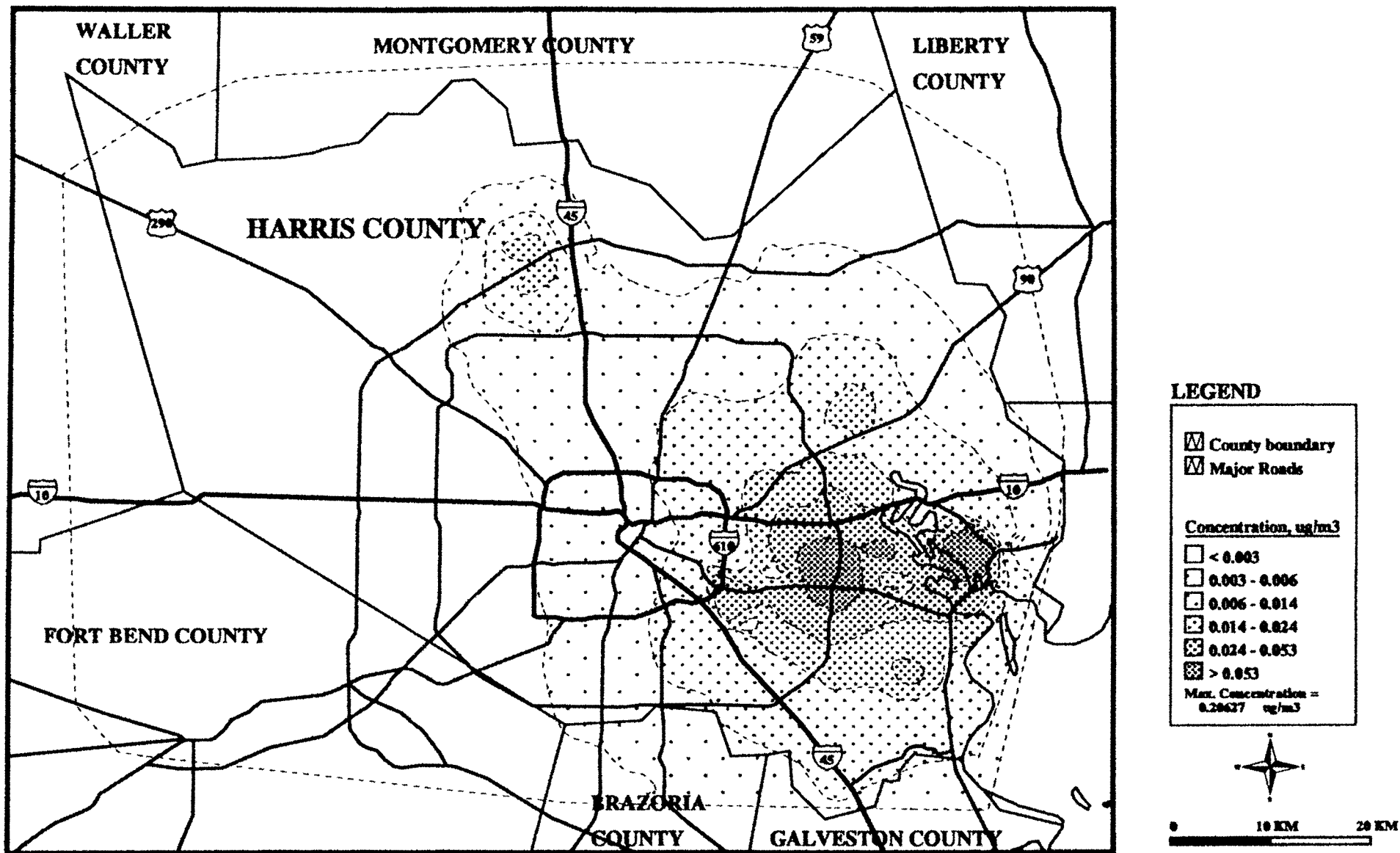


Figure B.3-12. Isopleths of Annual Average Concentrations, Houston, Texas  
Primary Formaldehyde, Major Sources (1987-1991)

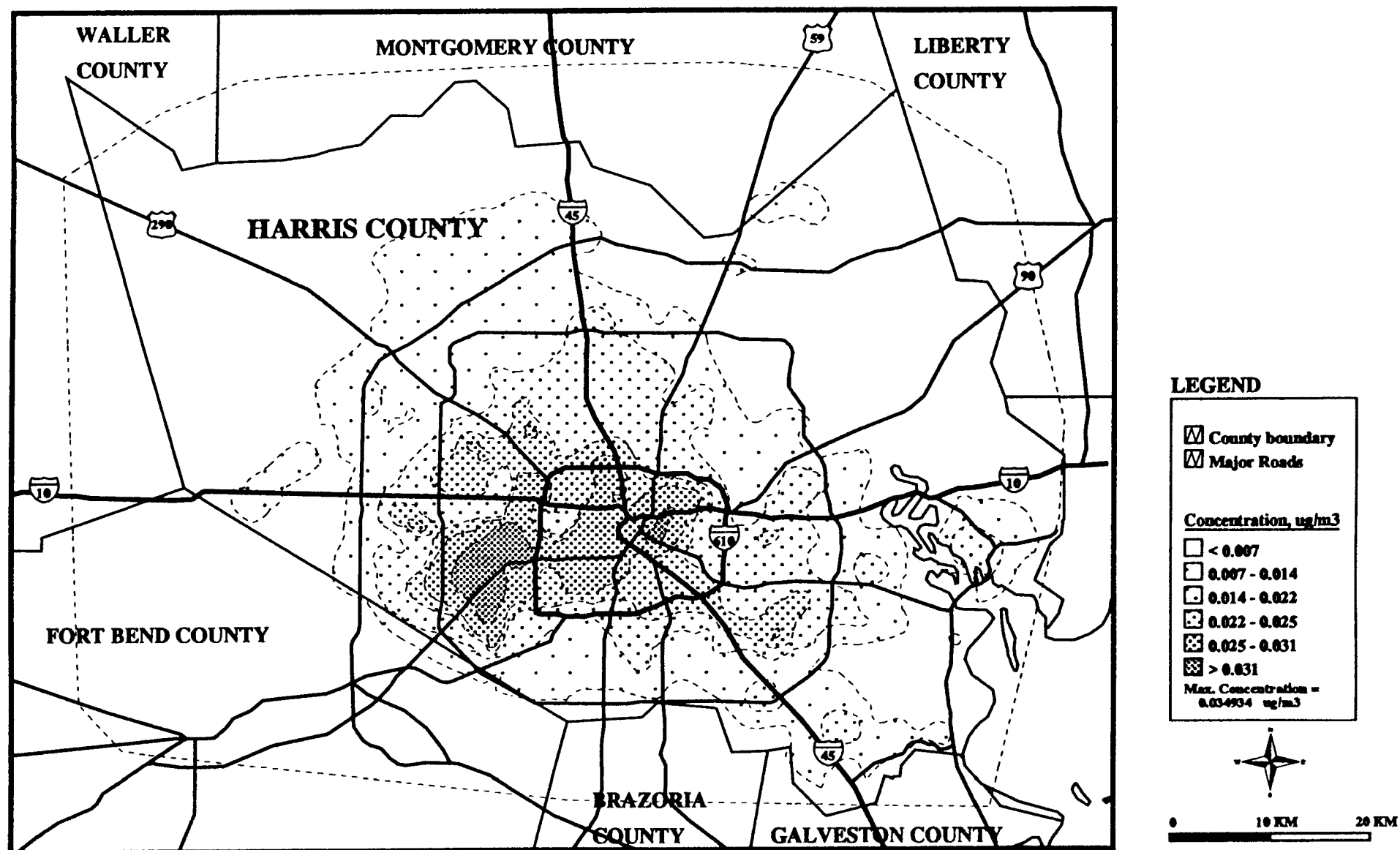


Figure B.3-13. Isopleths of Annual Average Concentrations, Houston, Texas  
Primary Formaldehyde, Area Sources (1987-1991)

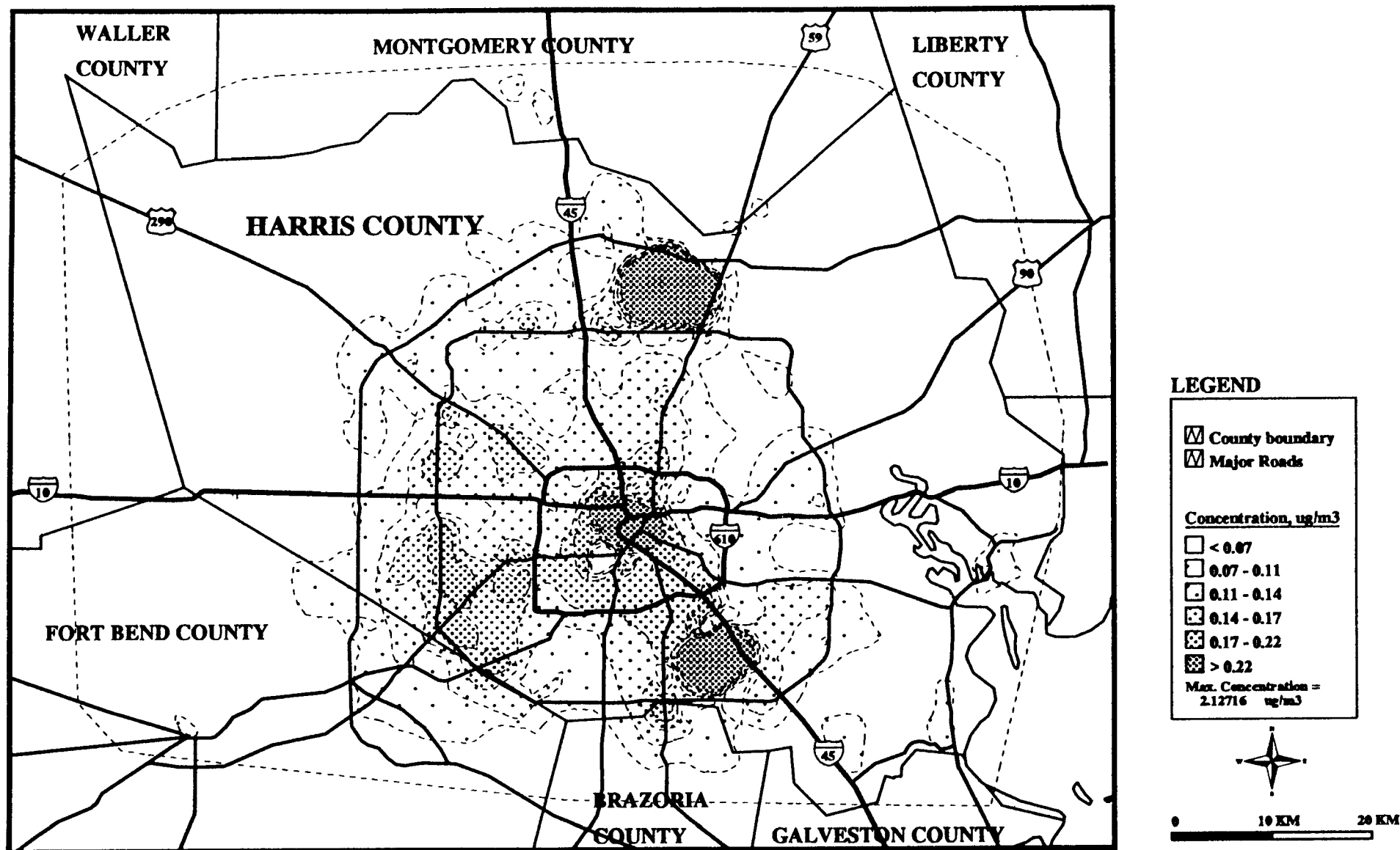


Figure B.3-14. Isopleths of Annual Average Concentrations, Houston, Texas  
Primary Formaldehyde, Mobile Sources (1987-1991)

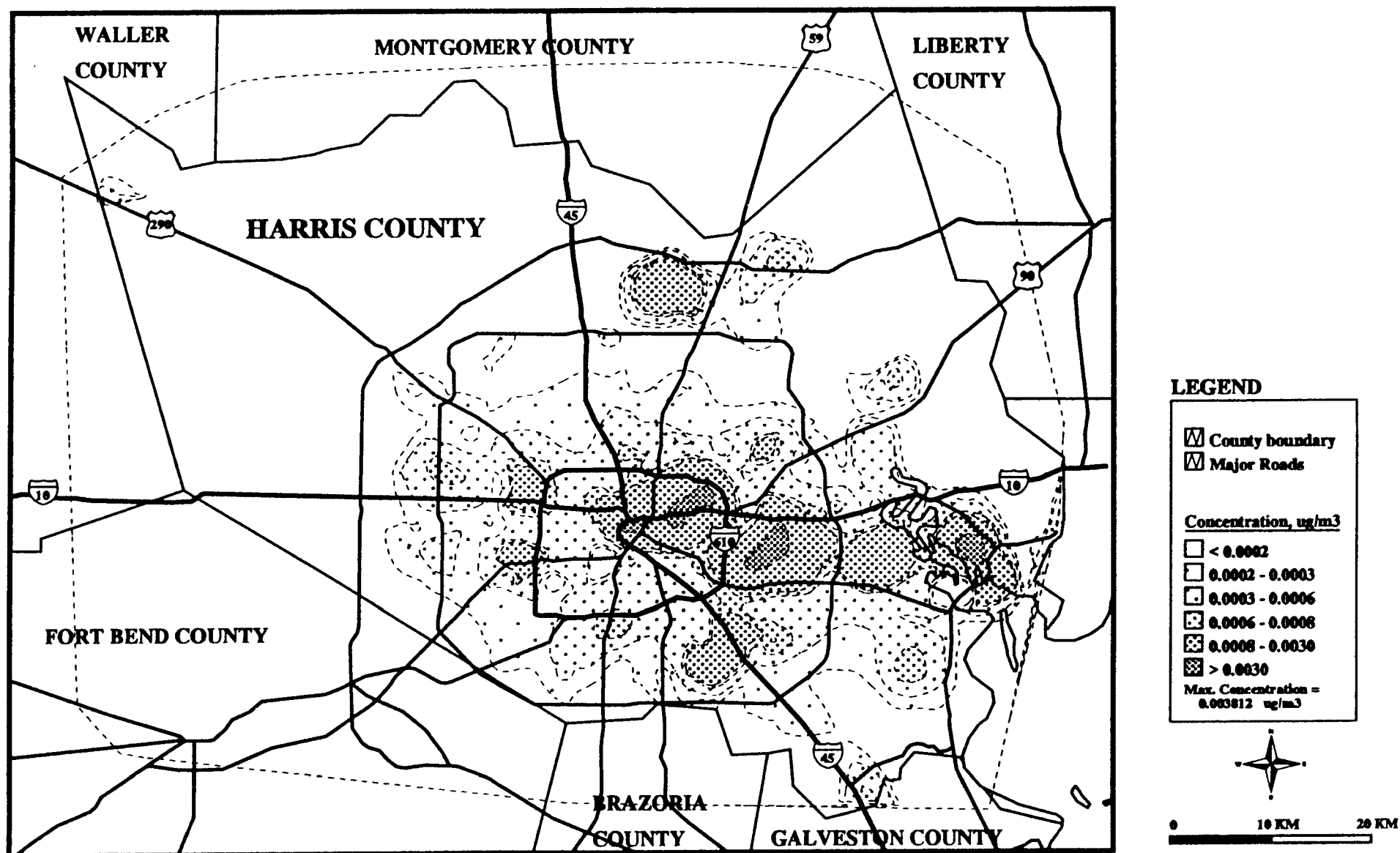


Figure B.3-15. Isopleths of Annual Average Concentrations, Houston, Texas  
POM, All Sources (1987-1991)

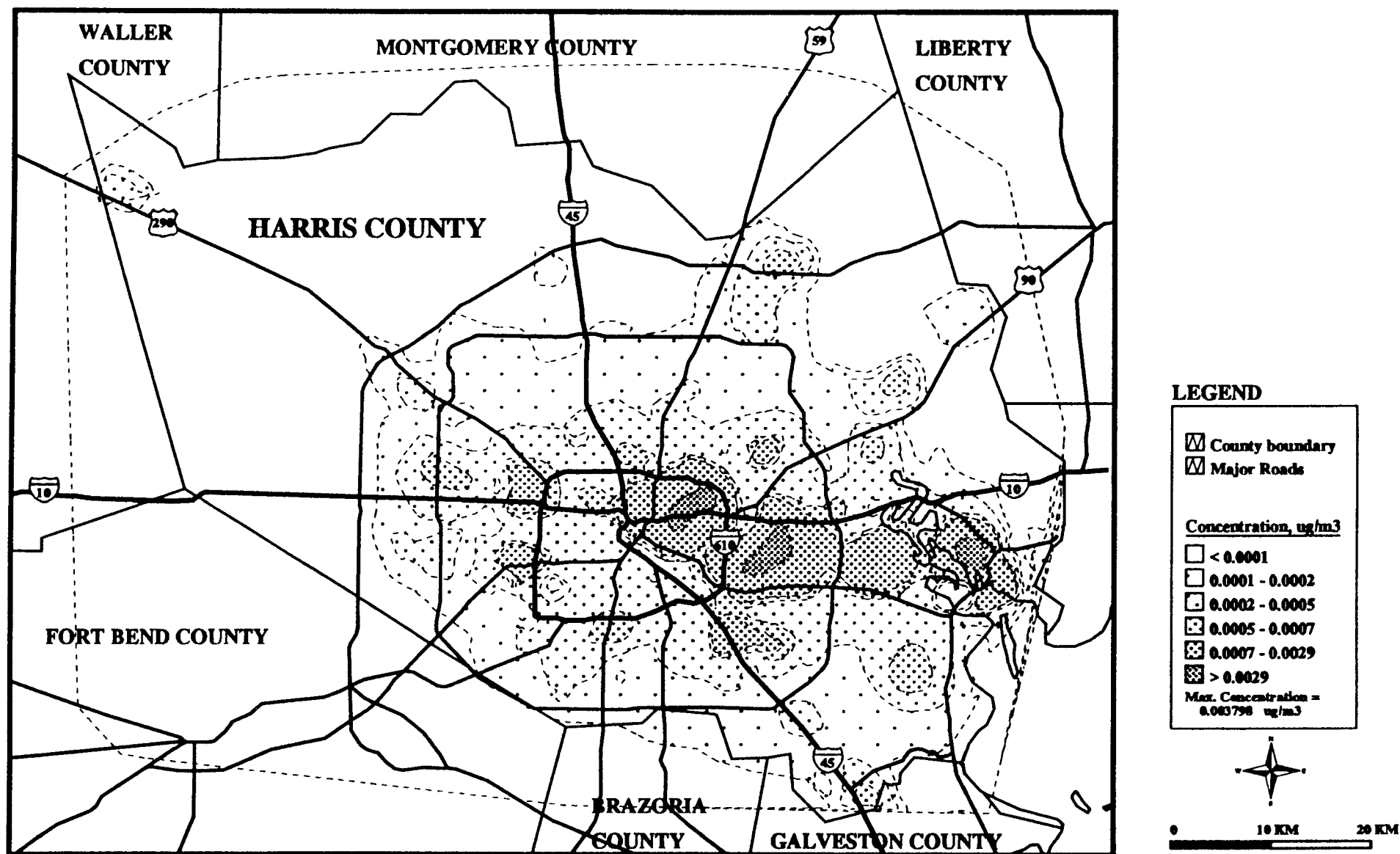


Figure B.3-16. Isopleths of Annual Average Concentrations, Houston, Texas  
POM, Area Sources (1987-1991)

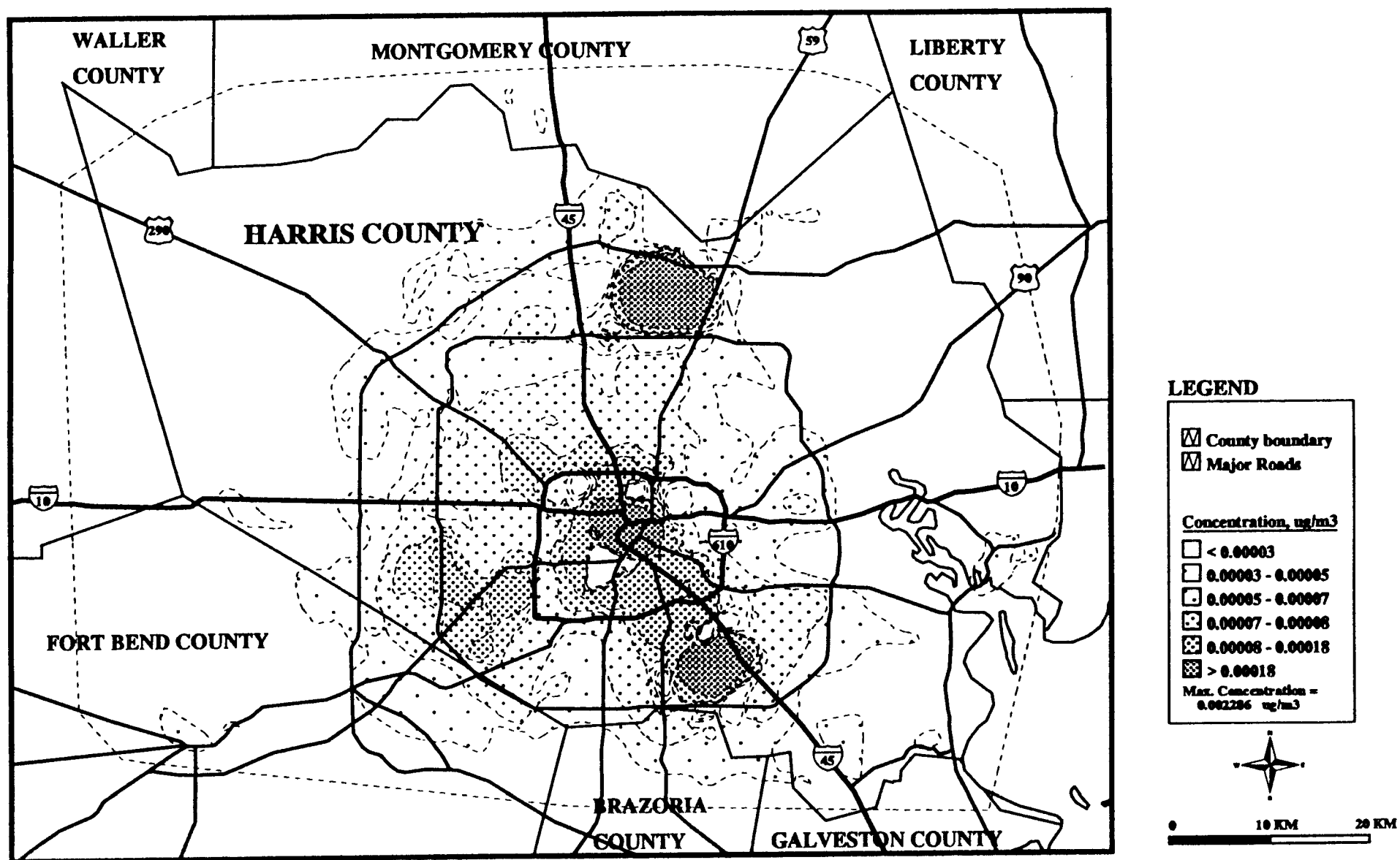


Figure B.3-17. Isopleths of Annual Average Concentrations, Houston, Texas  
POM, Mobile Sources (1987-1991)

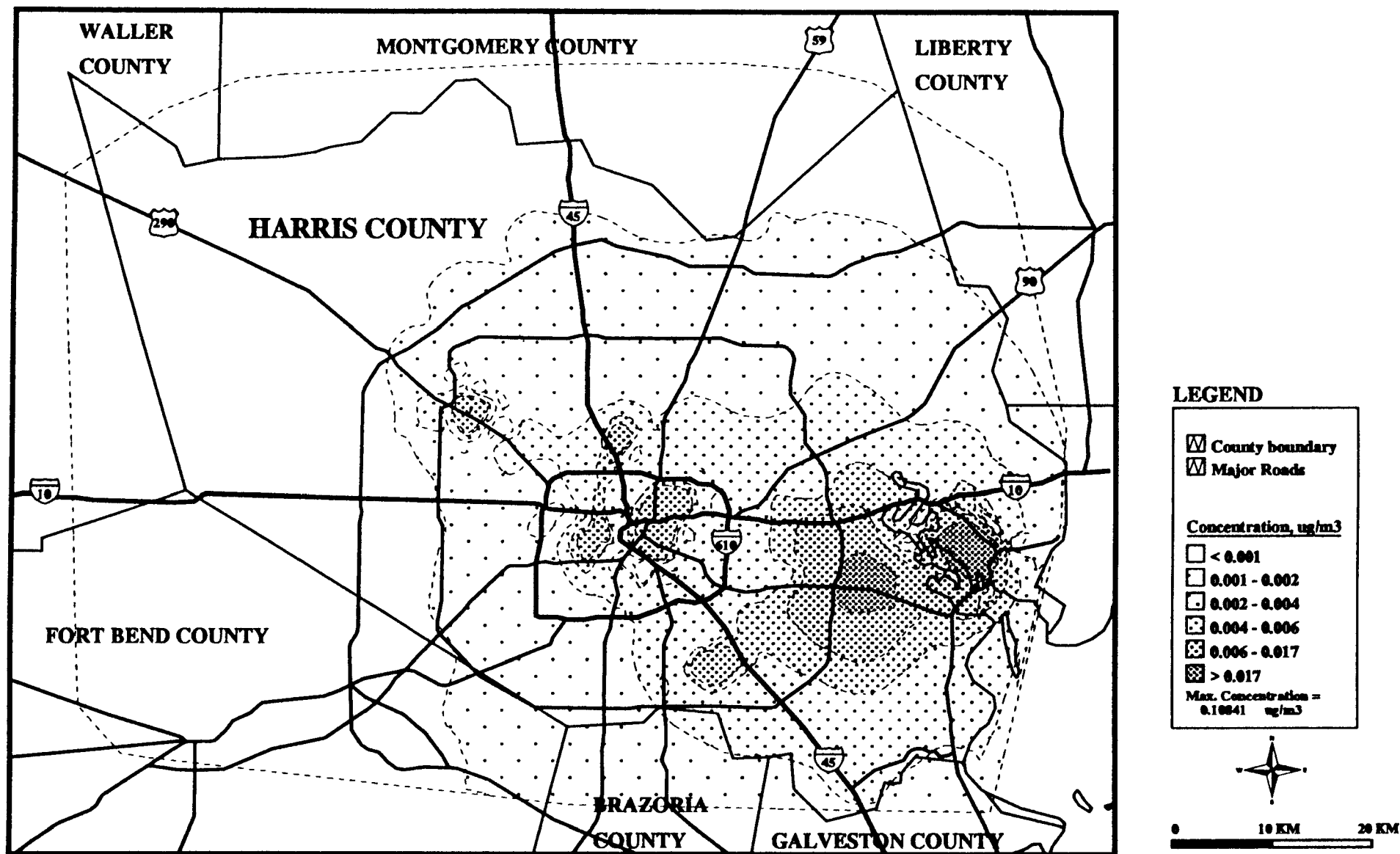


Figure B.3-18. Isopleths of Annual Average Concentrations, Houston, Texas  
Hexavalent Chromium, All Sources (1987-1991)

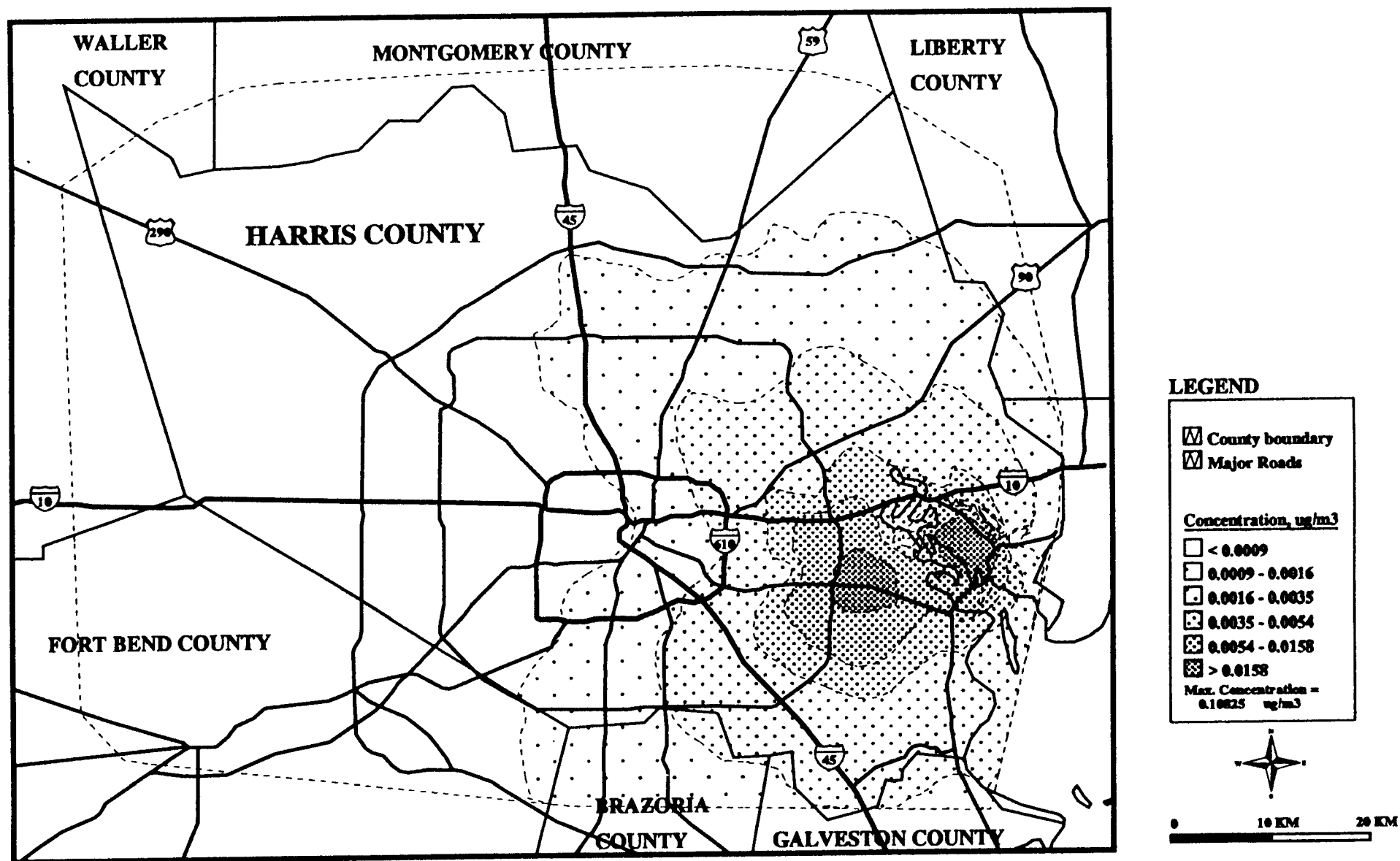


Figure B.3-19. Isopleths of Annual Average Concentrations, Houston, Texas  
Hexavalent Chromium, Major Sources (1987-1991)



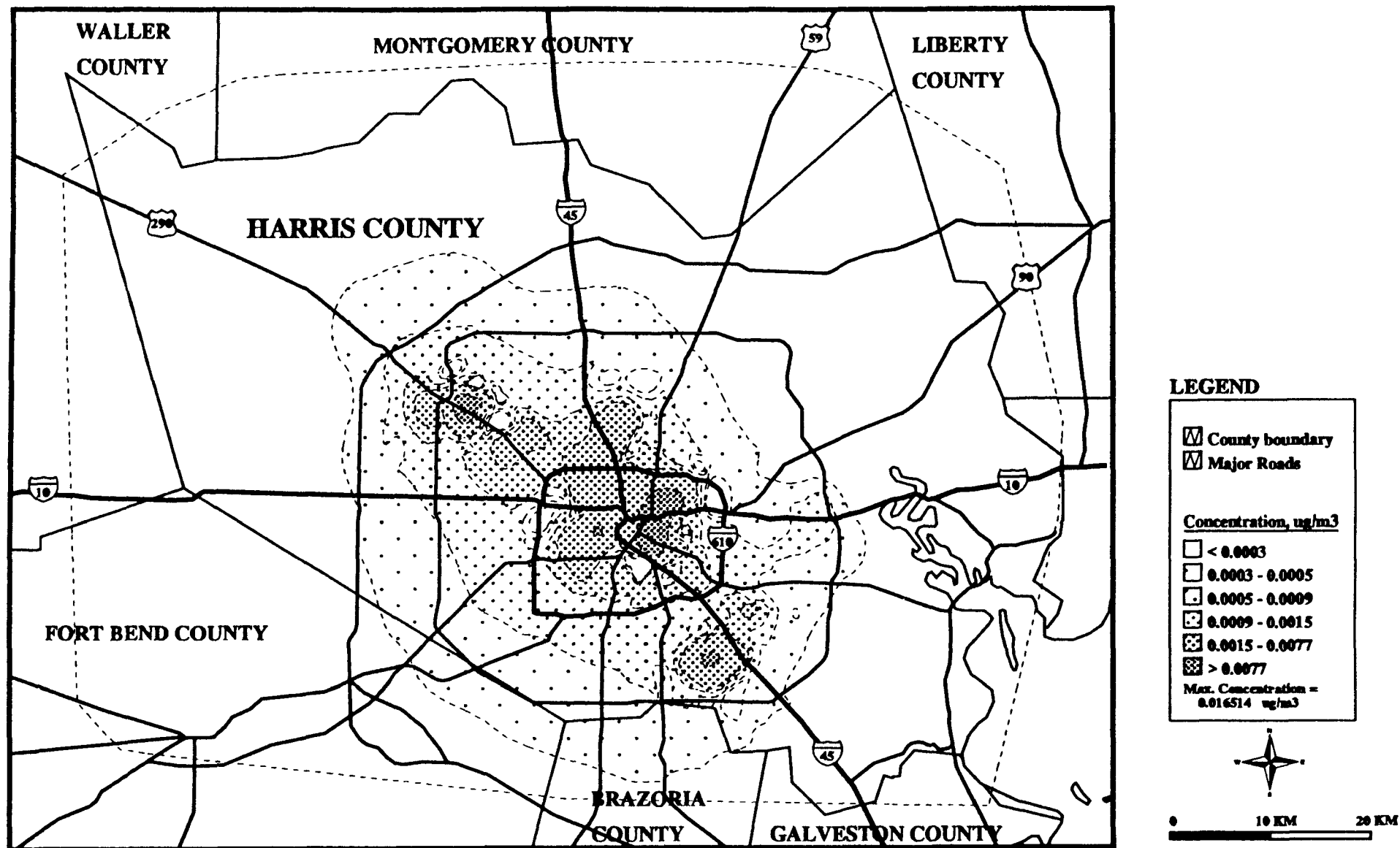


Figure B.3-20. Isopleths of Annual Average Concentrations, Houston, Texas  
Hexavalent Chromium, Area Sources (1987-1991)

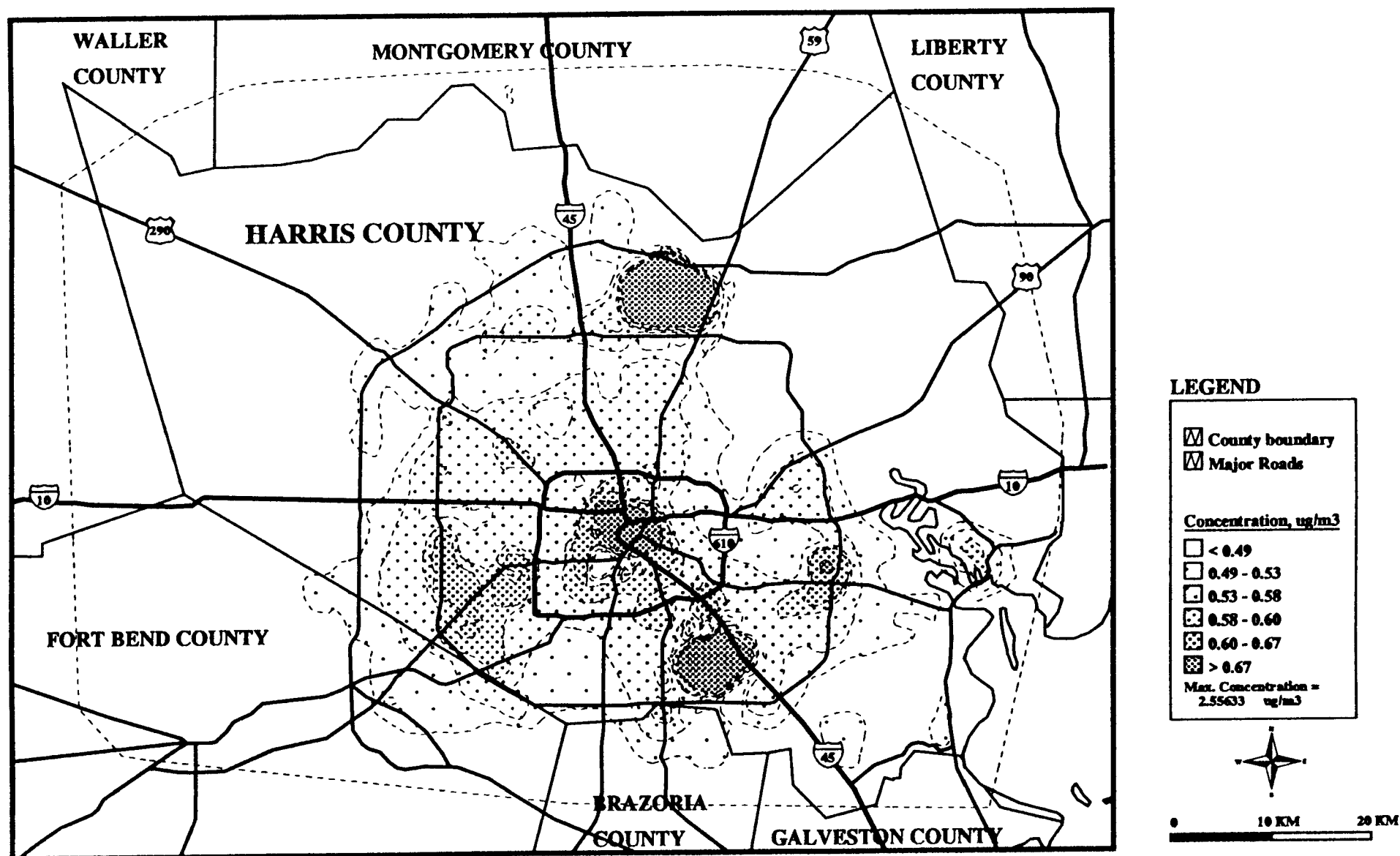
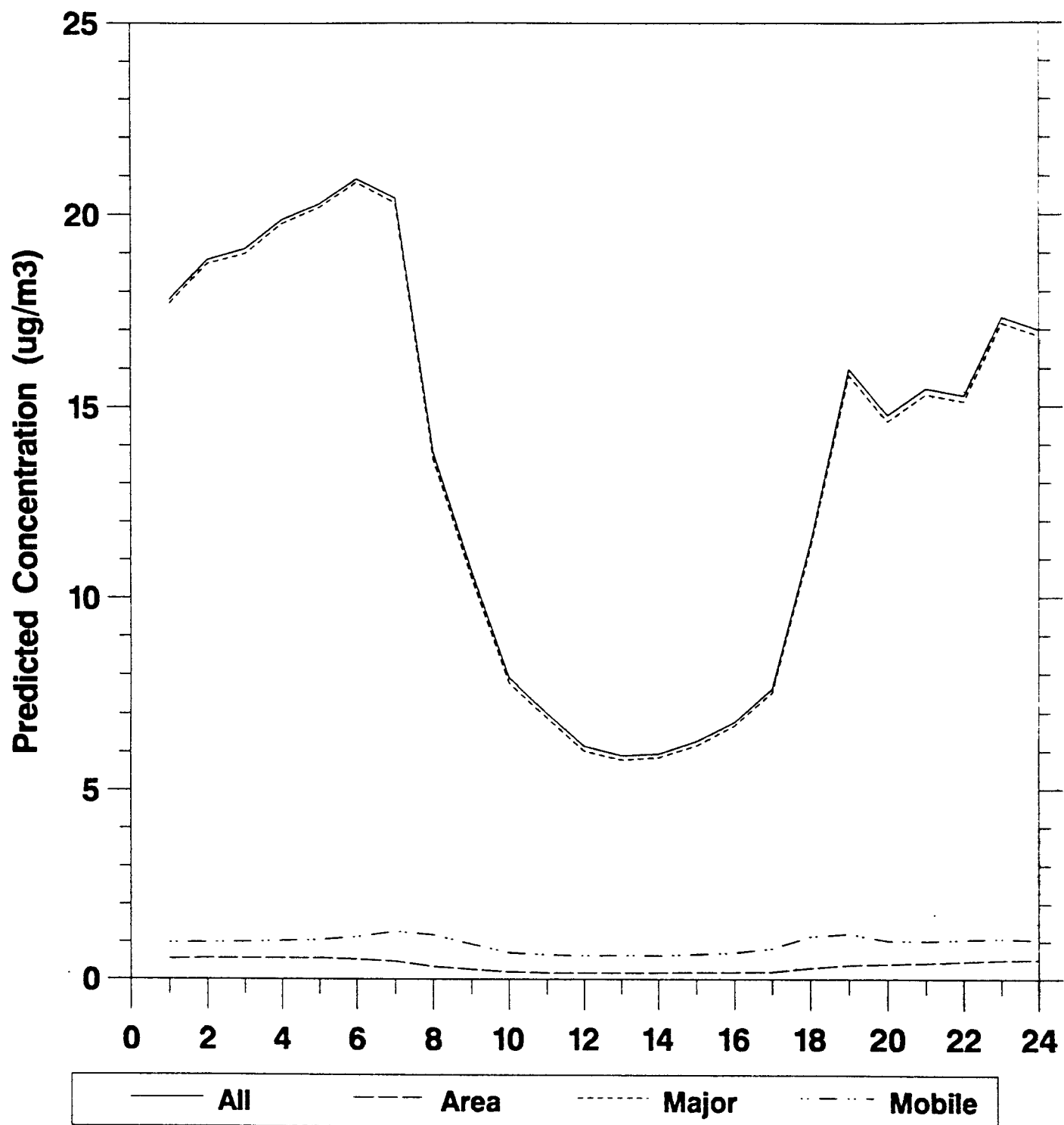
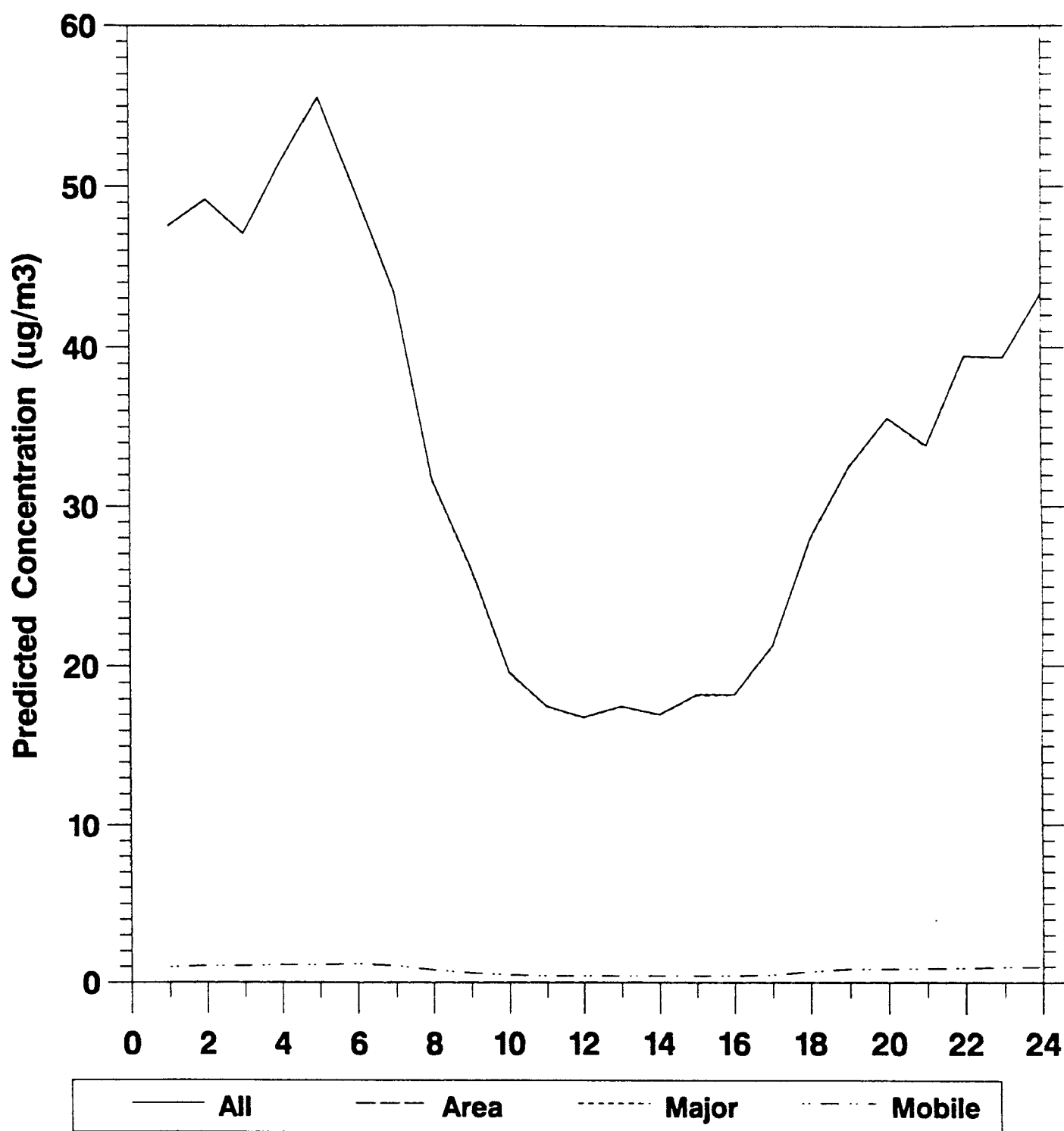


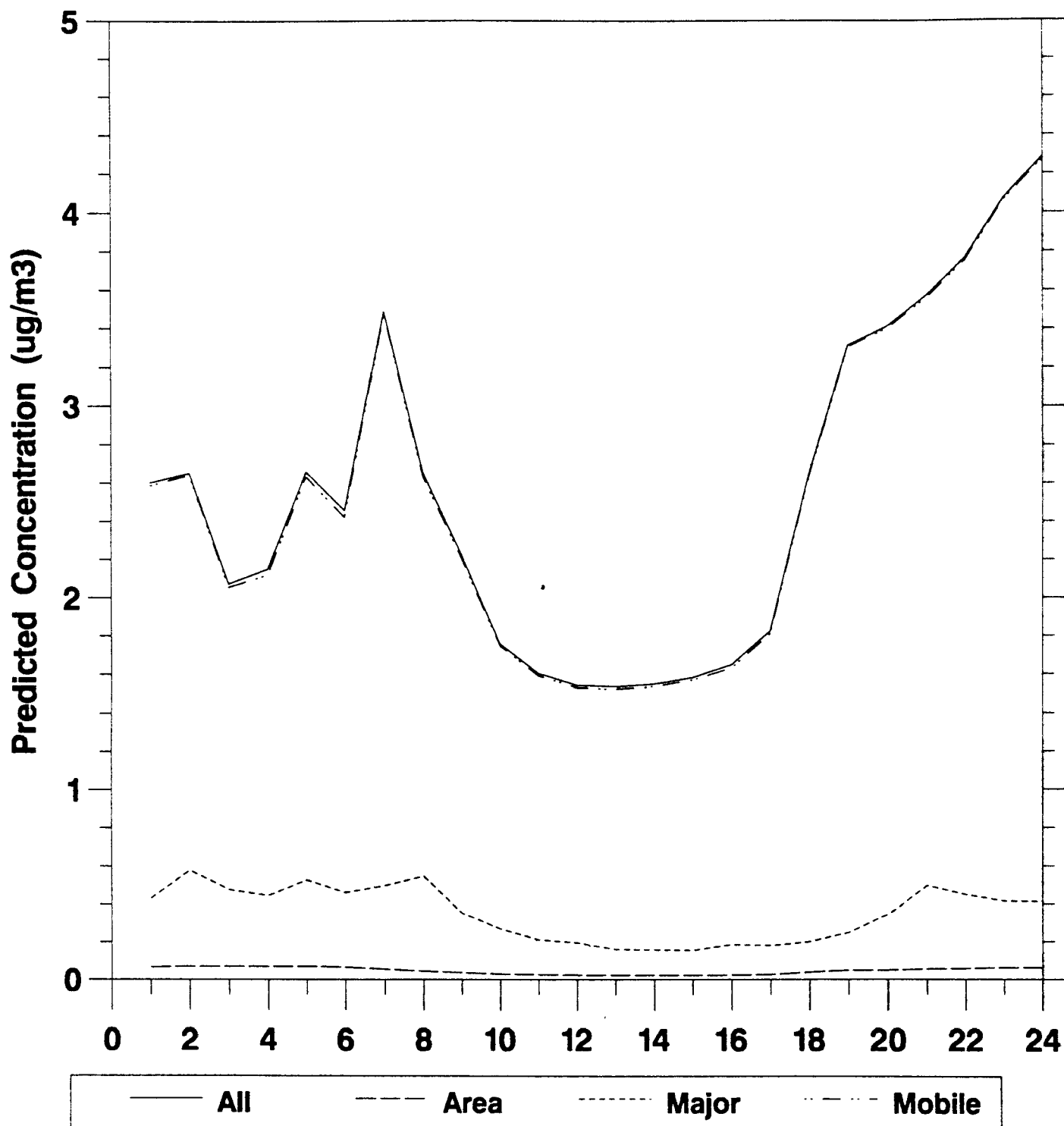
Figure B.3-21. Isopleths of Annual Average Concentrations, Houston, Texas  
Total Formaldehyde (Primary and Secondary), All Sources (1987-1991)



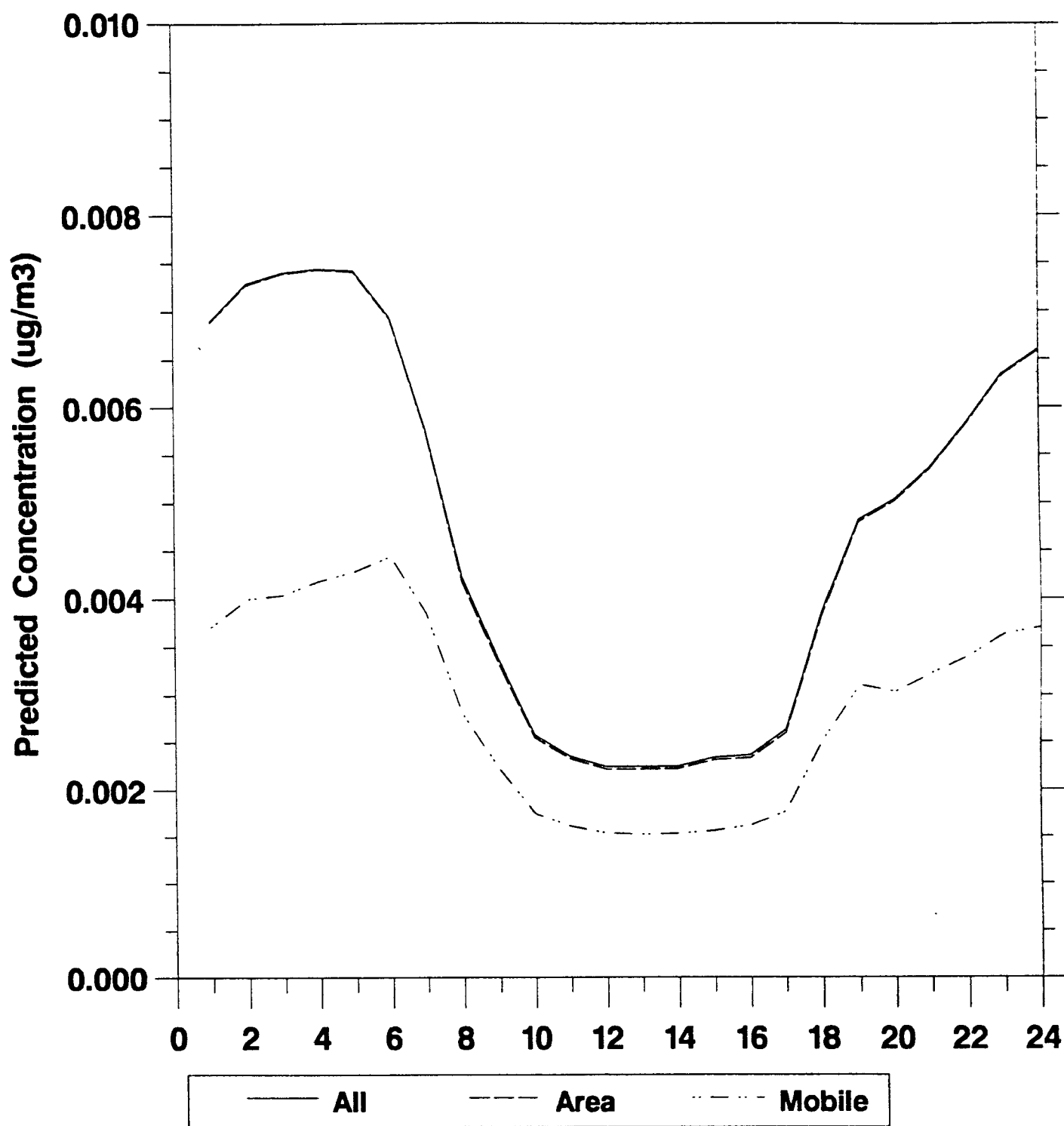
**Figure B.3-22. Urban Area Source Modeling, Houston, Texas  
Benzene Average Maximum Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



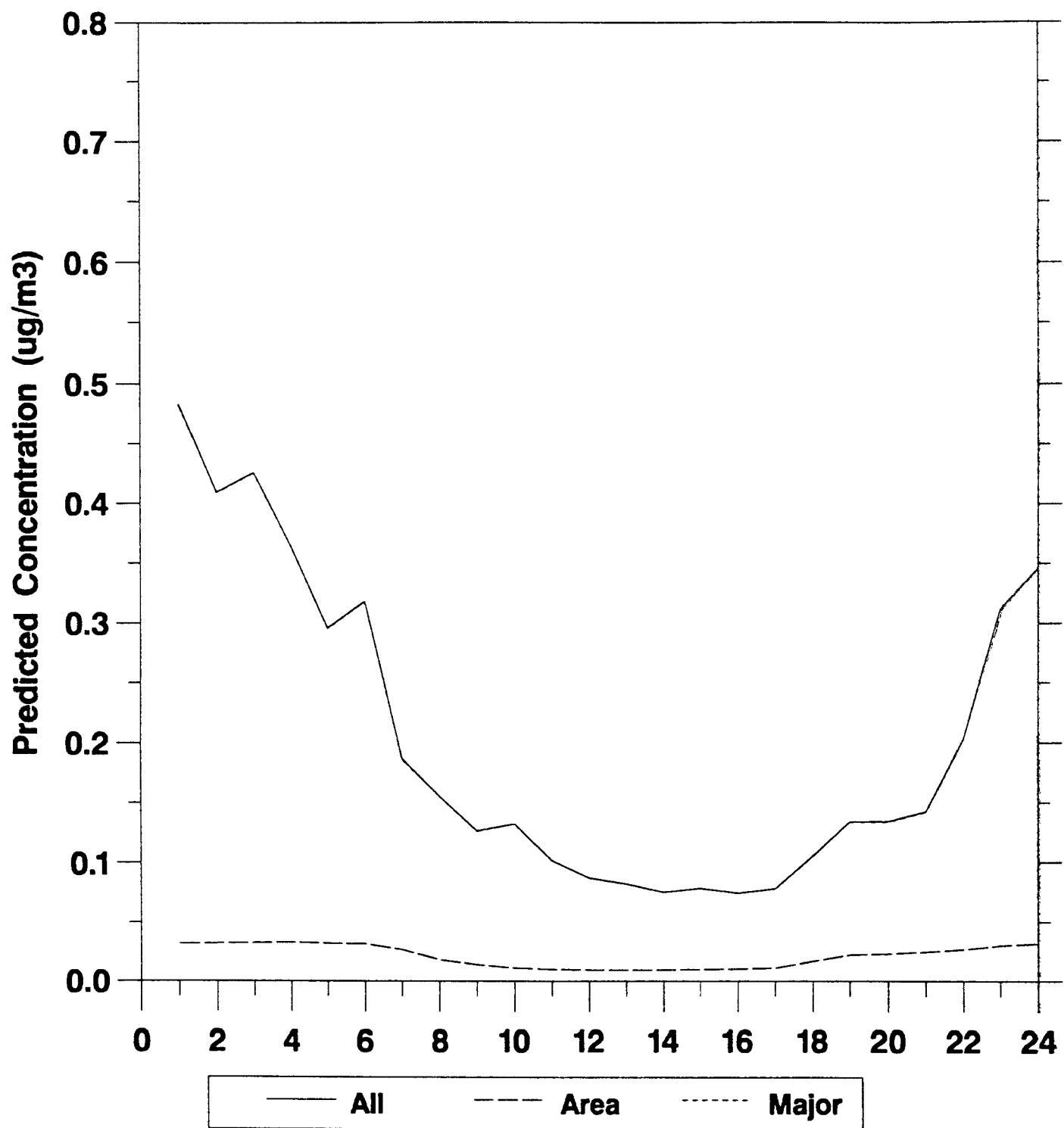
**Figure B.3-23. Urban Area Source Modeling, Houston, Texas**  
**1,3-Butadiene Average Maximum Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**



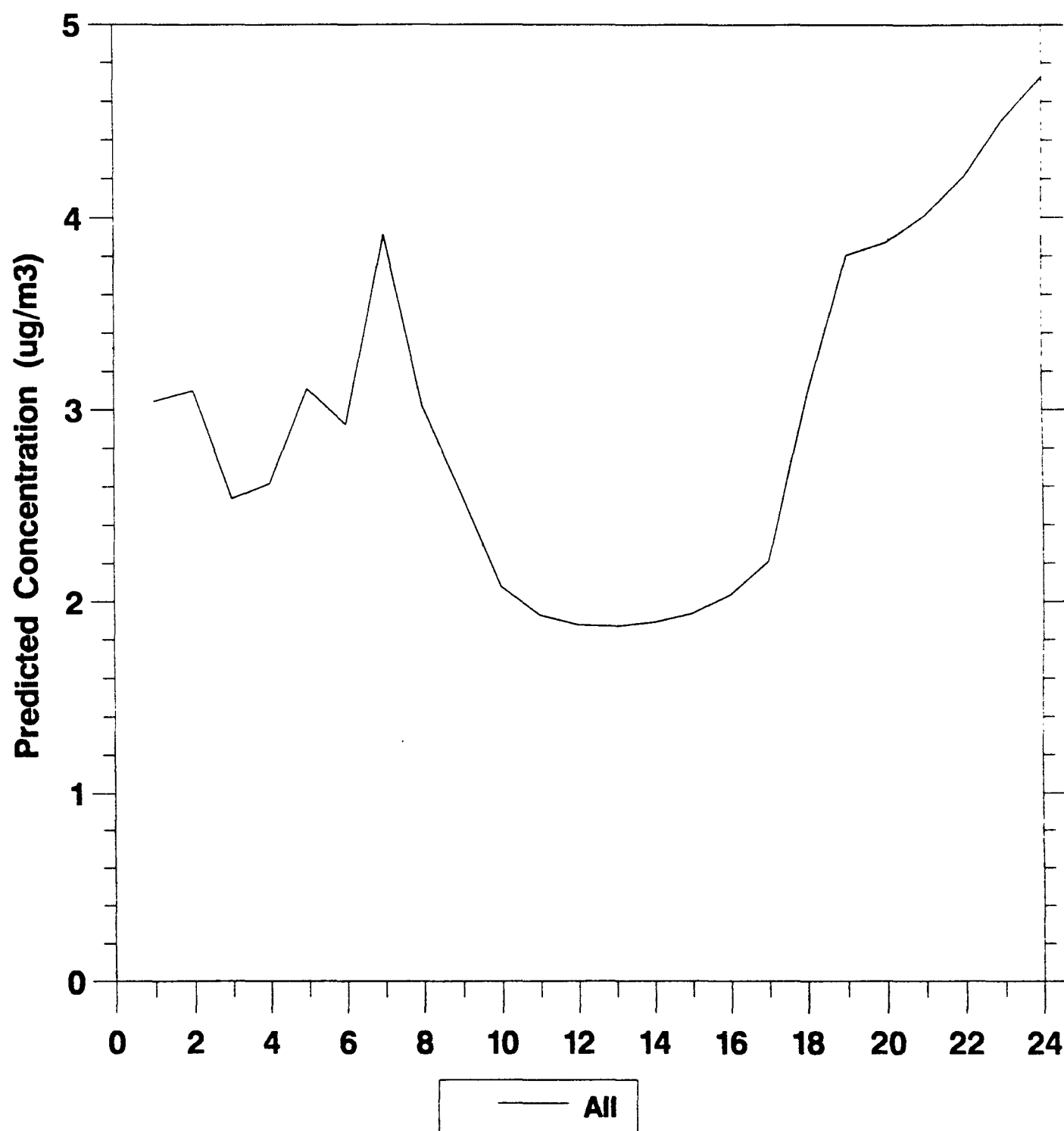
**Figure B.3-24. Urban Area Source Modeling, Houston, Texas**  
**Primary Formaldehyde Average Maximum Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**



**Figure B.3-25. Urban Area Source Modeling, Houston, Texas**  
**POM Average Maximum Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**

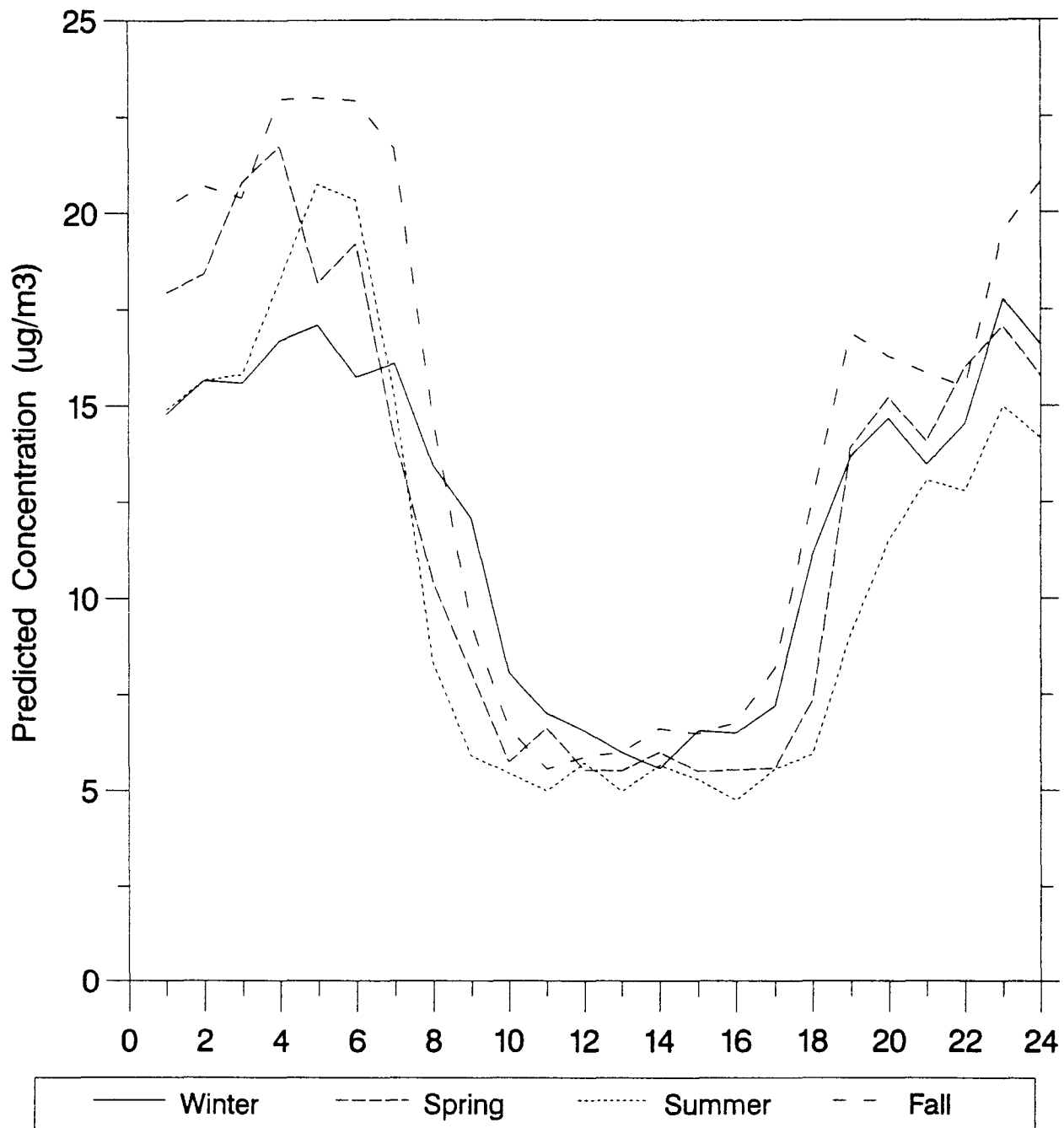


**Figure B.3-26. Urban Area Source Modeling, Houston, Texas**  
**Hexavalent Chromium Average Maximum Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**

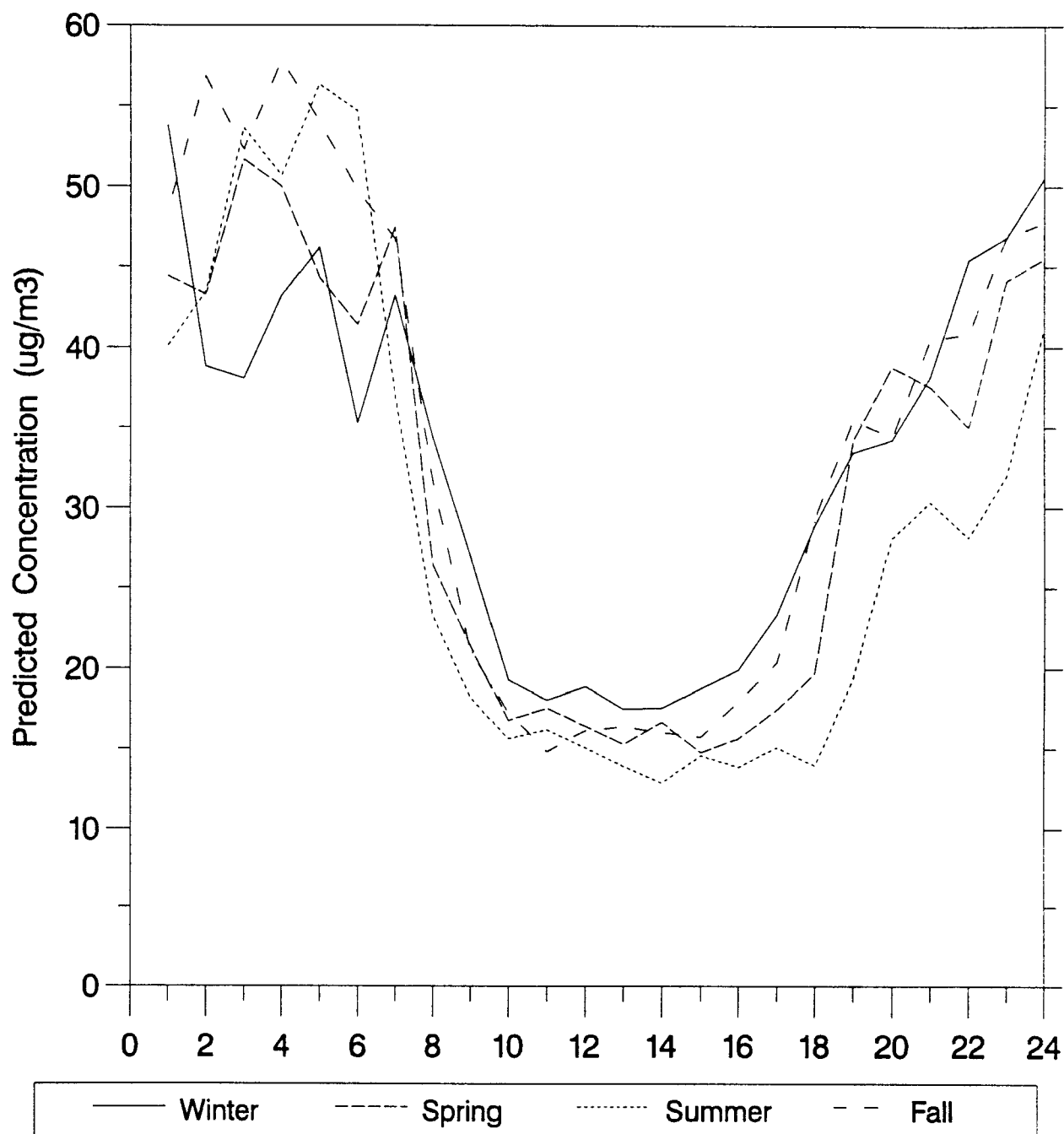


**Figure B.3-27. Urban Area Source Modeling, Houston, Texas**  
**Total Formaldehyde Average Maximum Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**

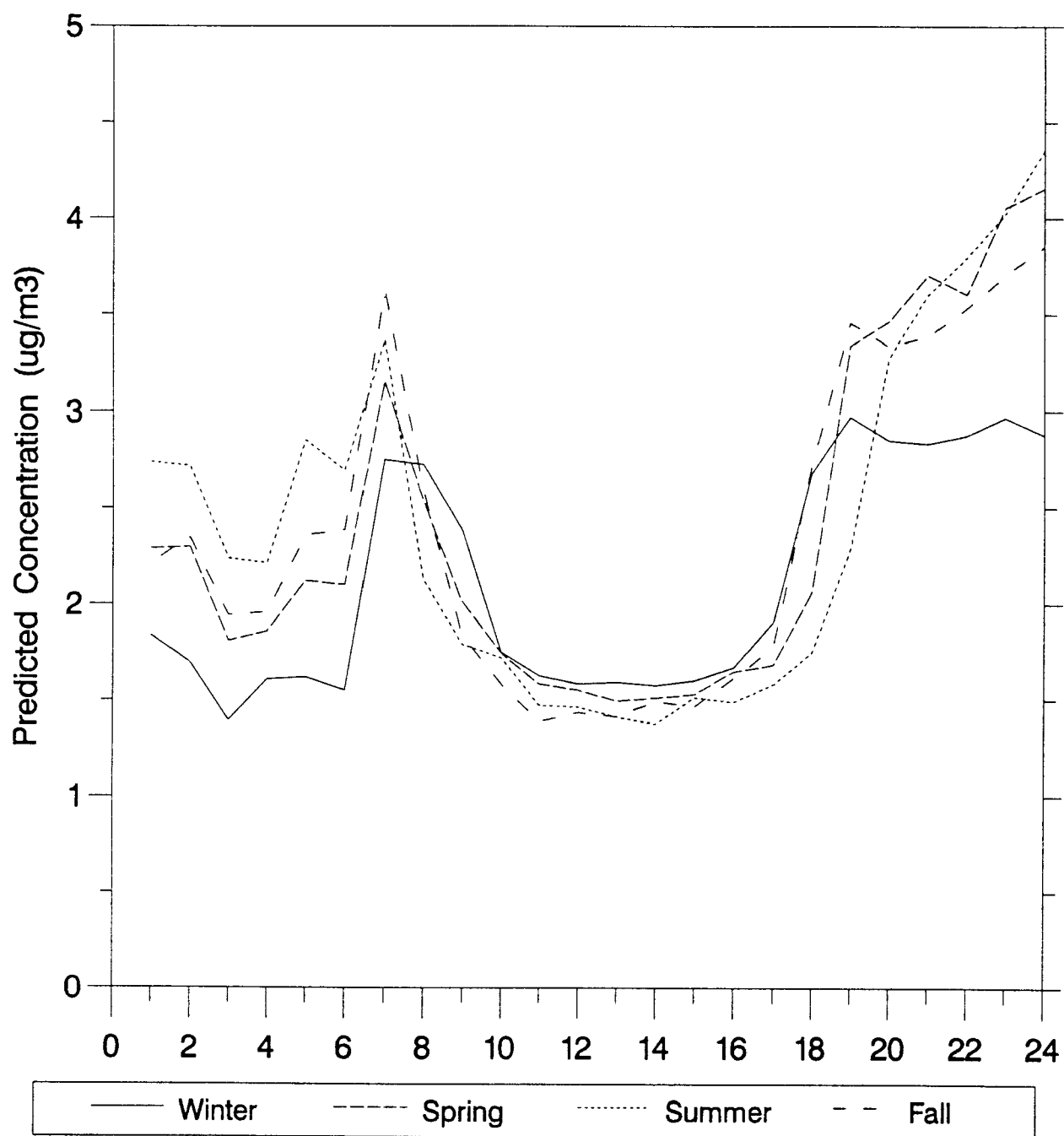




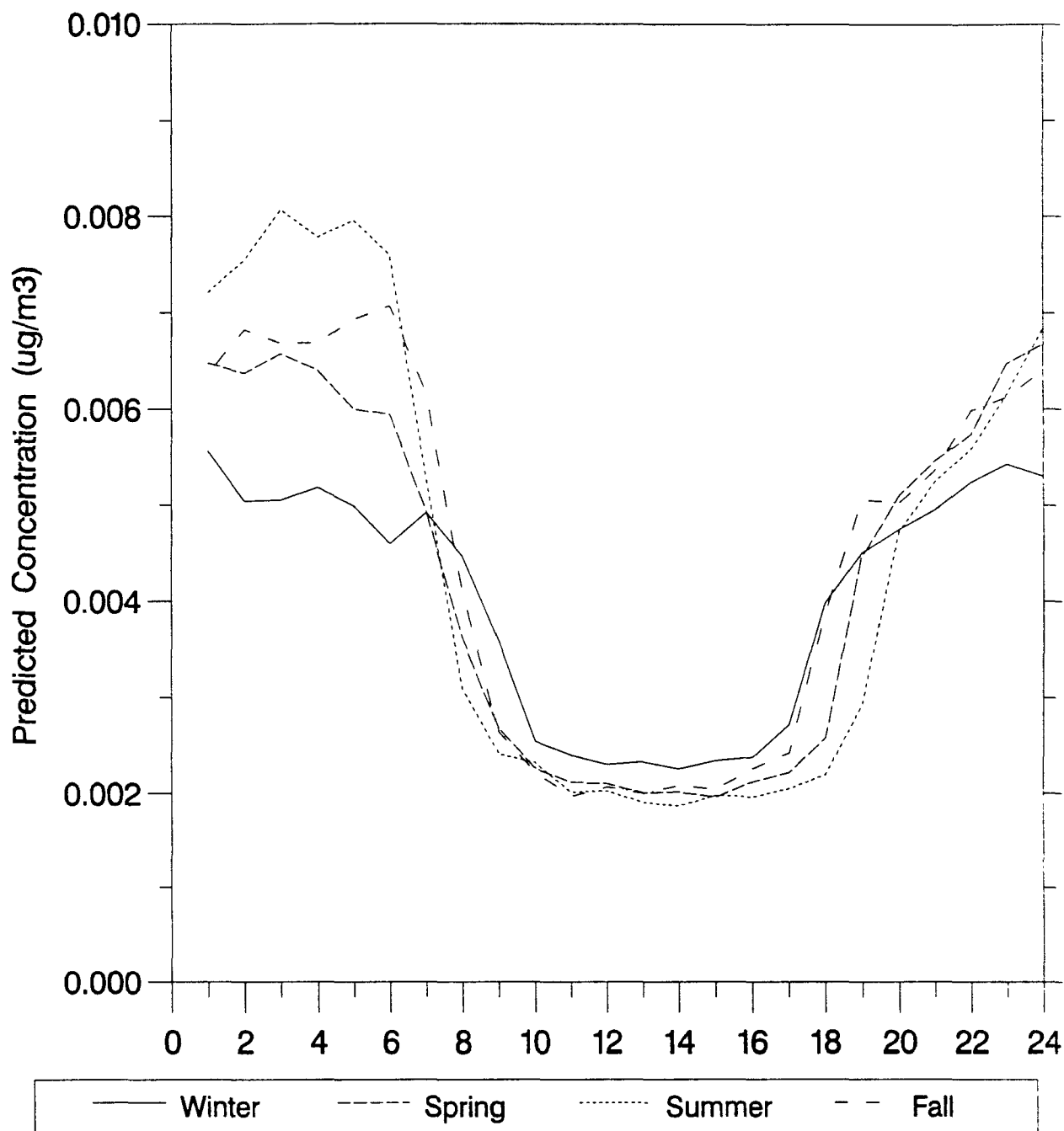
**Figure B.3-28. Urban Area Source Modeling, Houston, Texas  
Benzene Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



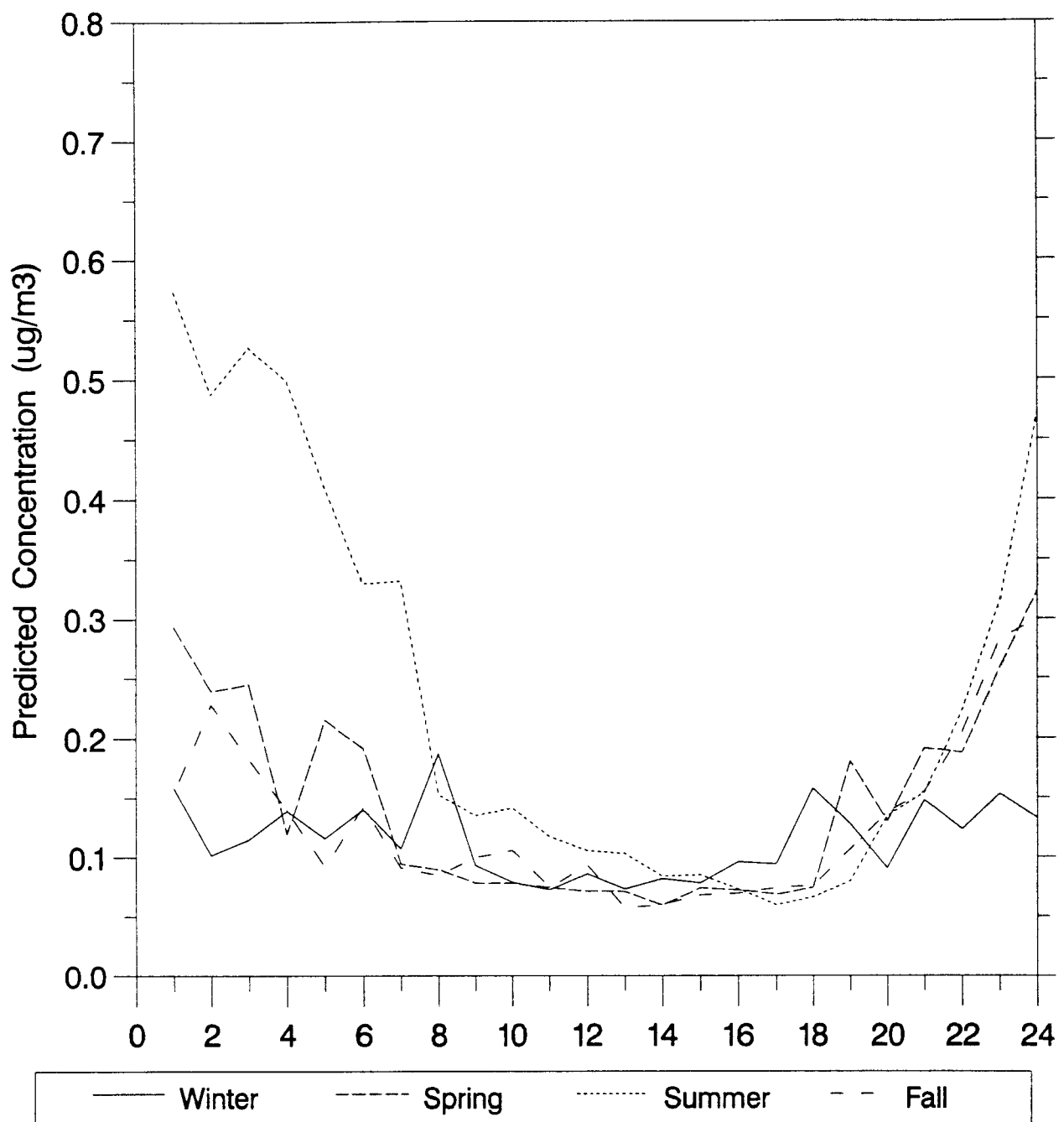
**Figure B.3-29. Urban Area Source Modeling, Houston, Texas  
1,3-Butadiene Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



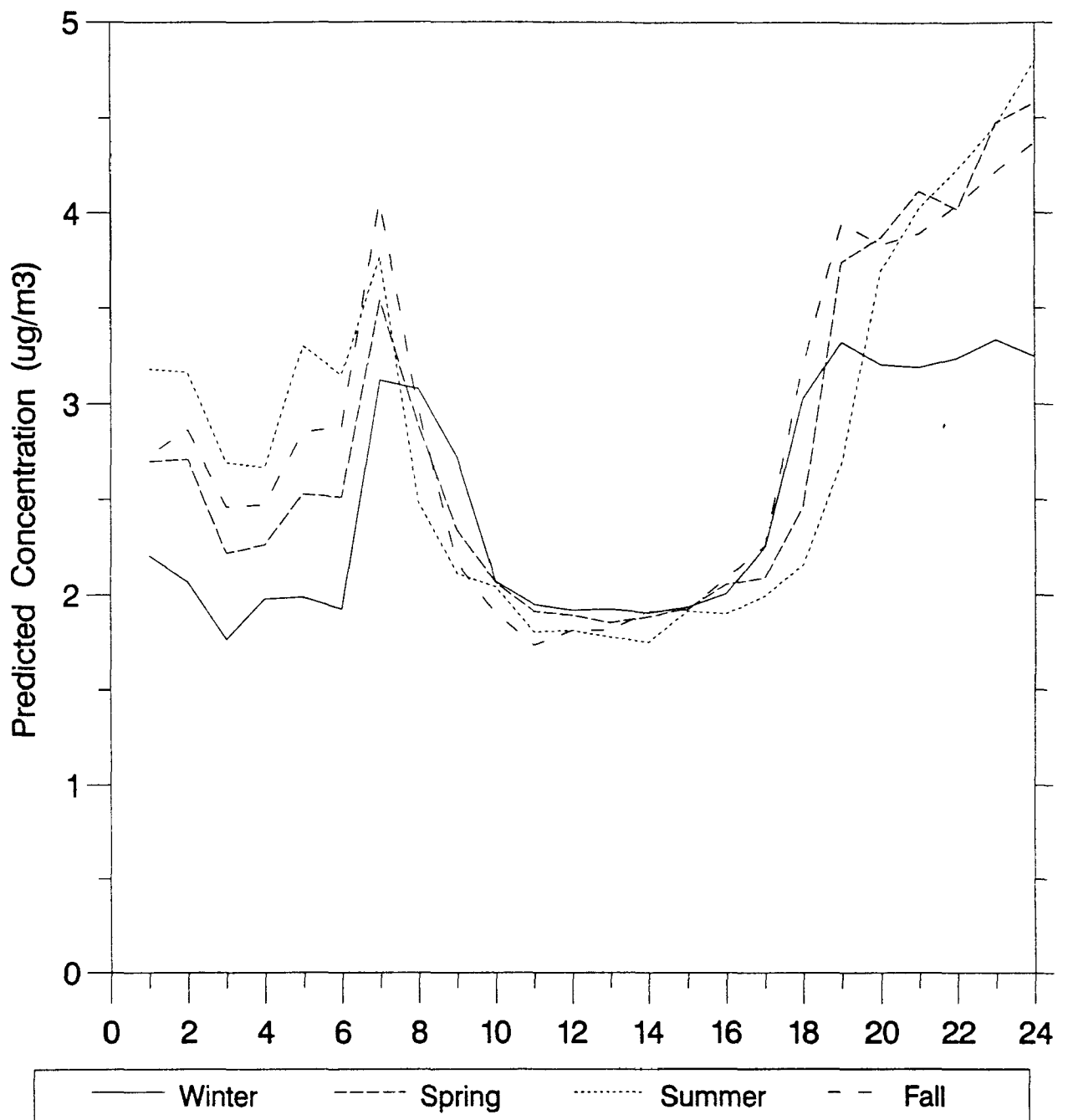
**Figure B.3-30. Urban Area Source Modeling, Houston, Texas**  
**Primary Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**



**Figure B.3-31. Urban Area Source Modeling, Houston, Texas  
POM Maximum Seasonal Average Concentrations by Hour of Day  
All Sources, 1987-1991 Meteorological Data**



**Figure B.3-32. Urban Area Source Modeling, Houston, Texas**  
**Hexavalent Chromium Maximum Seasonal Average Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**



**Figure B.3-33. Urban Area Source Modeling, Houston, Texas**  
**Total Formaldehyde Maximum Seasonal Average Concentrations by Hour of Day**  
**All Sources, 1987-1991 Meteorological Data**

## B.4 Preliminary Analysis of Air Quality Data and Modeled Estimates

As a preliminary, exploratory model evaluation analysis, available air quality data in the Houston area were obtained and compared with model estimated concentrations to determine how the model was performing. Ideally, the monitored and modeling data would be available for the same time period; however, in this exercise, the modeling period (1987-1991) did not coincide with the period (1993 -1994) for which monitored data exists or the emissions data period (1993 for point and area sources and 1990 for mobile sources). Although this disparity precludes day by day comparisons, annual averages of the monitored and modeling values were compared as a gross check on the model adequacy.

A review of available ambient air quality data in the Houston area indicated that 1,3-butadiene and benzene were monitored by the State of Texas at four locations in Harris County, Texas in 1993 and 1994. The location of the four monitored sites is shown in Figure B.3-1. Data were reported as 24-hour daily average values. For 1,3-butadiene, observed values were below the minimum detectable limit at many sites for many consecutive days. Thus, 1,3-butadiene data were excluded from further analysis.

Table B.4-1 provides various statistics for predicted and observed concentrations. Model receptors were located at the same coordinates as each of the four monitors. The number of modeled values represents the number of daily values in a 5-year meteorological data period, 1987 - 1991 (89 days X 5 years plus one leap day). The number of monitored observations ranges from 97 to 114, while modeled values are 1826. The mean observed benzene values range from 3.7 to 6.5 ug/m<sup>3</sup> and for predicted values from 0.78 to 1.58 ug/m<sup>3</sup>. Thus, on average, the modeled values under predict observed values at all four locations.

Figure B.4-1 provides graphical depictions of the observed to predicted values at the four monitoring sites for benzene. The figure shows that there is more variance in observed data than that in the modeled values. To investigate the degree of correlation between the monitored values, scatter plots comparing observed values among all benzene monitored sites are shown in Figure B.4-2 and show a slight positive association between the daily values between monitors 26 and 64. However, although the monitored values are in relatively close proximity to each, the degree of correlation among the monitored data appears low. This indicates that the monitors are influenced by proximity to nearby sources and micro meteorological conditions.

To investigate the degree of correlation between the modeled values, scatter plots comparing modeled concentrations at the same benzene monitoring sites are shown in Figure B.4-3. Similar to the monitoring data comparisons, these figures show that the degree of correlation among the modeled data at these locations is as low as that for the monitored data. Such a low correlation suggests that point sources dominate the variability in concentrations. Thus, the concentrations at these sites are influenced by proximity to nearby sources and micrometeorological conditions.

To further investigate this assumption, model prediction statistics by source type were calculated and are shown in Table B.4-3. These data also show that the predominant impact on

concentrations at these sites is from point sources. These data are also consistent with Table B.3-2 that show that the predominant impact for benzene is from point sources in Houston.

In conclusion, results show that modeled concentrations are lower than observed values for benzene and that proximity is important when dealing with point sources. Again, this is a preliminary, exploratory, analysis not to be confused with a model evaluation exercise.

**Table B.4-1 Statistics for Predicted and Observed Concentrations\***

Benzene (ug/m <sup>3</sup> )								
Monitor Site		N	Max.	Min.	10 <sup>th</sup> Percentile	Mean	Median	90 <sup>th</sup> Percentile
26	Predicted	1826	3.2	0.05	0.28	0.78	0.69	1.40
	Observed	98	10.11	0.02	1.66	3.70	3.58	6.37
64	Predicted	1826	5.90	0.08	0.20	0.99	0.68	2.23
	Observed	89	21.25	0.02	1.17	3.95	3.15	7.64
803	Predicted	1826	3.14	0.09	0.30	0.84	0.75	1.49
	Observed	114	114.27	0.02	1.59	6.51	3.79	11.11
1035	Predicted	1826	7.60	0.17	0.49	1.58	1.35	2.94
	Observed	97	37.07	0.02	1.36	5.84	4.55	8.12

\*Note: Model predictions are for all available days in 1987-1991 while monitored values are available only for selected days during 1993-1994.



**Table B.4-2 Model Prediction Statistics by Source Type**

<b>Benzene (ug/m<sup>3</sup>)</b>							
<b>Source Type</b>	<b>Monitor Site</b>	<b>Max.</b>	<b>Min.</b>	<b>10<sup>th</sup> Percentile</b>	<b>Mean</b>	<b>Median</b>	<b>90<sup>th</sup> Percentile</b>
Area	26	0.175	0.002	0.007	0.035	0.032	0.067
	64	0.285	0.008	0.019	0.060	0.055	0.108
	803	0.255	0.013	0.032	0.079	0.075	0.131
	1035	0.609	0.086	0.166	0.265	0.267	0.361
Mobile	26	1.227	0.022	0.056	0.127	0.108	0.220
	64	1.322	0.067	0.137	0.269	0.245	0.435
	803	1.241	0.041	0.092	0.182	0.167	0.287
	1035	1.353	0.057	0.131	0.232	0.216	0.339
Point	26	3.139	0.000	0.148	0.621	0.511	1.237
	64	5.285	0.000	0.000	0.659	0.281	1.869
	803	2.412	0.000	0.055	0.580	0.510	1.182
	1035	6.927	0.000	0.030	1.088	0.891	2.393
All	26	3.218	0.054	0.279	0.784	0.689	1.396
	64	5.898	0.076	0.197	0.988	0.684	2.229
	803	3.143	0.094	0.301	0.841	0.746	1.486
	1035	7.597	0.169	0.489	1.585	1.347	2.938

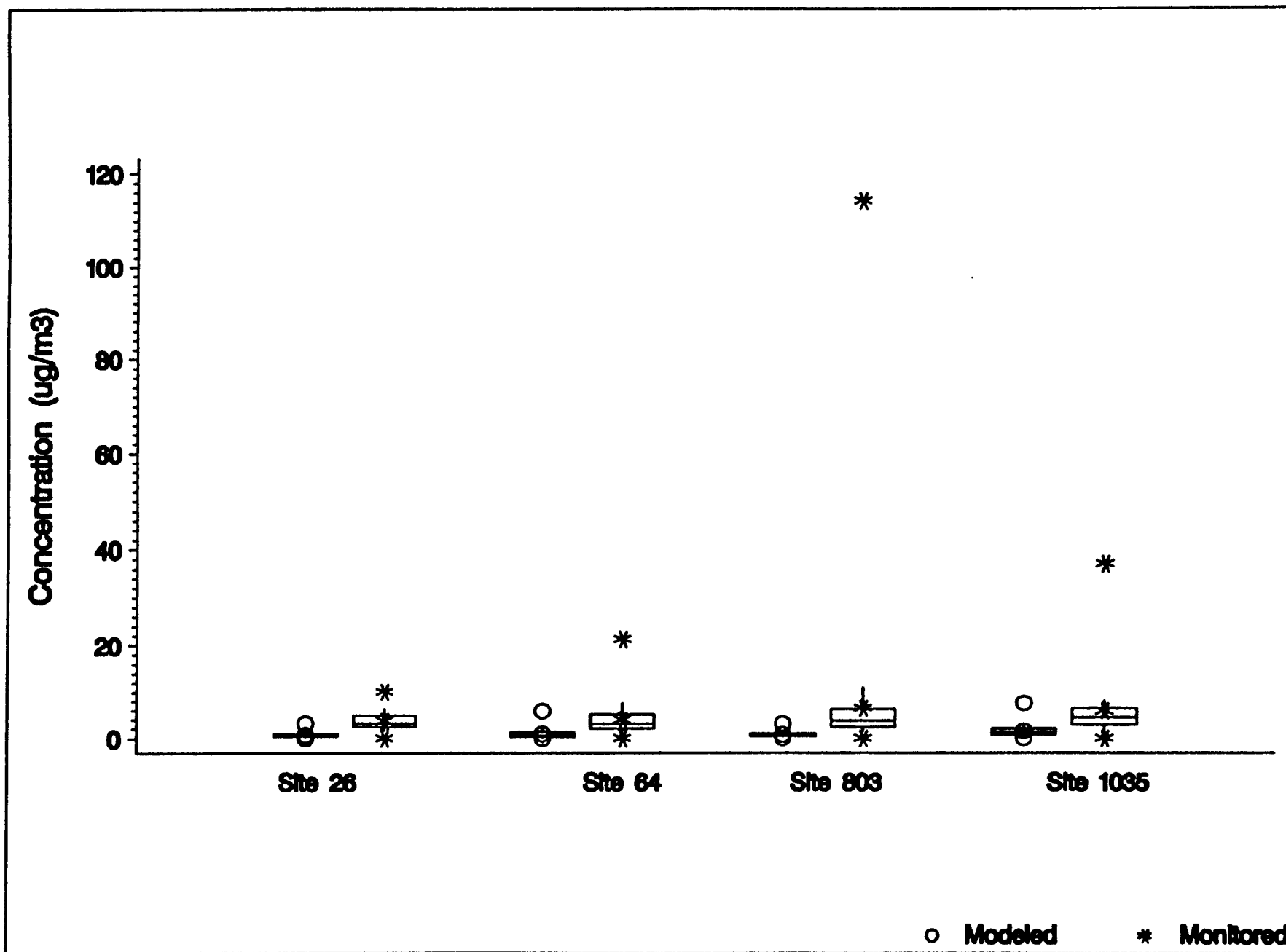


Figure B.4-1. Box Plot of Houston Benzene 1987-1991 Modeled vs. 1993-1994 Monitored

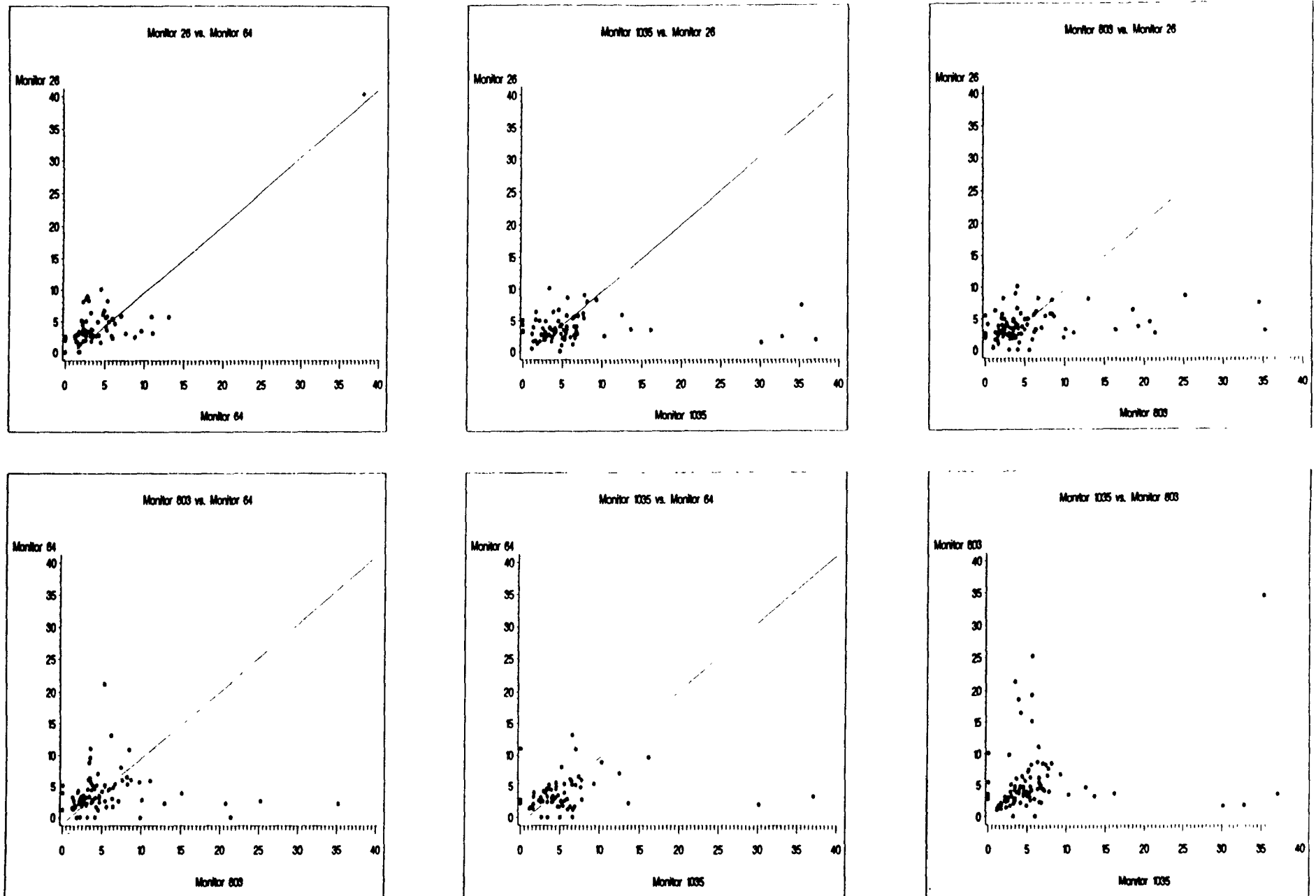


Figure B.4-2. Scatter Plots of Monitored Benzene Values ( $\mu\text{g}/\text{m}^3$ )

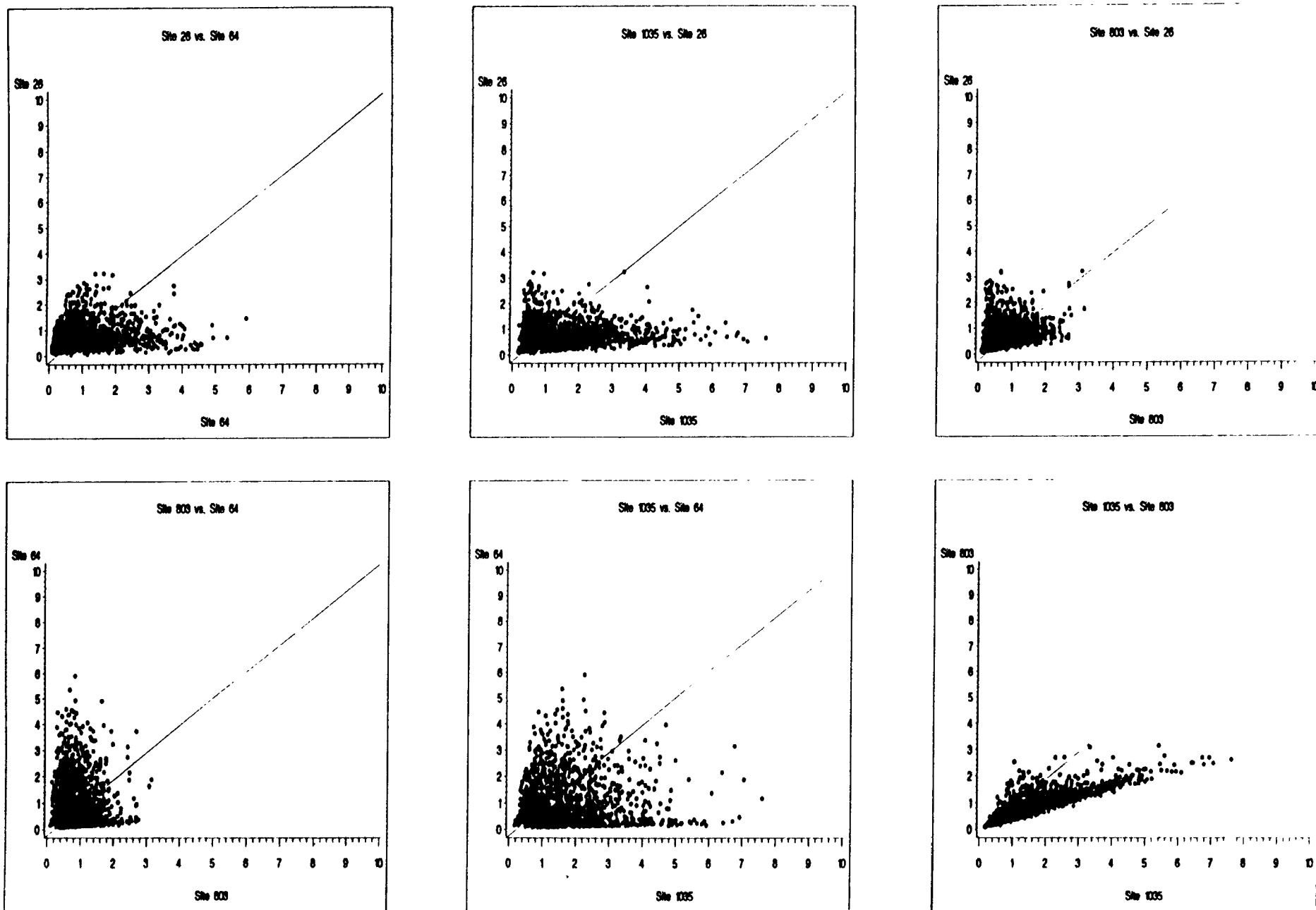


Figure B.4-3. Scatter Plots of Modeled Benzene Values ( $\mu\text{g}/\text{m}^3$ )

## B.5 SUMMARY AND CONCLUSIONS

This appendix presents the results of a study conducted in support of an air quality impact analysis for five toxic air pollutants emitted from major, area and mobile sources located in Houston, Texas. In this study, the ambient concentrations attributable to these sources were estimated through the application of the ISCST3 dispersion model.

Both the annual average concentrations, as well as the seasonal average concentrations by hour of day, were estimated. The results of the modeling study show that a majority of the total concentrations for Houston can be attributed to the major point sources. The mobile sources were found to be the next largest contributors. The area sources were found to contribute least to the total concentrations.

A study of the hour-by-day variations of seasonal average concentrations showed that the concentrations are higher during the morning and evening hours, and that the concentrations are generally higher during winter and fall seasons as compared to spring and summer seasons.

A preliminary comparison between available air quality data in Houston area with model estimated concentrations showed that on average the model predictions are lower than monitored values. Due to the differences in the meteorological, emissions and monitored value periods, additional analysis is needed to verify this conclusion.

In order to appropriately explain the observed patterns in concentrations, it is recommended that further detailed studies be conducted. The detailed studies should focus on an analysis of the temporal variations in emissions for various types of sources and their contributions to the predicted total concentrations. The studies should also take into consideration the effects of the meteorological conditions. It is also recommended that these studies be conducted using at least the same five years of meteorological data as used in this report. Long term trends could be identified using additional years of data.

## B.6 REFERENCES

- Radian Corp., 1995a. Development of the Houston Area Source Toxics Emissions (HASTE) Inventory. Prepared for Texas Natural Resources Conservation Commission.
- Radian Corp., 1995b. Air Quality Dispersion Modeling for the Houston Area Source Toxic Emissions (HASTE) Project. Prepared for Texas Natural Resources Conservation Commission.

## **APPENDIX C**

**PROPOSED METHODS FOR SELECTING RECEPTOR SAMPLES**

**FOR THE APPLICATION OF THE ISCST3 DISPERSION MODEL**

**TO URBAN AREAS**

## APPENDIX C - TABLE OF CONTENTS

C. PROPOSED METHODS FOR SELECTING RECEPTOR SAMPLES FOR THE APPLICATION OF THE ISCST3 DISPERSION MODEL .....	C-1
C.1 INTRODUCTION .....	C-1
C.2 INITIAL EVALUATION OF SAMPLING METHODS A AND B .....	C-3
C.2.1 Sampling Objectives .....	C-3
C.2.2 Population of Interest .....	C-3
C.2.3 Finite Population Survey Sampling versus Continuous Spatial Sampling .....	C-4
C.2.4 The General Sampling Design .....	C-4
C.2.5 Defining the Strata .....	C-4
C.2.6 Sample Size Determination .....	C-10
C.2.7 Selection of Samples by Method A .....	C-16
C.2.8 Selection of Samples by Method B .....	C-16
C.2.9 Decision to Perform Sensitivity Analyses .....	C-20
C.3. SUMMARY OF SAMPLING METHODS A AND B .....	C-21
C.3.1 Step-by-Step Procedure .....	C-21
C.3.2 Limitations of Methods A and B .....	C-22
C.4 SENSITIVITY ANALYSES OF SAMPLING METHODS A AND B .....	C-23
C.4.1 Evaluation of Sample Sizes Calculated by Methods A and B .....	C-24
C.4.2 Effects of Reducing Nominal Sample Sizes .....	C-28
C.5 REFINEMENTS TO SAMPLING METHODOLOGY BASED ON FINDINGS OF SENSITIVITY ANALYSES .....	C-36
C.5.1 Proposed Refinements .....	C-36
C.5.2 Revised Methodology Applied to Annual Average Data .....	C-37
C.5.3 Application of Method C to Data from Specific Hourly/Seasonal Combinations .....	C-39
C.6 APPLICATION OF SAMPLING METHOD C TO HOUSTON BENZENE CONCENTRATIONS .....	C-48
C.6.1 Houston Study Area .....	C-48
C.6.2 Defining the Strata for Method C .....	C-49
C.6.3 True Sampling Distributions .....	C-53
C.6.4 Method C Applied to Annual Average Data with Emissions-Defined Strata .....	C-53
C.6.5 Method C Applied to the Hourly/Seasonal Data with Emissions-Defined Strata .....	C-57
C.6.6 Results Using HAPEM-Like Strata .....	C-63
C.6.7 Forced Receptors - Certainty Units .....	C-63
C.6.8 Summary Discussion .....	C-71
C.7 STEP-BY-STEP GUIDELINES FOR USING SAMPLING METHOD C .....	C-73
C.8 REFERENCES .....	C-77



## APPENDIX C - LIST OF TABLES

Table C.2-1 The Distribution of Emissions Values Across the Census Block Groups Within Each Stratum .....	C-11
Table C.2-2 Number of Grid Cells and Total Population for Each Stratum .....	C-13
Table C.2-3 Sample Sizes Selected for Each Stratum in the Central Region .....	C-15
Table C.4-1 Comparison of Sample Size Calculation Results Using Concentration Values versus Using the Emissions Values .....	C-25
Table C.4-2 Results of 200 Monte Carlo Samples for Sampling Methods A and B Using Sample Sizes as Calculated in Table 2-3 .....	C-27
Table C.4-3 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 30 Drawn from Each Stratum .....	C-29
Table C.4-4 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 25 Drawn from Each Stratum .....	C-30
Table C.4-5 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 20 Drawn from Each Stratum .....	C-31
Table C.4-6 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 15 Drawn from Each Stratum .....	C-32
Table C.4-7 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 10 Drawn from Each Stratum .....	C-33
Table C.4-8 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Nominal Size 5 Drawn from Each Stratum .....	C-34
Table C.5-1 Results of 200 Monte Carlo Samples for Sampling Method A When Applied to Annual Average Data .....	C-38
Table C.5-2 Descriptive Statistics for the Winter, 7 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-40
Table C.5-3 Descriptive Statistics for the Winter, 12 p.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-41
Table C.5-4 Descriptive Statistics for the Winter, 12 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-42

Table C.5-5 Descriptive Statistics for the Summer, 7 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-43
Table C.5-6 Descriptive Statistics for the Summer, 12 p.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-44
Table C.5-7 Descriptive Statistics for the Summer, 12 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C .....	C-45
Table C.6-1 Sampling Results for Annual Average Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-56
Table C.6-2 Sampling Results for Winter 7 am Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-58
Table C.6-3 Sampling Results for Winter 12 pm Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-59
Table C.6-4 Sampling Results for Summer 7 am Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-60
Table C.6-5 Sampling Results for Summer 12 pm Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-61
Table C.6-6 Sampling Results for Summer 12 am Houston Benzene Concentrations Using the Emissions-Defined Strata .....	C-62
Table C.6-7 Sampling Results for Annual Average Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-64
Table C.6-8 Sampling Results for Winter 7 am Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-65
Table C.6-9 Sampling Results for Winter 12 pm Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-66
Table C.6-10 Sampling Results for Summer 7 am Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-67
Table C.6-11 Sampling Results for Summer 12 pm Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-68
Table C.6-12 Sampling Results for Summer 12 am Houston Benzene Concentrations Using the HAPEM-Like Strata .....	C-69

Table C.6-13 The Sampling Distribution for Stratum 3 of the Emissions-Defined Strata and Stratum 9 of the Hapem-like Strata after the Outlying Block Group at Utm Coordinates 283.735 East, 3289.09 North Is Forced into the Sample .....	C-70
Table C.7-1 Example Sas Code <sup>a</sup> for Generating a Sample of Receptor Points Using Method C .....	C-74
Table C.7-2 Sample Inputs for Step 5 Calculations (Exposure District No. 1 of Hypothetical Study Area) .....	C-75

## APPENDIX C - LIST OF FIGURES

Figure C.2-1 Contour map of the Phoenix study area showing cell emission totals and 4 km by 4 km emissions inventory grid cells .....	C-6
Figure C.2-2 Contour map of the Phoenix study area showing cell emission totals and the locations of all block group centroids .....	C-7
Figure C.2-3 Strata constructed for application of sampling procedure to central region of Phoenix .....	C-8
Figure C.2-4 Strata constructed for application of sampling procedure to central region of Phoenix with superimposed centroids of all block groups .....	C-9
Figure C.2-5 Location of block groups selected by Method A .....	C-17
Figure C.2-6 Location of block groups selected by Method B .....	C-19
Figure C.6-1 Contour map of the Phoenix region indicating cell emission totals .....	C-50
Figure C.6-2 Emissions-defined strata based on Phoenix contour map presented in Figure 6-1 .....	C-51
Figure C.6-3 HAPEM-like strata defined for Phoenix region .....	C-52
Figure C.6-4 Distribution of block-group centroids over emissions-defined strata .....	C-54
Figure C.6-5 Distribution of block-group centroids over HAPEM-like strata .....	C-55

## **C. PROPOSED METHODS FOR SELECTING RECEPTOR SAMPLES FOR THE APPLICATION OF THE ISCST3 DISPERSION MODEL**

### **C.1 INTRODUCTION**

The U.S. Environmental Protection Agency (EPA) is preparing tools to estimate the exposures to five toxic pollutants [benzene, butadiene, formaldehyde, hexavalent chromium, and polycyclic organic matter (POM)] for two cities (Houston, and Phoenix). It is likely that the Hazardous Air Pollutant Exposure Model (HAPEM) (Johnson, et.al., 1996) will be used to develop these estimates. This model requires input data showing the spatial pattern of outdoor concentrations for a particular pollutant across the specified study area. To achieve this objective, researchers apply a dispersion model to an emissions data base for each study area to produce estimates of pollutant concentration at selected "receptor points." These concentrations are then averaged over a set of user-defined "exposure districts." These districts are used in the HAPEM exposure assessments as potential locations for homes and work places. Exposures within each district are estimated as a function of (1) the average outdoor pollutant concentration determined for the district and (2) the microenvironment occupied by the exposed individual.

Because HAPEM uses census data in computing exposures, it is useful to define HAPEM districts as aggregates of census units such as blocks, block groups or census tracts. To better relate these districts to dispersion model estimates of outdoor pollution levels, analysts typically use the centroids of selected census units as the receptor points for the dispersion model runs. As the computational time required for a dispersion model run increases with the number of receptor points used, it is advantageous to minimize the number of census units used as receptor locations. This goal can be accomplished by (1) defining a population of possible receptor points that provides a good characterization of the residential patterns in the study area and then (2) using statistical techniques to select a representative sample from this population that meets the minimum requirements of the analyst.

Researchers developed two candidate schemes (Method A and Method B) for selecting this sample which can be generalized to other pollutants in other study regions. Section C.2 of this report describes these methods and provides examples of the application of each method to one pollutant (1,3-butadiene) in one study area (Phoenix). Section C.3 provides a step-by-step procedure for implementing each method and summarizes the principal limitations of Methods A and B.

Researchers performed a series of sensitivity analyses using the Phoenix butadiene data to test the statistical assumptions underlying Methods A and B. Section C.4 summarizes these analyses, identifies Method B as the superior method, and concludes that both methods tend to select samples which are larger than necessary to achieve a specified set of sampling goals. Section C.5 presents Method C, a revised version of Method B, which is capable of achieving the sampling goals using a significantly smaller sample. To illustrate the generalizability of Method C, researchers applied it to annual average data and to data representing various combinations of time of day and season. Section C.5 presents results of these analyses and discusses how the method may be used in future exposure analyses.

Analysts noted that the Phoenix results were representative of a region dominated by area sources and may not be indicative of results obtained for a region dominated by point sources. Consequently, the analysis was repeated using benzene data for Houston, an area with a greater density of point sources. Section C.6 presents the results of applying Method C to Houston benzene data based on two alternative schemes for defining sampling strata. Step-by-step guidelines for using Method C are presented in Section C.7.

## C.2 INITIAL EVALUATION OF SAMPLING METHODS A AND B

This section presents the objectives to be met by the sampling approach and provides two alternative methods for accomplishing these objectives. The application of each method to butadiene in Phoenix is described, and the results of these applications are compared.

### C.2.1 Sampling Objectives

The sampling methods were developed to accomplish the following objectives:

- Develop a set of sampling strata (geographic zones) which completely covers the designated study area. These strata may later be considered as potential exposure districts for the HAPEM model.
- Estimate the minimum sample size required to estimate the mean outdoor concentration (expressed as an annual average concentration) for each stratum to within 10 percent of the true mean concentration with 95 percent confidence.
- Choose a sample which is representative of the outdoor concentrations at population centers (as represented by the geographic centroids of census block groups).

In accomplishing these objectives, researchers assumed that the sample representing each combination of pollutant and study area would be selected independently.

### C.2.2 Population of Interest

Researchers defined the study population as consisting of all census block groups (BGs) in the designated study area according to the 1990 U.S. census. In the example considered here, the study area consisted of all BGs in a rectangular region surrounding Phoenix, Arizona. The universal transverse mercator (UTM) coordinates of the corners of this rectangle are listed below:

<u>Corner</u>	<u>UTM Zone</u>	<u>UTM (east)</u>	<u>UTM (north)</u>
Northeast	12	485 km	3763 km
Southeast	12	485 km	3663 km
Southwest	12	349 km	3663 km
Northwest	12	349 km	3763 km

The study area contained approximately 13,600 square kilometers of land area and 1894 BGs. The 1894 BGs were assumed to constitute the sampling frame for the study area (i.e., the total population of available sampling units). The target population for this study is the current population of Phoenix.

### C.2.3 Finite Population Survey Sampling versus Continuous Spatial Sampling

Although outdoor pollutant concentration is a continuous variable which is distributed spatially over a region, researchers assumed that the HAPEM exposure assessment would be concerned only with those areas in the region actually inhabited by people. Consequently, researchers decided to choose a survey sample from a finite population of points (the geographic centroids of the BGs) rather than using continuous-variable spatial sampling methods. In essence, the goal of the selection process was a sample that represented the outdoor pollutant concentrations "observed" by the population, rather than one representing the overall spatial pattern of outdoor pollutant concentrations. The population was assumed to be clustered at points located at the centroids of BGs.

In developing the sampling methods described below, researchers treated the outdoor concentration determined by the dispersion model for a particular receptor as an attribute of the population subgroup residing within the associated census unit.

### C.2.4 The General Sampling Design

The proposed sampling design is a stratified random sample from a finite population. According to this general approach, the study region is first divided into a number of smaller geographic zones or strata. This procedure helps to account for the spatial variation of the data. Next, a sample is taken from within each of the strata. Two methods can be used to obtain this sample. The first method draws a simple random sample without replacement from within each stratum. The second method stratifies each stratum further into 4 kilometer by 4 kilometer grid cells, and then takes a random sample from within each grid cell. In each method, analysts used computer-generated random numbers to draw random samples from the sampling frame. The following subsections describe how researchers applied each method to the sample task (butadiene in Phoenix) and provide a comparison of the results of each method.

### C.2.5 Defining the Strata

EPA compiled an emission inventory for each combination of pollutant (benzene, butadiene, formaldehyde, hexavalent chromium, or POM) and study area (Houston or Phoenix). To assist future HAPEM exposure assessments, analysts will apply a version of the ISCST3 dispersion model to each emission inventory using local meteorological data for the associated study area. Each run of the dispersion model will produce estimates of outdoor pollutant concentrations at a set of receptor points defined by the user. As the spatial pattern in concentration estimates is likely to be similar to the spatial pattern in the emissions data, it is reasonable to assume that the area-source emissions data in each inventory will provide a reasonable basis for defining the sampling strata. [Note that this assumption may not be valid for point-source emissions data. In such cases, the approach discussed in Section C.5 (Sampling Method C) is likely to produce superior results.]

In compiling the emissions database for Phoenix, researchers defined the emissions inventory area by a rectangular grid containing 850 cells (34 cells by 25 cells). Each cell



measured 4 km by 4 km. The grid included all of the Phoenix metropolitan area and some adjoining rural areas. Using the emissions inventory data, estimates were developed for emissions from various area-source types for each cell in the grid. (Point-source emission estimates were not included in the database.) The data from the various source types within each cell were summed to provide a total emissions value for the cell. The grid cell totals for butadiene emissions ranged from 0 to 5097.2 kg/yr with a mean of 315.5 kg/yr and a standard deviation of 854.9 kg/yr. Note that these data may be revised, as analysts are currently refining the emission estimates for Phoenix.

Figure C.2-1 presents an emissions contour map of the Phoenix region indicating the cell emission totals together with the 4 km by 4 km grid cells. The same map with the locations of the block group centroids is provided in Figure C.2-2. From the latter map, it can be seen that a very high proportion of the population is located in areas of relatively high butadiene emissions.

Because a large majority of the sampling units (i.e., the block group centroids) are located in the center of the region, the study region was divided into two large subregions -- the "central" region and the "outer" region. Researchers noted that points in the outer region were too widely spaced to attribute any type of "stratum mean" concentration to them. In a sense, each of these points (or in some cases, small clusters of points) formed its own stratum. Consequently, researchers recommended that every point in the outer region be included in the sample. These points can be treated as separate exposure districts or aggregated into small clusters at a later stage in the exposure assessment process.

In the central region, each 4 km x 4 km grid cell was assigned a classification corresponding to the total emissions for that cell. The classification categories were high (greater than 4000 kg/yr), moderate (between 1000 and 4000 kg/yr), or low (less than 1000 kg/yr) emissions. The "breakpoint" values of 1000 and 4000 were chosen somewhat subjectively, but were considered reasonable as they were approximately equal to the 25th and 75th percentiles, respectively, of the distribution of emissions over all sampling units.

After classification, the grid cells of the central region were placed into sampling strata according to two objectives:

- (1) each stratum would contain cells with similar emission levels and
- (2) each stratum would be a contiguous collection of cells with no stratum being completely surrounded by another stratum.

The 12 resulting strata are shown in Figure C.2-3. Here, Strata 1 and 2 are "high" strata, Strata 3 through 7 are "moderate" strata, and Strata 8 through 12 are "low" strata. Figure C.2-4 shows the distribution of sampling units across the 12 central strata.

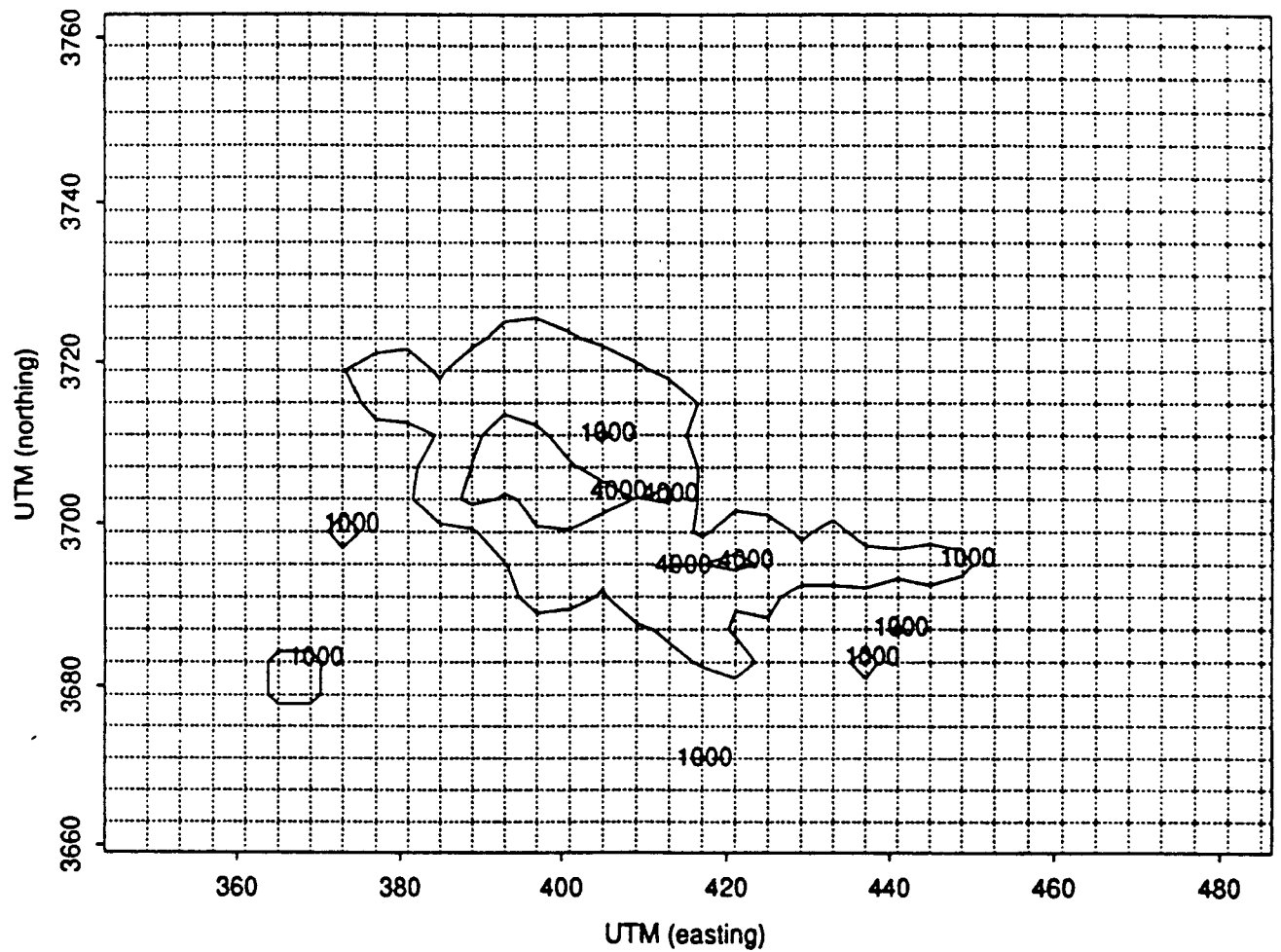


Figure C.2-1 Emissions contour map of the Phoenix study area showing cell emission contours of butadiene and 4 km by 4 km emissions inventory grid cells.

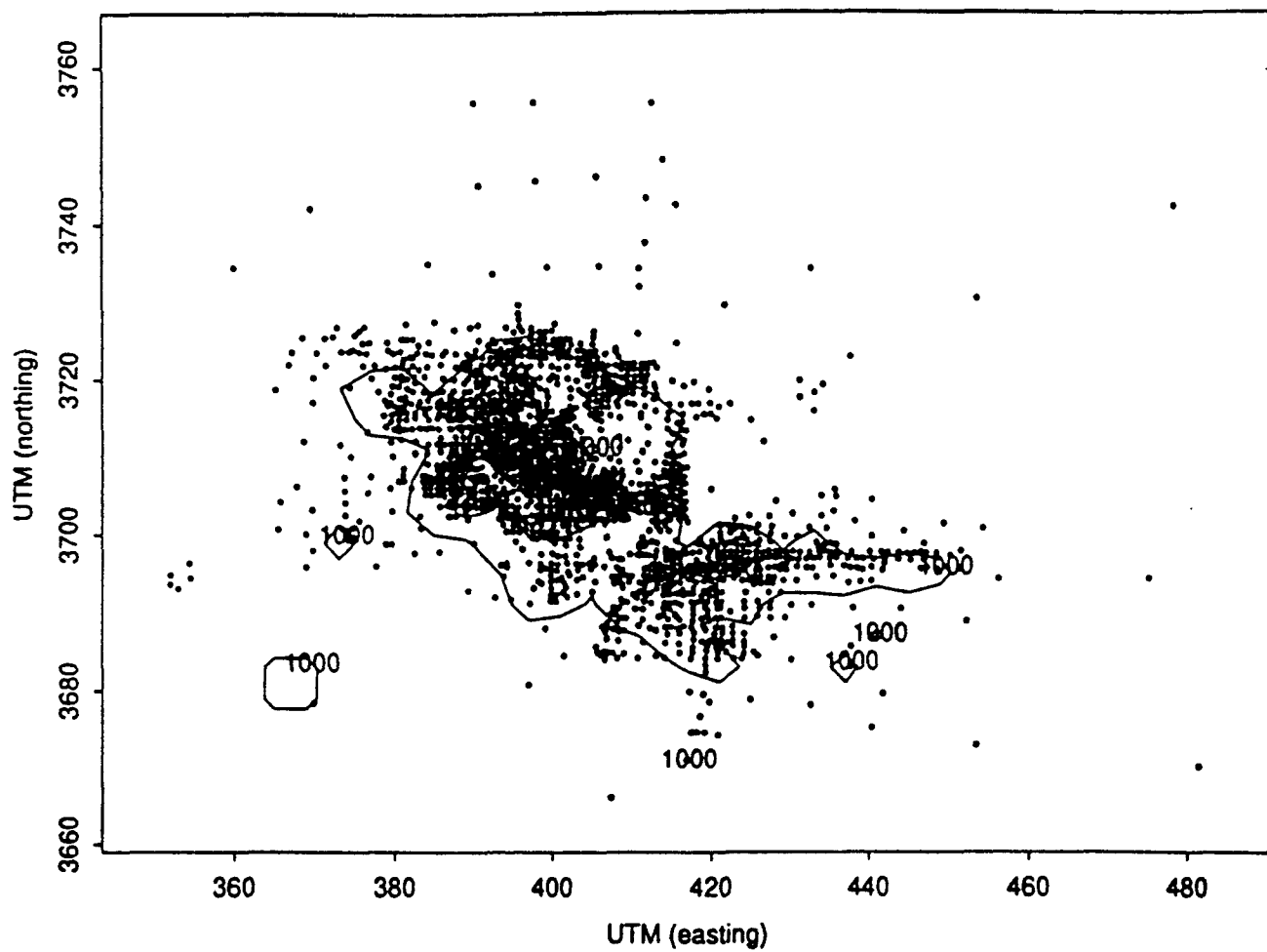


Figure C.2-2 Emissions contour map of the Phoenix study area showing cell emission contours of butadiene and the locations of all block group centroids.

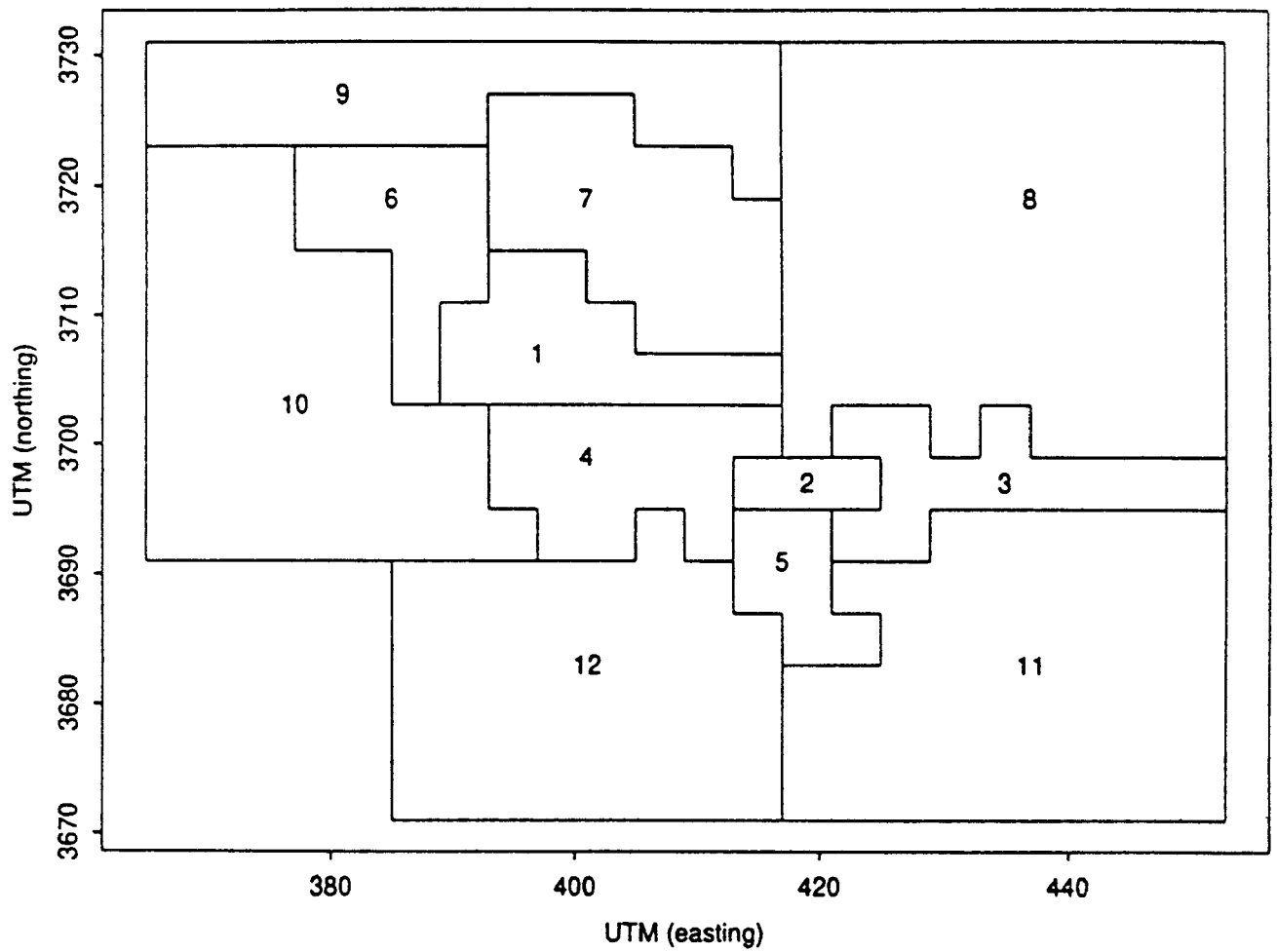


Figure C.2-3 Strata constructed for application of sampling procedure to central region of Phoenix.

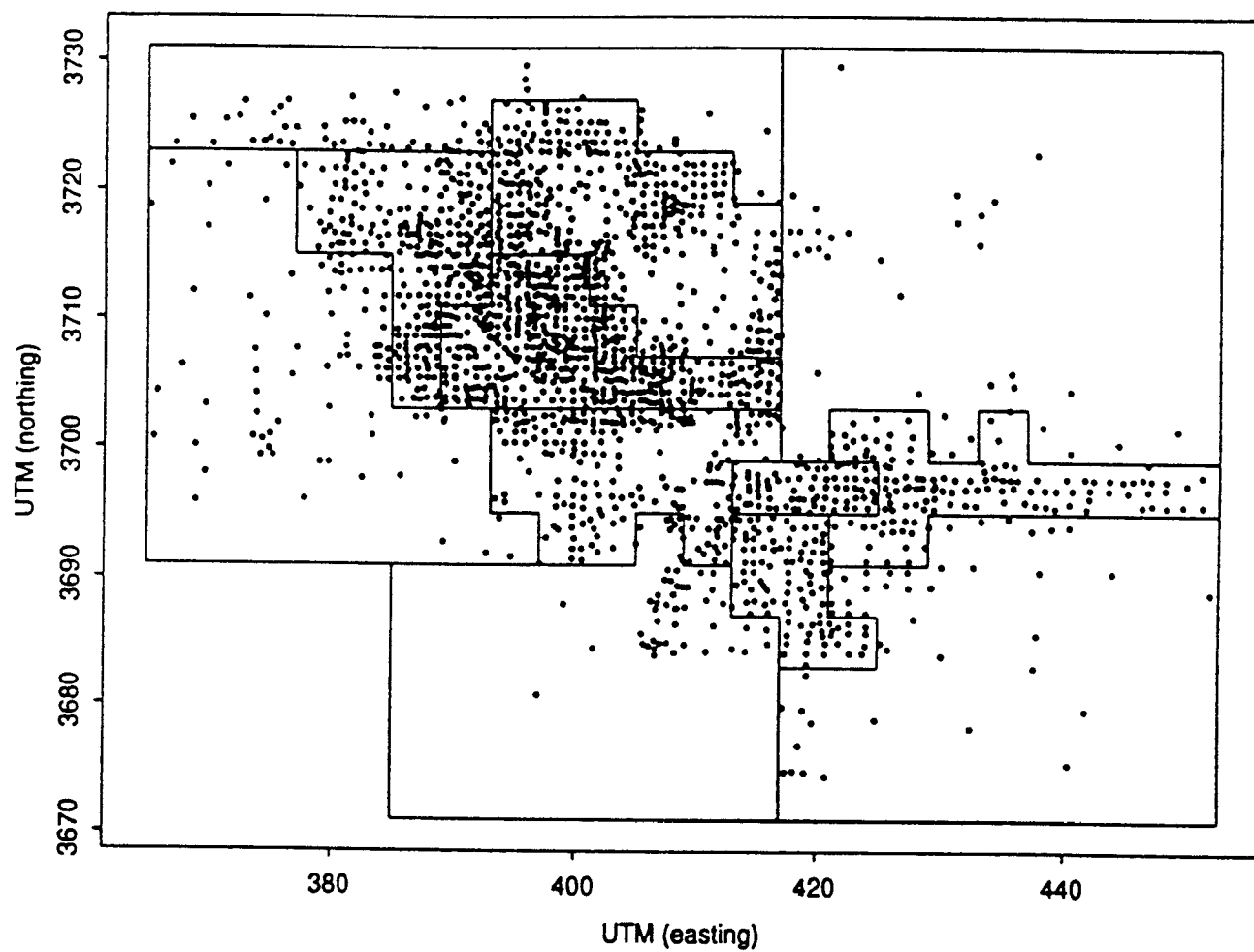


Figure C.2-4 Strata constructed for application of sampling procedure to central region of Phoenix with superimposed centroids of all block groups.

Table C.2-1 shows the distribution of the emissions values over the sampling units (i.e., the BGs) within each stratum. As expected from the definition of the classification categories, the means for Strata 1 and 2 are between 4000 and 5000 kg/yr, the means for Strata 3 through 7 are all near 2000 kg/yr, and the means for Strata 8 through 12 are approximately 500 kg/yr. The standard deviations for the high and low strata all fall between 165 and 430 kg/yr, while the standard deviations for the moderate strata range from 715 to approximately 890 kg/yr.

Table C.2-2 lists the number of grid cells and the total population within each stratum. The results in this table indicate that the 305 grid cells of the central region contain almost 98 percent of the study area population. The high and moderate strata (Strata 1 through 7) are also the most densely populated areas. Almost 25 percent of the study area population are located in the areas of highest butadiene emissions (Strata 1 and 2), and approximately 80 percent of the population are located in areas of at least moderate ( $> 1000$  kg/yr) butadiene emissions.

#### C.2.6 Sample Size Determination

In selecting a sample from a defined population, analysts typically begin the process by estimating the sample size needed to estimate a particular population parameter with some specified degree of precision. To meet the second objective (see Section C.2.1), researchers determined that the mean of each sampling stratum in the central region should be estimated to within 10 percent of the true mean with 95 percent confidence.

In calculating the sample size required to achieve a specified sampling objective, it is often necessary to obtain a reasonable approximation of the expected precision of the estimate being produced. This approximation could be an educated guess, an estimate obtained from prior studies, or an estimate obtained using a surrogate variable which has similar distributional characteristics to the variable under study. Because the pollutant concentrations to be estimated by the dispersion model are a function of the emissions data provided to the dispersion model, the concentration estimates are likely to be roughly proportional to the emissions data. Analysts incorporated this assumption of proportionality into the sample size calculations discussed below.

The emission values in each stratum have a mean ( $\mu_Q$ ) and a standard deviation ( $\sigma_Q$ ). If one can assume that the outdoor pollutant concentration,  $C$ , in each stratum is proportional to the emissions for that stratum,  $Q$ , then  $C = kQ$  for some constant  $k$ . (In reality,  $k$  is not a constant but is a function of meteorological parameters as well as the total emissions for the entire study region). If  $C = kQ$ , then the mean concentration ( $\mu_C$ ) is equal to  $k\mu_Q$  and the standard deviation of the concentrations ( $\sigma_C$ ) is equal to  $k\sigma_Q$ . The coefficient of variation,  $\gamma$ , is defined as the ratio of the standard deviation to the mean. Thus, under the assumption of proportionality, the coefficients of variation for the emissions and associated outdoor concentrations in each stratum are equal,

**Table C.2-1 The Distribution of Emissions Values Across the Census Block Groups Within Each Stratum**

Sampling area	Number of block groups in area	Distribution of emission values, kg/yr								
		Mean	Standard deviation	Skewness <sup>a</sup>	Kurtosis <sup>a</sup>	Minimum	25th Percentile	Median	75th Percentile	Maximum
All BGs	1894	2666.7	1512.9	-0.02	-1.15	0.82	1620.8	2552.8	3991.1	5097.2
BGs in Outer Region	33	52.2	77.8	1.72	2.92	0.82	1.60	4.75	89.5	319.6
BGs in Central Region	1861	2713.1	1485.3	-0.01	-1.15	19.94	1644.1	2609.7	3991.1	5097.2
Stratum 1	457	4587.8	428.9	-0.37	-1.16	3774.6	4285.4	4675.6	5031.8	5097.2
Stratum 2	75	4233.0	300.3	0.64	-1.61	3991.1	3991.1	4040.3	4641.6	4641.6
Stratum 3	128	2206.8	773.6	0.69	-0.45	1163.1	1734.1	1967.0	2629.4	3650.0
Stratum 4	201	2706.3	887.7	-0.07	-0.89	868.8	2229.3	2609.7	3434.8	3924.5
Stratum 5	100	1952.4	861.6	1.16	-0.22	1141.1	1278.1	1644.1	1822.5	3551.5
Stratum 6	246	2221.6	715.7	0.49	0.08	711.0	1768.1	2168.4	2453.4	3620.9

**Table C.2-1 The Distribution of Emissions Values Across the Census Block Groups Within Each Stratum (continued)**

Sampling area	Number of block groups in area	Distribution of emission values, kg/yr								
		Mean	Standard deviation	Skewness <sup>a</sup>	Kurtosis <sup>a</sup>	Minimum	25th Percentile	Median	75th Percentile	Maximum
Stratum 7	365	2513.7	756.0	-0.22	-0.80	535.2	1929.8	2552.8	3134.7	3648.2
Stratum 8	39	387.8	309.6	0.33	-1.46	34.6	70.9	331.3	715.7	851.7
Stratum 9	65	416.6	165.5	-1.00	-0.03	19.9	300.5	508.3	537.1	621.6
Stratum 10	86	511.4	415.7	1.71	2.51	31.0	257.7	427.6	683.2	1633.3
Stratum 11	46	594.1	406.1	0.36	-0.35	34.5	204.2	633.0	779.2	1561.2
Stratum 12	53	672.9	230.0	-0.97	0.23	28.5	430.1	732.2	890.5	904.9

<sup>a</sup>Dimensionless.



**Table C.2-2 Number of Grid Cells and Total Population for Each Stratum**

Sampling area	Number of grid cells	Number of block groups	Number of populated grid cells	Total Population	Percentage of total population
Entire study area	850	1894	206	2,138,258	100.0
Outer region	545	33	29	44,021	2.1
Central region	305	1861	177	2,094,237	97.9
Stratum 1	13	457	13	411,318	19.2
Stratum 2	3	75	3	116,821	5.5
Stratum 3	12	128	12	205,531	9.6
Stratum 4	14	201	14	223,354	10.4
Stratum 5	6	100	6	138,373	6.5
Stratum 6	12	246	12	247,370	11.6
Stratum 7	21	365	21	380,968	17.8
Stratum 8	69	39	21	73,510	3.4
Stratum 9	24	65	16	71,240	3.3
Stratum 10	43	86	31	93,570	4.4
Stratum 11	48	46	19	71,770	3.4
Stratum 12	40	53	9	60,452	2.8

$$\gamma_c = \frac{\sigma_c}{\mu_c} = \frac{k\sigma_c}{k\mu_c} = \frac{\sigma_o}{\mu_o} = \gamma_o \quad (1)$$

Consequently, it is reasonable to use the emissions data as a surrogate for the concentration data in estimating the sample size required to estimate the mean concentration in each stratum.

The general formula for calculating the sample size,  $n$ , for a simple random sample of each stratum is

$$n = [(r/Z_{\alpha/2} \gamma_c)^2 + (1/N)]^{-1}, \quad (2)$$

where  $r$  is the relative precision desired,  $Z_{\alpha/2}$  is the 100(1 -  $\alpha/2$ ) percentile from the standard normal distribution (1.96 for  $\alpha = 0.05$ ),  $\gamma_c$  is the unknown coefficient of variation for the concentrations, and  $N$  is the total number of sampling units in the stratum.<sup>2</sup> The goal of the sample is a precision of +/- 10 percent with 95 percent confidence. Consequently, the value of  $r$  is 0.10, the value of  $\alpha$  is 0.05, and the corresponding value of  $Z_{\alpha/2}$  is 1.96. Practically,  $n$  should always be an integer, and, to assure the proper coverage probability, should always be rounded up. In addition, this sample size formula is based on the Central Limit Theorem, which generally holds for samples of at least 30, regardless of the underlying distribution. Therefore, if any calculated sample size was less than 30, a sample of size 30 was applied.

In the above sample size formula,  $\gamma_c$  will be estimated by  $\gamma_o$  under the assumption of proportionality. Note from the formula that if  $\gamma_o$  substantially underestimates  $\gamma_c$ , the calculated sample size will be less than required for the desired amount of precision. However, if  $\gamma_o$  overestimates  $\gamma_c$ , oversampling will result, which is generally not a concern.

The sample size formula does not consider potential spatial correlations between population density and emission rates. The goal of the sample selection is to estimate the mean concentration at a specified set of locations (the census block centroids) and not to estimate the spatial trend in concentrations; consequently, it is not necessary to account for spatial correlations in the calculation of sample size or variance.

Table C.2-3 contains the sample sizes for each stratum calculated using Equation 2. The total sample size required is 470. The calculated sample sizes for Strata 1, 2, and 12 were all less than 30. Therefore, to meet the requirements of the Central Limit Theorem, samples of size 30 were imposed on each these strata. All other sample size calculations resulted in values between 30 and 45, with the exception of Stratum 10 where a sample of 64 is required.

**Table C.2-3 Sample Sizes Selected for Each Stratum in the Central Region**

Sampling area	Total number of block groups, N	Coefficient of variation, $\gamma_Q$	Required sample size, n	Sizes of samples taken	
				Method A	Method B
Stratum 1	457	0.0935	30 <sup>a</sup>	30	30
Stratum 2	75	0.0709	30 <sup>a</sup>	30	30
Stratum 3	128	0.3506	35	35	35
Stratum 4	201	0.3280	35	35	35
Stratum 5	100	0.4413	43	43	43
Stratum 6	246	0.3222	35	35	36
Stratum 7	365	0.3008	32	32	34
Stratum 8	39	0.7983	34	34	36
Stratum 9	65	0.3973	32	32	33
Stratum 10	86	0.8129	64	64	72
Stratum 11	46	0.6836	37	37	38
Stratum 12	53	0.3418	30 <sup>a</sup>	30	30
Total for central region	1861	-	437	437	452

<sup>a</sup>The calculated sample sizes for Strata 1, 2, and 12 were less than 30.

Two methods for obtaining samples of the specified sizes are proposed. The following two subsections describe these methods and present the results of applying each method to the sample task.

### C.2.7 Selection of Samples by Method A

In Method A, the analyst randomly selects the specified number of BGs from each stratum without regard to location of grid cells within the stratum. An unbiased estimate for the mean concentration at the block group centroids can be constructed for each stratum using the stratum sample mean  $\bar{Y}$ . The variance for this estimate is calculated by the equation

$$Var(\bar{Y}) = \frac{N-n}{N} \frac{\sigma_c^2}{n} \quad (3)$$

where  $\sigma_c^2$  is the variance of the concentrations within the stratum.<sup>2</sup>

For the test application (butadiene in Phoenix), analysts used Method A to select a stratified random sample of 470 BGs from the sampling frame of 1894 BGs. Of these 470 BGs, 437 were selected from the 12 strata in the central region as indicated in Table C.2-3. The remaining 33 BGs in the sample include all of the BGs located in the outer region. As previously discussed, analysts determined that the BGs in the outer region were too widely dispersed to be aggregated into multi-BG strata.

Figure C.2-5 shows the locations of the centroids of the selected BGs. In the proposed HAPEM analysis, these BGs can be used to construct as many as 45 "exposure districts," 12 corresponding to the 12 strata in the central region and 33 corresponding to the 33 individual BGs in the outer region. (The number will be less than 45 if analysts decide to aggregate groups of BGs in the outer region). The outdoor pollutant concentration in each district would be determined by averaging the dispersion model estimates for centroids of the BGs assigned to each district.

### C.2.8 Selection of Samples by Method B

As indicated by Figure C.2-5, random clustering of data can occur when Method A is used to draw the sample. There is no guarantee that a simple random sample will adequately cover the study area. To improve coverage, the first-stage strata can be further stratified using the 4 km by 4 km grid cells within each stratum. A random sample of block groups can then be taken from each grid cell. The sample sizes determined above (listed as "required sample size" in Table 2-3) can be divided among the grid cells using a proportional allocation. The number of BGs sampled from grid cell  $h$  would be calculated as

$$n_h = nN_h / N, \quad (4)$$

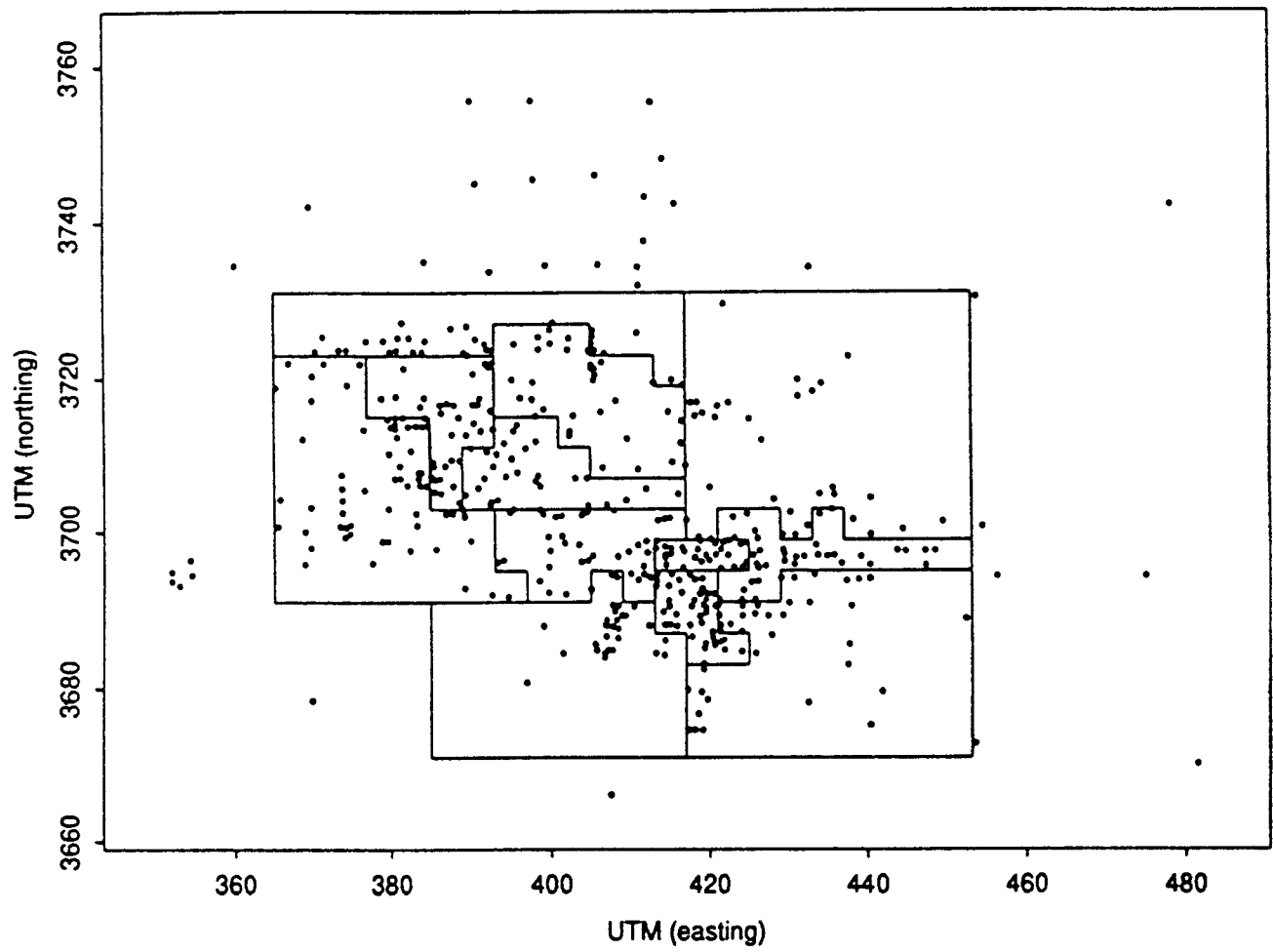


Figure C.2-5 Location of block groups selected by Method A.

where  $n$  is the total sample size in the particular stratum,  $N_h$  is the total number of BGs in cell  $h$ , and  $N$  is the total number of BGs in the stratum. This result is then rounded to the nearest integer. To guarantee that all populated grid cells are included in the sample, however, at least one BG must be sampled from each populated cell. Therefore, if  $n_h$  as calculated above would normally be rounded down to 0 (e.g.  $n_h = 2(35) / 246 \approx 0.28$ ), it is rounded up to 1. Note that this rounding scheme may actually result in larger sample sizes being drawn from some strata than those calculated for Method A. Unbiased estimates of the mean concentrations in each stratum can be constructed using a weighted sample mean of the form

$$\bar{y} = \frac{1}{N} \sum_{h=1}^L N_h \bar{y}_h , \quad (5)$$

where  $L$  is the total number of grid cells in the stratum,  $N_h$  is the number of block group centroids in cell  $h$ ,  $\bar{y}_h$  is the sample mean concentration in cell  $h$ , and  $N$  is the total number of BGs in the stratum (Thompson, 1992). Note that this formula places more weight on the more densely populated grid cells. The variance for this estimate is

$$Var(\bar{y}) = \sum_{h=1}^L \left[ \frac{N_h}{N} \right]^2 \left[ \frac{N_h - n_h}{N_h} \right] \frac{\sigma_h^2}{n_h} , \quad (6)$$

where  $\sigma_h^2$  is the concentration variance for grid cell  $h$ . However,  $\sigma_h^2$  will not be estimable for cells from which only one sample is taken. Consequently, estimates of  $Var(\bar{y})$  may be too low.

For the example application (1,3-butadiene in Phoenix), analysts used Method B to select a stratified random sample of 485 BGs (Figure C.2-6). Of these, 452 were selected from strata in the central region as shown in Table C.2-3. The remainder were selected by taking all 33 BGs in the outer region. Consistent with the results of Method A, the BGs selected by Method B can be assigned to 45 districts, 12 corresponding to the 12 strata in the central region and 33 corresponding to the individual BGs in the outer region. The outdoor pollutant concentration in each district would be determined by using the above formula to compute a weighted average of the dispersion model estimates for centroids of the BGs assigned to each district.

The locations of the BGs for one sample selected by Methods A and B are displayed in Figures C.2-5 and C.2-6, respectively. Both samples show good coverage of the study region, although some random clumping of data is present in both samples. Thus, analysts anticipate that Method A will provide results very similar to Method B.

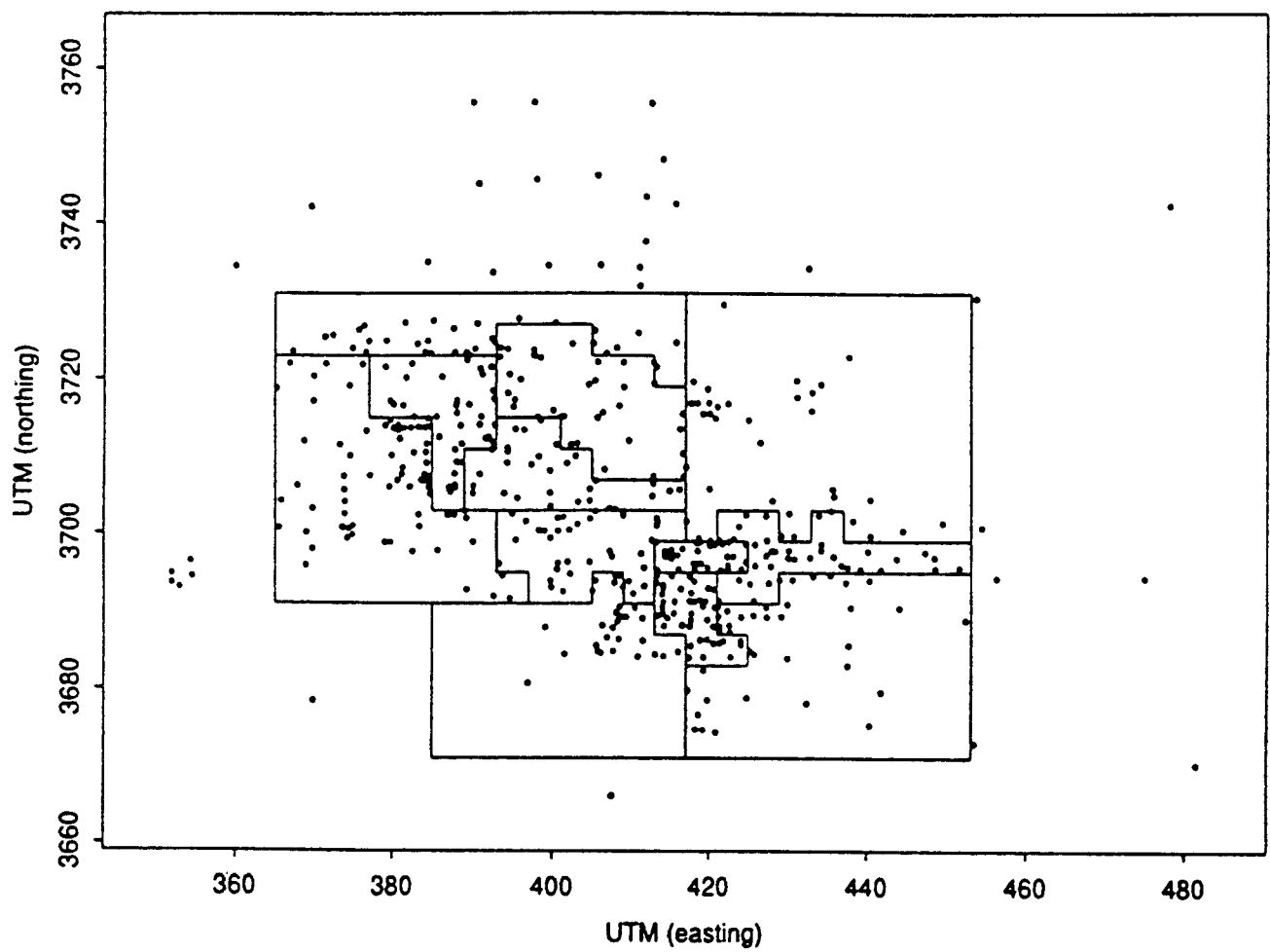


Figure C.2-6 Location of block groups selected by Method B.

### C.2.9 Decision to Perform Sensitivity Analyses

Based on these observations, researchers noted that Method A may be preferable to Method B because (1) it is easier to apply, (2) it requires a smaller sample size, and (3) it provides an unbiased estimate of the variance.

Section C.3 provides step-by-step procedures for implementing Methods A and B. Before making a final endorsement of one of these candidate sampling methods, researchers conducted a series of analyses to test the statistical assumptions underlying Methods A and B. Section C.4 summarizes these analyses, identifies Method B as the superior method in terms of statistical performance, and concludes that both methods tend to select samples which are larger than necessary to achieve the specified set of sampling goals. Sections C.5 and C.6 present Method C, a revised version of Method B, which is capable of achieving the sampling goals using a significantly smaller sample.



### C.3 SUMMARY OF SAMPLING METHODS A AND B

This section summarizes the step-by-step procedure for drawing samples using Methods A and B, the two methods proposed in the previous section. It also provides a summary of the principal limitations associated with these methods.

#### C.3.1 Step-By-Step Procedure

1. *Align the census population data and the emissions data for the particular pollutant and city.* For the Phoenix data, this step was implemented by first converting the block group centroid coordinates from latitude and longitude to UTM coordinates, and then converting the UTM coordinates to X and Y coordinates corresponding to the emissions grid.
2. *Define the sampling strata.* This step requires both subjective and objective decisions on the part of the sampler. For the Phoenix data, the study region was first divided into two large subregions, the central region and the outer region, based on the density of the population within each of these regions. (If the populations in the other cities in this study are more uniformly spread over the study region, this step may not be necessary). Next, contour maps of the emissions data were drawn with isopleths separating areas of "high", "moderate", and "low" total emissions. For Phoenix, analysts subjectively chose isopleth levels of 1000 kg/yr and 4000 kg/yr for butadiene. Each of the emissions grid cells was then categorized as "high", "moderate", or "low", corresponding to the emissions level within that grid cell. Smaller contiguous strata were formed by joining together neighboring grid cells of the same emissions category. These strata were further broken into smaller strata to avoid the "doughnut effect", in which one stratum is completely surrounded by another stratum. (The commuting algorithms of the HAPEM model do not allow for "doughnut shaped" exposure districts.) Additionally, extremely large strata were divided into smaller strata to help maintain the spatial characteristics of the concentration values.
3. *Determine the required sample sizes.* Each block group centroid was assigned an emissions value equivalent to the emissions value of the grid cell into which it fell. The distribution of emissions values across the block groups was then tabulated for each stratum. In particular, the coefficient of variation was determined for each stratum. Under the assumption of proportionality between the emissions and concentrations within each stratum, the sample size formula from Section C.2 was then utilized to determine the minimum sample size required to meet certain specifications. A minimum sample size requirement of 30 for each stratum was imposed to meet the large sample requirements necessary for the Central Limit Theorem.
4. *Select the sample using one of the two proposed methods.* Under Sampling Method A from Section C.2, simple random samples of block group centroids

were chosen from within each stratum to act as receptor points. Under Sampling Method B, each stratum was further stratified into grid cells, and a random sample of block group centroids was chosen from within each populated grid cell in each stratum. Simple random number generators were utilized to choose these samples, without replacement.

This approach is applicable to each combination of study area and pollutant.

### C.3.2 Limitations of Methods A and B

The sampling methods described above have the following limitations.

- The analyst must create a separate set of strata for each pollutant included in the analysis. Consequently, a distinct sample of block groups must be drawn for each pollutant.
- The method for defining the strata presented above will not be adequate when the pollutant under consideration is dominated by point-source emissions rather than area-source emissions.
- In determining sample size, emissions should not be used as a surrogate for the concentrations when point sources dominate the emissions for a particular pollutant.

These limitations will be addressed further in the development of a revised sampling method in Section C.5.

## C.4 SENSITIVITY ANALYSES OF SAMPLING METHODS A AND B

Previous sections of this appendix describe two methods for choosing a sample which is representative of the outdoor pollutant concentrations at population centroids. In both methods, the study region is first divided into a set of sampling strata which completely covers the study area. In sampling Method A, a simple random sample is taken from within each stratum for the purposes of estimating the mean concentration of the stratum. In sampling Method B, each stratum is further stratified into grid cells, and a weighted estimator is used to estimate the mean concentration of the stratum. One sampling objective was to estimate the mean outdoor pollutant concentration for each stratum to within 10 percent of the "true" mean concentration with 95 percent confidence. Note that here the "true" mean concentration would be the value obtained from the dispersion model estimates at each centroid within the stratum.

In calculating the sample sizes required to meet the sampling objectives for a simple random sample, analysts had to make two assumptions about the distribution of the concentrations. First, it was assumed that the coefficient of variation for the emission values was approximately the same as the coefficient of variation for the concentration values within a particular stratum. This is the "assumption of proportionality" discussed in previous sections. Second, it was assumed that the sample size for each stratum was "large enough" for the Central Limit Theorem to hold; i.e., the sampling distribution of the sample mean should be approximately normal.

Analysts conducted a series of sensitivity analyses with the following three general objectives in mind:

- To check the validity of the sample size assumptions.
- To compare sampling Method A with sampling Method B with regards to the accuracy and precision of the estimates.
- To identify deficiencies in the proposed sampling methods and develop modifications to address these deficiencies.

It will be shown through the sensitivity analyses that the methods previously proposed adequately meet the sampling objectives. However, modifications to these methods will be proposed to increase the efficiency and reduce the number of assumptions required.

Researchers obtained dispersion model estimates of the butadiene concentrations at each of the block group centroids in the Phoenix study region for 1991. Estimates were obtained for annual average concentrations as well as for 96 hourly/seasonal average concentrations. The sensitivity analyses will first focus on the annual average data. Subsequent analyses will then be performed on specific combinations of hour and season. Since all current analyses are being performed specifically on butadiene concentrations in the Phoenix study area, broad generalizations of these results to all pollutants in all cities must be approached cautiously.

#### C.4.1 Evaluation of Sample Sizes Calculated by Methods A and B

Table C.4-1 contains the true means and standard deviations of the annual average concentrations, by strata, as well as a comparison of the sample size calculations using the emissions values and the concentration values. The true mean concentrations for strata 1 and 2 are "high" relative to the other strata, the means for strata 3 through 7 are "moderate", and the means for strata 8 through 12 are "low". Recall that strata 1 and 2 were defined in Section C.2 as high emissions strata, strata 3 through 7 were defined as moderate emissions strata, and strata 8 through 12 were defined as low emissions strata. The strata mean concentrations range from about 0.04 to almost 0.18  $\mu/\text{m}^3$ . Additionally, the standard deviations of the concentrations within each stratum follow a similar pattern to that observed for the emission values in Section C.2. That is, the "high" and "low" strata tend to vary less than the "moderate strata".

Equation 2 in Section C.2 was used to calculate the sample size required to meet the sampling objectives for a simple random sample. Recall that  $Z_{\alpha/2}$  is 1.96 and  $r$  is 0.1 in this formula. Also recall that the result,  $n$ , should always be rounded up to the next positive integer. The coefficient of variation for the concentrations,  $\gamma_C$ , is shown in the fifth column of Table C.4-1, and is simply defined as the standard deviation divided by the mean. The coefficient of variation for the emissions values,  $\gamma_Q$ , which was used to approximate  $\gamma_C$  in the sample size calculations of Section C.2, is shown in column 7 of this table, and the ratio of these two values is presented in the ninth column. The sample sizes calculated from Equation 2 using  $\gamma_C$  and  $\gamma_Q$  are shown in columns 6 and 8 of Table C.4-1, respectively. The sample sizes in column 8 are the same as those presented in Table C.2-3, with the exception that sample sizes less than 30 were not rounded up here. The ratio of these sample sizes is shown in column 9.

In every stratum,  $\gamma_Q$  exceeds  $\gamma_C$ , which indicates that the required sample size was overestimated in each stratum when using the distribution of the emissions values to approximate the distribution of the concentration values. Sometimes, the use of  $\gamma_Q$  led to an estimated sample size which was actually more than twice the sample size required to meet the objectives. As discussed in Section C.2, oversampling is generally not a concern in estimation problems. On the other hand, oversampling results in longer dispersion model runs. Therefore, it may be beneficial to reduce the amount of oversampling. This point will be addressed further below.

The results presented in the last column of Table C.4-1 indicate that using the emissions values to represent the concentration values in calculating sample sizes would have been adequate, although somewhat inefficient, for the annual average 1,3-butadiene data. Analysts used Monte Carlo sampling methods to compare sampling Methods A and B with regards to the accuracy and precision of the estimated mean concentrations as well as to assess the validity of the normality assumption required for the sample size formula. Using the sample sizes presented in Table 2-3, analysts drew 200 samples of block group centroids using each method. The mean concentration of each stratum was then estimated by the concentrations at the block group centroids in each of the 200 samples. Recall that for Method A, the usual sample mean is used as the estimator, while for Method B the weighted sample mean in Equation 5 is used. Here, each estimator will simply be referred to as a sample mean.

**Table C.4-1 Comparison of Sample Size Calculation Results Using the Concentration Values  
Versus Using the Emissions Values**

Stratum	Total number of block groups, N	Descriptive statistics and sample size calculations for concentration values				Sample size calculations for emissions values (see Table 2-3)		Ratio of coefficients of variation, $\gamma_C / \gamma_Q$	Ratio of required sample sizes, $n_C / n_Q$
		True mean ( $\mu/m^3$ )	Standard deviation	Coefficient of variation, $\gamma_C$	Calculated sample size, $n_C$	Coefficient of variation, $\gamma_Q$	Calculated sample size, $n_Q$		
1	457	0.1786	0.0137	0.0767	3	0.0935	4	0.82	0.75
2	75	0.1601	0.0066	0.0415	1	0.0709	2	0.59	0.50
3	128	0.1033	0.0233	0.2259	18	0.3506	35	0.64	0.51
4	201	0.1191	0.0261	0.2190	17	0.3280	35	0.67	0.49
5	100	0.0895	0.0271	0.3025	27	0.4413	43	0.69	0.63
6	246	0.1137	0.0228	0.2008	15	0.3222	35	0.62	0.43
7	365	0.1175	0.0230	0.1956	15	0.3008	32	0.65	0.47
8	39	0.0540	0.0212	0.3928	24	0.7983	34	0.49	0.71
9	65	0.0496	0.0119	0.2391	17	0.3973	32	0.60	0.53
10	86	0.0549	0.0160	0.2916	24	0.8129	64	0.36	0.38
11	46	0.0434	0.0182	0.4185	28	0.6836	37	0.61	0.76
12	53	0.0503	0.0166	0.3293	24	0.3418	25	0.96	0.96
Total	1861	0.1211	0.0462	-	213	-	378	-	0.56

Each of the 200 Monte Carlo samples resulted in estimates of the true strata means. The distribution of these estimates (for each stratum) across the 200 samples is known as the sampling distribution. If the mean of the sampling distribution is equal to the true mean, then the estimator is unbiased. In addition, the variance of the sampling distribution can be estimated from these 200 samples. The square root of the variance, the standard error of the estimator, gives one an idea of the "spread" of the sampling distribution around the true mean.

Table C.4-2 contains the results of the 200 Monte Carlo samples obtained by applying sampling Methods A and B to the annual average butadiene data. Included in this table are the estimated means of the sampling distribution and the estimated standard errors of the sampling distribution for both sampling methods. Statistics for evaluating the attainment of the sampling objective and for testing the normality assumption are also provided.

Comparing the true means in Table C.4-1 with the estimated means in Table C.4-2 reveals that both sampling methods produce unbiased estimates of the mean concentration, as anticipated by analysts. However, the standard errors are consistently much smaller under Method B than under Method A. Typically, estimators from stratified random samples (e.g., Method B) have lower variances than those from simple random samples (e.g., Method A). Therefore, analysts anticipated that Method B would result in more precise estimates than Method A. However, it was not anticipated that Method B would outperform Method A to the degree observed here. As can be seen from Table C.4-2, the standard errors from Method A are generally between 2 and 7 times higher than those from Method B.

The reason for the large differences in precision between the two types of estimators as seen here is fairly obvious. The variance of an estimator from Method A is a function of the overall stratum concentration variance, as shown in Equation 3 in Section C.2. The variance of an estimator from Method B is a function of the variances within each of the grid cells within each stratum, as shown in Equation 6. Therefore, the large differences in precision are mostly attributable to the small within-cell concentration variances relative to the overall stratum concentration variances.

Recall that one of the primary sampling objectives was to estimate the true mean concentration to within 10 percent with 95 percent confidence. Table C.4-2 contains the proportion of sample means from the 200 Monte Carlo samples which were within 10 percent of the true strata concentrations for both sampling methods. Both methods easily met the sampling objectives. For Method A, at least 198 of the 200 (99 percent) sample means from each stratum were within 10 percent of the true mean. For Method B, every sample mean met the objective.

Equation 2 in Section C.2 is used for calculating the required sample sizes for simple random samples. This equation requires that the distribution of the unweighted sample mean from a simple random sample be at least approximately normal. To validate this assumption, the p-values from Shapiro-Wilks tests of the normality of the sampling distributions for Method A are included in Table C.4-2. In general, a p-value larger than 0.05 indicates that the normality of

**Table C.4-2 Results of 200 Monte Carlo Samples for Sampling Methods A and B Using Sample Sizes  
as Calculated in Table C.2-3**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	30	0.1786	0.0023	1.0	30	0.1786	0.0010	1.0	0.7733
2	30	0.1601	0.0010	1.0	30	0.1602	0.0006	1.0	0.7413
3	35	0.1032	0.0035	0.990	35	0.1032	0.0008	1.0	0.1547
4	35	0.1191	0.0039	0.990	35	0.1191	0.0010	1.0	0.9949
5	43	0.0892	0.0031	0.995	43	0.0895	0.0006	1.0	0.8108
6	35	0.1134	0.0033	1.0	36	0.1136	0.0010	1.0	0.6466
7	32	0.1174	0.0039	1.0	34	0.1175	0.0010	1.0	0.4202
8	34	0.0540	0.0013	1.0	36	0.0540	0.0004	1.0	0.2152
9	32	0.0495	0.0015	1.0	33	0.0496	0.0008	1.0	0.7461
10	64	0.0550	0.0010	1.0	72	0.0549	0.0003	1.0	0.8829
11	37	0.0434	0.0013	1.0	38	0.0434	0.0002	1.0	0.2394
12	30	0.0505	0.0020	0.995	30	0.0503	0.0005	1.0	0.5321

sampling distribution cannot be rejected. None of the p-values in this table indicate that the normality assumption should be questioned for samples of at least 30. Thus, the minimal sample size of 30 which was imposed on all strata is adequate. However, it will be shown below that it may be possible to reduce this minimal requirement.

#### C.4.2 Effects of Reducing Sample Sizes

Tables C.4-3 through C.4-8 contain similar statistics to those in Table C.4-2 for 200 Monte Carlo samples of nominal sizes 30, 25, 20, 15, 10, and 5, respectively. The sample sizes are “nominal” due to the rounding scheme employed in sampling Method B. Recall from Section C.2 that at least one block group centroid must be sampled from each populated grid cell under this method. Therefore, while the nominal sample size is always the size of the sample actually selected under Method A, samples selected under Method B will tend to be larger. In fact, under Method B, the minimum sample size drawn from any particular stratum equals the number of populated grid cells within that stratum. Recall that a “populated” grid cell was defined as one which contains at least one block group centroid.

Four of the p-values from the Shapiro-Wilks test of normality presented in Tables C.4-3 through C.4-8 are significant at the 0.05 level (Table C.4-3, stratum 1; Table C.4-4, stratum 2; Table C.4-6, stratum 4; and Table C.4-8, stratum 3). However, it is well known that the Shapiro-Wilks statistic is overly conservative (i.e., it rejects the hypothesis of normality too often). Therefore, in addition to performing these tests, normal probability plots and histograms were examined for each Monte Carlo sample. The results of these tests, as well as the plots, indicate that the sampling distribution is at least approximately normal in all strata for sample sizes as small as 5. Therefore, these data indicate that the minimal sample size requirement of 30 imposed in Section C.2 could be lowered to 5. This result only effects the sample sizes for strata 1, 2, and 12, as the calculated sample sizes from the other strata were all greater than 30. However, this would result in a total sample size requirement of 382 block group centroids being drawn from the central region under Method A, rather than 437 as calculated in Section C.2. This is a reduction of 55 block groups (13 percent) in the overall sample size.

Analysts also observed from Tables C.4-3 through C.4-8 that, although both sampling methods result in unbiased estimates for the true mean concentration, Method B continues to provide significantly more precise estimates than Method A, even when smaller sample sizes are used. The standard error of the mean estimator from Method A is generally 2 to 7 times higher than that of Method B for each stratum in all of these tables. Note that this comparison is not entirely fair, as larger samples are usually drawn using Method B, and larger samples generally lead to more precise estimates. However, the difference in standard errors is so great here that it is reasonable to conclude that Method B results in a much more precise estimator than Method A. As previously mentioned, the high precision of Method B is primarily a result of the low within-cell concentration variances observed in this data.



**Table C.4-3 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 30  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	30	0.1783	0.0025	1.0	30	0.1786	0.0010	1.0	0.0018 <sup>a</sup>
2	30	0.1601	0.0010	1.0	30	0.1602	0.0006	1.0	0.2204
3	30	0.1034	0.0035	0.995	33	0.1033	0.0008	1.0	0.7251
4	30	0.1193	0.0045	0.990	31	0.1191	0.0009	1.0	0.8368
5	30	0.0891	0.0044	0.950	31	0.0896	0.0008	1.0	0.6738
6	30	0.1136	0.0039	0.990	30	0.1137	0.0011	1.0	0.8122
7	30	0.1174	0.0044	0.990	31	0.1175	0.0011	1.0	0.7118
8	30	0.0539	0.0019	0.995	34	0.0539	0.0005	1.0	0.5961
9	30	0.0495	0.0015	1.0	33	0.0497	0.0008	1.0	0.2951
10	30	0.0549	0.0023	0.970	41	0.0549	0.0007	1.0	0.4006
11	30	0.0435	0.0018	0.980	33	0.0434	0.0004	1.0	0.7102
12	30	0.0503	0.0020	0.990	30	0.0503	0.0005	1.0	0.5466

<sup>a</sup> The results from a normal probability plot and a histogram indicate that the sampling distribution is at least approximately normal.

**Table C.4-4 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 25  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	25	0.1784	0.0026	1.0	25	0.1786	0.0011	1.0	0.9279
2	25	0.1602	0.0011	1.0	25	0.1600	0.0007	1.0	0.0408 <sup>a</sup>
3	25	0.1033	0.0042	0.990	26	0.1032	0.0011	1.0	0.3877
4	25	0.1182	0.0053	0.975	25	0.1192	0.0012	1.0	0.6035
5	25	0.0895	0.0046	0.930	25	0.0894	0.0009	1.0	0.1756
6	25	0.1137	0.0040	0.985	26	0.1136	0.0012	1.0	0.7960
7	25	0.1180	0.0044	0.990	28	0.1176	0.0011	1.0	0.5219
8	25	0.0541	0.0026	0.960	29	0.0540	0.0006	1.0	0.3657
9	25	0.0495	0.0018	0.985	29	0.0496	0.0009	1.0	0.8800
10	25	0.0551	0.0028	0.950	38	0.0548	0.0007	1.0	0.5963
11	25	0.0434	0.0023	0.945	29	0.0434	0.0005	1.0	0.3380
12	25	0.0502	0.0027	0.935	25	0.0503	0.0006	1.0	0.9830

<sup>a</sup> The results from a normal probability plot and a histogram indicate that the sampling distribution is at least approximately normal.

**Table C.4-5 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 20  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	20	0.1786	0.0029	1.0	20	0.1783	0.0012	1.0	0.0686
2	20	0.1600	0.0012	1.0	20	0.1600	0.0009	1.0	0.0654
3	20	0.1035	0.0046	0.985	20	0.1031	0.0011	1.0	0.3180
4	20	0.1190	0.0058	0.970	21	0.1189	0.0014	1.0	0.1508
5	20	0.0896	0.0056	0.880	20	0.0897	0.0012	1.0	0.1115
6	20	0.1136	0.0049	0.980	21	0.1138	0.0015	1.0	0.5078
7	20	0.1174	0.0050	0.975	23	0.1175	0.0015	1.0	0.2064
8	20	0.0541	0.0035	0.855	27	0.0539	0.0007	1.0	0.1491
9	20	0.0496	0.0022	0.975	24	0.0497	0.0012	1.0	0.4034
10	20	0.0549	0.0029	0.930	36	0.0548	0.0008	1.0	0.1935
11	20	0.0432	0.0030	0.945	24	0.0434	0.0007	1.0	0.9480
12	20	0.0505	0.0030	0.935	22	0.0503	0.0006	1.0	0.5187

**Table C.4-6 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 15  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	15	0.1782	0.0036	1.0	15	0.1785	0.0016	1.0	0.0689
2	15	0.1601	0.0016	1.0	15	0.1601	0.0010	1.0	0.7642
3	15	0.1037	0.0061	0.940	16	0.1034	0.0013	1.0	0.8156
4	15	0.1189	0.0066	0.955	16	0.1192	0.0018	1.0	0.0032 <sup>a</sup>
5	15	0.0892	0.0064	0.830	15	0.0895	0.0014	1.0	0.6856
6	15	0.1148	0.0057	0.935	18	0.1136	0.0016	1.0	0.3086
7	15	0.1179	0.0062	0.960	21	0.1176	0.0017	1.0	0.5515
8	15	0.0538	0.0042	0.800	25	0.0539	0.0009	1.0	0.0971
9	15	0.0497	0.0030	0.885	21	0.0495	0.0013	1.0	0.4799
10	15	0.0551	0.0034	0.890	33	0.0549	0.0009	1.0	0.6217
11	15	0.0431	0.0041	0.695	21	0.0433	0.0008	1.0	0.1890
12	15	0.0505	0.0036	0.840	17	0.0504	0.0009	1.0	0.4908

<sup>a</sup>The results from a normal probability plot and a histogram indicate that the sampling distribution is at least approximately normal.

**Table C.4-7 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 10  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	10	0.1790	0.0044	1.0	13	0.1785	0.0016	1.0	0.6434
2	10	0.1601	0.0020	1.0	10	0.1603	0.0013	1.0	0.7803
3	10	0.1039	0.0070	0.845	13	0.1033	0.0016	1.0	0.2148
4	10	0.1195	0.0076	0.895	14	0.1190	0.0017	1.0	0.7887
5	10	0.0907	0.0079	0.705	11	0.0897	0.0014	1.0	0.1144
6	10	0.1145	0.0072	0.870	12	0.1138	0.0022	1.0	0.3483
7	10	0.1175	0.0076	0.885	21	0.1174	0.0015	1.0	0.4185
8	10	0.0538	0.0055	0.650	22	0.0541	0.0014	1.0	0.6295
9	10	0.0494	0.0032	0.910	17	0.0496	0.0018	0.995	0.6832
10	10	0.0549	0.0047	0.760	31	0.0549	0.0012	1.0	0.3455
11	10	0.0433	0.0054	0.605	20	0.0435	0.0009	1.0	0.6603
12	10	0.0501	0.0050	0.665	12	0.0501	0.0011	1.0	0.7422

**Table C.4-8 Results of 200 Monte Carlo Samples for Sampling Methods A and B with a Sample of Nominal Size 5  
Drawn from Each Stratum**

Stratum	Sampling Method A				Sampling Method B				P-value for the Shapiro - Wilks test of normality
	Sample size drawn from each stratum	Estimated mean of the sampling distribution, ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	Sample size drawn from each stratum	Estimated mean of the sampling distribution, ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean	
1	5	0.1790	0.0062	0.995	13	0.1786	0.0016	1.0	0.2042
2	5	0.1603	0.0029	1.0	6	0.1602	0.0017	1.0	0.4813
3	5	0.1042	0.0106	0.665	12	0.1033	0.0016	1.0	0.0080 <sup>a</sup>
4	5	0.1187	0.0125	0.630	14	0.1189	0.0017	1.0	0.2860
5	5	0.0904	0.0115	0.560	6	0.0895	0.0023	1.0	0.8763
6	5	0.1136	0.0094	0.790	12	0.1135	0.0021	1.0	0.5681
7	5	0.1176	0.0110	0.690	21	0.1174	0.0014	1.0	0.0643
8	5	0.0541	0.0085	0.460	21	0.0539	0.0014	1.0	0.0541
9	5	0.0503	0.0051	0.660	16	0.0495	0.0021	0.995	0.5721
10	5	0.0549	0.0073	0.555	31	0.0549	0.0011	1.0	0.3890
11	5	0.0428	0.0072	0.435	19	0.0435	0.0011	1.0	0.1403
12	5	0.0501	0.0077	0.515	9	0.0503	0.0016	0.995	0.4410

<sup>a</sup> The results from a normal probability plot and a histogram indicate that the sampling distribution is at least approximately normal.

Tables C.4-3 through C.4-8 also indicate the proportion of sample means from the 200 Monte Carlo samples which actually fall within 10 percent of the true concentration mean for each stratum. In Table C.4-3, at least 95 percent of the sample means are within 10 percent of the true mean under Method A. Under Method B, every sample mean fell within 10 percent of the true mean. However, as the sample size drawn from each stratum decreases in the subsequent tables, the proportion of sample means from sampling Method A which meet the sampling objective decreases. When the sample sizes fall below those presented in Table C.4-1, the proportion of sample means within 10 percent of the true mean generally falls below 0.95 for Method A. Under Method B, this proportion is at least 0.995 in all cases.

Although the results of this section show that both methods described in Section C.2 would have been adequate for the annual average butadiene data, it is clear that Method B is superior to Method A. Method C, a revised version of Method B which significantly improves the sampling process, will be presented in the next section.

## C.5 REFINEMENTS TO SAMPLING METHODOLOGY BASED ON FINDINGS OF SENSITIVITY ANALYSES

As analysts observed in the previous section, the sampling methods of Section C.2 of this appendix produced adequate results for the annual average 1,3-butadiene data. However, as was noted in Section C.5, Method B produces significantly more precise estimates than Method A. Additionally, the methods used for computing required sample sizes tend to cause significant oversampling, resulting in less efficient dispersion model runs. Refinements to the above sampling methodology are proposed in this section which address both of these issues, resulting in a more general sampling method.

The results of the previous section demonstrate that Method B produces significantly more precise estimates than Method A, but that both methods produce unbiased estimators. In this sense, Method B is a better sampling method than Method A. Therefore, the focus in this section will be on revising sampling Method B exclusively.

### C.5.1 Proposed Refinements

Recall that Table C.4-8 contains the results from 200 Monte Carlo samples of "nominal" size 5 from each stratum. However, with the exception of stratum 2, each of the strata contains at least 5 populated grid cells. Therefore, due to the rounding scheme employed in Method B, taking a sample of "nominal" size 5 results in choosing one centroid from each populated grid cell within each stratum for the Phoenix study area. In stratum 2, which contains only 3 populated grid cells, two centroids were chosen from each grid cell.

The results presented in Table C.4-8 suggest that taking one centroid from each grid cell would produce unbiased and very precise estimators. However, it would not be possible using this method to estimate the within-cell variances. Consequently, it would not be possible to estimate the variance of the weighted sample mean (see Equation 6, Section C.2). Without an estimate of the variance of the estimator, researchers would have no indication of the precision of their estimator.

Analysts suggest the following modifications to Sampling Method B. Randomly choose two block group centroids from each of the populated grid cells within each stratum. If there is only one centroid in a grid cell, it is included in the sample automatically. The weighted sample mean in Equation 5, Section C.2, can be used as before to estimate the mean concentration within each stratum. With two centroids being selected from each grid cell, it is also possible to estimate the variance of the weighted sample mean. Researchers can simply substitute the within-grid-cell sample variances for the true variances in Equation 6 to obtain an unbiased estimate.

This revised version of Method B (hereafter referred to as "Method C") has several advantages over the methods suggested in Section C.2. First, an unbiased estimate of the variance of the sample mean can be formed, which was not the case for the original Method B. Second, analysts are required to make fewer assumptions about the distribution of the pollutant



concentrations. Since a fixed number of centroids are being selected from within each grid cell, it is not necessary to calculate the required sample sizes. Thus, the following assumptions that were used in calculating sample sizes are no longer required: (1) the proportionality between the emission values and the concentration values and (2) the normality of the sampling distribution.

Another, and perhaps the most important, advantage of Method C is that the sample size is not pollutant dependant. That is, regardless of the pollutant being analyzed, researchers could use the same number of receptor points for the dispersion model. In fact, it may be possible to use the same receptor points for each pollutant, which would be beneficial to the researchers performing the dispersion model runs.

The original strata served to divide the study area into subregions, or "exposure districts". In addition, the strata played a large role in the sampling methods described in Section C.2 of this appendix. However, under Method C, the original strata themselves play no role in the actual sampling. Researchers can perform the dispersion model runs at the chosen receptor points before the study region is divided into the smaller exposure districts. In fact, the concentration values themselves could be used to create the exposure districts, rather than using the emissions values as a surrogate.

#### C.5.2 Revised Methodology Applied to the Annual Average Data

Analysts drew 200 Monte Carlo samples from each stratum using the revised sampling method (Method C) proposed above. Each of the 200 Monte Carlo samples resulted in estimates of the true strata means. The distribution of these estimates (for each stratum) across the 200 samples is known as the sampling distribution. In addition, the variance of the sampling distribution can be estimated from these 200 samples. The square root of the variance, the standard error of the estimator, gives one an indication of the "spread" of the sampling distribution around the true mean.

Table C.5-1 contains the results of the 200 Monte Carlo samples for the annual average butadiene data. For strata 1 through 7, the sample sizes are equal to twice the number of populated grid cells within the strata. That is, all of the populated grid cells within the most heavily populated strata contain at least 2 block group centroids. In each of the more sparsely populated strata, strata 8 through 12, there is at least one grid cell which contains only one block group centroid. Thus, the sample sizes taken from these strata are all less than twice the number of populated grid cells. The total required sample size from the central region is 312 centroids, a reduction of 125 from the 437 centroids suggested in Section C.2. Including the 33 centroids in the outer region, the total number of receptor points for the dispersion model run in the Phoenix study area would then be 345.

**Table C.5-1 Results of 200 Monte Carlo Samples for Sampling Method C<sup>a</sup>  
When Applied to the Annual Average Butadiene Data**

Stratum	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	26	0.1786	0.0010	1.0
2	6	0.1602	0.0016	1.0
3	24	0.1034	0.0011	1.0
4	28	0.1190	0.0011	1.0
5	12	0.0896	0.0016	1.0
6	24	0.1136	0.0015	1.0
7	42	0.1174	0.0011	1.0
8	29	0.0539	0.0008	1.0
9	26	0.0496	0.0014	1.0
10	50	0.0549	0.0007	1.0
11	30	0.0433	0.0007	1.0
12	15	0.0501	0.0009	1.0
Total	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

The results in Table C.5-1 indicate that this method provides unbiased estimates of the strata mean concentrations, as expected. A comparison of the standard errors in Table C.5-1 with those for Method A in Table C.4-2 reveals that Method C provides, in general, more precise estimators than Method A, even with a large reduction in sample size. Every sample mean from the 200 Monte Carlo samples was within 10 percent of the true stratum mean.

### C.5.3 Application of Method C to Data from Specific Hourly/Seasonal Combinations

In addition to the annual average concentrations, dispersion model estimates were obtained for the average concentration at each hour of the day during each season of the year. In all, there are 96 such hourly/seasonal combinations (24 hours x 4 seasons). Some of these combinations were found to have concentration estimates which varied to a much greater extent than the annual average concentrations. Other hourly/seasonal combinations had concentration estimates which varied little. In order to assess the generalizability of the revised sampling method, analysts applied the method to several hourly/seasonal combinations representing concentration estimates with differing levels of variance.

Various population activities and meteorological patterns throughout the day have effects on the levels of pollutant concentrations and the degree to which they vary. For this reason, researchers chose to test Method C on a time during rush hour, 7 a.m., a time at mid-day, 12 p.m., and a time at night, 12 a.m. Researchers found that the rush hour concentrations tended to be higher than average and vary greatly, the mid-day concentrations tended to be very similar to the annual average data, and the night time concentrations tended to be low and vary little. Researchers also found that the variability of the concentration estimates was similar in the winter and fall seasons as well as in the spring and summer seasons. Method C was applied to the 6 combinations of two seasons (winter and summer) and three times of day (7 a.m., 12 p.m., and 12 a.m.).

Descriptive statistics and results from 200 Monte Carlo samples for each of the 6 hourly/seasonal combinations are presented in Tables C.5-2 through C.5-7. Analysts noted the following points concerning the descriptive statistics in these six tables:

- The winter concentrations tended to be substantially higher than the corresponding summer concentrations at 7 a.m. and 12 p.m. At 12 a.m., the results from the two seasons were relatively similar.
- The 7 a.m. mean concentrations were higher than the other two hours in both seasons. This pattern is probably due to rush hour traffic. The winter mean concentrations at 7 a.m. ranged from 0.1021  $\mu/\text{m}^3$  in stratum 8 to 0.5516  $\mu/\text{m}^3$  in stratum 1. The summer mean concentrations at 7 a.m. ranged from 0.0502  $\mu/\text{m}^3$  in stratum 11 to 0.2417  $\mu/\text{m}^3$  in stratum 1.

**Table C.5-2 Descriptive Statistics for the Winter, 7 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean concentration ( $\mu/m^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.5516	0.0772	0.1400	26	0.5513	0.0052	1.0
2	75	0.4681	0.0260	0.0556	6	0.4683	0.0073	1.0
3	128	0.2649	0.0714	0.2695	24	0.2654	0.0046	1.0
4	201	0.3930	0.0764	0.1943	28	0.3927	0.0045	1.0
5	100	0.3015	0.0898	0.2978	12	0.3011	0.0060	1.0
6	246	0.4971	0.0814	0.1637	24	0.4972	0.0052	1.0
7	365	0.3328	0.0859	0.2580	42	0.3326	0.0044	1.0
8	39	0.1021	0.0724	0.7094	29	0.1020	0.0030	1.0
9	65	0.2158	0.0775	0.3590	26	0.2159	0.0066	1.0
10	86	0.2671	0.1042	0.3901	50	0.2668	0.0039	1.0
11	46	0.1032	0.0536	0.5188	30	0.1033	0.0022	1.0
12	53	0.2079	0.0767	0.3689	15	0.2079	0.0063	0.995
Total	1861	0.3863	0.1580	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

**Table C.5-3 Descriptive Statistics for the Winter, 12 p.m., Concentration Values and Results of 200 Monte Carlo Samples Using the Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean concentration ( $\mu/\text{m}^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.1585	0.0119	0.0751	26	0.1585	0.0008	1.0
2	75	0.1316	0.0063	0.0476	6	0.1316	0.0014	1.0
3	128	0.0763	0.0185	0.2427	24	0.0763	0.0008	1.0
4	201	0.1136	0.0285	0.2508	28	0.1136	0.0010	1.0
5	100	0.0670	0.0219	0.3273	12	0.0669	.00011	1.0
6	246	0.0977	0.0203	0.2078	24	0.0978	.00013	1.0
7	365	0.0922	0.0238	0.2577	42	0.0923	0.0008	1.0
8	39	0.0351	0.0142	0.4026	29	0.0351	0.0006	1.0
9	65	0.0409	0.0069	0.1700	26	0.0409	0.0009	1.0
10	86	0.0600	0.0125	0.2083	50	0.0600	0.0005	1.0
11	46	0.0356	0.0122	0.3430	30	0.0356	0.0003	1.0
12	53	0.0441	0.0133	0.3013	15	0.0440	0.0009	1.0
Total	1861	0.1019	0.0440	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

**Table C.5-4 Descriptive Statistics for the Winter, 12 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using the Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean concentration ( $\mu/m^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.0068	0.0006	0.0914	26	0.0067	0.0001	1.0
2	75	0.0047	0.0003	0.0698	6	0.0047	0.0001	1.0
3	128	0.0034	0.0006	0.1886	24	0.0034	0.0001	1.0
4	201	0.0061	0.0030	0.5018	28	0.0061	0.0006	0.920
5	100	0.0042	0.0005	0.1177	12	0.0042	0.0001	1.0
6	246	0.0049	0.0006	0.1179	24	0.0050	0.0001	1.0
7	365	0.0048	0.0008	0.1671	42	0.0047	0.0001	1.0
8	39	0.0030	0.0008	0.2527	29	0.0030	0.0001	1.0
9	65	0.0034	0.0003	0.0978	26	0.0034	0.0001	1.0
10	86	0.0042	0.0010	0.2303	50	0.0042	0.0001	1.0
11	46	0.0033	0.0014	0.4223	30	0.0033	0.0001	1.0
12	53	0.0038	0.0004	0.1021	15	0.0038	0.0001	1.0
Total	1861	0.0051	0.0017	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

**Table C.5-5 Descriptive Statistics for the Summer, 7 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using the Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean concentration ( $\mu/m^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.2417	0.0311	0.1288	26	0.2416	0.0016	1.0
2	75	0.2312	0.0078	0.0370	6	0.2313	0.0027	1.0
3	128	0.1324	0.0331	0.2498	24	0.1325	0.0015	1.0
4	201	0.1787	0.0324	0.1815	28	0.1786	0.0017	1.0
5	100	0.1371	0.0417	0.3045	12	0.1368	0.0025	1.0
6	246	0.1958	0.0387	0.1977	24	0.1959	0.0018	1.0
7	365	0.1421	0.0340	0.2392	42	0.1421	0.0016	1.0
8	39	0.0507	0.0285	0.5626	29	0.0507	0.0010	1.0
9	65	0.0672	0.0234	0.3485	26	0.0673	0.0022	1.0
10	86	0.1243	0.0311	0.2505	50	0.1242	0.0013	1.0
11	46	0.0502	0.0245	0.4884	30	0.0503	0.0011	1.0
12	53	0.0955	0.0327	0.3422	15	0.0955	0.0022	1.0
Total	1861	0.1686	0.0617	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

**Table C.5-6 Descriptive Statistics for the Summer, 12 p.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean Concentration ( $\mu/\text{m}^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/\text{m}^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.1280	0.0119	0.0926	26	0.1280	0.0008	1.0
2	75	0.1058	0.0081	0.0763	6	0.1058	0.0011	1.0
3	128	0.0596	0.0175	0.2940	24	0.0596	0.0006	1.0
4	201	0.0823	0.0238	0.2888	28	0.0823	0.0008	1.0
5	100	0.0495	0.0205	0.4143	12	0.0494	0.0009	1.0
6	246	0.0632	0.0162	0.2570	24	0.0633	0.0009	1.0
7	365	0.0772	0.0220	0.2853	42	0.0773	0.0007	1.0
8	39	0.0264	0.0107	0.4037	29	0.0264	0.0005	1.0
9	65	0.0282	0.0067	0.2384	26	0.0283	0.0007	1.0
10	86	0.0229	0.0092	0.4013	50	0.0229	0.0004	1.0
11	46	0.0250	0.0093	0.3734	30	0.0250	0.0002	1.0
12	53	0.0249	0.0074	0.2958	15	0.0248	0.0004	1.0
Total	1861	0.0776	0.0389	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.



**Table C.5-7 Descriptive Statistics for the Summer, 12 a.m., Concentration Values and Results of 200 Monte Carlo Samples Using Sampling Method C<sup>a</sup>**

Stratum	Number of block group centroids within stratum	Descriptive Statistics			Sampling Results (Method C)			
		Mean concentration ( $\mu/m^3$ )	Concentration standard deviation	Coefficient of variation	Sample size drawn from each stratum	Estimated mean of the sampling distribution ( $\mu/m^3$ )	Estimated standard error of the sample mean	Proportion of sample means within 10% of the true mean
1	457	0.0046	0.0010	0.2157	26	0.0046	0.0001	1.0
2	75	0.0062	0.0004	0.0671	6	0.0062	0.0002	1.0
3	128	0.0054	0.0007	0.1396	24	0.0054	0.0001	1.0
4	201	0.0068	0.0037	0.5459	28	0.0069	0.0009	0.785
5	100	0.0059	0.0010	0.1644	12	0.0059	0.0001	1.0
6	246	0.0032	0.0003	0.0834	24	0.0032	0.0001	1.0
7	365	0.0035	0.0006	0.1714	42	0.0035	0.0001	1.0
8	39	0.0040	0.0006	0.1550	29	0.0040	0.0001	1.0
9	65	0.0023	0.0005	0.2127	26	0.0023	0.0001	1.0
10	86	0.0032	0.0012	0.3828	50	0.0032	0.0001	1.0
11	46	0.0045	0.0012	0.2697	30	0.0045	0.0001	1.0
12	53	0.0059	0.0009	0.1503	15	0.0059	0.0001	1.0
Total	1861	0.0045	0.0019	-	312	-	-	-

<sup>a</sup>Sampling Method C consists of drawing two BG centroids at random from each populated grid cell.

- The 12 p.m. winter concentrations ranged from  $0.0351 \mu/\text{m}^3$  in stratum 8 to  $0.1585 \mu/\text{m}^3$  in stratum 1. The 12 p.m. summer concentrations ranged from  $0.0229 \mu/\text{m}^3$  in stratum 10 to  $0.1280 \mu/\text{m}^3$  in stratum 1. The 12 a.m. winter concentrations ranged from  $0.0030 \mu/\text{m}^3$  in stratum 8 to  $0.0068 \mu/\text{m}^3$  in stratum 1. The 12 a.m. summer concentrations ranged from  $0.0023 \mu/\text{m}^3$  in stratum 9 to  $0.0068 \mu/\text{m}^3$  in stratum 4.
- The concentration standard deviations were significantly higher for both seasons at 7 a.m. than at the other two hours. The winter values varied substantially more than the summer values at 7 a.m. At the other two hours, the concentrations from the two seasons displayed about the same degree of variation.
- At 7 a.m. and 12 p.m., the concentration means for both seasons followed a similar pattern as the annual average data with regards to the high, moderate, and low strata. However, at 12 a.m., the mean concentrations were very nearly uniform over the study area.
- The concentration values from stratum 4 varied considerably more than those of the other strata at 12 a.m. This important finding will be discussed further below.

Coefficients of variation are presented in Tables C.5-2 through C.5-7 merely for completeness. They actually play no role in the revised sampling method. However, comparison of these coefficients of variation with those for the emissions values in Table C.4-1 reveals that sampling Method A would not meet the sampling objective of estimating the mean to within 10 percent with 95 percent confidence in several instances. For example, the coefficient of variation in stratum 4 for the winter, 12 a.m., data is 0.5018, which is much higher than the 0.3280 from the emissions values. This difference would result in the under-estimation of the required sample size by almost half. In most other cases, the coefficient of variation for the emissions is still higher than that of the concentrations, which would result in oversampling, as before.

The results of 200 Monte Carlo samples using Method C are also presented in Tables C.5-2 through C.5-7. Each of the 200 samples included the same block group centroids for all combinations of hour and season. Because the estimated mean of the sampling distribution is approximately equal to the true mean in all cases, analysts determined that the method produces unbiased results.

In almost every case, all, or nearly all, of the 200 sample means are within 10 percent of the true mean concentration. The only two exceptions both occur in stratum 4 at 12 a.m. For the winter data, 191 of 200 (92 percent) sample means met the 10 percent objective, while for the summer data, only 157 out of 200 (78.5 percent) sample means met the 10 percent objective. As discussed above, the standard deviation of the concentrations in stratum 4 at 12 a.m. was considerably higher than those of all other strata for both seasons. This relatively large variance was apparently caused by the variance within one particular grid cell (UTM coordinates 397 Easting, 3700 Northing), perhaps due to a point source within this grid cell. For the winter, concentrations ranged from  $0.0064$  to  $0.0395 \mu/\text{m}^3$ , while the summer concentrations ranged

from 0.0056 to 0.0430  $\mu\text{m}^3$  within this grid cell. The variances in the other grid cells of stratum 4 were not considerably different from each other. Because the concentrations are generally low at 12 a.m., and because the population tends to be indoors at this hour, this particular decrease in precision may not significantly affect population exposure estimates based on these estimates of ambient air quality.

Method C has been shown to be superior to the other sampling methods proposed in this report in several respects:

- The assumptions proposed in Section C.2 are not required for this method.
- The "exposure districts" can be defined after the dispersion model runs have been completed.
- The method produces unbiased estimates with higher precision based on smaller sample sizes. This results in faster dispersion model runs.

Although the proposed method worked very well for 1,3-butadiene in the Phoenix study area, it is possible that the corresponding estimates for other pollutants of interest may not exhibit the same degree of statistical precision. By estimating the variance of the estimator, researchers can obtain information regarding the precision of the estimator. One quick way to check that the sampling objectives are being met is to form an interval by taking the estimator plus and minus two standard errors (the square root of the variance). If the width of this interval is less than 20 percent of the estimate, then researchers can be somewhat assured that the objectives are being met. If the precision requirements are not being met, larger sample sizes may be necessary.

In evaluating the results of applying Method C to 1,3-butadiene in the Phoenix study area, analysts noted that the butadiene emissions in Phoenix were dominated by area sources. To determine whether Method C would perform well for a pollutant emitted largely by point sources, the method was subsequently applied to benzene emissions in Houston, an area where almost half of the total benzene emissions derive from point sources. Section C.6 summarizes the results of this application.

## C.6 APPLICATION OF SAMPLING METHOD C TO HOUSTON BENZENE CONCENTRATIONS

Section C.5 of this appendix describes sampling Method C, a stratified random sampling method for obtaining receptor points for use in the ISCST3 dispersion model. The results presented in Section C.5 demonstrate that this method works adequately for estimating the mean butadiene concentrations at the census block group centroids in a number of pre-defined exposure districts in Phoenix. However, the butadiene emissions in Phoenix were dominated by area sources, and the question remains as to how well Method C would work for a pollutant which is produced largely by point sources. In this section, the quality of the estimates obtained from sampling Method C is examined for Houston benzene concentrations, where almost half of the total emissions derive from point sources.

Section C.2 of this appendix describes a method for defining sampling strata to be used as exposure districts. However, the results of the analyses described in Section C.5 suggest that the exposure districts can actually be defined after the sample has been drawn and the dispersion model runs have been completed. This approach would enable researchers performing population exposure assessments to define the exposure districts to best suit their needs. Analysts "test" this claim in this section using two sets of sampling strata; the first set will be defined using the emissions values, as in Section C.2, and the second set will be defined as exposure districts used in the HAPEM model might be defined.

As in Section C.5, the sampling methods will be applied to the annual average benzene concentrations and to specific hourly/seasonal combinations. The same combinations of two seasons (winter and summer) and three times of day (7 a.m., 12 p.m., and 12 a.m.) that were used in Section C.5 for 1,3-butadiene in Phoenix will be considered here. However, complete data were not available for the winter - 12 a.m. combination, so that particular combination will not be considered here.

Unlike the results presented in Section C.5, the method of Monte Carlo sampling will not be used in this section. Instead, exact variances of the sample mean will be calculated using Equation 6 in Section C.2. These variances will then be used, together with the normal probability distribution, to estimate the probability that the sample mean will lie within 10 percent of the true mean for each stratum.

### C.6.1 Houston Study Area

The Houston study area consists of all census block groups in a rectangular region surrounding Houston, Texas, according to the 1990 U.S. census. The UTM coordinates of the corners of this rectangle are listed below:

<u>Corner</u>	<u>UTM Zone</u>	<u>UTM (east)</u>	<u>UTM (north)</u>
Northeast	15	316 km	3342 km
Southeast	15	316 km	3266 km
Southwest	15	214 km	3266 km
Northwest	15	214 km	3342 km

The study area contains approximately 7,752 square kilometers of land area and 2265 BGs. The 2265 BGs were assumed to constitute the sampling frame for the study area (i.e., the total population of available sampling units). The target population for this study is the current population of Houston. (See Section C.2 for an explanation of why survey sampling methods were chosen over spatial sampling methods).

#### C.6.2 Defining the Strata for Method C

For convenience, analysts utilized the emissions inventory grid to divide the study area, consistent with the approach used for Phoenix. However, it should be noted that Method C does not require the use of the emissions inventory grid. Unlike the grid used in Phoenix, which was comprised of 4 km by 4 km cells, the emissions inventory grid for Houston is comprised of 2 km by 2 km cells. This would require a sample size of 1330 receptor points under Method C, which would result in unsatisfactory dispersion model run times. Since the 4 km by 4 km grid cells worked well in Phoenix, analysts chose to aggregate the 2 by 2 grid cells into 4 by 4 grid cells for Houston, which reduced the sample size required for Method C to 573 receptor points. Note that the grid cell size being used here is in no sense "optimal", but it does work adequately for the data analyzed thus far. It is not clear how an "optimal" grid cell size could be determined without the use of trial-and-error methods, and would undoubtedly depend on the data at hand.

The results of previous analyses suggest that Method C will perform well even when the sampling strata are defined after the dispersion model runs are completed. The only restriction on the strata is that they must be comprised of aggregations of the original grid cells. As a "test" of this assertion, analysts created two sets of sampling strata. The first set was created following the methods outlined in Section C.2 using the emissions values for each grid cell. Breakpoint values of 500 and 2000 were subjectively chosen to define these strata. Because the block group centroids were more uniformly located over the Houston study area than they were in Phoenix, analysts determined that it was not necessary to first divide the region into outer and central regions. Figure C.6-1 presents a contour map of the Phoenix region indicating the cell emission totals together with the 4 km by 4 km grid cells for the Houston study area. Figure C.6-2 shows the 17 strata which were created using the contour map as a guide, as in Section C.2. These strata shall hereafter be referred to as the emissions-defined strata.

The other set of strata were defined in a manner similar to the method commonly used to define exposure districts for HAPEM. These strata were created by aggregating grid cells, and are thus formed by concentric squares rather than the "bulls-eye" pattern of concentric circles which is typically used in HAPEM. Figure C.6-3 presents the 17 strata which were formed in this manner, and which will hereafter be referred to as the HAPEM-like strata.

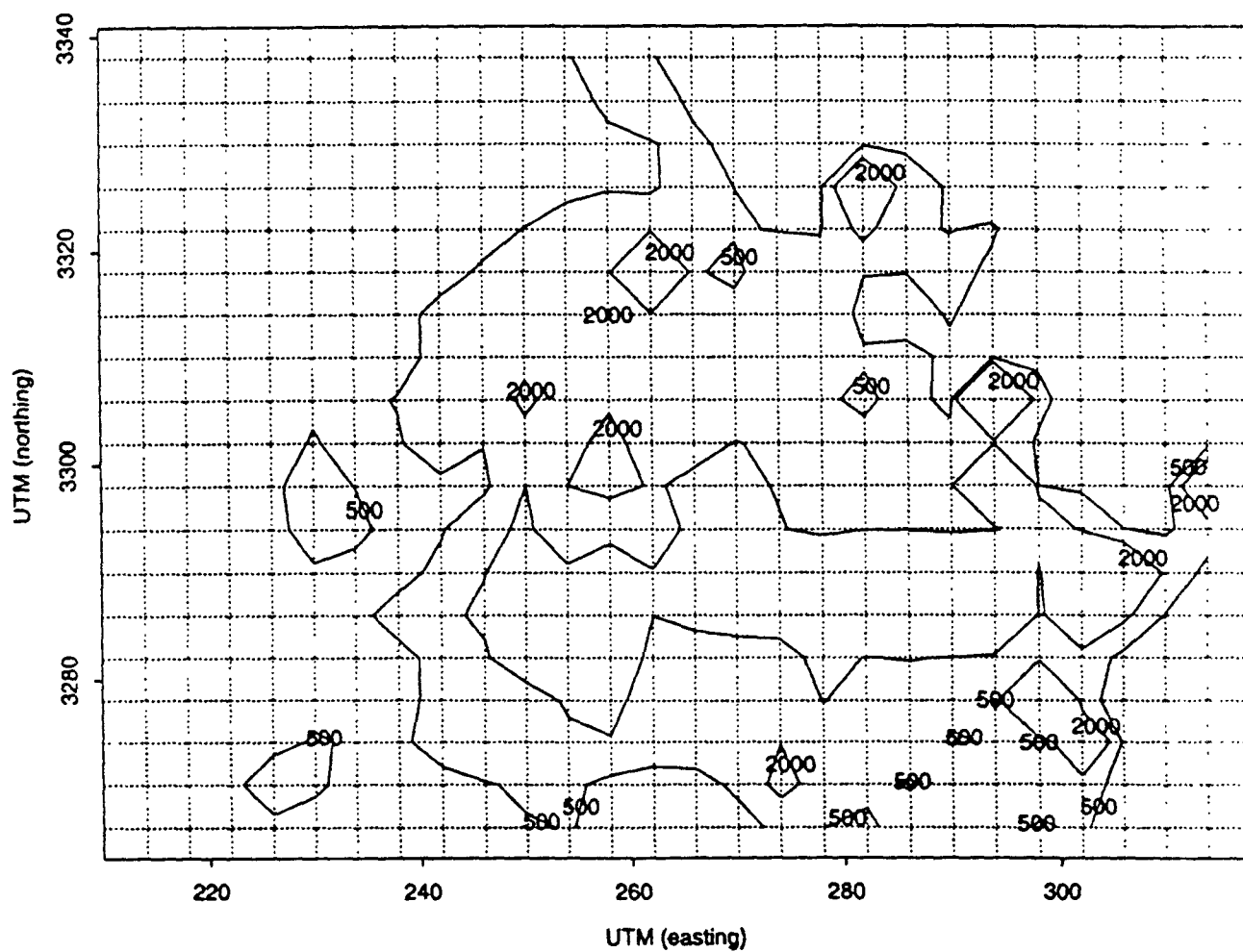


Figure C.6-1 Benzene emissions contour map of the Phoenix region.

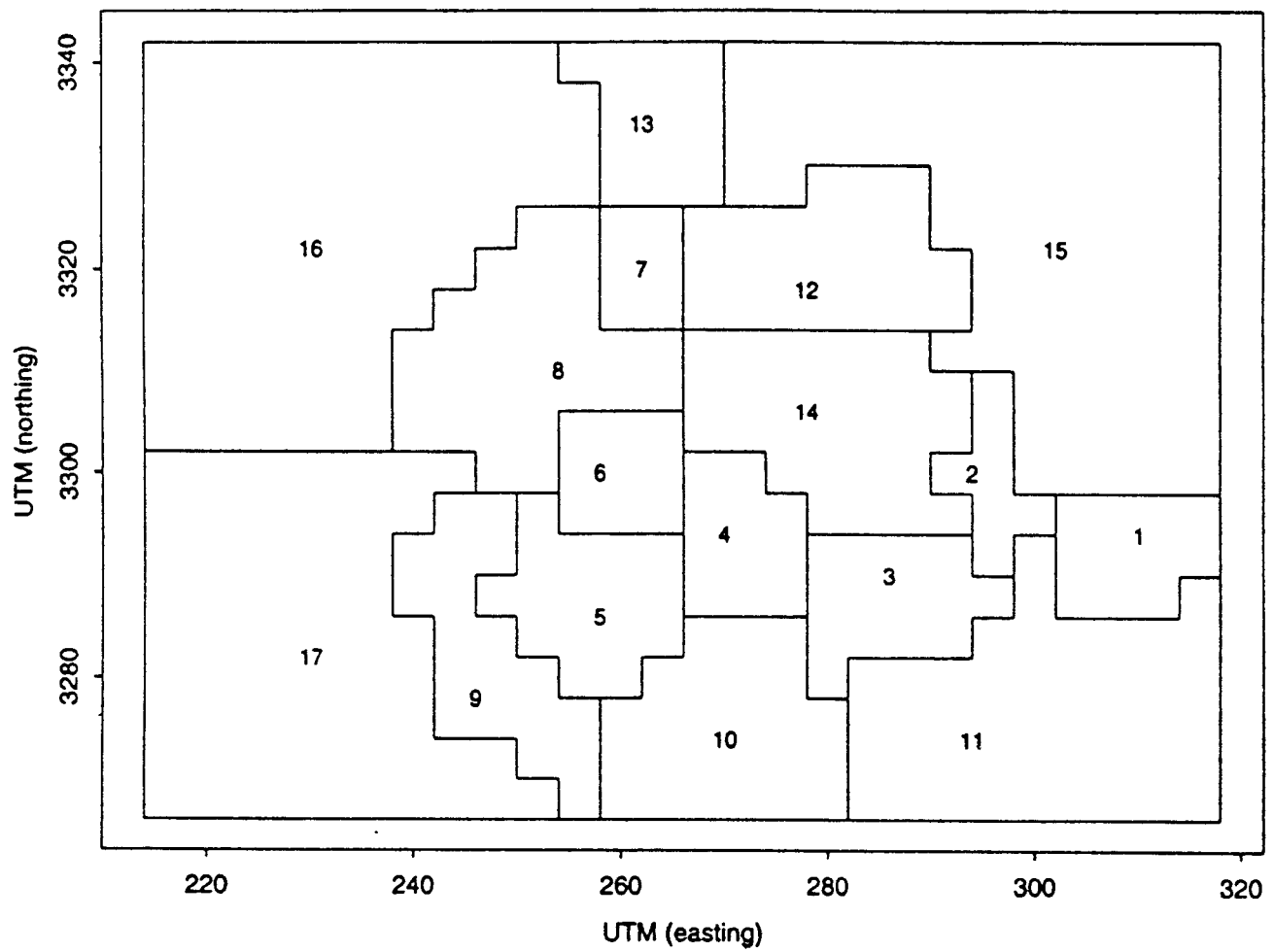


Figure C.6-2 Emissions-defined strata based on Phoenix contour map presented in Figure C.6-1.

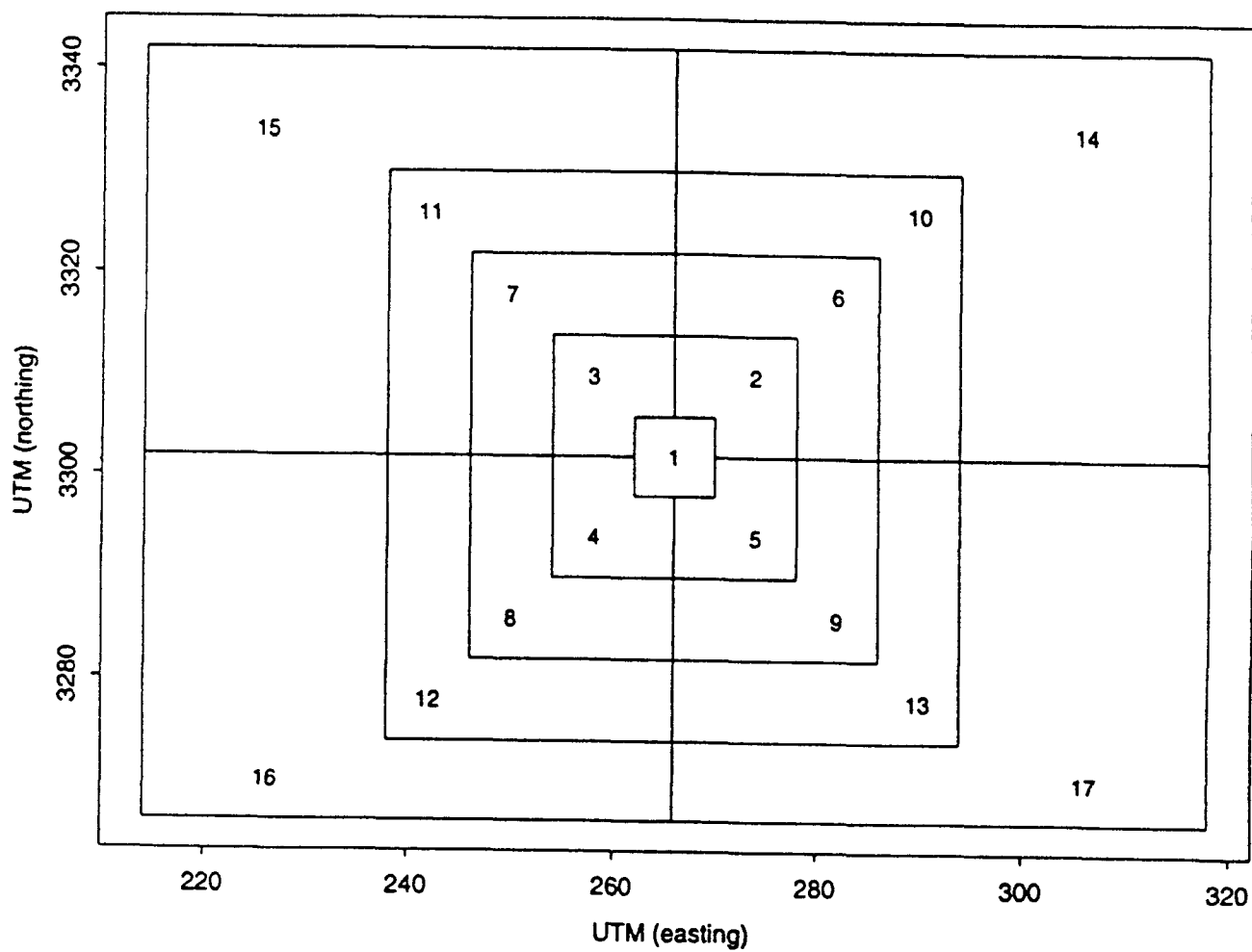


Figure C.6-3 HAPEM-like strata defined for Phoenix region.



Figures C.6-4 and C.6-5 show the distribution of the BG centroids over the emissions-defined strata and the HAPEM-like strata, respectively.

### C.6.3 True Sampling Distributions

In Section C.5, analysts used Monte Carlo sampling to determine the distributional properties of the sample mean for each stratum. This time-consuming process is unnecessary, however, since the distributional properties of estimators from stratified random samples are well known (Thompson, 1992). The estimators are unbiased (i.e., the mean of the estimator is equal to the value being estimated). Furthermore, the true benzene concentrations are available for all of the block groups in Houston. Consequently, the true variances are known. Therefore, the true variances (or standard errors) of the sample means can be calculated using Equation 6 from Section C.2. Assuming that the true distributions of the sample means are at least approximately normal, analysts can then estimate the probability that the sample mean will be within 10 percent of the true mean for each stratum.

### C.6.4 Method C Applied to Annual Average Data with Emissions-Defined Strata

Table C.6-1 contains the descriptive statistics and sampling distributions for the annual average benzene data using the emissions-defined strata. As noted previously, using Method C with 4 km by 4 km grid cells requires 573 receptor points for Houston. The annual average stratum means range from 0.19  $\mu\text{g}/\text{m}^3$  in Stratum 16 to 1.28  $\mu\text{g}/\text{m}^3$  in Stratum 2. The concentration standard deviations range from 0.04 in Strata 16 and 17 to 2.65 in Stratum 3; the standard deviations in strata 2 and 3 are each more than two times higher than that of any other stratum. For most of the strata, the probability is approximately 1.0 that the sample mean will be within 10 percent of the true mean. In fact, this probability is less than 95 percent only for strata 1 and 3. The problems encountered in these two strata will be discussed next.

For the annual average data, two block groups have concentrations which stand out as outliers. The first grid cell is located (UTM coordinates 297 east, 3300 north) in stratum 2 of the emissions-defined strata. The annual average concentration at this point is 10.31  $\mu\text{g}/\text{m}^3$ , while the mean of stratum 2 is only 1.28  $\mu\text{g}/\text{m}^3$  and the next highest concentration in this stratum is 1.75  $\mu\text{g}/\text{m}^3$ . However, this block group is in a grid cell by itself and would be included as a receptor point in every sample. Therefore, this receptor point does not contribute to the variance of the sample mean, and does not cause a problem. However, if the grid over the study area were to be shifted, this may no longer be true, and this particular outlier would then pose a precision problem. In this case, the discussion which follows for the other outlier would apply here as well.

The other outlier is located at UTM coordinates 2834 east, 3289 north in stratum 3 of the emissions-defined strata. The annual average concentration at this point is 36.14  $\mu\text{g}/\text{m}^3$ , while the next highest concentration is 3.30  $\mu\text{g}/\text{m}^3$  and the stratum mean is 1.15  $\mu\text{g}/\text{m}^3$ . This block group is not located in a grid cell by itself, and therefore does contribute heavily to the variance of the sample mean. This block group does cause a problem with the precision of the sample

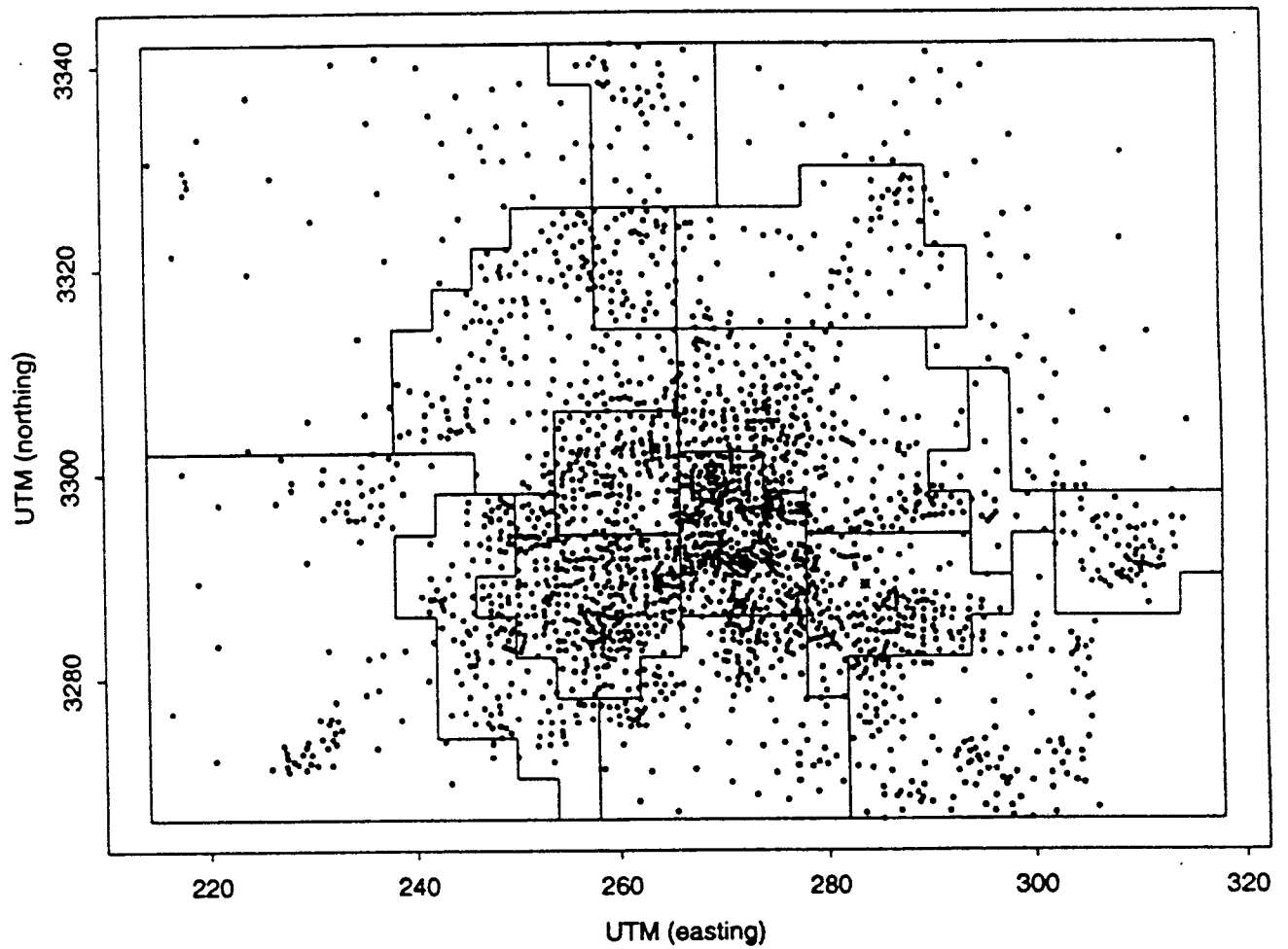


Figure C.6-4 Distribution of block-group centroids over emissions-defined strata.

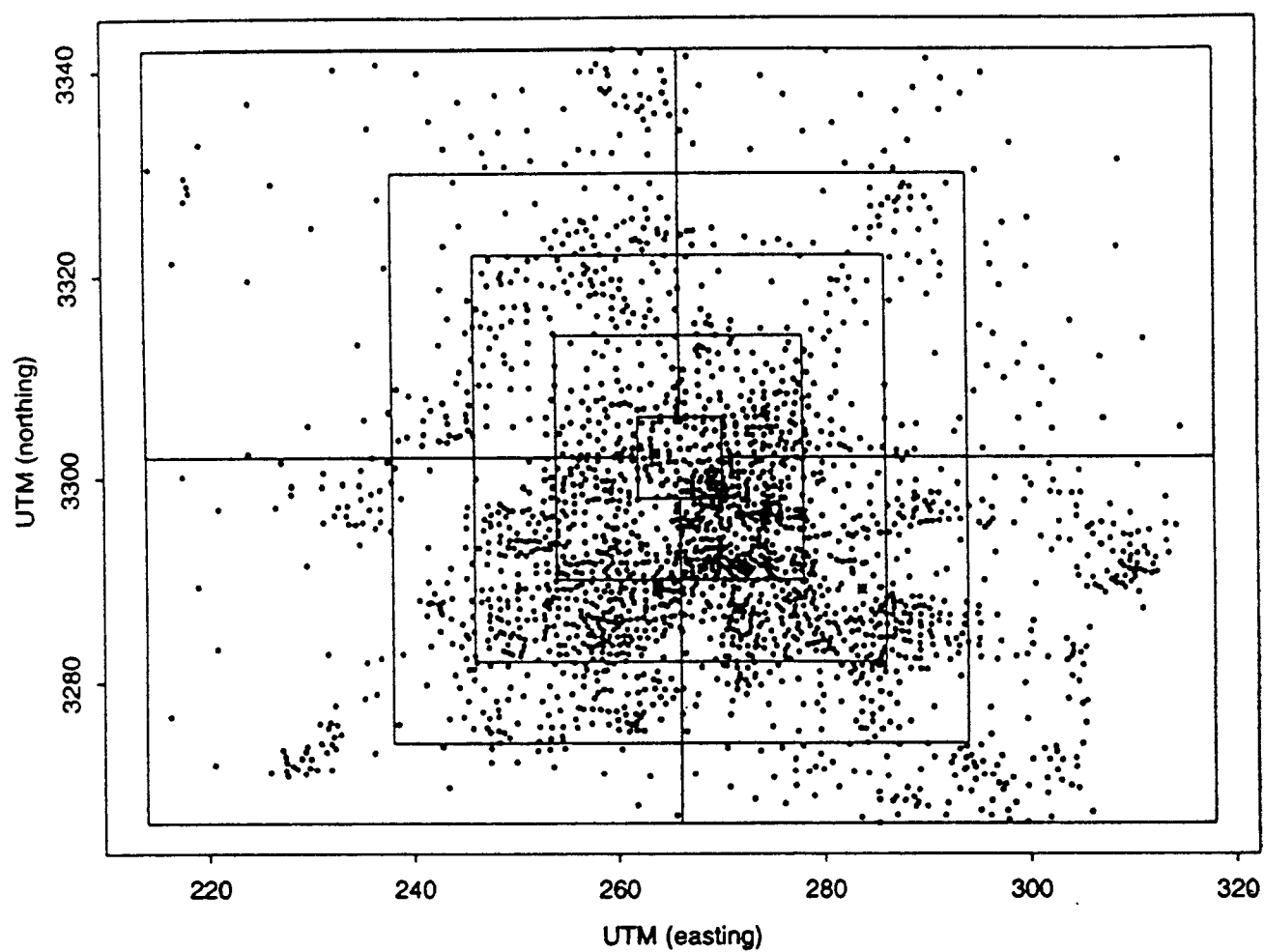


Figure C.6-5 Distribution of block-group centroids over HAPEM-like strata.

**Table C.6-1 Sampling Results for Annual Average Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration, $\mu\text{g}/\text{m}^3$	True standard deviation of concentrations, $\mu\text{g}/\text{m}^3$	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.46	0.42	18	0.0449	0.7000
2	20	1.28	2.10	12	0.0130	1.0000
3	179	1.15	2.65	28	0.4377	0.2079
4	356	0.66	0.09	22	0.0103	1.0000
5	325	0.52	0.07	32	0.0094	1.0000
6	146	0.48	0.05	18	0.0056	1.0000
7	52	0.51	0.11	12	0.0257	0.9524
8	160	0.40	0.08	55	0.0046	1.0000
9	123	0.36	0.06	34	0.0049	1.0000
10	143	0.44	0.10	43	0.0081	1.0000
11	148	0.43	0.20	52	0.0136	0.9984
12	70	0.36	0.11	38	0.0101	0.9996
13	39	0.28	0.05	22	0.0050	1.0000
14	256	0.57	0.12	48	0.0062	1.0000
15	62	0.22	0.09	54	0.0025	1.0000
16	45	0.19	0.04	43	0.0003	1.0000
17	79	0.20	0.04	42	0.0025	1.0000
Total	2265	0.54	0.81	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

mean, as is obvious in the results for Stratum 3 in Table C.6-1. The estimated probability that the sample mean will be within 10 percent of the true mean is only 0.21 in this stratum. One possible remedy for this problem will be discussed in the subsection "Forced Receptors - Certainty Units" below.

The lack of precision observed in Stratum 1 was not caused by any one outlying block group concentration. Instead, it was caused by the relatively high concentration variance within one particular grid cell (UTM coordinates 302 east, 3290 north). The only obvious pre-sampling solution for this problem would be to use smaller grid cells, which would, of course, result in larger sample sizes. The objective that the mean concentration should be estimated to within 10 percent with 95 percent confidence was not met in Stratum 1. Instead, the mean concentration can be estimated to within approximately 19 percent with 95 percent confidence.

#### C.6.5 Method C Applied to the Hourly/Seasonal Data with Emissions-Defined Strata

Tables C.6-2 through C.6-6 contain descriptive statistics and sampling distributions for the concentrations from the five different hourly/seasonal combinations for the emissions-defined strata. Analysts noted the following points concerning the descriptive statistics in these five tables:

- The winter and summer benzene concentration patterns were very similar, which was not the case for butadiene in Phoenix.
- The 7 a.m. and 12 a.m. benzene concentrations tended to be high relative to the 12 p.m. concentrations. The 12 a.m. concentrations were very similar to the 7 a.m. concentrations, an unexpected result.
- The 7 a.m. winter concentrations ranged from 0.15  $\mu\text{g}/\text{m}^3$  in stratum 16 to 1.34  $\mu\text{g}/\text{m}^3$  in stratum 2. The 7 a.m. summer concentrations ranged from 0.14  $\mu\text{g}/\text{m}^3$  in stratum 16 to 1.78  $\mu\text{g}/\text{m}^3$  in stratum 3. The 12 p.m. winter concentrations ranged from 0.18  $\mu\text{g}/\text{m}^3$  in stratum 15 to 0.78  $\mu\text{g}/\text{m}^3$  in stratum 2. The 12 p.m. summer concentrations ranged from 0.11  $\mu\text{g}/\text{m}^3$  in stratum 17 to 0.58  $\mu\text{g}/\text{m}^3$  in stratum 2. The 12 a.m. summer concentrations ranged from 0.13  $\mu\text{g}/\text{m}^3$  in stratum 16 to 2.48  $\mu\text{g}/\text{m}^3$  in stratum 2.
- The concentration standard deviations for all strata were similar for all time periods.

In the majority of the strata, the probability that the sample mean will be within 10 percent of the true stratum mean is approximately 1.0 for all hourly/seasonal combinations. There are a few situations where this probability is less than 95 percent. However, as will be shown below, in most of these cases the sampling method performed reasonably well.

**Table C.6-2 Sampling Results for Winter 7 Am Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.47	0.39	18	0.0448	0.7069
2	20	1.34	2.14	12	0.0211	1.0000
3	179	1.23	1.65	28	0.2685	0.3537
4	356	0.77	0.11	22	0.0127	1.0000
5	325	0.70	0.10	32	0.0127	1.0000
6	146	0.56	0.05	18	0.0075	1.0000
7	52	0.52	0.15	12	0.0341	0.8702
8	160	0.40	0.10	55	0.0060	1.0000
9	123	0.48	0.08	34	0.0066	1.0000
10	143	0.61	0.13	43	0.0123	1.0000
11	148	0.55	0.19	52	0.0140	0.9999
12	70	0.39	0.12	38	0.0138	0.9954
13	39	0.30	0.06	22	0.0039	1.0000
14	256	0.58	0.13	48	0.0083	1.0000
15	62	0.26	0.09	54	0.0028	1.0000
16	45	0.15	0.05	43	0.0005	1.0000
17	79	0.26	0.05	42	0.0030	1.0000
Total	2265	0.62	0.58	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-3 Sampling Results for Winter 12 Pm Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.31	0.29	18	0.0345	0.6297
2	20	0.78	1.20	12	0.0083	1.0000
3	179	0.75	1.41	28	0.2339	0.2517
4	356	0.63	0.07	22	0.0072	1.0000
5	325	0.53	0.07	32	0.0091	1.0000
6	146	0.53	0.04	18	0.0054	1.0000
7	52	0.50	0.07	12	0.0166	0.9972
8	160	0.45	0.07	55	0.0041	1.0000
9	123	0.37	0.08	34	0.0054	1.0000
10	143	0.39	0.06	43	0.0055	1.0000
11	148	0.38	0.12	52	0.0069	1.0000
12	70	0.36	0.11	38	0.0091	0.9999
13	39	0.33	0.06	22	0.0062	1.0000
14	256	0.53	0.07	48	0.0057	1.0000
15	62	0.18	0.05	54	0.0013	1.0000
16	45	0.26	0.05	43	0.0003	1.0000
17	79	0.21	0.06	42	0.0025	1.0000
Total	2265	0.49	0.44	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-4 Sampling Results for Summer 7 Am Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.87	0.52	18	0.0657	0.8162
2	20	1.73	3.15	12	0.0235	1.0000
3	179	1.78	1.71	28	0.2657	0.4973
4	356	0.91	0.15	22	0.0188	1.0000
5	325	0.86	0.12	32	0.0158	1.0000
6	146	0.61	0.05	18	0.0074	1.0000
7	52	0.60	0.21	12	0.0393	0.8731
8	160	0.47	0.13	55	0.0085	1.0000
9	123	0.67	0.14	34	0.0079	1.0000
10	143	0.99	0.21	43	0.0208	1.0000
11	148	0.70	0.25	52	0.0171	1.0000
12	70	0.44	0.18	38	0.0189	0.9788
13	39	0.25	0.06	22	0.0065	0.9999
14	256	0.66	0.17	48	0.0105	1.0000
15	62	0.27	0.12	54	0.0029	1.0000
16	45	0.14	0.05	43	0.0005	1.0000
17	79	0.38	0.13	42	0.0050	1.0000
Total	2265	0.79	0.70	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.



**Table C.6-5 Sampling Results for Summer 12 Pm Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.32	0.31	18	0.0398	0.5727
2	20	0.58	0.83	12	0.0073	1.0000
3	179	0.49	1.28	28	0.2142	0.1792
4	356	0.39	0.05	22	0.0057	1.0000
5	325	0.31	0.06	32	0.0076	1.0000
6	146	0.30	0.04	18	0.0045	1.0000
7	52	0.29	0.04	12	0.0092	0.9987
8	160	0.25	0.05	55	0.0033	1.0000
9	123	0.19	0.05	34	0.0040	1.0000
10	143	0.22	0.06	43	0.0055	0.9999
11	148	0.20	0.11	52	0.0086	0.9809
12	70	0.26	0.07	38	0.0070	0.9998
13	39	0.21	0.04	22	0.0040	1.0000
14	256	0.34	0.05	48	0.0037	1.0000
15	62	0.17	0.04	54	0.0009	1.0000
16	45	0.13	0.03	43	0.0003	1.0000
17	79	0.11	0.03	42	0.0020	1.0000
Total	2265	0.30	0.39	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-6 Sampling Results for Summer 12 Am Houston Benzene Concentrations  
Using the Emissions-Defined Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True variance of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	62	0.83	0.55	18	0.0773	0.7149
2	20	2.48	2.50	12	0.0549	1.0000
3	179	1.66	5.49	28	0.9144	0.1438
4	356	0.73	0.14	22	0.0174	1.0000
5	325	0.56	0.09	32	0.0116	1.0000
6	146	0.51	0.05	18	0.0081	1.0000
7	52	0.72	0.34	12	0.0725	0.6764
8	160	0.39	0.11	55	0.0062	1.0000
9	123	0.35	0.07	34	0.0059	1.0000
10	143	0.47	0.15	43	0.0156	0.9972
11	148	0.47	0.73	52	0.0633	0.5386
12	70	0.52	0.12	38	0.0111	1.0000
13	39	0.36	0.09	22	0.0055	1.0000
14	256	0.82	0.40	48	0.0195	1.0000
15	62	0.57	0.27	54	0.0095	1.0000
16	45	0.13	0.06	43	0.0004	1.0000
17	79	0.18	0.05	42	0.0030	1.0000
Total	2265	0.67	1.63	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

In all cases, the sample mean from stratum 3 has very low precision. The same block group which had an unusually large annual average concentration has a relatively large concentration at every hourly/seasonal combination. It is exclusively this outlier which causes the precision in stratum 3 to be so low; this point will be explored further in a later subsection.

The results from Tables C.6-2 through C.6-6 indicate that the sample mean from stratum 1 fails to meet the precision objective in all cases. The cause of this lack of precision is the same as was discussed previously for the annual average data. The sample mean will not be within 10 percent of the true stratum 1 mean at least 95 percent of the time for any of these hourly/seasonal combinations. However, in all of these cases, the sample mean will be within 25 percent of the true mean at least 95 percent of the time.

The sample mean from stratum 7 fails to meet the precision objectives for the winter 7 a.m., summer 7 a.m., and summer 12 a.m. data sets. In all cases, however, the sample mean from stratum 7 will be within 20 percent of the true mean at least 95 percent of the time. The sample mean from stratum 11 for the summer 12 a.m. data will be within approximately 27 percent of the true mean approximately 95 percent of the time.

#### C.6.6 Results Using HAPEM-Like Strata

The set of HAPEM-like strata, shown in Figure C.6-3, are being used to demonstrate the flexibility available in sampling Method C. It was claimed in Section C.5 that Method C would produce precise estimates, no matter how the grid cells were aggregated into strata (i.e., exposure districts). The results in this section are being presented in support of this claim.

Tables C.6-7 through C.6-12 contain descriptive statistics and sampling distributions for the annual average concentrations and the five different hourly/seasonal combinations for the HAPEM-like strata. In almost every case, the sample mean will be within 10 percent of the true stratum mean concentration at least 95 percent of the time (this probability is actually close to 1.0 in most cases). The outlying block group concentration described in the previous subsection is now located in stratum 9. The sample mean from this stratum has low precision for all data sets; this point will be explored further in the following subsection. The only other case where the level of precision is lower than prescribed is in stratum 17 for the summer 12 a.m. data. Here, the sample mean is within 10 percent of the true mean with approximately 85 percent confidence.

#### C.6.7 Forced Receptors - Certainty Units

In the previous subsections, one very extreme outlier was consistently noted in the Houston benzene concentrations. The UTM coordinates for the block group centroid where these high concentrations were observed are 2834 east, 3289 north. This point is marked with an asterisk in Figures C.6-4 and C.6-5. Table C.6-13 contains the benzene concentrations observed at this centroid in each of the data sets. The stratum means for both types of strata are also included in this table for comparative purposes. The concentration observed at this one block group is always between 13 and 45 times higher than the mean concentration of the entire stratum! There are no other block groups within 1 km of this block group. There are 5 block

**Table C.6-7 Sampling Results for Annual Average Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.55	0.05	8	0.0091	1.0000
2	108	0.52	0.05	16	0.0084	1.0000
3	72	0.48	0.05	16	0.0060	1.0000
4	148	0.50	0.07	16	0.0094	1.0000
5	277	0.68	0.09	16	0.0133	1.0000
6	62	0.44	0.08	27	0.0059	1.0000
7	94	0.43	0.10	32	0.0113	0.9999
8	283	0.50	0.07	31	0.0098	1.0000
9	255	0.91	2.24	32	0.3069	0.2322
10	55	0.34	0.12	30	0.0119	0.9961
11	76	0.36	0.10	31	0.0130	0.9949
12	141	0.37	0.08	41	0.0062	1.0000
13	185	0.69	0.26	46	0.0208	0.9991
14	53	0.23	0.21	51	0.0005	1.0000
15	67	0.23	0.07	49	0.0029	1.0000
16	82	0.20	0.05	44	0.0025	1.0000
17	217	0.48	0.74	87	0.0156	0.9980
Total	2265	0.54	0.81	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-8 Sampling Results for Winter 7 am Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.56	0.07	8	0.0158	0.9996
2	108	0.53	0.06	16	0.0108	1.0000
3	72	0.50	0.07	16	0.0095	1.0000
4	148	0.62	0.08	16	0.0128	1.0000
5	277	0.75	0.10	16	0.0145	1.0000
6	62	0.45	0.09	27	0.0074	1.0000
7	94	0.43	0.14	32	0.0165	0.9915
8	283	0.68	0.12	31	0.0132	1.0000
9	255	1.01	1.40	32	0.1879	0.4095
10	55	0.39	0.13	30	0.0169	0.9780
11	76	0.36	0.11	31	0.0145	0.9865
12	141	0.51	0.11	41	0.0076	1.0000
13	185	0.83	0.26	46	0.0258	0.9988
14	53	0.27	0.20	51	0.0008	1.0000
15	67	0.21	0.09	49	0.0022	1.000
16	82	0.27	0.06	44	0.0030	1.0000
17	217	0.55	0.74	87	0.0151	0.9998
Total	2265	0.62	0.58	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-9 Sampling Results for Winter 12 pm Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.58	0.04	8	0.0103	1.0000
2	108	0.53	0.05	16	0.0088	1.0000
3	72	0.53	0.05	16	0.0056	1.0000
4	148	0.54	0.05	16	0.0083	1.0000
5	277	0.65	0.06	16	0.0085	1.0000
6	62	0.43	0.08	27	0.0059	1.0000
7	94	0.46	0.07	32	0.0079	1.0000
8	283	0.52	0.07	31	0.0098	1.0000
9	255	0.66	1.18	32	0.1640	0.3111
10	55	0.32	0.10	30	0.0105	0.9977
11	76	0.40	0.07	31	0.0082	1.0000
12	141	0.35	0.06	41	0.0056	1.0000
13	185	0.52	0.12	46	0.0120	1.0000
14	53	0.19	0.10	51	0.0006	1.0000
15	67	0.29	0.07	49	0.0036	1.0000
16	82	0.21	0.06	44	0.0024	1.0000
17	217	0.36	0.43	87	0.0107	0.9993
Total	2265	0.49	0.44	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-10 Sampling Results for Summer 7 am Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.63	0.07	8	0.0131	1.0000
2	108	0.61	0.08	16	0.0142	1.0000
3	72	0.59	0.07	16	0.0102	1.0000
4	148	0.71	0.14	16	0.0182	0.9999
5	277	0.90	0.16	16	0.0223	1.0000
6	62	0.49	0.14	27	0.0144	0.9994
7	94	0.53	0.19	32	0.0231	0.9790
8	283	0.81	0.14	31	0.0158	1.0000
9	255	1.33	1.49	32	0.1845	0.5298
10	55	0.41	0.16	30	0.0214	0.9417
11	76	0.36	0.11	31	0.0111	0.9989
12	141	0.77	0.21	41	0.0096	1.0000
13	185	1.25	0.45	46	0.0441	0.9955
14	53	0.27	0.24	51	0.0005	1.0000
15	67	0.18	0.08	49	0.0038	1.0000
16	82	0.42	0.15	44	0.0049	1.0000
17	217	0.79	1.06	87	0.0212	0.9998
Total	2265	0.79	0.70	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-11 Sampling Results for Summer 12 pm Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.34	0.03	8	0.0073	1.0000
2	108	0.33	0.03	16	0.0051	1.0000
3	72	0.30	0.04	16	0.0048	1.0000
4	148	0.32	0.05	16	0.0070	1.0000
5	277	0.40	0.05	16	0.0073	1.0000
6	62	0.29	0.06	27	0.0049	1.0000
7	94	0.26	0.05	32	0.0041	1.0000
8	283	0.30	0.06	31	0.0081	0.9998
9	255	0.41	1.08	32	0.1503	0.2164
10	55	0.24	0.06	30	0.0078	0.9979
11	76	0.23	0.06	31	0.0058	0.9999
12	141	0.19	0.05	41	0.0046	1.0000
13	185	0.31	0.11	46	0.0073	1.0000
14	53	0.18	0.10	51	0.0003	1.0000
15	67	0.16	0.06	49	0.0024	1.0000
16	82	0.11	0.03	44	0.0020	1.0000
17	217	0.25	0.33	87	0.0127	0.9554
Total	2265	0.30	0.39	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.



**Table C.6-12 Sampling Results for Summer 12 am Houston Benzene Concentrations  
Using the HAPEM-Like Strata**

Stratum	Descriptive Statistics			Sampling Distribution for Method C		
	Number of BGs in stratum, N	True mean concentration ( $\mu\text{g}/\text{m}^3$ )	True standard deviation of concentrations ( $\mu\text{g}/\text{m}^3$ )	Sample size, n	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
1	90	0.58	0.05	8	0.0118	1.0000
2	108	0.58	0.07	16	0.0089	1.0000
3	72	0.51	0.06	16	0.0071	1.0000
4	148	0.56	0.10	16	0.0115	1.0000
5	277	0.77	0.13	16	0.0222	0.9995
6	62	0.64	0.16	27	0.0218	0.9968
7	94	0.46	0.19	32	0.0228	0.9582
8	283	0.53	0.09	31	0.0125	1.0000
9	255	1.27	4.62	32	0.6416	0.1572
10	55	0.57	0.21	30	0.0140	1.0000
11	76	0.45	0.35	31	0.0421	0.7177
12	141	0.36	0.10	41	0.0068	1.0000
13	185	0.94	0.58	46	0.0393	0.9832
14	53	0.64	0.95	51	0.0012	1.0000
15	67	0.22	0.12	49	0.0030	1.0000
16	82	0.18	0.05	44	0.0030	1.0000
17	217	0.70	1.08	87	0.0484	0.8530
Total	2265	0.67	1.63	573	-	-

<sup>a</sup> Assuming approximate normality of the sampling distribution. Values listed as 1.0000 have been rounded to four decimal places.

**Table C.6-13 The Sampling Distribution for Stratum 3 of the Emissions-Defined Strata and Stratum 9 of the HAPEM-Like Strata after the Outlying Block Group at UTM Coordinates 283.735 East, 3289.09 North Is Forced into the Sample**

Data set	Concentration in outlying BG, $\mu\text{g}/\text{m}^3$	Emissions-Defined Stratum 3 Sampling Distribution			HAPEM-Like Stratum 9 Sampling Distribution		
		Stratum Mean Concentration ( $\mu\text{g}/\text{m}^3$ )	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>	Stratum Mean Concentration ( $\mu\text{g}/\text{m}^3$ )	True standard error of sample mean	Probability sample mean is within 10% of true mean <sup>a</sup>
Annual Average	36.14	1.15	0.0450	0.9897	0.91	0.0285	0.9986
Winter, 7 am	22.69	1.23	0.0499	0.9864	1.01	0.0322	0.9983
Winter, 12 pm	19.32	0.75	0.0268	0.9949	0.66	0.0177	0.9998
Summer, 7 am	23.04	1.78	0.0821	0.9699	1.33	0.0510	0.9910
Summer, 12 pm	17.47	0.49	0.0118	1.0000	0.41	0.0078	0.9979
Summer, 12 am	73.93	1.66	0.0894	0.9363	1.27	0.0597	0.9669

<sup>a</sup> Assuming approximate normality of the sampling distribution.

groups within 2 km of this block group, and none of them has an annual average concentration higher than  $3.3 \mu\text{g}/\text{m}^3$ .

Because this outlying block group is not in a grid cell by itself, it has great influence on the variance of the sample mean. Recall that the variance of the sample mean (Equation 6) depends only on the within grid cell variances. The presence of an outlier of this magnitude in any grid cell will substantially increase the variance within that grid cell, which will then inflate the true variance of the sample mean.

Unfortunately, there is not a simple method that would help to alleviate the problems caused by such an outlier. One way to eliminate the influence from this one point would be to force it into the sample. In survey sampling terminology, any sampling unit which is included in the sample automatically is known as a "certainty unit." Certainty units do not contribute to the variance of the sample mean because they are included in every sample, and are not selected randomly. In effect, under Method C, any block group which is located in a cell either alone or with only one other block group is a certainty unit. Therefore, treating an outlier as a certainty unit would be equivalent to treating it as if it were contained within its own grid cell. Of course, treating this outlier as a certainty unit would require researchers to have the foresight to know, before drawing a sample, that this point would probably have extremely high concentrations. While this is probably not a realistic suggestion in general, it may have been possible for researchers to predict before hand that this particular point would have high concentrations; there are 108 point sources of benzene within 1 km of the outlying block group. Identifying this type of situation before hand would mean looking at the location of every block group in the frame with respect to its proximity to point sources, which would be a very time consuming operation. Furthermore, it is not recommended that researchers include points as certainty units without good justification.

Forcing the outlying block group into the sample would eliminate the precision problems which were observed in stratum 3 of the emissions-defined strata and stratum 9 of the HAPEM-like strata. The sampling distribution obtained from these strata after this outlier is forced into the sample are shown in Table C.6-13. In all cases, the standard errors of the sample mean are substantially reduced. Forcing this block group into the sample would increase the sample size in each of these strata by 1. Thus, the results in Table C.6-13 are for sample sizes of 29 and 33 from stratum 3 and stratum 9, respectively.

#### C.6.8 Summary Discussion

The results of this section indicate that sampling Method C works reasonably well, even for an air pollutant which is produced largely by point sources. In most cases, the method produces estimates which adequately satisfy the sampling objectives prescribed in Section C.1. In fact, in most cases the estimates are actually "too precise", which indicates that larger grid cells and smaller sample sizes could be used. However, "optimizing" the size of the grid cell would have to be accomplished by trial-and-error methods for each pollutant, and it is not clear how this could be done accomplished to obtaining estimates of the grid cell variances. For the two pollutants analyzed to date (1,3-butadiene and benzene), precise estimators could have been

obtained by choosing only one receptor point per grid cell. This would have substantially reduced the sample size. However, this sample size reduction would make it impossible to obtain an unbiased estimate for the standard error of the sample mean. It is necessary for researchers to have such an estimate in order to have an indication of the precision of their estimates.

## C.7 STEP-BY-STEP GUIDELINES FOR USING SAMPLING METHOD C

This section summarizes the step-by-step procedure for drawing samples of receptor points using sampling Method C, as proposed in Section C.5, and then defining exposure districts for use in HAPEM and similar exposure assessment models. Steps 1 through 3 below define the method for obtaining the sample of receptor points. Exposure districts are created in Step 4, and estimates of the mean and variance of the ambient concentration across each district are calculated in Step 5. Note that Steps 4 and 5 can be performed after the sample of receptor points has been selected.

1. *Define the study area and obtain a listing of all census block group centroids within the study area.* The user will need to have the UTM coordinates of all of the block group centroids within the study area.
2. *Define a grid over the entire study area.* In this report, analysts used a regular grid of 4 km by 4 km cells, created using the emissions inventory grid. This grid does not need to be regular, however, and the cell width of 4 km is not necessarily the "optimal" grid size.
3. *Randomly choose 2 block group centroids from each of the populated grid cells. If there is only one centroid in a grid cell, it is included in the sample automatically.* This step can be performed prior to defining the "exposure districts" (see the next step). Recall that a "populated" grid cell was defined as one which contains at least one block group centroid. Table C.7-1 contains example SAS code for performing Steps 2 and 3.
4. *Define the "exposure districts".* This step can be performed using the emissions values in the same manner as the sampling strata were defined in Section C.2, or it can be accomplished using the actual concentration values after the sampling has been completed and the dispersion model has been run. Consequently, Method C provides the researchers performing the exposure assessments with greater flexibility in defining the exposure districts. The individual grid cells can be aggregated in any way desirable to form the exposure districts, given that the final exposure districts are formed simply by aggregating grid cells.
5. *Obtain estimates for mean concentrations in each exposure district.* This step can be accomplished using Equation 5. In addition, unbiased estimates of the variances of these sample means should be obtained using Equation 6.

This approach is independent of the pollutant and is applicable to all cities of interest. The same set of receptors could be used for each pollutant within a given city.

**Table C.7-1 Example SAS Code<sup>a</sup> for Generating a Sample of Receptor Points  
Using Method C**

```

DATA ONE;
  INFILE 'BGS.DAT';          *File containing block group information;
  INPUT EAST NORTH;          *May need to define additional variables;
  U = RANUNI(-1);            *Generates uniform (0,1) random number;

*** Creates 4 km by 4 km grid over area with southwest corner at UTM coordinates 214 east, 3266 north, as in
Houston. Assigns each bg to a grid cell;

  X = FLOOR((EAST - 214)/4);
  Y = FLOOR((WEST - 3266)/4);

  ID = X*100 + Y;            *Creates pseudo id for each grid cell;

PROC SORT;
  BY ID U;

***Now, two data sets will be created. The first will contain the block group with the smallest random number
from each grid cell. This block group will be included as a receptor point. The second data set will contain all
the rest of the block groups. Then, the block group with the next highest random number is selected, and the two
data sets are merged together into the final data set;

DATA TWO THREE;
  SET ONE;
  BY ID;
  IF (FIRST.ID) THEN OUTPUT TWO;
  ELSE OUTPUT THREE;

DATA FOUR;
  SET THREE;
  BY ID;
  IF (FIRST.ID);

DATA FIVE;                  *The final data set containing selected receptors;
  SET TWO FOUR;

```

<sup>a</sup>It is assumed that the user has a file ('BGS.DAT') which contains the listing of all block group centroids in the study area. This file should contain at least the UTM coordinates of the block group centroids.

To illustrate the procedures to be used in making these calculations, we have constructed a hypothetical example in which analysts used the sample SAS code in Table 7-1 to implement Steps 1 through 3 with respect to Study Area X. The effort produced a grid of 4 km by 4 km cells over the study area and then randomly selected a maximum of two BGs from each grid cell. In Step 4, the analysts defined 20 exposure districts as contiguous collections of grid cells based on a review of the emissions data.

Exposure District No. 1 was defined as the aggregation of 10 populated grid cells (identified as  $h = 1$  through 10 in Table C.7-2). Table C.7-2 lists values for

$N_h$  = total number of BGs in cell  $h$ , and  
 $n_h$  = number of BGs selected from cell  $h$ .

For each grid cell,  $N_h > 2$ . Consequently, analysts were able to randomly select two BGs from each cell (i.e.,  $n_h = 2$ ).

Step 5 was the only step in the methodology which required statistical calculations. Each selected BG was associated with a concentration value obtained from the dispersion model run. The mean ( $\bar{y}_h$ ) and standard deviation ( $\sigma_h$ ) statistics listed in Table C.7-2 for each cell were calculated from these values using the standard formulas. The mean concentration for the entire exposure district ( $\bar{y}$ ) was then calculated by the expression

$$\bar{y} = \frac{1}{N} \sum_{h=1}^L N_h \bar{y}_h ,$$

in which

$L$  = total number of grid cells in the exposure district = 10, and  
 $N$  = total number of BGs in the exposure district = 171.

The variance of  $\bar{y}$  was calculated by the expression

$$Var(\bar{y}) = \sum_{h=1}^L \left[ \frac{N_h}{N} \right]^2 \left[ \frac{N_h - n_h}{N_h} \right] \frac{\sigma_h^2}{n_h} ,$$

in which all terms have been previously defined. Making the indicated substitutions from Table 7-2, the analysts obtained  $\bar{y} = 0.596$  and  $var(\bar{y}) = 0.003$  for Exposure District No. 1. The same procedure was then applied to each of the other exposure districts in the hypothetical study area.

**Table C.7-2 Sample Inputs for Step 5 Calculations  
(Exposure District 1 of Hypothetical Study Area)**

Grid cell ID (h)	Total number of BGs in grid cell ( $N_h$ )	Number of selected BGs in grid cell ( $n_h$ )	Mean of concentrations estimated for selected BGs ( $\bar{y}_h$ ), $\mu\text{g}/\text{m}^3$	Standard deviation of concentrations estimated for selected BGs ( $\sigma_h$ ), $\mu\text{g}/\text{m}^3$
1	16	2	0.59	0.20
2	18	2	0.40	0.24
3	15	2	0.32	0.10
4	21	2	0.39	0.35
5	14	2	0.62	0.29
6	17	2	0.63	0.29
7	18	2	0.85	0.28
8	16	2	0.95	0.16
9	19	2	0.60	0.22
10	17	2	0.65	0.16



## C.8 REFERENCES

Johnson, T., Warnasch, J., McCoy, M., Capel, J., and Riley, M. 1996. Developmental Research for the Hazardous Air Pollutant Exposure Model (HAPEM) as Applied to Mobile Source Pollutants. IT Air Quality Services, Durham, North Carolina. February.

Thompson, S. K. 1992. Sampling. Wiley, New York.

## **APPENDIX D**

### **DEFAULT STACK PARAMETERS**

**OBTAINED FROM OZONE TRANSPORT ASSESSMENT GROUP**

**FOR SUBSTITUTION OF MISSING DATA**

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
101001	332	6.23	423	52.83
101002	483	15.91	299	63.79
101003	410	19.05	271	79.75
101004	252	12.18	343	53.60
101005	308	13.78	394	66.36
101006	185	10.14	348	41.27
101007	134	8.80	413	55.69
101008	398	15.69	329	67.72
101009	182	5.43	365	55.83
101010	357	15.55	315	36.25
101012	212	8.24	337	62.75
101013	342	13.58	333	66.38
102001	151	7.74	356	29.45
102002	165	7.08	387	35.18
102003	217	11.63	283	35.62
102004	95	5.18	413	30.13
102005	74	3.96	395	29.80
102006	68	4.27	396	27.27
102007	134	6.94	480	31.78
102008	194	8.13	361	34.50
102009	102	4.76	380	34.54
102010	67	3.49	361	23.55
102011	80	6.56	208	28.31
102012	176	6.72	328	43.60
102013	125	6.01	430	30.78
102014	171	8.03	456	43.47
102999	25	1.20	205	0.00
103001	141	6.05	416	18.33
103002	153	6.79	394	25.46
103003	104	4.89	381	30.24
103004	98	5.19	417	25.72
103005	60	4.13	383	30.45
103006	75	3.73	382	23.94
103007	61	3.20	325	19.06
103009	82	3.99	380	49.59
103010	42	1.90	346	34.99
103012	126	6.81	282	36.85
103013	74	5.94	433	27.02
105001	40	2.07	294	27.47
105002	32	1.40	449	32.58
201001	55	7.47	647	69.49
201002	55	7.32	651	67.07
201007	38	3.28	375	31.19
201008	28	3.28	949	89.45
201009	30	12.58	791	152.12

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
201900	125	1.31	131	0.00
202001	32	2.26	552	86.47
202002	33	1.92	729	75.22
202003	16	0.69	414	26.50
202004	30	1.70	635	21.35
202005	151	7.52	374	53.81
202009	40	3.74	689	87.90
202010	31	0.87	674	28.23
203001	30	1.60	786	78.14
203002	28	1.87	706	95.23
203003	57	1.38	700	69.81
203010	26	2.15	393	16.91
204001	50	14.39	291	39.09
204002	0	0.00	0	0.00
204003	67	8.82	449	83.83
204004	37	1.83	348	31.64
288888	52	2.81	334	67.34
301001	91	1.46	270	52.03
301003	101	5.99	405	75.24
301005	78	2.81	464	45.00
301006	60	3.39	480	29.53
301007	26	1.00	82	26.06
301008	51	6.76	156	24.98
301009	52	2.47	117	15.72
301010	43	1.57	153	15.85
301011	36	1.42	100	42.37
301012	80	1.39	214	20.15
301013	82	2.98	346	72.30
301014	28	3.58	102	13.68
301015	43	1.24	149	21.02
301016	82	5.46	120	41.38
301017	62	2.59	154	66.47
301018	48	2.33	180	38.73
301019	85	4.28	251	37.37
301020	34	1.76	127	14.34
301021	97	2.69	117	13.13
301022	38	2.39	171	25.43
301023	140	4.90	145	50.05
301024	65	3.51	178	31.40
301025	136	2.42	96	34.69
301026	48	2.47	158	27.04
301027	89	3.15	139	47.72
301028	106	3.55	118	56.52
301029	102	4.97	118	38.30
301030	107	5.84	128	56.58

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
301031	97	0.96	151	32.26
301032	162	3.83	719	36.45
301033	67	1.59	229	34.23
301034	49	2.43	259	12.60
301035	93	2.35	146	28.60
301038	112	2.06	234	66.21
301039	66	1.46	642	1.97
301040	110	7.78	147	70.06
301041	87	1.00	203	26.24
301042	132	3.28	146	44.86
301045	36	0.00	109	0.00
301060	50	1.85	124	15.71
301066	0	0.00	0	0.00
301070	42	2.41	139	7.61
301091	54	1.44	294	30.32
301099	0	0.00	0	0.00
301100	49	3.54	272	39.83
301112	104	5.12	580	44.02
301120	46	3.24	178	37.31
301121	52	2.61	150	27.38
301124	100	0.00	77	0.00
301125	46	2.60	129	31.43
301126	34	0.81	107	0.00
301127	74	1.92	283	44.62
301130	63	1.42	75	53.05
301132	49	6.01	274	4.97
301133	40	0.77	552	22.44
301137	34	0.90	89	9.04
301140	54	22.39	425	12.02
301152	210	3.50	0	0.00
301153	42	1.66	567	2.01
301156	43	1.29	95	1.53
301157	45	34.10	309	0.03
301158	34	0.34	130	103.27
301167	50	3.31	176	16.02
301169	45	0.85	131	0.94
301174	53	10.14	344	34.47
301176	0	0.00	0	0.00
301181	34	0.87	122	18.16
301190	72	6.69	438	7.48
301195	0	0.00	0	0.00
301197	74	10.78	460	27.69
301202	48	1.64	122	24.48
301205	43	1.82	116	54.92
301206	52	13.58	322	8.75

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
301210	27	2.76	146	0.93
301211	85	84.06	400	0.14
301250	36	2.51	189	16.98
301251	45	4.10	235	27.67
301252	55	0.98	201	48.36
301253	38	0.33	104	0.95
301254	73	6.74	253	22.41
301258	48	4.15	187	6.41
301301	31	1.86	93	11.79
301303	51	2.50	130	48.35
301304	70	3.40	126	55.49
301305	30	0.79	141	63.70
301800	27	2.64	115	3.65
301810	34	4.51	107	24.52
301820	15	2.98	82	9.22
301830	30	2.61	153	0.49
301840	42	2.11	189	11.66
301870	24	1.62	107	1.76
301875	29	3.02	100	0.26
301885	31	2.16	287	10.53
301888	26	2.21	104	28.37
301900	94	7.89	790	26.13
301999	45	2.04	184	23.07
302001	39	3.70	157	35.74
302002	85	2.43	690	44.81
302003	160	3.75	210	33.14
302004	43	1.58	68	21.00
302005	74	2.94	89	44.59
302006	36	3.84	79	38.50
302007	66	2.95	113	44.02
302008	63	2.43	86	39.16
302009	87	2.42	130	21.48
302010	161	4.18	87	25.84
302012	133	5.93	151	182.54
302013	67	3.04	156	21.79
302014	52	1.58	109	46.45
302015	84	5.52	251	43.72
302016	79	6.32	213	53.64
302017	69	0.00	350	0.00
302018	40	3.12	263	24.32
302019	35	1.61	101	48.20
302022	39	2.24	68	21.83
302026	38	3.42	71	9.96
302030	63	2.69	156	58.96
302031	72	7.00	80	22.16

**Appendix D.        DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
302032	43	1.42	318	26.01
302033	59	3.68	117	40.71
302036	23	1.55	235	23.51
302038	0	0.00	0	0.00
302040	58	2.04	251	53.47
302888	104	2.00	117	27.72
302900	68	3.05	222	36.26
302999	46	1.92	132	19.60
303000	49	1.66	139	17.37
303001	56	4.53	140	47.85
303002	144	4.23	294	59.51
303003	204	8.18	298	26.15
303005	261	9.25	184	27.53
303006	79	9.49	299	98.70
303007	95	16.46	142	0.00
303008	150	7.40	459	28.31
303009	119	9.52	455	24.17
303010	132	8.32	89	35.21
303012	134	2.00	0	0.00
303014	80	3.23	206	0.00
303023	112	4.71	107	54.06
303024	33	3.40	72	112.22
303030	97	4.11	135	44.84
303888	74	7.66	112	60.93
303900	155	6.43	640	51.65
303999	37	2.63	181	34.51
304001	50	3.33	429	30.51
304002	52	3.54	367	40.51
304003	49	3.93	169	53.49
304004	69	2.45	208	39.88
304005	29	2.09	164	73.15
304006	50	6.00	97	45.58
304007	47	4.46	214	38.61
304008	41	2.67	281	30.84
304009	25	1.08	538	8.60
304010	71	2.50	151	11.58
304020	57	8.24	258	30.99
304022	32	2.17	193	50.06
304049	0	0.00	0	0.00
304050	29	0.86	183	24.75
304888	36	3.48	155	34.15
304900	66	3.43	593	33.37
304999	55	2.54	294	19.89
305001	37	3.45	268	44.07
305002	34	4.44	230	50.71

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
305003	41	2.72	329	42.46
305004	91	5.59	403	26.88
305005	44	2.87	244	19.69
305006	86	3.99	167	47.86
305007	105	5.05	167	39.50
305008	38	2.03	150	42.55
305009	51	3.00	160	52.08
305010	72	5.27	98	50.47
305011	31	3.13	81	25.09
305012	58	3.77	231	46.46
305013	85	2.61	143	32.14
305014	90	4.77	399	45.81
305015	56	2.65	216	39.15
305016	68	5.12	214	39.41
305017	54	6.03	186	42.61
305018	38	4.00	403	66.58
305019	60	3.95	105	37.34
305020	29	3.45	86	55.33
305021	36	1.35	80	0.00
305025	41	2.16	107	57.35
305026	32	0.95	73	6.52
305030	29	1.83	120	38.11
305032	49	1.70	70	52.87
305033	0	0.00	77	0.00
305040	34	4.94	146	40.74
305100	103	1.88	91	54.62
305101	52	2.43	78	56.01
305102	72	1.85	82	26.41
305103	38	4.00	77	3.28
305104	34	1.83	101	43.73
305105	58	1.74	79	60.05
305150	41	1.05	176	49.40
305888	26	1.89	117	13.98
305900	58	4.29	384	48.86
305999	57	2.36	197	48.52
306001	107	5.28	577	22.39
306002	129	6.50	428	60.46
306003	94	6.96	480	59.65
306004	112	11.00	705	28.48
306005	15	11.33	151	18.64
306006	49	10.93	224	58.06
306007	34	11.97	115	10.98
306008	35	13.16	198	24.46
306009	142	14.49	1168	22.95
306010	61	4.23	235	15.59



**Appendix D.        DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
306011	49	3.08	755	30.07
306012	91	5.50	305	35.04
306014	113	10.41	787	41.03
306016	199	2.50	1400	40.74
306099	189	3.67	872	58.46
306100	24	2.53	142	7.30
306888	28	52.11	149	5.36
306999	31	8.79	255	22.19
307001	174	5.54	204	50.59
307002	105	2.63	190	45.76
307003	150	3.60	138	42.44
307004	53	5.24	138	10.80
307005	19	4.06	120	9.60
307007	56	3.64	174	48.15
307008	58	2.89	148	34.81
307011	48	1.07	200	0.04
307013	42	6.67	145	27.24
307020	34	3.39	87	34.18
307030	41	2.91	101	45.28
307888	33	3.04	144	29.66
307900	97	4.45	231	40.02
307999	40	3.68	115	24.10
308001	41	2.23	92	40.78
308005	60	1.28	74	44.45
308006	36	2.23	127	32.85
308007	34	2.16	116	37.04
308008	31	2.07	110	54.51
308009	291	2.04	76	54.12
308010	35	1.70	124	45.27
308900	35	1.67	287	21.12
308999	37	2.00	122	34.29
309001	28	2.52	183	27.57
309002	27	2.09	86	36.32
309003	35	2.11	127	6.07
309006	0	0.00	0	0.00
309010	32	2.77	110	35.18
309011	33	2.13	116	41.93
309015	33	2.92	102	30.52
309016	36	2.47	90	19.21
309020	29	3.50	142	14.36
309025	35	2.40	468	36.22
309030	30	2.23	71	14.23
309040	24	1.72	97	58.45
309060	29	2.73	77	39.47
309888	39	1.42	122	27.12

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
309900	39	2.26	318	27.52
309999	36	2.00	154	28.88
310001	92	4.71	241	11.22
310002	40	4.38	632	49.50
310004	30	1.87	587	18.23
310888	9	7.37	98	13.67
311001	20	0.00	77	0.00
312999	33	2.27	111	16.64
313010	30	2.00	77	63.66
313030	68	1.41	86	28.95
313065	40	3.05	251	22.28
313070	32	1.83	723	23.00
313900	28	0.95	313	11.37
313999	35	3.17	99	33.40
314009	20	2.20	85	28.50
314010	18	1.50	575	37.73
314011	40	3.75	93	31.13
314015	0	0.00	0	0.00
314999	36	4.65	99	27.02
315010	45	1.50	70	3.77
315020	47	3.49	163	10.32
320999	27	2.29	97	31.97
330001	40	2.80	248	51.66
330002	35	3.11	193	48.85
330003	18	3.56	77	2.99
330004	40	3.28	209	72.58
330005	0	1.00	95	0.00
330888	29	3.00	108	29.81
360001	40	2.00	600	63.66
385001	27	16.29	83	25.67
390001	65	7.00	1850	1.30
390002	110	8.86	302	41.91
390004	116	6.26	440	32.77
390005	69	5.19	339	53.89
390006	62	4.10	373	36.95
390007	159	7.29	585	13.57
390008	95	5.77	280	51.88
390009	79	3.96	194	41.09
390010	48	4.01	309	39.50
390012	140	7.12	354	36.82
390013	124	7.34	342	28.19
399900	74	3.38	327	48.06
399999	41	4.35	162	18.83
401001	26	2.39	125	23.83
401002	30	2.42	92	29.21

**Appendix D.        DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
401003	56	2.72	101	26.57
401004	0	0.00	0	0.00
401005	45	1.30	120	44.36
401010	0	0.00	0	0.00
401888	38	2.47	124	31.75
401999	10	0.00	70	0.00
402001	38	3.97	125	26.08
402002	37	2.82	178	29.93
402003	32	2.23	123	35.92
402004	33	2.58	92	28.05
402005	34	2.77	116	31.51
402006	39	2.88	112	33.41
402007	38	2.25	163	34.25
402008	40	2.26	254	25.71
402009	40	2.59	133	37.70
402010	61	2.78	187	34.50
402011	38	3.12	199	34.22
402012	32	2.67	138	1.46
402013	43	2.89	277	42.80
402014	56	3.78	124	28.19
402015	34	1.33	420	36.81
402016	85	3.66	119	38.29
402017	45	2.58	299	29.41
402018	45	2.81	476	50.86
402019	33	3.61	75	44.20
402020	34	3.17	106	33.32
402021	27	2.62	88	54.83
402022	46	2.54	88	40.76
402023	32	4.96	74	65.05
402024	36	3.74	78	31.18
402025	40	2.92	116	36.42
402026	45	2.02	153	16.99
402099	0	0.00	0	0.00
402888	35	3.08	116	26.30
402900	55	4.06	370	37.78
402999	38	2.94	112	35.31
403001	29	3.55	92	1.57
403002	41	2.97	73	0.01
403003	0	0.00	77	0.00
403010	30	4.10	96	1.10
403011	43	9.87	82	0.51
403012	59	6.92	355	4.62
403888	22	4.17	108	2.42
403999	34	6.01	76	6.06
404001	29	3.50	97	0.84

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
404002	25	3.08	121	2.15
404003	54	1.53	95	1.21
404004	20	3.19	87	20.27
405001	37	2.32	327	39.88
405002	32	2.41	244	40.76
405003	29	2.13	178	40.56
405004	36	2.43	297	38.93
405005	38	3.15	198	42.54
405006	37	3.45	237	36.76
405007	28	2.45	73	13.63
405008	18	0.69	75	0.00
405888	26	1.73	106	44.20
406001	20	2.27	123	10.98
406002	21	11.24	92	4.75
406003	12	0.32	66	0.10
406004	12	0.31	68	0.00
406888	15	2.91	119	9.97
407004	26	1.46	131	0.01
407008	27	1.36	107	29.97
407016	27	3.40	90	1.16
407020	47	8.86	389	0.15
407032	28	1.94	124	0.07
407036	29	5.27	115	1.55
407040	27	2.06	116	0.12
407044	32	1.94	266	1.76
407048	44	5.95	238	0.07
407052	27	2.74	83	0.41
407056	28	2.88	89	1.69
407060	34	4.09	113	5.91
407064	31	0.75	79	0.01
407068	31	2.18	111	0.03
407076	35	3.27	108	0.08
407080	36	2.56	93	0.08
407084	25	2.48	135	0.22
407158	0	0.00	0	0.00
407172	34	2.47	131	0.13
407176	31	3.30	78	0.01
407180	35	22.35	72	0.02
407208	33	2.52	73	0.09
407220	37	1.78	73	11.52
407228	29	3.04	85	0.39
407232	22	0.00	77	0.00
407816	45	11.39	504	10.90
407820	54	7.22	556	8.15
407832	18	1.28	112	0.01

**Appendix D.            DEFAULT STACK PARAMETERS OBTAINED FROM OZONE TRANSPORT  
ASSESSMENT GROUP FOR SUBSTITUTION OF MISSING DATA**

SCC	Stack Height (ft)	Stack Diameter (ft)	Exit Temperature (°F)	Exit Velocity (ft/s)
407848	8	0.36	81	0.00
407860	34	0.57	87	0.00
407864	16	0.17	100	0.00
407872	16	0.34	216	40.53
407999	28	3.40	169	0.98
408999	20	3.73	213	3.95
490001	70	2.61	148	26.00
490002	33	1.89	120	38.82
490003	24	19.37	451	9.43
490004	11	2.63	81	0.04
490005	33	2.13	170	19.57
490900	52	9.72	1232	31.76
490999	34	2.58	122	26.35
501001	167	5.79	462	50.91
501002	0	0.00	0	0.00
501004	0	0.00	69	0.00
501005	93	3.98	395	37.28
501006	3	0.00	77	0.00
501007	18	2.67	111	0.24
501900	79	4.62	487	17.05
502001	62	2.87	760	31.25
502002	19	5.11	1317	12.98
502003	83	3.28	174	0.00
502005	64	2.91	762	39.01
502006	25	0.70	96	0.00
502900	64	4.02	418	49.08
503001	58	3.33	652	32.29
503002	51	2.78	439	25.03
503005	91	4.19	526	35.02
503006	38	3.24	593	63.71
503007	48	2.57	238	9.21
503008	19	2.40	81	28.15
503900	75	3.36	811	24.05
625400	86	4.83	177	38.70

## **APPENDIX E**

### **PARAMETERS RELATING TO THE FATES OF SELECT ATMOSPHERIC POLLUTANTS**

## APPENDIX E - TABLE OF CONTENTS

E. PARAMETERS RELATING TO THE FATES OF SELECT ATMOSPHERIC	
POLLUTANTS .....	E-1
E.1 GAS-PHASE ORGANIC REACTIONS .....	E-1
E.1.1 Photolysis of Formaldehyde and Acetaldehyde .....	E-2
E.1.2 Summary of Atmospheric Half-lives .....	E-2
E.2 REACTIONS OF POLYCYCLIC AROMATIC HYDROCARBON (PAH) ....	E-4
E.3 WET AND DRY DEPOSITION .....	E-5
E.4 GAS-PARTICLE PARTITIONING .....	E-7
E.5 REFERENCES .....	E-14

## APPENDIX E - LIST OF TABLES

Table E.1 Parameters Used in Determining the Atmospheric Fate of Hazardous Gas Phase Pollutants .....	E-8
Table E.2 Average 24 Hour OH, NO <sub>3</sub> , and O <sub>3</sub> Concentrations in a Moderately Polluted Atmosphere .....	E-11
Table E.3 Estimation of Photo-induced Decay Rate Constants for PAHs Under Summer and Winter Conditions .....	E-11
Table E.4 Comparison of the Cumulative Deposition of Particles During 100 Seconds by Diffusion and Gravitational Settling .....	E-11
Table E.5 Particle Phase Size Distributions of High Molecular Weight PAH at 298K .....	E-12
Table E.6 Particle Phase Size Distributions of Trace Metals at 298K, 1atm Pressure .....	E-12
Table E.7 Predicted Partitioning of PAHs Based on Equation 21 Model .....	E-12
Table E.8 Predicted Partitioning of Dioxins and Dibenzofurans Based on Equation 22 Model .....	E-13
Table E.9 Partitioning Potential of Other Hazardous Air Pollutants .....	E-13



## APPENDIX E

### PARAMETERS RELATING TO THE FATES OF SELECT ATMOSPHERIC POLLUTANTS

This appendix contains a summary of a literature survey conducted by Fletcher, et al. (1997). The purpose of this literature survey was to provide best estimates of some parameters used in the determination of the fates of selected hazardous air pollutants (HAPs). These data would then serve as inputs to EPA's Industrial Source Complex Short Term Model (ISCST3). An estimate of the atmospheric half-lives for gas-phase organic compounds reacting with hydroxyl radicals (OH), the nitrate radical (NO<sub>3</sub>), and ozone (O<sub>3</sub>) is presented first. The photolysis of formaldehyde and acetaldehyde is then discussed. This is followed by a discussion of polycyclic aromatic hydrocarbon (PAH) half-lives as developed from reactions in sunlight. Parameters used in calculating wet and dry deposition velocities, namely particle density and size, molecular diffusivities in air and water, and Henry's law constants are then presented. Finally, particle-gas phase distribution for semi-volatile contaminants (PAHs and trace metals such as chromium) is discussed.

#### E.1 GAS-PHASE ORGANIC REACTIONS

Second order reaction rate constants at 298K for 22 volatile organic compounds from reactions with OH, NO<sub>3</sub>, O<sub>3</sub> are summarized in Table E.1. References used to generate this table are given at the end of this report. Temperature dependent OH rate constants exist for a number of the compounds in Table E.1. Generally, reaction rates will show an exponential temperature dependence over a narrow temperature range and can be estimated by the Arrhenius equation (Atkins, 1990)

$$k = A * e^{-E_a/RT} \quad (\text{Eq. 1})$$

where  $k$  is the rate of reaction,  $E_a$  is the activation energy for the reaction (J mol<sup>-1</sup>),  $R$  is the gas constant (8.31 J K<sup>-1</sup> mol<sup>-1</sup>), and  $T$  is the temperature of the system in Kelvin. The coefficient  $A$  is a preexponential factor that is independent of, or only slightly dependent on, the temperature of the system. If the temperature dependence of the preexponential factor is not small, typically with small activation energies, a different form relating temperature to the reaction rate needs to be employed (Finlayson-Pitts, 1986)

$$k = B T^n e^{-E_a/RT} \quad (\text{Eq. 2})$$

In this form of the Arrhenius equation the preexponential factor is separated into a temperature-independent constant,  $B$ , and a temperature-dependent constant,  $T^n$ , where the exponent  $n$  is a number chosen to best fit the data and  $T$  is the temperature in Kelvin. Table E.1 lists temperature-dependent rate equations for the reaction of several hazardous contaminants with the

OH radical. Although temperature dependent equations of the rate constant for reactions with the NO<sub>3</sub> radical and O<sub>3</sub> exist, they are not provided in this report.

#### E.1.1 Photolysis of Formaldehyde and Acetaldehyde

Rate constants for the photolysis of formaldehyde and acetaldehyde were estimated by integrating the product of the absorption cross section,  $\sigma(\lambda)$ , the quantum yield,  $\phi(\lambda)$ , and the actinic flux,  $J(\lambda)$ , at each wavelength,  $\lambda$  (Finlayson-Pitts, 1986)

$$k_{phot} = \int_{\lambda=290nm} \sigma(\lambda) \phi(\lambda) J(\lambda) d\lambda \quad (\text{Eq. 3})$$

The photolysis rate constants listed in Table E.1 were calculated at 40° latitude for wintertime (January 1) and summertime (June 1) conditions for three times of the day: 9 am, 12 noon, and 3 pm. These values are based on absorption cross section, quantum yield, and the actinic flux data available in the literature (Atkinson, 1989; Atkinson, 1997; and Finlayson-Pitts, 1986).

#### E.1.2 Summary of Atmospheric Half-lives

From the rate constant data in Table E.1, half-lives were computed for each pollutant. For pollutants, here referred to as [A], that react with oxidants OH, NO<sub>3</sub> or O<sub>3</sub>, the half-life ( $t_{1/2}$ ) can be calculated by (Atkins, 1990)

$$t_{1/2} = \frac{\ln(2)}{k_A [B]} \quad (\text{Eq. 4})$$

where  $k_A$  is the second order rate constant. For the gas-phase reactions of interest in this report, [B] is the concentration of the oxidizing species, and fluctuates according to temperature and level of contamination in the atmosphere by the oxidant. Average concentrations of oxidants, [B], in a relatively polluted atmosphere are given in Table E.2. For a first order reaction such as photolysis, the reaction rate depends only on the concentration of A. The equation to calculate the half-life of a first order reaction is simply (Atkins, 1990)

$$t_{1/2} = \frac{\ln(2)}{k_1} \quad (\text{Eq. 5})$$

Reaction with OH radicals is the primary loss pathway for the majority of the compounds considered here. The NO<sub>3</sub> radical readily photolyzes and will have a relatively low steady-state concentration during daylight hours. Overnight it can be assumed that all OH radicals are reacted, and the steady-state concentration is approximately zero. However, NO<sub>3</sub> concentrations increase overnight, and this becomes the main route of decay of atmospheric contaminants. O<sub>3</sub>

can react with gas-phase compounds during the day and night, however, a closer inspection of the ozone rate constants in Table E.1 reveals that these reactions are, in general, extremely slow and will usually have little effect on the overall half-lives of atmospheric contaminants.

The overall half-lives for selected HAPs were calculated from equation 6 and are listed in Table E.1. It is cautioned that formaldehyde and acetaldehyde are also produced *via* photochemical processes and hence only introducing a decay constant will underestimate

formaldehyde and acetaldehyde concentrations (Finlayson-Pitts, 1986)

$$\frac{1}{t_{1/2, overall}} = \frac{1}{t_{1/2, OH}} + \frac{1}{t_{1/2, NO_3}} + \frac{1}{t_{1/2, O_3}} + \frac{1}{t_{1/2, phot}} \quad (\text{Eq. 6})$$

## E.2 REACTIONS OF POLYCYCLIC AROMATIC HYDROCARBON (PAH)

PAHs on particles react *via* photo-induced processes with gas phase  $O_3$ ,  $NO_2$ , nitric acid ( $HNO_3$ ), and dinitrogen pentoxide ( $N_2O_5$ ), however, reactions in sunlight are the fastest. In the late spring through the early fall months, when the daily average total solar radiation ranges from 0.25 to 0.4 calories  $cm^{-2} min^{-1}$ , photolysis reactions will be more important than  $O_3$ -PAH reactions. During the winter months, with a daily average  $O_3$  concentration of 0.02,  $O_3$  reactions may become more important.

Estimates of the first order rate constants of PAH decay in this report were developed from outdoor chamber studies of PAHs on wood soot particles (Kamens, 1988). From these rate constants half-lives were estimated for summer and winter time conditions and are tabulated in Table 3. In these chamber experiments a reduction in the rate of reaction is typically observed after two half-lives. Although this is not always the case, reducing the rate constants by a factor of 2 after two half-lives of reaction is recommended. Given the current very limited data base, losses due to PAH reaction with  $O_3$  were not included in half-life estimates. Because a reduction in the reaction rates entered in ISCST3 is not possible after two half-lives, ISCST3 users should use the rates and half-lives as listed in Table E.3.

The rates of photolysis of PAHs are highly dependent upon the types of particles onto which the PAH adsorbs. For combustion particles such as wood soot or diesel particles, the half-life of benzo- $\alpha$ -pyrene (BaP) is predicted here. However, if the particle were dark fly ash, the half-life of BaP is on the order of a few days rather than a few hours (Atkinson, 1990). The rate constants and half-lives presented in this report are applicable to atmospheres that have significant wood and diesel combustion particle emissions. We would expect that atmospheres dominated by coal power plants and incinerator emissions would have PAH half-lives which are considerably longer than those presented here (Behymer, 1988; Pennise, 1996; and Wehry, 1990).

### E.3 WET AND DRY DEPOSITION

Wet and dry deposition are important sinks for atmospheric pollutants such as PAHs and trace metals bound to aerosol particles. Wet deposition is generally considered to be a more efficient removal process relative to dry deposition (Steiger, 1989). Key parameters which influence wet and dry deposition in the ISC are particle densities and diameters, Henry's law constants, and molecular diffusivities of selected hazardous air pollutants.

**Particle size.** Particle size is a major factor in determining deposition velocities due to the gravitational settling velocity. For example, although a larger percentage of PAHs are associated with the smaller, high surface area fine particles (generally greater than 75%), the deposition velocities of particle-bound PAHs are overwhelmingly controlled by the gravitational settling velocity of the larger coarse particles (Suman, 1989).

Table E.4 illustrates the relation between particle size and cumulative deposition (Finlayson-Pitts, 1986). Gravitational settling velocity is much more important than diffusion for larger particles. The fine particles remain suspended in the atmosphere, can possibly travel much longer distances, and participate in a large number of atmospheric reactions. In the absence of wet deposition, submicrometer particles ( $<1.0 \mu\text{m}$ ) have atmospheric residence times that vary between 100 - 1000 hours while particles with diameters greater than  $1 \mu\text{m}$  have much shorter residence times on the order of 10-100 hours (Finlayson-Pitts, 1986).

**Henry's law values.** The rate of volatilization of the contaminant is dependent on the value of its Henry's law constant (HLC) and is controlled by its molecular diffusion through air. Henry's law values on the order of  $10^{-7} \text{ atm m}^3 \text{ mol}^{-1}$  indicate that the substance is relatively nonvolatile. Henry's law values between  $10^{-5}$  and  $10^{-3} \text{ atm m}^3 \text{ mol}^{-1}$  indicate that both the gas and aqueous phases play a significant role in determining into which medium the compound will partition. When HLC are relatively high (greater than  $10^{-3} \text{ atm m}^3 \text{ mol}^{-1}$ ) the compound is only slightly soluble (Lyman, 1996). Table E.1 lists Henry's law constants for several atmospheric pollutants.

**Molecular diffusivity.** Molecular diffusion is the net transport of a molecule through a gaseous or liquid medium by Brownian motion. The molecular diffusivity has units of  $\text{cm}^2 \text{ sec}^{-1}$ . A comparison of experimentally determined molecular diffusivities to calculated diffusivities for this report gave deviations of less than 10% in most cases (Schwarzenbach, 1993). Table E.1 lists the molecular diffusivity of select atmospheric pollutants at 298K in both air and water,  $D_{\text{air}}$  and  $D_{\text{w}}$ , respectively, as calculated by simplified equations which only require the molecular mass,  $m$ , to be known (Schwarzenbach, 1993)

$$D_{\text{air}} = \frac{1.55}{m^{0.65}} \quad (\text{Eq. 7})$$

$$D_{\text{w}} = \frac{2.7 \times 10^{-4}}{m^{0.71}} \quad (\text{Eq. 8})$$

**Particle size distributions.** Miguel and Friedlander (1978) measured the concentration of high molecular weight PAHs (specifically BaP and coronene) adsorbed onto several different sizes of particles for two temperatures. Table E.5 quantifies the high molecular weight PAH concentration to particle size distribution.

**Trace metal distribution.** Relatively little work has been published concerning the partitioning of trace metals, or more specifically the partitioning of chromium, onto particles. Whitby (1977) reported approximate total particulate mass size distributions which have been used by modelers to estimate chromium and other trace metal particle size distributions (U.S. EPA, 1997). Such a distribution is given in Table E.6. Despite the lack of data related to chromium particle size distributions, the assumed values used by modelers closely match experimental results of Steiger *et al.* (Steiger, 1989) for the distributions of lead and vanadium in the particle phase.

## E.4 GAS-PARTICLE PARTITIONING

The phase in which a compound exists in the atmosphere is largely dependent on its vapor pressure. Compounds with low volatility, such as large PAHs (6 rings), exist almost exclusively in and on particulate matter (i.e., in the “particle phase”), whereas highly volatile compounds, such as small PAHs (2 rings, i.e., naphthalene), remain mostly in the gas phase. However, semi-volatile compounds, which have ambient vapor pressures of approximately  $10^{-3}$  to  $10^{-7}$  torr, demonstrate significant partitioning between the gas and particle phases. The most straightforward model for predicting the partitioning of PAHs was presented by Yamasaki (1982)

$$\log \frac{C_g}{C_p / TSP} = \log \frac{1}{K_p} = -\frac{A}{T} + B \quad (\text{Eq. 9})$$

where  $C_g$  is the gas-phase concentration ( $\text{ng}/\text{m}^3$ ),  $C_p$  is the particle-phase concentration ( $\text{ng}/\text{m}^3$ ), TSP is the concentration of total suspended particulates ( $\text{ng}/\text{m}^3$ ),  $K_p$  is the equilibrium partition coefficient,  $T$  is the temperature (K), and  $A$  and  $B$  are empirically determined parameters. Table E.9 shows the values of  $A$  and  $B$  determined by Yamasaki for samples taken in Osaka, Japan (17). In order to determine the percent of PAH mass present on particulate matter, a value of  $25 \mu\text{g m}^{-3}$  was used for TSP. This value represents a reasonable urban particle load (U.S. EPA, 1993).

Yamasaki's model can only be used when the values of  $A$ ,  $B$ ,  $K_p$ , and  $T$  are known. Unfortunately the coefficients  $A$  and  $B$  have not been determined for many hazardous pollutants. Therefore, a slightly more complicated model for predicting gas-particle partitioning has been proposed by Pankow and Bidleman (1992)

$$\log K_p = m_r \log p_L^\circ + b_r \quad (\text{Eq 10})$$

where  $p$  is the sub-cooled liquid vapor pressure (torr) for the compound of interest at a given temperature, and  $m_r$  and  $b_r$  are empirical parameters. This model requires the determination of  $p$  values at different temperatures. Various methods for estimating  $p$  values have been proposed (1,100). Values of  $m_r = -1.15$  and  $b_r = -9.70$  have been reported by Pankow and Bidleman (1992) for dioxins and dibenzofurans. The resulting  $K_p$  values, and the percent in the particle phase at 298K and  $TSP = 25 \mu\text{g}/\text{m}^3$  are shown in Table E.7. Values of  $p$  were taken from Eitzer and Hites (1989).

Table E.8 classifies other hazardous air pollutants between those that have the potential to partition between the gas and particle phases and those that exist almost exclusively in the gas phase. This classification is based on the hazardous air pollutants' vapor pressure and a simplified partitioning model (U.S. EPA, 1993).

**Table E.1 Parameters Used in Determining the Atmospheric Fate of  
Hazardous Gas Phase Pollutants**

Compound	$k_{OH}$	temp dependent $k_{OH}$	$k_{NO_3}$	$k_{O_3}$	$k_{phot}$		$t_{1/2}$		diffusiv- ity in air	diffusiv -ity in water	Henry's law constant $\text{Pa m}^3 \text{mol}^{-1}$
					winter	summer	winter	summer			
acetaldehyde	$1.6 \times 10^{-11}$	$5.6 \times 10^{-12} \cdot \exp(310/T)$ $240 < T < 530$	$2 \times 10^{-15}$	$< 10^{-21}$	$2.4 \times 10^{-7}$	$2.4 \times 10^{-7}$	4 d	9 h	0.1325	$1.84 \times 10^{-5}$	7.8
acrolein	$2.0 \times 10^{-11}$	nr	$1.1 \times 10^{-15}$	nr	na	na	4 d	8 h	0.1131	$1.54 \times 10^{-5}$	13.3
acrylamide	nr	nr	nr	nr	na	na	nc	nc	0.097	$1.31 \times 10^{-5}$	$1.4 \times 10^{-4}$
acrylonitrile	$4.8 \times 10^{-12}$	nr	$5.1 \times 10^{-17}$	$< 1 \times 10^{-18}$	na	na	17 d	28 h	0.1173	$1.61 \times 10^{-5}$	10
benzene	$1.2 \times 10^{-12}$	$7.57 \times 10^{-12} \cdot \exp(-529/T)$ $T \leq 325$	$< 3 \times 10^{-17}$	$2 \times 10^{-23}$	na	na	65 d	6 d	0.0912	$1.22 \times 10^{-5}$	543
bis(2-ethylhexyl) phthalate	nr	nr	nr	nr	nr	nr	nc	nc	0.0326	$3.98 \times 10^{-6}$	nr
1,3 - butadiene	$6.7 \times 10^{-11}$	$1.48 \times 10^{-11} \cdot \exp(448/T)$ $250 < T < 425$	$1.0 \times 10^{-13}$	$6.3 \times 10^{-18}$	na	na	8 h	2 h	0.1158	$1.59 \times 10^{-5}$	7180
carbon tetrachloride	$< 5 \times 10^{-16}$	$< 1 \times 10^{-12} \cdot \exp(-2260/T)$ $250 < T < 300$	nr	nr	na	na	440 y	37 y	0.0587	$7.56 \times 10^{-6}$	2600
chloroform	$1.0 \times 10^{-13}$	$3.3 \times 10^{-12} \cdot \exp(-1030/T)$ $340 < T < 300$	$2.6 \times 10^{-17}$	nr	na	na	2 y	64 d	0.0692	$9.05 \times 10^{-6}$	379
1,4-dichlorobenzene	$3.2 \times 10^{-13}$	nr	$< 4 \times 10^{-17}$	nr	na	na	218 d	20 d	0.0605	$7.81 \times 10^{-6}$	322
1,1-dichloroethene	$8.1 \times 10^{-12}$	nr	$1 \times 10^{-15}$	$3.7 \times 10^{-21}$	na	na	9 d	19 h	0.0793	$1.05 \times 10^{-5}$	19250 <sup>a</sup>
1,2-dichloropropane	$\leq 4.4 \times 10^{-14}$	nr	nr	nr	na	na	182 d	15 d	0.0718	$9.41 \times 10^{-6}$	258



**Table E.1 Parameters Used in Determining the Atmospheric Fate of  
Hazardous Gas Phase Pollutants (continued)**

Compound	$k_{OH}$	temp dependent $k_{OH}$	$k_{NO_3}$	$k_{O_3}$	$k_{phot}$		$t_{1/2}$		diffusiv- ity in air	diffusiv- ity in water	Henry's law constant $Pa\ m^{-1}\ mol^{-1}$
					winter	summer	winter	summer			
1,3-dichloroprope ne	nr	nr	nr	nr	na	na	nc	nc	0.0726	$9.53 \times 10^{-6}$	132*
ethyl acrylate	nr	nr	nr	nr	na	na	nc	nc	0.073	$9.59 \times 10^{-6}$	nr
ethylene dibromide	$2.3 \times 10^{-13}$	nr	nr	nr	na	na	349 d	29 d	0.0516	$6.56 \times 10^{-6}$	6902
ethylene dichloride	$2.2 \times 10^{-13}$	nr	nr	nr	na	na	365 d	30 d	0.0782	$1.03 \times 10^{-5}$	120
ethylene oxide	$8 \times 10^{-12}$	nr	nr	nr	na	na	10 d	20 h	0.1323	$1.84 \times 10^{-5}$	nr
formaldehyde	$1.0 \times 10^{-11}$	$8.8 \times 10^{-12} * \exp(25/T)$ $240 < T < 300$	$5.8 \times 10^{-16}$	$\leq 2.1 \times 10^{-20}$	$3.2 \times 10^{-5}$	$1.1 \times 10^{-4}$	6 h	2 h	0.1698	$2.41 \times 10^{-5}$	0.032
hydrazine	$6.5 \times 10^{-11}$	nr	nr	$\sim 3 \times 10^{-17}$	na	na	7 h	2 h	0.1211	$1.67 \times 10^{-5}$	nr
methyl chloride	$4.3 \times 10^{-14}$	$1.8 \times 10^{-12} * \exp(-1115/T)$ $240 < T < 300$	$10^{-17}$	nr	na	na	4 y	148 d	0.0864	$1.15 \times 10^{-5}$	900
quinoline	nr	nr	nr	nr	na	na	nc	nc	0.0756	$9.97 \times 10^{-6}$	nr
styrene	$5 \times 10^{-12}$	nr	$1.5 \times 10^{-13}$	nr	na	na	10 h	5 h	0.0555	$7.11 \times 10^{-6}$	250
1,1,2,2- tetrachloroet hane	nr	nr	nr	nr	nr	nr	nc	nc			45.3
tetrachloroet hylene	$1.7 \times 10^{-13}$	$9.4 \times 10^{-12} * \exp(-1200/T)$ $300 < T < 420$	$< 10^{-17}$	$\leq 10^{-21}$	na	na	1 y	39 d	0.0559	$7.17 \times 10^{-6}$	1518

**Table E.1 Parameters Used in Determining the Atmospheric Fate of  
Hazardous Gas Phase Pollutants (continued)**

Compound	$k_{OH}$	temp dependent $k_{OH}$	$k_{NO_3}$	$k_{O_3}$	$k_{phot}$		$t_{1/2}$		diffusiv- ity in air	diffusiv -ity in water	Henry's law constant $Pa\ m^3\ mol^{-1}$
					winter	summer	winter	summer			
1,1,2-trichloroethane	$3 \times 10^{-13}$	$5.0 \times 10^{-13} \cdot \exp(445/T)$ 230 < T < 420	nr	nr	na	na	267 d	22 d	0.065	$8.46 \times 10^{-6}$	901
trichloroethylene	$2.2 \times 10^{-12}$	$1.63 \pm 0.22 \cdot (T/300)^{2.64} \cdot \exp[(70 \pm 55)/T]$ 295 < T < 850	$2.9 \times 10^{-16}$	$\leq 5 \times 10^{-20}$	na	na	28 d	3 d	0.0644	$8.36 \times 10^{-6}$	95.5
vinyl chloride	$6.6 \times 10^{-12}$	$1.14 \times 10^{-12} \cdot \exp(1045/RT)$ 299 < T < 426	$4 \times 10^{-16}$	nr	na	na	11 d	1 d	0.1054	$1.43 \times 10^{-5}$	8207 <sup>a</sup>

"nr" = no recommended value.

"na" = does not react in sunlight or no experimental data reported.

"nc" = not computed.

<sup>a</sup> at low ionic strengths and 293-298 K.

**Table E.2 Average 24 Hour OH, NO<sub>3</sub>, and O<sub>3</sub> Concentrations in a Moderately Polluted Atmosphere**

radical	concentration (molecules cm <sup>-3</sup> )		reference
	summer	winter	
OH	1.2 × 10 <sup>6</sup>	1.0 × 10 <sup>5</sup>	48
NO <sub>3</sub>	2.4 × 10 <sup>8</sup>	1.2 × 10 <sup>8</sup>	48
O <sub>3</sub>	1.1 × 10 <sup>12</sup>	6.8 × 10 <sup>11</sup>	49
temperature (K)	324.8	284.3	50

**Table E.3 Estimation of Photo-induced Decay Rate Constants for PAHs Under Summer and Winter Conditions**

PAH		Summer		Winter	
		k <sub>phot</sub> (sec <sup>-1</sup> )	t <sub>1/2</sub> <sup>a</sup> (hr)	k <sub>phot</sub> (sec <sup>-1</sup> )	t <sub>1/2</sub> (hr)
cyclopenta(c,d)pyrene	CpC	1.66 × 10 <sup>-4</sup>	1.2	3.23 × 10 <sup>-5</sup>	6.0
benz(a)anthracene	BaA	8.8 × 10 <sup>-5</sup>	2.2	8.5 × 10 <sup>-6</sup>	22.7
chrysene and triphenylene	Chry	3.3 × 10 <sup>-5</sup>	5.8	2.63 × 10 <sup>-6</sup>	73.0
benzo(b)fluoranthene	BbF	2.62 × 10 <sup>-5</sup>	7.3	4.27 × 10 <sup>-6</sup>	45.1
benzo(k)fluoranthene	BkF	3.35 × 10 <sup>-5</sup>	5.7	3.7 × 10 <sup>-6</sup>	52.0
benz(a)pyrene	BaP	7.62 × 10 <sup>-5</sup>	2.5	8.52 × 10 <sup>-6</sup>	22.6
indeno(1,2,3-cd)pyrene	Ind	4.65 × 10 <sup>-5</sup>	4.1	1.65 × 10 <sup>-6</sup>	116.7
benzo(ghi)perylene	BghiP	3.73 × 10 <sup>-5</sup>	5.2	2.95 × 10 <sup>-6</sup>	65.4

<sup>a</sup> t<sub>1/2</sub> = ln(2)/k<sub>phot</sub>. Computed rate constants were divided by a factor of two to account for PAH formed on higher temperature particles.

**Table E.4 Comparison of the Cumulative Deposition of Particles During 100 Seconds by Diffusion and Gravitational Settling<sup>a</sup>**

Diameter μm	Cumulative Deposition	
	Diffusion (number cm <sup>-2</sup> )	Gravitational Settling (number cm <sup>-2</sup> )
0.001	2.5	6.5E-5
0.01	0.26	6.7E-4
0.1	2.9E-2	8.5E-3
1.0	5.9E-3	0.35
10	1.7E-3	31
100	5.5E-4	2500

<sup>a</sup> Assume unit particle densities and deposition onto a horizontal surface from unit aerosol concentrations. Adapted from Finlayson-Pitts and Pitts (1986).

**Table E.5 Particle Phase Size Distributions of High Molecular Weight PAH at 298K**

Compound	Particle Diameter ( $\mu\text{m}$ )			
	<0.26	0.26 - 1.0	1.0 - 4.0	>4.0
High molecular weight PAH (MW >220 g mol <sup>-1</sup> )	70-75%	10-15%	5-10%	5%

**Table E.6 Particle Phase Size Distributions of Trace Metals at 298K, 1atm Pressure**

Compound	Particle Diameter ( $\mu\text{m}$ )				
	0.1	0.3	1.0	3.0	10.0
Trace metals	20%	50%	20%	5%	5%

**Table E.7 Predicted Partitioning of PAHs Based on Equation 21 Model**

Compound	number of rings	A	B	$K_p$ (m <sup>3</sup> $\mu\text{g}^{-1}$ )	percent of PAH mass associated with particles at 298K and TSP = 25 $\mu\text{g m}^{-3}$
phenanthrene & anthracene	3	4117	21.45	$2.3 \times 10^{-5}$	0.057
methylphenanthrene & methylanthracene	3	3365	18.46	$6.7 \times 10^{-5}$	0.17
fluoranthene	4	4421	21.52	$2.0 \times 10^{-4}$	0.51
pyrene	4	4183	20.52	$3.4 \times 10^{-4}$	0.84
benzo(a)fluorene & benzo(b)fluorene	4	4554	21.49	$6.1 \times 10^{-4}$	1.50
chrysene, benz(a)anthracene, & triphenylene	4	5826	24.89	$4.5 \times 10^{-3}$	10.1
benzofluoranthene	5	5693	23.24	$7.2 \times 10^{-2}$	64.1
benzo(a)pyrene & benzo(e)pyrene	5	4864	19.99	$2.1 \times 10^{-1}$	84.0

**Table E.8 Predicted Partitioning of Dioxins and Dibenzofurans  
Based on Equation 22 Model**

No. of chlorine atoms	p (torr) at 298K	$K_p$ ( $\text{m}^3 \mu\text{g}^{-1}$ )	percent of compound mass associated with particles at 298K and TSP = $25 \mu\text{g m}^{-3}$
4	$7.2 \times 10^{-7}$ to $1.7 \times 10^{-6}$	$2.3 \times 10^{-3}$ to $8.6 \times 10^{-4}$	5.5 to 2.1
5	$1.1 \times 10^{-7}$ to $4.8 \times 10^{-7}$	$2.0 \times 10^{-2}$ to $3.7 \times 10^{-3}$	33.4 to 8.4
6	$3.1 \times 10^{-8}$ to $9.8 \times 10^{-8}$	$8.6 \times 10^{-2}$ to $2.3 \times 10^{-2}$	68.3 to 36.4
7	$7.6 \times 10^{-9}$ to $1.4 \times 10^{-8}$	$4.3 \times 10^{-1}$ to $2.1 \times 10^{-1}$	91.6 to 84.3
8	$1.9 \times 10^{-9}$ to $2.0 \times 10^{-9}$	2.1 to 2.0	98.2 to 98.1

**Table E.9 Partitioning Potential of Other Hazardous Air Pollutants**

Hazardous air pollutants that may partition between gas and particle phases or exist exclusively in the particle phase	Hazardous air pollutants that exist almost exclusively in the gas phase
arsenic compounds beryllium compounds cadmium compounds chromium compounds manganese compounds nickel compounds lead compounds mercury compounds bis(2-ethylhexyl)phthalate coke oven emissions	acetaldehyde acrolein acrylonitrile benzene 1,3-butadiene carbon tetrachloride (tetrachloromethane) chloroform (trichloromethane) ethylene dichloride (1,2-dichloromethane) formaldehyde methylene chloride (dichloromethane)trichloroethylene tetrachloroethylene (perchloroethylene) vinyl chloride (chloroethene) 1,4-dichlorobenzene (p-dichlorobenzene) ethylene dibromide (1,2-dibromoethane) 1,1,2,2-tetrachloroethane acrylamide 1,3-dichloropropane 1,1-dichloroethene (vinylidene chloride) 1,2-dichloropropane (propylene dichloride) ethyl acrylate ethylene oxide hydrazine methyl chloride (chloromethane) quinoline styrene 1,1,2-trichloroethane

## E.5 REFERENCES

- Atkins, P. W., 1990. Physical Chemistry 4<sup>th</sup> ed, W. H. Freeman and Company, New York.
- Atkinson, R., Baulch, D. L., Cox, R. A., Hampson, Jr., R. F., Kerr, J. A., Troe, J., 1989. *J. Phys. Chem. Ref. Data* 18, 881.
- Atkinson, R., 1990. Final Report to California Air Resources Board, Contract No. A732-107 March.
- Atkinson, R., Baulch, D. L., Cox, R. A., Hampson, Jr., R. F., Kerr, J. A., Rossi, M. J., 1997. *J. Phys. Chem. Ref. Data* 26, 521.
- Behymer, T. D., Hites, R. A., 1988. *Environ. Sci. and Technol.* 22, 1311.
- Eitzer, E. D., Hites, R. A., 1989. *Environ. Sci. Technol.* 23, 1389.
- Finlayson-Pitts, B. J., Pitts, Jr, J. N., 1986. *Atmospheric Chemistry: Fundamentals and Experimental Techniques*, John Wiley and Sons, New York.
- Fletcher, K., Strommen, M., Kamens, R., 1997. Final Report on Parameters Relating to the Fates of Select Atmospheric Pollutants. Prepared for Jawad S. Touma, U. S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC.
- Kamens, R. M., Guo, Z., Fulcher, J. N., Bell, D. A., 1988. *Environ. Sci. & Technol.* 22, 103.
- Lyman, W. J., Reehl, W. F., Rosenblatt, D. H., 1996. *Handbook of Chemical Property Estimation Methods*, American Chemical Society, Washington, DC.
- Miguel, A. H., Freidlander, S. K., 1978. *Atmos. Environ.* 12, 2407.
- Pankow, J. F., Bidleman, T. F., 1992. *Atmos. Environ.* 26A, 1071.
- Pennise, D., Kamens, R. M., 1996. *Environ. Sci. Technol.* 30, 2832.
- Schwarzenbach, R. P., Gschwend, P. M., Imboden, D. M., 1993. *Environmental Organic Chemistry* John Wiley and Sons, New York.
- Steiger, M., Schultz, M., Schwikowski, M., Naumann, K., Dannecker, W., 1989. *J. Aerosol Sci.* 20, 1229.
- Suman, D., 1989. *Aerosol Sci and Technol.* 10, 131.

- U. S. EPA (Environmental Protection Agency), 1993. *Simultaneous Control of PM-10 and Hazardous Air Pollutants: Rationale for Selection of Hazardous Air Pollutants as Potential Particulate Matter or Associated with Particulate Matter at Source Conditions*; Research Triangle Park, NC. EPA/452-R-93-013.
- U. S. EPA (Environmental Protection Agency), 1997. Personal Communication with Dr. Russ Bullock. Office of Research and Development. Research Triangle Park, NC.
- Wehry, E. L., 1990. "Adsorption and photochemical transformations of PAHs sorbed on coal fly ash." Atmospheric Processes of Organic Toxic Pollutants and Their Role in current Environmental Issues Workshop, Canada Institute for Research in Atmospheric Chemistry.
- Whitby, K. T., 1977. *Atmos. Environ.* 12, 135.
- Yamasaki, H., Kuwata, K., Miyamoto, H., 1982. *Environ. Sci. Technol.* 16, 189.

## **APPENDIX F**

### **ADDITIONAL INFORMATION ON HOW TO USE OZIPR**



## APPENDIX F - TABLE OF CONTENTS

F. ADDITIONAL INFORMATION ON HOW TO USE THE OZIPR .....	F-1
F.1 THE OZIPR INPUT FILE .....	F-1
F.2. SAMPLE SIMULATION 1: HOUSTON TEXAS , SUMMER .....	F-10
F.3 RUNNING ADDITIONAL 24 HOUR SIMULATIONS .....	F-12
F.4 REFERENCES .....	F-18

## APPENDIX F - LIST OF EXHIBITS

Exhibit F.1 { <b>SAMPLE INPUT FILE</b> houtx2a.inp } .....	F-13
Exhibit F.2 { <b>SAMPLE CB-4 REACTIVITIES FILE</b> cb4_form.rea } .....	F-14
Exhibit F.3 { <b>SAMPLE BOUNDARY FILE</b> cb4sum.BOUND_a } .....	F-15
Exhibit F.4 { <b>SAMPLE EMISSIONS (MASS) FILE</b> cb4sum.MASS } .....	F-16
Exhibit F.5 { <b>SAMPLE METEOROLOGY FILE</b> summer.met } .....	F-17

## APPENDIX F

### ADDITIONAL INFORMATION ON HOW TO USE THE OZIPR

#### F.1 THE OZIPR INPUT FILE

To run OZIPR, one needs an input file, *filename.inp*, which is a text file consisting of a series of OPTIONS and COMMANDS. (See Sample Input File.) The program is started by typing "ozipr *filename.inp*". A standard output file, *filename.out* is created.

The input file can contain the names of external files to be included in the text of the input file. The character "@" followed by the name of the external file to be included will read the external file text line by line into its position in the input file. There are three required external data sets that are included at the beginning of the input file: a chemical mechanism, a zenith set, and reactivities. Any of the input can be isolated in external files if so desired, or simply included in the input file.

General rules for the input file are:

6. Usually, the first four letters of a command option name are significant.
7. Comments can be included following the first character " ! " or bracketed between { } or ( ).
8. Top level commands are followed by a " > ", the option commands, and ends with a " < ".
9. Case is not important for this version.

Sample files are included at the end of this section.

A line by line description of the sample input file follows.

Line 1 of the sample input file is:

**@cb4jsox\_form.mec**

The chemical mechanism external file includes information about the number of carbons for some organic species, and the list of reactions with rate constants used for the model run. For the purpose of determining secondarily produced formaldehyde, the chemical mechanism must differentiate between the primary formaldehyde and the secondary formaldehyde produced in situ. Secondarily produced formaldehyde is added as a new species, FRMS, while the symbol used for formaldehyde, FORM, now will refer to only primarily produced formaldehyde.

The number of carbons for organic species is included under "cnum", and must include the new species FRMP:

```
cnum = par, 1.0, eth, 2, ole, 2, tol, 7,  
xyl, 8, form, 1, ald2, 2, nr, 1,  
frmp, 1;
```

The chemical mechanism used in this study is the standard Carbon Bond IV mechanism with minor modifications to allow it to differentiate between primary and secondary formaldehyde production. This includes:

1. In the reaction equations that produce formaldehyde (formaldehyde on the right side of the equation), the symbol FORM is changed to FRMP in reactions 45, 46, 49, 50, 51, 56, 57, 58, 59, 60, 61, 62, 70, 71, 76, and 77 in the reactions list.
2. For each of the reactions where formaldehyde is depleted (formaldehyde on the left side of the equation), similar equations must be added that differ only by depleting FRMP instead of FORM. These additional equations can simply be added at the bottom of the list of chemical reactions in the mechanism file, as equations 86-90.

Line 2 of the sample input file is:

```
@jspec640.zen
```

The second line refers to a zenith angle file, specific for the Carbon-Bond Mechanism IV chemical mechanism file, and is used in interpolation of photochemical rates. This file will not need to be edited.

Line 3 of the sample input file is:

```
@cb4_form.rea
```

This line refers to an external file containing the fractional speciations of VOC species and is also specific to the Carbon-Bond Mechanism IV chemical mechanism. This file will need to be edited to include FRMP in background air. This file can also be edited if it is desirable to alter the VOC speciation.

There are 4 columns under reactivity. The first is the organic species name. The last three columns indicate the fraction of total VOC that comprises each species in: 1) the initial mixture ad/or emissions (VOC entered under CALC, EMIS, or MASS); 2) the aloft VOC (VOCALOFT under TRANSPORT); and 3) the transported surface-layer VOC (VOCSURFACE under TRANSPORT). Initial VOC concentrations will be defined as VOCSURFACE air (in this case), and therefore initial speciation is defined in the last column of the file. The FRMP to FORM ratio assumed for a 6AM initialization is 9:1. To approximate the ratio for a different

initialization hour, look at the seasonal tables of hourly FORM and FRMP for Houston. The initial ratio of FRMP to FORM will not be crucial, because of the relatively short lifetime of formaldehyde.

**BOUNDARY >**  
**REACTIVITY =**

par,	0.5640,	0.4980,	0.4980,
eth,	0.0370,	0.0340,	0.0340,
ole,	0.0350,	0.0200,	0.0200,
tol,	0.0890,	0.0420,	0.0420,
xyl,	0.1170,	0.0260,	0.0260,
form,	0.0210,	0.0700,	0.0070,
frmp,	0.0000,	0.0000,	0.0630,
ald2,	0.0520,	0.0370,	0.0370,
nr,	0.0850,	0.2730,	0.2730;

<

Line 4:

**MODIFY> ACC = 0.0001; <**

This line modifies the accuracy in the tolerance of the numerical integration routine.

Lines 5-14:

**TITLE > HOUSTON, TX Design Day Sim SUMMER <**  
**PLACE>**

**CITY = HOUSTON, TX;**  
**LAT = 29.8,**  
**LONG = 95.2,**  
**TZone = 5.0,**  
**Year = 1988,**  
**Month = 7,**  
**Day = 16; <**

**TIME> 0600, 2900 <**

These lines define the place and time for the run. TITLE is the title chosen for the run, up to 72 characters. PLACE is used in determining the zenith angle of the sun throughout the day. CITY name can be up to 24 characters. LATITUDE is latitude in degrees north, and LONGITUDE is the longitude in degrees west. TZONE is the local time zone, Eastern Daylight time = 4, Pacific Daylight Time = 7. YEAR, is the year, MONTH is the number for month, and DAY is the day of the month. TIME is the range of hours to run the simulation. In this example, the model will run from 6 am on July 16 until 6 am on July 17.

Line 16:

**@summer.MET**

Meteorology is defined in an external include file which includes pressure, mixing heights, air temperature, and air moisture. This file should be edited for meteorology specific to the region and time of year.

In this example, the hourly mixing heights were entered explicitly using the MIXING option:

```
MIX[24] =  
          598., 697., 797., 897., 996.,  
1096., 1195., 1295., 1395., 1395., 1395., 1395., 1395., 1392., 1278.,  
1106., 924., 738., 576.,  
567., 562., 558., 562., 566.;
```

As these are needed at the beginning and end of each hour, the number of mixing heights should equal one plus the number of hours for the simulation. If the number of values is less than this, the last value entered will be used for the remaining hours.

Temperature is entered in a similar manner:

```
TEMP[24,K] =  
          296.7, 298.0, 300.1, 301.8, 303.1,  
304.1, 304.9, 305.3, 305.5, 305.5, 305.1, 304.5, 303.6, 302.3, 301.1,  
300.3, 300.0, 299.0, 298.5,  
298.0, 297.6, 297.2, 296.9, 296.7;
```

As for the mixing height, the number of values should equal one plus the number of hours for the simulation. The units for temperature are entered as "K" for Kelvin, "C" for Centigrade, and "F" for Fahrenheit.

Pressure is entered as a constant value of 29.75 inches of mercury for this example:

```
pres[in] = 29.75;
```

In this example, air moisture is entered as relative humidity:

```
RH[24] =  
          92.0, 86.5, 81.0, 75.5, 70.0,  
64.5, 59.0, 60.8, 62.7, 64.5, 66.3, 68.2, 70.0, 71.8, 73.7,  
75.5, 77.3, 79.2, 81.0,  
82.8, 84.7, 86.5, 88.3, 90.2;
```

The hourly values in average relative humidity are entered for the beginning and end of each hour.

DILUTION option. In this example, the hourly mixing heights were entered explicitly using the MIXING option:

```
MIX[24] =  
          598., 697., 797., 897., 996.,  
1096., 1195., 1295., 1395., 1395., 1395., 1395., 1395., 1392., 1278.,  
1106., 924., 738., 576.,  
567., 562., 558., 562., 566.;
```

As these are needed for the beginning and end of each hour, the number of mixing heights should equal one plus the number of hours for the simulation.

Temperature is entered in a similar manner:

```
TEMP[24,K] =  
          296.7, 298.0, 300.1, 301.8, 303.1,  
304.1, 304.9, 305.3, 305.5, 305.5, 305.1, 304.5, 303.6, 302.3, 301.1,  
300.3, 300.0, 299.0, 298.5,  
298.0, 297.6, 297.2, 296.9, 296.7;
```

The number of values input should equal one plus the number of hours for the simulation. If the number of values is less than this, the last value entered will be used for the remaining hours. The units for temperature are entered as "K" for Kelvin, "C" for Centigrade, and "F" for Fahrenheit.

In this example air moisture is entered as RH (relative humidity):

```
RH[24] =  
          92.0, 86.5, 81.0, 75.5, 70.0,  
64.5, 59.0, 60.8, 62.7, 64.5, 66.3, 68.2, 70.0, 71.8, 73.7,  
75.5, 77.3, 79.2, 81.0,  
82.8, 84.7, 86.5, 88.3, 90.2;
```

The hourly values in average relative humidity are entered for the beginning and end of each hour. The total number of values equals the total simulation hours plus 1. If wnum is less than this, the last value entered will be used for the remaining hours.

In line 17 of the input file, boundary conditions are defined in an external include file for the sample input file:

Line 17:

```
@ cb4sum.BOUND_f
```

Boundary conditions include the assumed ratios of NO<sub>2</sub>/NO<sub>x</sub> in any NO<sub>x</sub> entered as initial conditions or as emissions. Also included are deposition velocities for selected chemical species, and initial concentrations assumed for selected chemical species.

The TRANSPORT option allows predetermining the concentrations of selected species for both surface and aloft air. The top level command here is BOUNDARY:

```
BOUND >  
conditions....  
<
```

IFRACTION\_NO<sub>2</sub>, is the NO<sub>2</sub>/NO<sub>x</sub> fraction for the NO<sub>x</sub> in controllable emissions at the start of the simulation.

```
IFRaction_NO2 = .05;
```

Hourly deposition velocities in cm/s for NO<sub>2</sub>, O<sub>3</sub>, HNO<sub>3</sub>, H<sub>2</sub>O<sub>2</sub>, and PAN are entered using the DEPOSITION option and the format:

```
DEPO [24] =  
NO2 = .24, .36, .48, .54, .60,  
.60, .60, .60, .60, .60, .54, .48, .36, .24, .24,  
.24, .24, .24, .24,  
.24, .24, .24, .24, .24,
```

TRANSPORT is used to define the initial concentrations for O<sub>3</sub>, NO<sub>x</sub>, VOC, and CO that are due to transport.

```
TRANSPORT =  
O3SURFACE = .021, { W=.021, SP=.033, SU=.024, A=.026 }  
O3ALOFT = 100, 0.04,  
VOCSURFACE = .000,  
VOCALOFT = .030,  
NOXSURFACE = .000,  
NOXALOFT = .002,  
COSURFACE = .000,  
COALOFT = .500;
```

Surface layer concentrations entered under TRANSPORT can be used for the species as initial conditions. Values are in ppm (ppmc for VOC). Input for O3ALOFT includes two values, the first of which is the value in meters above which height the O3ALOFT value should be used. The second value is the O3ALOFT concentration in ppm.

VOC species entered under TRANSPORT will be speciated as designated in the boundary conditions for REACTIVITY. These are species specific to the chemical mechanism used and are not defined in this external file, but are defined in the external reactivity file, cb4\_form.rea. This external file was included in line 3 of the input file. There are 4 columns under reactivity. The first is the organic species name. The last three columns indicate the fraction of total VOC that comprises each species in the initial mixture and/or emissions (VOC entered under CALC, EMIS, or MASS), the aloft VOC (VOCALOFT under TRANSPORT), and the transported surface-layer VOC (VOCSURFACE under TRANSPORT). So when initializing VOC using VOCSURFACE, it is the last column under REACTIVITY that is accessed. From the external file cb4\_form.rea :

**BOUNDARY >**  
**REACTIVITY =**

par,	0.5640,	0.4980,	0.4980,
eth,	0.0370,	0.0340,	0.0340,
ole,	0.0350,	0.0200,	0.0200,
tol,	0.0890,	0.0420,	0.0420,
xyl,	0.1170,	0.0260,	0.0260,
form,	0.0210,	0.0700,	0.0070,
frmp,	0.0000,	0.0000,	0.0630,
ald2,	0.0520,	0.0370,	0.0370,
nr,	0.0850,	0.2730,	0.2730;

<

Note that REACTIVITY in its external file is nested in its top level command, BOUNDARY.

Finally, the boundary file also can contain initial concentrations of species other than CO, and O<sub>3</sub> using the INITIALIZE option. If initial concentrations of O<sub>3</sub> are desired, it is not necessary to initialize any species this way. However, this option allows you to input realistic non-zero concentrations for initializing a simulation. In this example, we have initialized the concentrations of all species to equal the 6 a.m. concentrations from a previous day's simulation:

**INIT =**

no2	=	0.000044041,
no	=	0.000000121,
o	=	0.000000000,
no3	=	0.000000007,
.		
.		
isop	=	0.000964346,
nr	=	0.068530180,
apin	=	0.000000000,
unkn	=	0.000000000;



In line 18 of the input file, an external file which defines the hourly emissions is included.

Line 18:

**@ cb4sum.MASS\_f**

In this file, the hourly emission mass for up to 10 species is entered in units of kg/km<sup>2</sup> per hour. The species name is followed by its molecular weight in brackets. An initial concentration is also entered (in units of ppm). For species other than VOC, NOX, and CO, the initial concentration specified here is added on to concentrations specified in the INIT option. After the concentration, hourly emission values are entered for each hour in the format:

MASS[24] >

VOC [ 14.50] = .5530,  
3.936, 4.894, 6.044, 7.217, 7.746,  
8.077, 8.318, 8.377, 8.759, 9.120,  
9.282, 9.198, 7.868, 6.221, 5.035,  
4.201, 3.613, 3.368, 2.780,  
2.528, 2.370, 2.432, 2.520, 2.883,

Line 19:

After adding the emissions via the external file, we finish with the CALCULATE option.

In this option, the initial concentrations due to emissions for VOC, NO<sub>x</sub>, and CO are input. In this example we have already specified initial concentrations under the BOUNDARY options, so we have set them equal to zero here.

**VOC = 0.000;**  
**NOX = 0.000;**  
**CO = 0.000;**

Also included under CALCULATE is the PRINT option, which is used to select species concentrations to be output.

For this example, we output only concentrations for 5 species:

**PRINT[FULL] =**  
**NAMES[5] = form, frmp, O<sub>3</sub>, NO, NO<sub>2</sub>**  
**CDUMP[5] = form, frmp, O<sub>3</sub>, NO, NO<sub>2</sub>**  
**NODUMP;**

There are many other options available for use with OZIPR. However, these are not necessary for this application and are not reviewed here. For further information, the Users Guide (Gery and Crouse 1991) should be consulted.

## F.2. SAMPLE SIMULATION 1: HOUSTON TEXAS , SUMMER

The analysis will be specific to both a given city and season. Houston, Texas has been used in the development of this process. A prototypical day has been determined for each season: winter, spring, summer, and autumn. Seasonally averaged hourly meteorology data for each city are used. For example, Houston's 8 AM temperature values for each spring day are averaged to create the 8 AM temperature value used in the simulation. The emissions data for each season are handled in a like fashion.

Files needed for this simulation include the input file:

**houtx2a.inp**

There are three external files specific to the cb-4 mechanism that must be made available:

**cb4jsox\_form.mec** ( the chemical mechanism file)

**jspec640.zen** ( the zenith angle file)

**cb4\_form.rea** ( the reactiities file)

In addition, additional input for the simulation is included in these other external files:

**summer.MET** ( meteorology )

**cb4sum.BOUND\_a** ( boundary conditions )

**cb4sum.MASS** ( emissions )

Local meteorology information is included in the file "**summer.MET**". Houston temperature and mixing height values were taken from 5 year averages of the hourly data. (1987-1991). Morning and afternoon relative humidities from "Comparative Climatic Data for the United States" (NCDC, 1984) were used and data for all hours were linearly interpolated from the two values.

Background concentrations are entered in the external file "**cb4sum.BOUND\_a**". The background concentrations for O<sub>3</sub>, VOC, NO<sub>x</sub>, and CO used in this simulation are from 1997 Photochemical Ambient Monitoring Site (PAMS) data. These species are initialized by entering them as surface values of transported air:

**O3SURFACE = .024,**  
**VOCSURFACE = .553,**  
**NOXSURFACE = .000,**  
**COSURFACE = .911,**

Emissions for the simulation are entered in "**cb4sum.MASS\_a**". For this simulation, biogenic emissions for VOC and NO<sub>x</sub> were provided by running PC-Beis2 for Harris County, Texas using temperature data from the ISCST3 model inputs. This was run for July 14, 1990. The PC

Beis2 program and associated files can be obtained from the EPA SCRAM web pages (<http://www.epa.gov/scram001/index.htm>). Daily anthropogenic emission values for NO<sub>x</sub>, CO, and reactive organic gases (ROG) were extracted from the OTAG base emissions inventory for Harris County Texas.

Once all the input is gathered, the model is begun by typing "ozi`pr`.x *filename*.inp", or in this case, "ozi`pr`.x houtx2a.inp".

### F.3 RUNNING ADDITIONAL 24 HOUR SIMULATIONS

The results for determining secondary formaldehyde based on one 24 hour run can be very dependent on initial conditions, and perhaps not provide the best estimate for secondary formaldehyde. Additional consecutive 24 hour simulations may be performed to the point where the 24 hour curve for formaldehyde values (primary and secondary) remains approximately the same from day to day. The output file, "*filename.out*", contains information about the species concentrations throughout the run. The model can be run for several 24 hour interactions, or until both primary and secondary formaldehyde concentrations have approximately the same concentrations for each species at the beginning and end of the 24 hour run. In order to set up a simulation to continue from the end of the last simulation, there will be some differences in the input. The boundary file in Simulation 1, **cb4sum.BOUND\_a**, can be adapted for a second run with the boundary file **cb4sum.BOUND\_b** using the final species concentrations in the output file, "*houtx2a.out*". First, the VOCSURFACE and NOXSURFACE concentrations under TRANSPORT must be set to 0.0, as their constituents (organic species, NO and NO<sub>2</sub>) will be initialized individually. O<sub>3</sub> and CO from the output file are used to initialize O3SURFACE and COSURFACE under TRANSPORT. All other species will be initialized under the INIT option (also part of the boundary conditions). A new input file, *houtx2b.inp*, should be created that includes the external boundary file **cb4sum.BOUND\_b**. This process of using the final species concentrations as input to a subsequent run can be continued for additional 24 hour periods.

Exhibit F.1 { **SAMPLE INPUT FILE** houtx2a.inp }

```
@/home/wendy/models/ozipr/mec/cb4jsox_form.mec
@/home/wendy/models/ozipr/mec/jspec640.zen
@/home/wendy/models/ozipr/mec/cb4_form.rea
TITLE > HOUSTON, TX Design Day Sim SUMMER <
MODIFY> ACC = 0.0001; <
PLACE>
  CITY = HOUSTON, TX;
  LAT = 29.8,
  LONG = 95.2,
  TZone = 5.0,
  Year = 1988,
  Month = 7,
  Day = 16; <
TIME> {100, 2400} 0600, 2900 <
@/home/wendy/models/ozipr/sims/summer.MET
@/home/wendy/models/ozipr/sims/cb4sum.BOUND_a
@/home/wendy/models/ozipr/sims/cb4sum.MASS_a
CALC>
  VOC = .000; { W=.302, SP=.427, SU=.553, A=.507 }
  NOX = .000; { W=.001, SP=.001, SU=.000, A=.000 }
  CO = .000; { W=.716 SP=.721 SU=.911, A=.896 }

PRINT[FULL] =
  NAMES[5] = form,frmp,o3,no,no2,
  CDUMP[5] = form,frmp,o3,no,no2,
  NODUMP;
{
  PRINT[FULL] =
  NODUMP;
}

<
END.
```

Exhibit F.2 { **SAMPLE CB-4 REACTIVITIES** FILE cb4\_form.rea }

boundary >

reac =

par,	0.5640,	0.4980,	0.4980,
eth,	0.0370,	0.0340,	0.0340,
ole,	0.0350,	0.0200,	0.0200,
tol,	0.0890,	0.0420,	0.0420,
xyl,	0.1170,	0.0260,	0.0260,
form,	0.0210,	0.0700,	0.0070,
frmp,	0.0000,	0.0000,	0.0630,
ald2,	0.0520,	0.0370,	0.0370,
nr,	0.0850,	0.2730,	0.2730;

{reactivity = site specific reactivity of the 6-9 am mix as nmoc fractions}

!=====

=

< {boun}

Exhibit F.3 { **SAMPLE BOUNDRY FILE** cb4sum.BOUND\_a }

```
BOUND>          { SUMMER }
DEPO [24] =
NO2 =           .24, .36, .48, .54, .60,
               .60, .60, .60, .60, .60, .54, .48, .36, .24, .24,
               .24, .24, .24, .24,
               .24, .24, .24, .24, .24,
O3 =           .30, .50, .60, .70, .80,
               .80, .80, .80, .80, .80, .70, .60, .50, .30, .30,
               .30, .30, .30, .30,
               .30, .30, .30, .30, .30,
HNO3 =         2.60, 2.60, 3.00, 3.30, 3.50,
               3.50, 3.50, 3.50, 3.50, 3.50, 3.30, 3.20, 3.00, 2.60, 2.60,
               2.60, 2.60, 2.60, 2.60,
               2.60, 2.60, 2.60, 2.60, 2.60,
H2O2 =         1.60, 1.70, 1.80, 1.90, 2.00,
               2.00, 2.00, 2.00, 2.00, 2.00, 1.90, 1.80, 1.70, 1.60, 1.60,
               1.60, 1.60, 1.60, 1.60,
               1.60, 1.60, 1.60, 1.60, 1.60,
PAN =          .24, .36, .48, .54, .60,
               .60, .60, .60, .60, .60, .54, .48, .36, .24, .24,
               .24, .24, .24, .24,
               .24, .24, .24, .24, .24;
IFRaction_NO2 = .050;
TRANSPORT =
O3SURFACE =    .024, { W=.021, SP=.033, SU=.024, A=.026 }
O3ALOFT  =    100, 0.04,
VOCSURFACE =   .553, { W=.302, SP=.427, SU=.553, A=.507 }
VOCALOFT  =    .030,
NOXSURFACE =   .000, { W=.001, SP=.001, SU=.000, A=.000 }
NOXALOFT  =    .002,
COSURFACE =   .911, { W=.716 SP=.721 SU=.911, A=.896 }
COALOFT   =    .000;
{ INITial Concentrations = }
```

<



Exhibit F.4 { **SAMPLE EMISSIONS (MASS) FILE** cb4sum.MASS }

{cb4sum.MASS}

MASS[24] >

VOC [ 14.50] = .5530, {from Allan, beis,anthro}  
3.936, 4.894, 6.044, 7.217, 7.746,  
8.077, 8.318, 8.377, 8.759, 9.120,  
9.282, 9.198, 7.868, 6.221, 5.035,  
4.201, 3.613, 3.368, 2.780,  
2.528, 2.370, 2.432, 2.520, 2.883,

NO [ 30.00] = .0000000001, {from Allan, beis}  
.0070, .0075, .0083, .0097, .0109,  
.0115, .0125, .0129, .0129, .0125,  
.0121, .0118, .0118, .0112, .0106,  
.0103, .0100, .0100, .0094,  
.0078, .0078, .0076, .0072, .0072,

NOX [ 46.00] = .0000000001, {from Allan, anthro}  
4.702, 5.496, 5.896, 6.343, 6.650,  
6.734, 6.802, 6.842, 6.984, 7.350,  
7.451, 7.317, 6.686, 5.657, 4.751,  
4.200, 3.814, 3.596, 3.194,  
2.940, 2.772, 2.864, 3.080, 3.604,

ISOP[ 68.13] = .0000010, {from Allan, beis}  
0.484, 0.822, 1.122, 1.899, 2.567,  
2.999, 3.585, 3.809, 3.692, 3.239,  
2.669, 1.492, 0.133, 0.000, 0.000,  
0.000, 0.000, 0.000, 0.000,  
0.000, 0.000, 0.000, 0.000, 0.088,

CO [ 28.00] = .91100, {from Allan anthro}  
13.995, 19.234, 21.852, 24.906, 27.148,  
28.435, 29.318, 29.916, 31.692, 34.627,  
35.833, 35.150, 29.333, 21.523, 15.068,  
11.170, 8.331, 6.781, 4.949,  
3.544, 2.094, 2.720, 3.570, 6.899;

<

Exhibit F.5 { **SAMPLE METEOROLOGY FILE** summer.met }

```
MET> { start at proper hour, see TIME above }
MIX[24] =
{ 567., 562., 558., 562., 566., 598., 697., 797., 897., 996.,}
      598., 697., 797., 897., 996.,
1096., 1195., 1295., 1395., 1395., 1395., 1395., 1395., 1392., 1278.,
1106., 924., 738., 576.,
567., 562., 558., 562., 566.;
TEMP[24,K] =
{ 298.0, 297.6, 297.2, 296.9, 296.7, 296.7, 298.0, 300.1, 301.8, 303.1,}
      296.7, 298.0, 300.1, 301.8, 303.1,
304.1, 304.9, 305.3, 305.5, 305.5, 305.1, 304.5, 303.6, 302.3, 301.1,
300.3, 300.0, 299.0, 298.5,
298.0, 297.6, 297.2, 296.9, 296.7;
PRES[IN] = 29.75;
RH[24] =
{ 82.8, 84.7, 86.5, 88.3, 90.2, 92.0, 86.5, 81.0, 75.5, 70.0,}
      92.0, 86.5, 81.0, 75.5, 70.0,
64.5, 59.0, 60.8, 62.7, 64.5, 66.3, 68.2, 70.0, 71.8, 73.7,
75.5, 77.3, 79.2, 81.0,
82.8, 84.7, 86.5, 88.3, 90.2;
<
```

#### F.4 REFERENCES

- Gery, M. W. and Crouse, R. R., 1991. "User's Guide for Executing OZIPR", U.S. Environmental Protection Agency, Research Triangle Park, NC.
- NCDC (National Climatic Data Center), 1984. "Comparative Climatic Data for the United States", Asheville, NC.

# **TECHNICAL REPORT DATA**

*(Please read Instructions on reverse before completing)*

1. REPORT NO EPA-454/R-99-021		2.		3. RECIPIENT'S ACCESSION NO	
4. TITLE AND SUBTITLE  Air Dispersion Modeling of Toxic Pollutants in Urban Areas; Guidance, Methodology and Example Applications				5. REPORT DATE July 1999	
				6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S)				8. PERFORMING ORGANIZATION REPORT NO	
9. PERFORMING ORGANIZATION NAME AND ADDRESS  EC/R Incorporated 1129 Weaver Dairy Road Chapel Hill, NC 27514				10. PROGRAM ELEMENT NO	
				11. CONTRACT/GRANT NO EPA Contract No. 68D98006	
12. SPONSORING AGENCY NAME AND ADDRESS U.S. Environmental Protection Agency Office of Air Quality Planning and Standards Emissions, Monitoring & Analysis Division Research Triangle Park, NC 27711				13. TYPE OF REPORT AND PERIOD COVERED Final Report	
				14. SPONSORING AGENCY CODE	
15. SUPPLEMENTARY NOTES EPA Work Assignment Manager: Jawad S. Touma					
16. ABSTRACT  This report provides guidance on how to use the Industrial Source Complex Short Term (ISCST3) model to estimate concentrations of air toxic pollutants for urban-wide analyses. Urban areas contain major sources and numerous smaller, area sources. As a result modeling analyses poses special challenges. Section 1 provides guidance and recommendations on specific issues for urban-wide analyses of air toxics. Section 2 provides an overview of two applications of the Industrial Source Complex Short Term (ISCST3) model to urban-wide studies. Appendices A and B provide demonstration of the methodology for two example applications.					
KEY WORDS AND DOCUMENT ANALYSIS					
a. DESCRIPTORS		b. IDENTIFIERS/OPEN ENDED TERMS		c. COSATI Field/Group	
Air Pollution Air Quality Dispersion Models Meteorology Air Toxics Urban Area Modeling					
18. DISTRIBUTION STATEMENT  Release Unlimited		19. SECURITY CLASS (Report) Unclassified		21. NO OF PAGES 284	
		20. SECURITY CLASS (Page) Unclassified		22. PRICE	