Final Technical Report

Analysis of Carbon Monoxide Exposure for Fourteen Cities using HAPEM-MS3

AEAR WA II-48

Michael Zelenka
Work Assignment Manager

National Exposure Research Laboratory
U.S. Environmental Protection Agency

Prepared by:

Graham Glen, Project Scientist
Doug Shadwick, Project Scientist

ManTech Environmental Technology, Inc.
Research Triangle Park, NC 27709

EPA Contract No. #68-D5-0049
Work Assignment 048

Foreword

This report presents the results of work performed by ManTech Environmental Technology, Inc., under Work Assignment II-48 of Contract 68-D5-0049 for the National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC. This report has been reviewed by ManTech Environmental Technology, Inc., and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Table of Contents

List of Appendices

List of Tables

CHAPTER 1

EXPOSURE ESTIMATES USING THE HAPEM-MS3 MODEL

Introduction

The purpose of this report is to present the findings from an application of the Hazardous Air Pollutant Exposure Model for Mobile Sources, Version 3 (hereafter HAPEM-MS3, or HAPEM) to fourteen urban areas for the calendar year 1990. This report contains a description of the HAPEM model, model results and discussion, an uncertainty analysis, and a quality assurance section. The Appendices contain detailed tables of model results for each of fourteen study areas, and a set of maps (one per study area). The principal results are tables of exposure estimates to ambient carbon monoxide (CO) by demographic group, quarter of the year, and county. Other tables summarizing exposure by hour, air district, and microenvironment are also presented, along with tables of percentiles and fractiles of the exposure distribution within each study area. A companion report presents a sensitivity analysis, which investigates the dependence of the exposure estimates on the set of microenvironmental factors used for the analysis.

The CO exposure estimates were calculated for a common set of 23 demographic groups in each study area, using a common set of model parameters. This allows for the direct comparison of different study areas. The entire set of output tables is very extensive and has been placed on a CD-ROM. The additional tables on the CD-ROM not printed in this report include hourly and microenvironmental breakouts for each demographic group in each study area. The printed tables in this report only contain these results for the demographic group of ‘all persons’. The complete set contains an additional 3 x 22 x 14 = 924 pages of tables beyond those printed in this report.

Description of HAPEM

HAPEM calculates exposures to carbon monoxide (CO), or other mobile source pollutants, over a year, for all the people in a given metropolitan area. HAPEM does not estimate the exposure of individuals, but of an entire demographic group in a particular air district. In the model, each person in the group is effectively interchangeable with any other person in the group, meaning that their activities are all selected with the same probabilities from the same database. This results in HAPEM providing an overall population-weighted average exposure for the demographic group. Due to the method of analyzing air quality data, the model does not provide estimates for time periods shorter than a calendar quarter. HAPEM is therefore a long-term, population level exposure model and should not be applied either to individuals or to short duration events. Due to the large data files and long execution time, HAPEM is run on the EPA's IBM mainframe computer.

There are four main types of data used as inputs to HAPEM:

- CO Monitoring Data
- Time-Activity Data
- Microenvironmental Data
- Population Data

The output from HAPEM provides electronic files and tables summarizing the quarterly and annual exposure for each demographic group. These tables and files have been extended to include new types of output such as summaries at the county level as well as microenvironmental exposures. The details of these enhancements are provided later in this report.

CO Monitoring Data:

The study area is divided into districts, which are roughly circular areas of diameter 20 km around each ambient CO monitor. Districts are actually composed of a whole number of census tracts. The data from the monitors is extracted from the AIRS database prior to running HAPEM. With N monitors, there are N+1 districts. For the 23 demographic groups there are ($5 N^2 + 33 N + 28$) cohorts (defined by their demographic groups, home locations, and work locations) . From 1 to 18 monitors per study area can be accommodated in the current version of the model. An extra district (number 19 in the tables) is created in each study area and it consists of those areas not close to (i.e. within 10 km of) any ambient monitor. The concentration in district 19 is just set to the average concentration of all the monitors in the study area. Because this concentration is assumed rather than measured, exposure estimates for district 19 are not as reliable as those for other districts. In the results, a flag has been created to indicate counties with estimates that are based largely on district 19 values.

In HAPEM, the CO monitoring data is processed by the two programs TSERIES and AQAVG. The TSERIES program smooths the data and fills in missing values by first performing a Fourier analysis on the year-long time series of hourly values at each monitor. The most significant terms are retained. TSERIES uses a second-order autoregressive technique incorporating a random factor to estimate the missing data. This means that the results from TSERIES will differ slightly from run to run if there are any missing data in the original time series. The AQAVG program then averages the smoothed time series into 24 hourly averages for each quarter of the year, at each monitor. The analysis by both programs is done independently for each monitor, so the inclusion or exclusion of other monitors in the study area does not affect the results. These programs have not been functionally changed for the current set of HAPEM runs. The average annual concentration at each monitor as calculated by AQAVG is given in the last column of Table 1, for comparison to the mean of the observed (non-missing) data. Small differences in the means are expected since the filled values do not necessarily have the same average as the rest of the data.

<u>Time Activity Data:</u>

The time-activity data originally comes from the three-city database (data from Denver, Washington DC, and Cincinnati, collected in 1982-1985). The records are specially processed for use in HAPEM. The file currently contains 3568 person-days of data, divided into 8 states (combinations of three binary divisions: winter/summer, weekend/weekday, and warm/cool weather), with information on the time spent during each clock hour in each of 37 microenvironments. The warm/cool division for the states is set at 84°F in June, July and August, and at 55°F during the remaining months. This database has not been altered from the one used in previous HAPEM-MS2 and HAPEM-MS3 runs.

Although the time-activity database has not changed, the method of extracting patterns from it has been improved. Up until the present runs, HAPEM used stochastic sampling to extract records from this activity database, selecting a random time-activity day from the appropriate subgroup for each day of the year. These patterns in effect are realizations of annual time-activity patterns created by repeated independent random sampling of daily patterns. The exposure estimates therefore apply to each particular realization, rather than a true mean over all possible activity patterns. To obtain a better estimate of the mean, and also to estimate the size of this stochastic variation, it was customary to repeat the entire HAPEM runs typically ten times each. With the changes and enhancements to HAPEM (described in the next section), this feature was replaced with an exact calculation of the mean and variance associated with this method of creating annual activity time-series. A single HAPEM run now provides all the information (and more) than the ten runs did previously.

The time-activity database (stored on the EPA's IBM mainframe computer under the file name MEDUR.DATA) records the time spent in each of the 37 HAPEM microenvironments during each clock hour of the day. The data for the same hours on all the days in each calendar quarter

are averaged together before being combined with concentration data, since a similar averaging was already carried out for the concentration data by the AQAVG program.
A new program called DURAVG has been added to HAPEM to average the patterns that apply to a single demographic group. These averages are used for the hourly and microenvironmental exposure calculations.

Microenvironmental Data:

The microenvironmental factors are entered into HAPEM as data, and can be varied by the user. The values used in the current runs are on the file MEFILE5. These factors are derived from the factors used in previous HAPEM runs, with the main difference that the additive terms are all set to zero. This choice was motivated by the argument that the multiplicative factors represent the penetration of ambient CO into other microenvironments, while the additive terms represent sources within the given microenvironment. The goal of the present model runs was to estimate the exposure to ambient CO rather than total CO. There was one other change from the factors used in previous runs: the four in-home microenvironments (#13-16) were combined into a single microenvironment, since these four differ only in the potential for local (in-home) emissions. A list of the factors used for the current analysis is given in Table 2. A further discussion of the choice of microenvironmental factors and their impact on the exposure estimates is presented in the sensitivity analysis report.

The MECONC program is used to calculate the microenvironmental concentrations. It must be run after the two air quality programs (TSERIES and AQAVG) and requires a file of microenvironmental factors. The file of factors is named MEFILE5. The results from MECONC are stored on the file MECONC5. The MECONC program must be run before the EXPCODI program.

Population

The 1990 Census provides excellent detail on home location and on demographic groups, and somewhat less information on work location. The primary census unit used in the HAPEM model is the census tract, which typically contains about 5000 people. The study areas used for the present runs contain anywhere from 100 to 4000 tracts. The census provides the number of people in each demographic group living in each tract, and also the distribution of commuting times for each tract. For previous model runs, the POP90 program aggregated the population data directly to the HAPEM district (air monitor) level. Since the current set of runs require the new feature of summarizing exposure by county, the POP90 program has been modified to provide summaries of populations for each (county x district) combination.

Two programs are used in the commuting calculations: TVLTIME and ODEST. The TVLTIME program reads the census commuting times and creates an array of them for later use by ODEST. The ODEST program has been rewritten slightly and now combines the functions of three earlier programs (DIST, ODEST, and TTFRA). ODEST calculates the number of people who commute from any one census tract to any other. This is an iterative calculation that looks for a solution consistent with the commuting time totals from the census. Once found, the results are aggregated up to the HAPEM air district level and stored in the HOMEWORK file.

All of the above pieces can be considered as data pre-processing. These are necessary steps prior to the actual exposure calculations. Previously, the exposure calculations were carried out by the MERGE and GRAPH programs. For the current runs, these have been replaced by the new programs EXPCODI and EXPMEHR. The EXPCODI program calculates exposure by county and also by air district, for each demographic group and quarter of the year. The EXPMEHR program calculates average exposures by hour of the day and by microenvironment for each demographic group, on a city-wide basis.

Changes and enhancements to HAPEM

HAPEM currently is comprised of ten Fortran programs as outlined in the previous section. The changes made to HAPEM for the current runs do not affect the nature of the model or any of the model assumptions, but serve the following main purposes:

1) to add new features to the output tables (including county and microenvironmental information, as well as more detailed percentile and fractile tables),
2) to give exact mean and variance calculations instead of stochastic estimates,
3) to obtain results for all demographic groups in a single run, and
4) to speed up program execution.

The results from the modified programs were compared to previous HAPEM runs. This was done for the HAPEM-MS3 runs for San Francisco carried out in 1996. All demographic groups were run for the study year 1990, using the same monitors and microenvironmental factors as in the previous runs (presented in a report to EPA prepared by IT Corporation entitled ‘Development and Evaluation of Enhancements to the Hazardous Air Pollutant Exposure Model HAPEM-MS3'). The results from the two sets of runs agreed within the stated uncertainties. The results of this comparison are presented in greater detail in the quality assurance section. In addition, a separate calculation was carried out in SAS to test the algorithms for the mean and variance calculations carried out in the new Fortran routines added to HAPEM.

.

The ten programs in HAPEM can be divided into three groups: those that did not change; those with minor changes; and new programs.

PROGRAM	CHANGES	SUMMARY DESCRIPTION OF CHANGES
TSERIES	none	
AQAVG	none	
DIST90	none	
TVLTIME	none	
ODEST	minor	DIST, ODEST, TTFRA combined (to run faster)
POP90	minor	County level information retained
MECONC	new	Used to be part of MERGE, now separate
DURAVG	new	Part of mean and variance calculations
EXPCODI	new	Creates tables by county and air district
EXPMEHR	new	Creates tables by microenvironment and hour

Detailed Description of Program Changes

ODEST

Previously, a sequence of four programs (TVLTIME, DIST, ODEST, and TTFRA) had to be run to carry out the commuting algorithms in HAPEM. While no changes in functionality have been made to this section of HAPEM, it was found that for large cities these programs were very slow in execution. Investigation showed that both DIST and ODEST wrote very large files that were subsequently read back in to the next program and were not needed by any other HAPEM programs. These files each had a number of records equal to the square of the number of census tracts in the study region. By combining the code of the last three commuting programs, a substantial reduction in computing time was achieved (for New York the reduction is from nearly one hour on the IBM mainframe to just four minutes). The program now called ODEST uses the same input files (TVLTIME and DISTRICT) and output file (HOMEWORK), with the same format, as the old set of three programs used, but without creating any intermediate files.

POP90

Two features have been added to the POP90 program: the retention of county information and the ability to process data for all demographic groups in one run. The program reads in census tract information which includes the state and county FIPS (Federal Information Processing Standard) codes as part of the tract ID. A list of distinct counties is created and sorted in FIPS order. In order to process all the demographic groups, the majority of the POP90 program was put into a long loop which is run once for each group. The final step in this loop is to rewind all the input files for re-use for the next demographic group. This was the easiest way (i.e. least likely to introduce new programming errors) to modify the program, although it would be more efficient to read the files just once and simultaneously apply the information to all the demographic groups.

MECONC

This is a short program that combines the ambient air data from AQAVG with the microenvironmental factors to calculate CO concentrations in each microenvironment. This task was previously carried out in the MERGE program, which is no longer used. It would also be possible in future to combine this task with the EXPCODI program, thereby reducing the number of programs by one. This was not done at present because of the need to test the new programs using a special MECONC file. This file (called DUMMY) had all microenvironmental concentrations set to a single value (1.0 ppm), so that errors in the exposure calculation could be easily detected. The results from this testing are described in the quality assurance chapter.

DURAVG

This program averages the microenvironmental durations of all individuals in a demographic group. The criteria for demographic group membership are defined in this program and applied to the records in the time-activity database. The results are used by the EXPMEHR program to calculate exposures at the hourly and microenvironmental level. Note that DURAVG is the only HAPEM program that needs to be run only once, as it applies to all study areas as long as the set of demographic groups does not change. In fact, the output file (stored as AVMEDUR.DATA) can simply be regarded as an additional part of the activity database.

EXPCODI

This program first calculates exposure-days of data for each air district for each person-day in the time-activity database. Separate home and work exposures are calculated for each air district, which are later combined with the commuting data to produce cohort exposures. This procedure reduces the number of exposures calculated for each person in the time-activity database from the square of the number of districts to just twice the number of districts. In previous sets of runs, HAPEM used the person-day as the unit for assembling a year-long time-activity time series by stochastic resampling. The EXPCODI program allows the results of this process to be evaluated by direct calculation, since the mean and variance of a random sample can be calculated directly from the characteristics of the universe that the sample is drawn from. These steps obviate the need to perform repeated runs of HAPEM, thus saving time, computer resources, and eliminating the need for extensive post-processing.

The daily exposures are then summed by demographic group, home district, work district, quarter of the year, and time-activity state (weekend/weekday, etc.). The data are combined with the output from the POP90 program (the POPALL files) and the results of the commuting algorithm (the HOMEWORK files) to determine the number of people in each category. This allows both the creation of the exposure tables (by county and district) and the calculation of population-weighted averages. Each table reports the mean and standard deviation for the average exposure rate (in micrograms per cubic meter) for each quarter and annually, for all the demographic groups. There is one such table for each county and one for each air district in the study area. This program also produces a similar single city-wide average table. The tables are stored electronically and can easily be imported into a word processor for manipulation and printing. The county tables are stored in a directory named EXPCO and the district tables are stored in

EXPDI.

EXPMEHR

This program is similar to the EXPCODI program, except that instead of printing tables by geographical divisions, it creates tables of exposure for all demographic groups by temporal (hour of the day) and microenvironmental divisions. For the hourly tables, the mean exposure and the percentage of daily exposure obtained each hour are reported. There are two types of microenvironment tables: one reporting the total accumulated exposure (in $\mu g/m^3$-year) along with the percentage of the total obtained in each microenvironment; and the other reporting the average concentration and duration in each of the microenvironments. Both of these tables pertain to all persons in the study area. Note that due to the way in which year-long time-activity sequences are assembled in HAPEM, the calculation of a variance for these tables would be difficult since there is significant covariance between different hours or microenvironments. That problem is beyond the scope of the current project. The hourly tables are stored in a directory named EXPHR and the microenvironmental tables are stored in EXPME.

Job Control Language (JCL) Files

To perform the HAPEM runs for the fourteen study areas, each of the 9 HAPEM programs (not counting DURAVG) had to be run 14 times each, a total of 126 small jobs on the IBM. (DURAVG was only run once). Also, each of the programs needs compilation before being run the first time. These small jobs can be batched together in one of two ways: either a single program is compiled and run 14 times (once per study area) in a single job, or else all the programs could be applied to one study area in a single batch job. The former method is preferable if the programs are still under development and testing. The latter method would be preferable if the code is not being changed, and the user wished to apply HAPEM to another city. Since several of the programs had to be modified and checked during the course of this project, the former method of organization is used. All of the programs except ODEST and EXPCODI can process all 14 study areas in less than five minutes of CPU, which is a system cutoff used for the fast processing queue. ODEST can be run for all 13 cities excluding New York in less than five minutes, and New York alone takes another four minutes. The EXPCODI program takes about 65 seconds per study area (independent of population), so four study areas can be run at a time in the fast queue. All told, the 14 study areas require a total of 37 minutes of CPU to compile and run the complete set of programs.

Data Handling and Data Management

The HAPEM model requires several types of data as inputs to the exposure calculations. Hourly ambient CO monitoring data, hourly ambient temperature data, U.S. census population data, and duration of activity by microenvironment data from the time-activity database are processed by HAPEM programs into forms that are used in the exposure calculations. The sources, handling, and management of this raw data are discussed in this section.

File Structure

With fourteen different study areas and the potential for running multiple years within study areas (not carried out under the present task), the problem of organizing data is crucial. The data that vary with study area or time were placed in partitioned data sets on the IBM mainframe. Each study area was labeled with a three character code derived from its name, and for time varying data the year was added as a suffix. For example, the data for Denver for 1990 are found in members named (Den90). All HAPEM files now follow this convention.

STUDY AREA	STATE(S)	ABBREVIATION
Baltimore	MD	BAL
Boston	MA,NH	BOS
Chicago	IL,IN	CHI
Denver	CO	DEN
Houston	TX	HOU
Los Angeles	CA	LAX
Minneapolis / St.Paul	MN,WI	MSP
New York City	NY,NJ,CT	NYC
Philadelphia	PA,NJ,DE	PHL
Phoenix	AZ	PHX
San Francisco	CA	SFB
Spokane	WA	SPO
St. Louis	MO,IL	STL
Washington D.C.	DC,VA,MD	WDC

Air Quality Data

Hourly average CO monitoring data were obtained from the U.S. EPA Aerometric Information and Retrieval System (AIRS) by means of interactive menus on the U.S. EPA mainframe IBM computer. Initially, all of the hourly average CO monitoring data for the entire U.S. for the calendar year 1990 was extracted from AIRS in EBCDIC files maintained on the mainframe computer. The extracted files contain coordinate information on the geographic location of monitors that was used to obtain specific monitoring data for a particular urban area. A summary of the hourly average CO monitoring data for calendar year 1990 that was used in each urban area is shown in Table 1.

Table 1 shows the urban area, AIRS monitor ID, the UTM zone and coordinates, the distance form the city center, the number of valid hourly average CO values for calendar year 1990, the minimum, mean, and maximum hourly average concentrations for the year, and the average concentration as calculated by the AQAVG program in HAPEM. The first line for each city (monitor 0) is not a monitor but is the location of the designated city center. The city center locations were generally the same ones used in earlier HAPEM runs (HAPEM-MS2 runs for the year 1988). The San Francisco city center was moved slightly to allow inclusion of the monitors in Contra Costa county, and new city centers were designated for Baltimore and Chicago, which were not run previously. The number of valid observations and the maximum hourly average concentration were compared to values in the AIRS data base using an online browse facility in AIRS. These matched exactly in the extracted data and the data shown in the AIRS browse facility. Only those monitors with at least 75% data capture were used in the HAPEM runs. Also, in the case of co-located monitors, only the monitor with the fewer number of missing values was used. The deleted monitors are marked with SITENUM='x' and AQAVG='xxxx'. The AIRS formatted EBCDIC files of hourly average CO data are used directly in the program TSERIES in HAPEM. The AQAVG results are printed here in part as a check that the data really are for the intended monitor. The monitor numbers and locations must be specified in the

DIST90 program, which does not examine the AIRS files directly. The TSERIES program processes all data on the AIRS files without checking the monitor numbers or locations. It is therefore very easy to create incompatible lists in the two sections of HAPEM. The user must therefore be careful to check that the monitor lists agree in both number and order. A mistake in number is usually manifest as a monitor with all zero concentrations. A mistake in order results in two monitors having each other's data (and hence mean concentration). These problems were checked for and eliminated from the model runs.

Microenvironmental Factors

The microenvironmental factors are used to estimate concentration for each of the 37 HAPEM microenvironments, which are modeled as linear functions of the ambient concentration. The factors used for the present HAPEM runs were a modified version of factors used in earlier runs. The major change is that the additive terms were all set to zero in the current runs, to eliminate non-ambient sources. A second change from previous sets of factors is that the four residential microenvironments have been combined, using a common factor which is the average of the four separate factors used previously. The factors used for the current set of model runs are listed in Table 2.

Meteorological Data

Hourly average temperature data was obtained in hourly format for each of the fourteen urban areas for calendar year 1990. A PC-SAS program called MET.SAS was written to calculate daily mean and maximum and write these values to files in the format used by HAPEM. The files were then transferred to the IBM mainframe as members of a partitioned data set.

The only function of the meteorological data in HAPEM is to select activity patterns based on a distinction between 'warm' and 'cool' days. The winter and summer seasons have different definitions: warm days start at 55°F and 84°F in winter and summer, respectively. The frequency of warm and cool days for each calendar quarter is given in Table 3 for each of the fourteen study areas.

Population Data

The 1990 U.S. tract level population data was provided by the EPA in the form of three files per study area. (There were too many variables for their software to handle on a single file). A PC-SAS program called ALLCITIES.SAS was written to read these files and reorganize the data into 25 output directories as required by HAPEM. Each of these directories contained fourteen files (one file per study area). Upon transfer to the IBM, each directory became a partitioned data set and each file became a member.

The population files did not always completely cover the intended study area. The three main areas that were missing were Union County NJ (part of NYC); Alexandria and Fairfax cities in Virginia (part of WDC); and the Illinois counties near St. Louis (STL). The HAPEM model processed all available data. The missing data do not affect the estimates for the other counties in their study areas at all for the 18 non-commuting demographic groups, and have at most a small effect on the commuting patterns for the remaining five demographic groups since the central city areas were not missing in any of these cases.

The populations of specific demographic groups appear in the county and district tables in Appendices A1-A14. These values are consistent with the census definitions for all but the three demographic groups (outdoor children, outdoor workers, and heart and respiratory problems) that are not directly based on census data. Some explanatory notes regarding the calculation and meaning of these populations are provided in the results section.

Table 4 summarizes the population data used in the model runs. For each county, the total population from the census, and the population in the portion of the county inside the study area are given. A flag is also provided. There are two criteria which may set the flag: a) if HAPPOP is less than half of the census population, and b) if the population in HAPEM monitor districts 1-18 totals less than a quarter of the census population. If either condition is met then the flag is set to one. Table 4 is also used to report average exposure level by county, which is a modeling result and is discussed in a later section of the report.

Time-Activity Data

The time-activity database containing duration by microenvironment was supplied with the current version of HAPEM. No additional activity survey data were added (or deleted) for the current runs, but the existing data were reorganized into array form of microenvironment by hour by home-work, while standardizing the record lengths for all the person-days. This allowed for a simplification of the programs to read and process data that were already in array form. There are 3568 person-days of data. The time-activity file is named MEDUR.DATA. As one of the first steps in the model runs, the program DURAVG was run to calculate averages for each demographic group. This summarized activity data is stored on the file AVMEDUR.DATA.

Results

The foregoing sections of this report detail the changes and enhancements to HAPEM that were carried out for the calculation of population exposure to CO in fourteen urban areas. The major result of the modifications to HAPEM are twofold. The range of calculations that can be carried out by HAPEM has been extended in certain directions (e.g. calculation of exposure by microenvironment). The second result is that file management of HAPEM output files can be carried out in an efficient manner. Since there are no longer multiple runs for simulation of specific sequences of daily activity patterns over a quarter or year, the number of output files to be managed in any run of HAPEM has been reduced significantly. Results for multiple cities can be shared in partitioned data sets so that single data sets are used to accumulate outputs from HAPEM.

Detailed CO exposure results from the HAPEM runs are given in tables in Appendices A1 - A14 to this report; one section for each of the fourteen study areas. For every study area the following tables are provided:

1) city-wide exposure by quarter and annually for all demographic groups,
2) county level exposure by quarter and annually for all demographic groups,
3) air district tables of exposure by quarter and annually for all demographic groups,
4) a table of percentile distribution of annual citywide exposure for all persons,
5) a table of fractiles of annual citywide exposure for all persons,
6) a table of citywide accumulated exposure by microenvironment for all persons
7) a table of citywide average concentration and duration by microenvironment,
8) a table of citywide average hourly exposure for all persons (diurnals).

In addition, a map of each study area is provided in Appendix B.

Except for the tables of accumulated exposure by microenvironment, all the other exposure tables report exposure in units of micrograms per cubic meter (µg/m3). This is not strictly an exposure, which should have the units of (concentration x time), but an exposure rate. It represents the time-weighted average concentration to which the group is exposed over the time period of interest. HAPEM has traditionally expressed exposures in this way, which allows for easier comparisons between different time periods, such as hourly, daily, quarterly, and annual exposures. The exception to this is the table of accumulated microenvironmental exposures, which depend heavily on duration and indicate the relative contributions of each microenvironment to the total. The second table showing microenvironmental breakouts indicates the average concentrations and durations that compose the microenvironmental exposure.

Exposure Results

Table 4 contains county level populations and annual exposure levels for all persons. The census population (for the entire county) is given in the CENSUS column. The population of the part of the county included in the HAPEM run is in the column labeled HAPPOP. For counties which are only partly in the study area this number is less than the total county population. The flag indicates whether or not the county passes both of the population criteria. The first of these is that at least half of the county population be included in the study area. The second is that at least a quarter of the county population be within 10 km of an ambient monitor (i.e. in districts 1-18). If both conditions are met then the FLAG variable is set to zero and the exposure estimates on the maps in Appendix B are colored in green. If either condition is not met then FLAG=1, and the bar on the map is colored yellow.

City-Wide and County-Wide Tables

The city-wide and county-wide tables share the same format; both provide the mean and the standard deviation of CO exposures in micrograms per cubic meter (µg/m3) by calendar quarter and annually for the twenty three demographic groups considered for this report. In addition, the population numbers as used in HAPEM for each demographic group are listed in a separate column of the output. In the header on each page are the total county population, the population in the area used by HAPEM, and the population assigned to district 19 in HAPEM (which is the district composed of all tracts not within 10 km of any air monitor). The larger the HAPEM population and the smaller the District 19 population, the more reliable are the exposure estimates. At the end of the header the estimates either pass or do not pass the population criteria (this is the flag discussed in Table 4). At the bottom of each page the final five lines identify the demographic group with the highest exposure in each quarter and annually. There is one city-wide table for each study area, and one table per county (from 1 to 18 additional tables per study area).

Air District Tables

The HAPEM model divides the study area into a set of non-overlapping air districts, which are composed of census tracts located within 10 km of an ambient carbon monoxide monitor. There is one air district per monitor, plus an additional district (# 19) which is assigned the city-wide average concentration (i.e. the average of all the monitors). The tables by air district have the same format as the city-wide and county level tables, showing the mean and standard deviation of exposure both quarterly and annually, along with the identification of the demographic groups with the highest exposures. There are from five to seventeen air districts in each study area.

Hourly (Diurnal) Exposure Tables

A table of the calendar quarter and annual average CO exposures by hour of the day is included for each study area. The table applies to all persons in the study area.. The table reports the average hourly exposure in µg/m3 and the percent of the daily total exposure obtained during that hour. If exposure were constant, each hour would contribute 4.17% of the daily total. Hour 1 in the table corresponds to the hour between midnight and 1 a.m., and so on. All times are local and refer to daylight savings time when applicable.

The hourly tables do not include an estimation of the variance due to selection of activity patterns. Unlike the quarterly and annual totals, it is difficult to calculate the variation for the hourly exposures due to the significant covariance between hours. If estimates of variance were desired, then the easiest way would be to use a Monte Carlo approach involving multiple runs to gauge the extent of this variation.

Microenvironmental Exposure Tables

The exposures per quarter and annually by microenvironment are presented in two tables for each study area. The first table gives the accumulated exposure and the percentage of the total across all microenvironments. This table is useful in distinguishing the relative importance of the various microenvironments to the long-term total exposure. The second table reports the average concentrations and durations in each microenvironment. As discussed in the uncertainty section of the report, these tables should be used with caution as the microenvironments with long total durations are based on relatively more sample data, and hence have more reliable durations and concentrations than do those with short total duration. The HAPEM model is designed to estimate the total exposure over a year (or quarter), and these totals have less relative error than the total for any one microenvironment. As for the hourly tables, due to covariance in exposure among microenvironments, it is difficult to calculate variances computationally. If an estimate were necessary, then a Monte Carlo approach could be used although this would substantially increase the computer time needed for the modeling.

Percentiles and Fractiles

Characterizations of the distribution of exposure to CO across a study area for a single demographic group (All persons) are presented in tables of selected percentiles and fractiles. Since the city-wide annual CO exposure represents a population weighted average across all air districts, the variation in annual CO exposure of differences between air districts due to air quality are to a large extent averaged out. This leads to a very small standard deviation for the city-wide annual CO exposure. Thus, the distribution based on the city-wide annual CO exposure is very narrow and does not capture the range of annual CO exposures seen across air districts. To better characterize the distribution of annual CO exposures across air districts in a study area, the following procedure was adopted.

A normal distribution of annual CO exposure was constructed for each air district in a study area using the mean and standard deviation of annual CO exposure as calculated for the air district by HAPEM. The total mass of the normal distribution for a particular air district was set to be proportional to the percent of the total HAPEM population represented by the air district. The sum of the proportionality factors across all air districts (i.e. for the study area) was set to unity. Thus a population weighted distribution for annual CO exposure was constructed for the whole study area.

The population weighted distribution for annual CO exposure was used to calculate percentiles and fractiles for annual CO exposure in the usual manner. That is, for the fractile calculation, the level of the annual CO exposure was set and the percent of the population weighted distribution at or below the fixed level was computed by summing the contributions (in percent) from the portions of the individual air district distributions at or below the fixed level of annual CO exposure. Similarly for the percentile calculation, the desired percentile was fixed. The total percent of the annual CO exposure was accumulated across air districts for increasing levels of annual CO exposure until it matched the fixed percentile. The value of the annual CO exposure at this point is taken as the value of the fixed percentile.

The procedure described above for the construction of percentiles and fractiles for annual CO exposure across a study area is not to be interpreted as characterizing the distribution of annual CO exposure across individuals in the study area. The calculations pertain to variations between the annual CO exposures computed from the annual activity sequences as composed in HAPEM. As discussed in the chapter on uncertainty analysis, other ways to construct annual activity patterns can be considered that better represent activity patterns for individuals.

Percentile Tables

A table of percentiles of the city-wide annual exposure distribution for each demographic group is presented for each study area. The values in these tables indicate the average concentrations below which certain proportions of the demographic group population are exposed. The points represented in these tables are the quartiles and the 90th, 95th, and 99th percentiles of the distribution. Most of the variation within a demographic group arises from the distribution of the group among the various monitor districts. Within any one district, HAPEM does not adequately capture the true extent of the variation among individuals, for reasons discussed elsewhere in this report. It is not correct to conclude that 99% of the individuals in a study area are exposed below the level in the table; rather that for a particular demographic group, assigned to a particular air district with a probability based on the relative population of that district, there is a 99% chance that the exposure for that group would be below the stated concentration.

Fractile Tables

A table of fractiles of the city-wide annual exposure distribution for each demographic group is presented for each study area. These differ from the percentile tables in that the fractile tables show the percentage of the demographic group exposed at or below fixed points in exposure, whereas the percentile tables showed the exposure levels for fixed percentage points in the distribution. The fractile tables indicate the fractions at or below each 100 $\mu g/m^3$ increment. All the groups in all study areas were exposed at or below 2000 $\mu g/m^3$. The remarks regarding the individual variation within monitor districts made in the above paragraph on percentile tables also pertain to the fractile tables.

Data Visualization

A series of fourteen maps (one per study area) are presented in Appendix B. Each map indicates the extent of the study area (generally, census tracts within 50 km of the center of the study area), the population density (in shades of red), the monitor locations (numbers in blue boxes), the extent of the monitor districts (outlined in blue), the county boundaries, major roads and bodies of water, and the annual average exposure level for all persons by county (vertical bars). The exposure levels use green bars for estimates that pass the population criteria and yellow bars for those that do not pass (these are usually counties further from the city center). All of the maps are drawn to the same scale and use the same shadings for population density and the same height scale for the exposure bars, to allow for comparisons between maps. The center of the study area is marked and a 50 km radius circle is drawn to indicate the potential extent of the study area. Not all counties within 50 km of the city center were included in each designated study area, usually because the census definition of the MSA or CMSA did not include them. In a few cases (e.g. Union county in NJ and Alexandria VA), areas which should have been included were not because the tract level population data were not available at the time of the model runs.

Additional Results on the CD-ROM

In addition to the tables in this report, the CD-ROM contains tables of hourly exposures, microenvironmental exposures, concentrations, and durations, and percentile tables and fractile tables for each demographic group in each study area. All of the tables are provided in ASCII files, without headers and page breaks, ready to read into a program or spreadsheet. In addition, all of the tables printed in Appendices A1-A14 of this report are available in electronic form on the CD-ROM as WordPerfect files. Descriptions of the directory structure and file formats are to be found in README files on the CD-ROM.

Discussion

Results on population exposure to carbon monoxide from ambient air were generated using the HAPEM-MS3 model as applied to 23 demographic groups in 14 urban areas. The changes made in the HAPEM computational method have allowed for a more efficient use of computer resources, with the result that a very extensive set of output tables can now be generated with relative ease. The difficulty has now shifted from the production of the results to the interpretation of the results. It is not possible to examine all the tables in detail; it is not even practical to print all of the tables that were generated. However, the salient points relating to each type of table are discussed below.

City-Wide and County-Wide Tables

The city-wide tables provide a comparison of results across demographic groups and quarters. A general result is that the outdoor worker group usually had the highest exposure in all study areas. This is reasonable, since these model runs were intended to represent exposure to ambient air, and this exposure is naturally higher outdoors than indoors. Another general result is that the first and fourth quarters have higher exposures than the second and third quarters. This is a well known property of carbon monoxide concentrations (due to slower dispersion in winter) and it is expected that exposure patterns would be similar. In all cases, the computed variances in exposure are found to be smaller than those from previous runs of the HAPEM model. This is due to the choice of microenvironmental factors and is discussed at greater length in the sensitivity analysis report. In short, the absence of additive factors results in less difference in concentration between two microenvironments, and hence less variation in exposure due to differences in activity patterns.

All the population numbers in the tables except for three specialized demographic groups (outdoor children, outdoor workers, and heart and respiratory problems) are derived directly from the 1990 census. The numbers pertain to those tracts included in the study area, not necessarily the entire Metropolitan Statistical Area (MSA). There are two main breakouts which should include each person exactly one: the five income categories, and the eleven categories from children age 0-17 to women age 65+. The totals in these categories are very close to (but not always exactly equal) the population of all persons. Possible reasons for this are a) the manner in which missing responses are handled, b) the extrapolation of information obtained only on the long census form (given to 1/6 of the households) to the rest of the population, and c) rounding errors when percentages are applied to the population. The first two demographic groups are Caucasians and African-Americans. The census also considers other races such as Asian, American Indian, and Other, but HAPEM does not have enough time-activity data for these groups to calculate a valid exposure estimate. The third group in the tables is the Hispanic group. The census bureau asks respondents if they consider themselves Hispanic in a separate question from the question on race. Therefore, Hispanics might also be counted in one of the race categories (e.g. Caucasian, or African-American, or Other) as well. Due to the missing race groups, and the possible double counting of Hispanics, the first three demographic groups may numerically add up to either less or more than the all persons group. The three groups not supplied by the census (outdoor children, outdoor workers, and heart and respiratory problems) were defined for previous HAPEM runs carried out by IT Corporation. The fractions of outdoor children and of people with heart and respiratory problems are functions of age. These are aggregated over the known age distributions for each study area. The fraction of Outdoor Workers as a function of occupation was also determined by IT Corporation, using data for Los Angeles in 1994. These fractions were applied to the census data on occupational category for each study area.

The county level tables exhibit the same patterns in exposure as the city-wide tables. The primary difference between counties is the overall level of exposure. The relative ranking of demographic groups and of quarters is generally the same across counties. This is due to the use of the same activity patterns, the same microenvironmental factors, and the same meteorological data in all counties in a study area. The only difference is in the ambient air quality, which acts as an overall scale factor which applies to all microenvironments and demographic groups. Some differences between demographic groups arise at the county level due to different populations per air district (hence different county-wide averages), but the general rankings of the demographic group exposures are similar in most counties.

The demographic group with the highest exposure was the Outdoor Worker group, while the Heart and Respiratory Problem group generally had the lowest exposure. These results are due to differences in the activity patterns. The microenvironmental factors are generally higher for outdoor locations as compared to indoor ones. Comparisons between groups based on race or income differences tend to be small in counties with only one air district, indicating that exposure differences due to activity patterns alone are not very great. For counties with two or more air districts, larger differences are seen, reflecting the geographical separation of the groups. Both the individual counties and the citywide averages show that males have consistently higher exposures than females. This is almost certainly due to the fact that men spend more time outdoors and in travel than women do on average. Finally, for suburban counties such as Adams and Arapahoe in the Denver study area, the working groups had higher exposures than the similar non-working groups. This is to be expected since many of the workers commute to the relatively higher exposure downtown area.

As an example of county comparisons within a study area, in Denver the central county (Denver county) has the highest exposure for all demographic groups. The lowest exposures are found in Boulder county, which is at a higher elevation and is less urbanized than Denver. Adams, Arapahoe, and Jefferson counties all extend from suburban and rural areas almost to the heavily urbanized downtown section of Denver. As might be expected, these counties had similar exposures, which were lower than those for central Denver. The sixth county in the study area was Weld county. Only a small portion of this county was included in the HAPEM run, all in district #19, so the estimate is close to the citywide average. This is due to HAPEM's internal assumptions. The true exposure level for Weld county is not known. For the Denver study area there was only one county predominantly in district #19, but in other study areas there were often several such counties. A flag has been created to indicate such counties, and applied to the tables and maps. In total, 47 out of 102 counties pass this test (indicated using green bars on the maps).

Air District Tables

Most of the comments made regarding the city-wide and county-wide tables also apply to the air district tables. Within air districts there are no longer any differences between demographic groups based on location (as there were in the county and city tables), so the ranking of the demographic groups is even more similar across air districts than the ranking across counties. Some differences in ranking still exist because the groups average a different amount of time in each microenvironment, and each microenvironment has a different dependence on the ambient concentration. The results for air districts show (as for counties) that the Outdoor Workers were usually the group with the highest exposure and the Heart and Respiratory Problem group had the lowest exposure.

Hourly (Diurnal) Tables

The sequence of 24 hourly exposures characterizes the diurnal exposure pattern. While the CD-ROM contains one table for each demographic group, for brevity only the table for all persons is printed in this report. The diurnal patterns generally follow a standard pattern, with a sharp peak around 8-9 a.m. and a somewhat lower but broader peak from 7-11 p.m. These patterns resemble standard diurnal CO ambient concentration profiles primarily due to traffic patterns. The two traffic driven peaks are stronger in the winter (first and fourth quarters) than in the summer. In the second quarter the evening peak almost vanishes. This is due to the relatively late sunsets in May and June, which allow time for the evening rush hour emissions to disperse before the strong solar-driven convective mixing ceases.

Microenvironmental Tables

As with the hourly tables, the full set of microenvironmental tables is available on the CD-ROM. The tables for the group of all persons are printed in this report. There are two tables for each study area: a table of contributions to total (accumulated) exposure by microenvironment, and a table of average concentrations and durations in each microenvironment. It should be noted that the tables reflect a population weighted average for all the people in the demographic group, not for any one individual.

As expected, the highest proportion of total exposure comes from the in-home microenvironments, due to the great amount of time spent there. Next in importance are travel (mostly by car) and exposure in an office or school. The relative contributions of the various microenvironments indicate the ones that dominate the determination of HAPEM results. Therefore, any attempt to reduce the uncertainty in the HAPEM model by improving the characterization of the microenvironments will have a larger impact if it focuses on these dominant microenvironments. The differences in relative contributions (i.e. percent of total) by quarter are not very large.

The second table illustrates the relative contribution of concentration and duration to the microenvironmental exposure. Generally, the microenvironments with high exposure also have long durations, meaning that people spend a lot of time there. The variations in concentration across microenvironments are relatively small, and are mostly due to the differences in their microenvironmental factors. The most dominant microenvironment in accumulated exposure is the home, despite having a moderate average concentration. The high concentration microenvironments generally had relatively short durations and they did not dominate the exposure results. A caution should be made here against relying on the accuracy of either the concentration or the duration for those microenvironments with small accumulated exposures: the time-activity database poorly samples those microenvironments and the microenvironmental factors are generally based on fewer data points than for the more common microenvironments. The numbers which are most reliable are those which contribute more heavily to the overall total. Note also that these concentrations reflect only the portion of CO in that microenvironment attributable to the ambient air and does not include local sources.

Visualization of Output

For each of the study areas a map is presented in Appendix B displaying the population density at the tract level, the outlines of the counties and the air districts, with a vertical bar in each county representing the annual exposure estimate for all persons. The maps generally show that the population density has declined to low levels by the 50 km boundary except for the very largest cities (Los Angeles and New York). In a few cases (particularly for Union County NJ, Alexandria VA, and the counties in Illinois near St. Louis), population data were not available for areas that would otherwise have been included in the studies. These areas have no impact on the exposure estimates for other counties in the same study areas, apart from a slight effect on the five commuting demographic groups.

A general feature of the fourteen areas is the concentration of monitor locations near the city center. This is pronounced in all cities except New York, Los Angeles, and Philadelphia. In Baltimore and Washington, for example, all the monitors are close to the central cities and there are no monitors at all between the two cities. The effect of this is that the exposure estimates for the central counties tend to be based primarily on monitor data (green exposure bars), whereas the estimates for the outer counties are often based on only a portion of the tracts in the county, and often these tracts are not particularly close to an ambient monitor (hence they are put in district 19). The result is that most of the outer counties in each study area have estimates which probably do not well represent the counties. These have been flagged and colored with yellow exposure bars on the maps. Due to the predominance of district 19 in these counties, the estimates are often at or very close to the overall city-wide mean exposure. In order to make exposure estimates for the population in these (sometimes extensive) regions far from ambient monitors, a method either of interpolation between monitors (or of modeling the distribution of CO in each study area) would be required. That task is beyond the scope of this report.

Summary

The enhancements to the running of the HAPEM model now allow the analyst to generate a very large amount of data in a relatively short time. This is reflected in the number of output tables from the model runs. From this point on, the interpretation of the results may consume more time than the production of the actual runs themselves. This would be the case not only for the present set of fourteen study areas, but also for any new study areas, despite the need of the latter for data pre-processing prior to running the model. Running HAPEM for other years besides 1990 would not be difficult, as long as the same (1990) census data could be used. The main limitation to applying the model to other study areas is that most new areas would have fewer CO monitors than the fourteen areas already examined. There are over 3100 counties in the country but only around 500 ambient CO monitors, so most of the counties are not monitored at all.

The results generally show consistent patterns across most study areas. The first and fourth quarters (winter) have higher exposures than the other two quarters. The Outdoor Workers generally are the group with the highest exposure in each study area. The central cities usually have higher exposures than the suburbs and are also more heavily monitored. The hourly exposure profiles generally indicate traffic patterns, with peaks at and following the morning and evening rush hours. The microenvironmental tables indicate the home, office, school and travel are the major sources of CO exposure for the demographic groups considered in this report.

TABLE 1

List of ambient CO monitors (from AIRS) operating in 1990 within 50 km of the designated city center.
Monitors with less than 75% data capture (6570 hourly values) were not used in HAPEM runs.
For co-located monitors, the one with greater data capture was used.
The average concentration of the filled-in time series used by HAPEM is given in under AQAVG.
Monitors not used in HAPEM runs have SITENUM=x and AQAVG=xxxx.
All concentrations reported in parts per million (ppm).

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
BALTIMORE	1	CENTER	0	18	4350.1	360.7	0.0	0	0	0	0	
BALTIMORE	1	240053001	1	18	4352.10	372.880	12.3429	6866	0.0	0.88080	15.00	0.89
BALTIMORE	1	245100018	2	18	4352.67	360.911	2.5816	7275	0.0	1.04396	10.80	1.05
BALTIMORE	1	245100034	3	18	4350.05	361.008	0.3120	7223	0.1	2.19569	16.30	2.18
BALTIMORE	1	245100040	4	18	4350.84	361.621	1.1821	7084	0.0	1.26128	12.40	1.26
BOSTON	2	CENTER	0	19	4691.5781	330.6902	0.0	0	0	0	0	
BOSTON	2	250170007	1	19	4723.77	310.489	38.0054	8084	0.0	1.25803	15.00	1.26
BOSTON	2	250250002	2	19	4690.37	327.095	3.7918	7842	0.0	1.30635	11.00	1.30
BOSTON	2	250250016	3	19	4692.50	332.000	1.6017	8309	0.0	1.18332	9.00	1.18
BOSTON	2	250250021	4	19	4693.53	333.008	3.0308	8646	0.0	0.86173	12.60	0.86
BOSTON	2	250250038	5	19	4691.50	330.840	0.1689	7800	0.0	1.27108	11.20	1.27
CHICAGO	3	CENTER	0	16	4631.00	444.500	0.0	0	0	0	0	
CHICAGO	3	170310037	1	16	4647.52	444.485	16.5180	7546	0.0	0.95400	17.45	0.96
CHICAGO	3	170310063	2	16	4636.10	447.365	5.8461	8642	0.0	0.96092	8.30	0.97
CHICAGO	3	170311002	3	16	4607.11	453.538	25.5453	8320	0.0	0.63834	9.62	0.64
CHICAGO	3	170313101	4	16	4645.04	426.538	22.7969	8527	0.0	0.93852	9.50	0.94
CHICAGO	3	170314002	5	16	4633.76	437.541	7.4874	8276	0.0	0.78852	8.87	0.79

TABLE 1 (continued)

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
CHICAGO	3	170314004	6	16	4649.87	427.539	25.3723	8327	0.0	0.73015	12.40	0.73
CHICAGO	3	170316004	7	16	4635.69	431.200	14.1044	8298	0.0	1.54791	10.04	1.54
CHICAGO	3	180890015	8	16	4608.45	461.570	28.2823	7339	0.3	0.84928	10.60	0.85
CHICAGO	3	180890021	9	16	4603.90	471.920	38.5521	8149	0.3	1.03940	8.90	1.04
CHICAGO	3	170313102	x	16	4646.61	427.385	23.1679	2086	0.1	0.77013	5.20	xxxx
CHICAGO	3	170313601	x	16	4649.58	427.154	25.4163	2152	0.0	0.65493	5.00	xxxx
CHICAGO	3	170314005	x	16	4650.54	424.731	27.7947	2012	0.0	0.60224	4.10	xxxx
CHICAGO	3	170431003	x	16	4644.16	423.025	25.1881	2142	0.0	0.64683	4.50	xxxx
DENVER	4	CENTER	0	13	4399.1328	500.446	0.0	0	0	0	0	
DENVER	4	080013001	1	13	4409.70	504.364	11.2730	8242	0.0	1.02224	11.80	1.02
DENVER	4	080050002	2	13	4379.69	503.673	19.7078	8587	0.0	0.47296	5.30	0.47
DENVER	4	080050003	3	13	4389.52	500.161	9.6210	8686	0.0	0.96513	12.50	0.97
DENVER	4	080130009	4	13	4446.05	491.315	47.7975	8552	0.0	1.28352	14.90	1.28
DENVER	4	080131001	5	13	4429.02	477.219	37.8547	7859	0.0	0.65810	14.80	0.67
DENVER	4	080310002	6	13	4399.95	501.084	1.0383	8594	0.0	1.52999	37.90	1.53
DENVER	4	080310013	7	13	4398.56	505.195	4.7833	8523	0.0	1.64006	19.40	1.64
DENVER	4	080310014	8	13	4400.00	497.360	3.2055	8158	0.0	1.21747	12.70	1.22
DENVER	4	080590002	9	13	4405.40	491.500	10.9229	8684	0.0	1.08588	13.40	1.09
HOUSTON	5	CENTER	0	15	3294.033	271.4929	0.0	0	0	0	0	
HOUSTON	5	482010024	1	15	3309.92	275.287	16.3377	7419	0.0	1.06419	8.30	1.06
HOUSTON	5	482010047	2	15	3302.86	259.485	14.9044	8124	0.2	1.24307	14.60	1.23
HOUSTON	5	482011034	3	15	3295.34	285.150	13.7199	7599	0.0	0.66434	9.60	0.66
HOUSTON	5	482011035	4	15	3291.19	281.710	10.6050	8489	0.0	0.60836	8.40	0.61
HOUSTON	5	482011037	5	15	3293.36	271.646	0.6863	8225	0.2	1.17012	10.50	1.16
LOS ANGELES	6	CENTER	0	11	3769.4038	384.9932	0.0	0	0	0	0	
LOS ANGELES	6	060370002	1	11	3777.41	414.916	30.9746	8327	0.0	1.18770	7.00	1.18
LOS ANGELES	6	060370113	2	11	3768.55	365.679	19.3330	8024	0.0	1.26682	15.00	1.26
LOS ANGELES	6	060371002	3	11	3782.27	378.704	14.3211	8356	0.0	2.28303	16.00	2.28
LOS ANGELES	6	060371103	4	11	3770.11	385.394	0.8129	8361	0.0	1.89128	13.00	1.89
LOS ANGELES	6	060371201	5	11	3785.14	358.722	30.6220	8161	0.0	1.79292	19.00	1.81
LOS ANGELES	6	060371301	6	11	3754.77	388.186	14.9820	8117	0.0	2.64137	24.00	2.65

TABLE 1 (continued)

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
LOS ANGELES	6	060371601	7	11	3763.95	402.174	18.0269	8345	0.0	1.72654	13.00	1.74
LOS ANGELES	6	060371701	8	11	3769.64	430.714	45.7214	8367	0.0	2.03645	13.00	2.03
LOS ANGELES	6	060372005	9	11	3771.76	397.769	12.9904	8332	0.0	1.76776	16.00	1.76
LOS ANGELES	6	060372401	10	11	3753.97	405.280	25.4933	8340	0.0	1.76978	12.00	1.77
LOS ANGELES	6	060374002	11	11	3743.04	390.002	26.8354	8278	0.0	1.95603	11.00	1.96
LOS ANGELES	6	060375001	12	11	3755.11	373.475	18.3610	8334	0.0	1.49700	19.00	1.53
LOS ANGELES	6	060376002	13	11	3805.99	359.011	44.8751	8289	0.0	1.09555	11.00	1.09
LOS ANGELES	6	060590001	14	11	3742.44	415.475	40.6956	8323	0.0	1.64941	17.00	1.65
LOS ANGELES	6	060595001	15	11	3754.18	412.113	31.1026	8312	0.0	1.86020	19.00	1.86
LOS ANGELES	6	061112002	16	11	3794.01	344.915	47.0311	6812	0.0	0.89298	10.00	0.90
LOS ANGELES	6	060374101	x	11	3810.28	362.549	46.6362	710	0.0	0.73803	5.00	xxxx
MINNEAPOLIS	7	CENTER	0	15	4979.8633	477.2979	0.0	0	0	0	0	
MINNEAPOLIS	7	270530056	1	15	4978.82	478.903	1.9155	8689	0.0	0.74286	7.70	0.74
MINNEAPOLIS	7	270530059	2	15	4977.06	477.500	2.8106	8430	0.0	2.02154	15.00	2.02
MINNEAPOLIS	7	270530954	3	15	4980.43	478.360	1.2038	6958	0.0	1.30813	8.90	1.31
MINNEAPOLIS	7	271230048	4	15	4977.05	493.012	15.9643	8020	0.0	1.18384	11.60	1.19
MINNEAPOLIS	7	271230050	5	15	4977.84	488.554	11.4360	8684	0.0	1.92433	22.50	1.93
MINNEAPOLIS	7	271230044	x	15	4977.85	486.870	9.7815	6133	0.0	1.65677	14.40	xxxx
MINNEAPOLIS	7	271230052	x	15	4981.79	486.934	9.8278	1402	0.2	1.61648	8.60	xxxx
NEW YORK	8	CENTER	0	18	4510.8672	585.7268	0.0	0	0	0	0	
NEW YORK	8	340030004	1	18	4522.85	586.935	12.0436	8394	0.1	1.80901	11.94	1.80
NEW YORK	8	340035001	2	18	4525.91	580.790	15.8360	8503	0.1	1.23407	10.79	1.23
NEW YORK	8	340130011	3	18	4508.57	572.280	13.6416	8405	0.1	0.95691	6.81	0.95
NEW YORK	8	340171002	4	18	4509.11	578.299	7.6326	8259	0.1	2.03861	12.14	2.04
NEW YORK	8	340232003	5	18	4484.32	561.990	35.6131	8091	0.1	1.13954	8.24	1.14
NEW YORK	8	340270003	6	18	4516.20	543.700	42.3638	8329	0.1	1.73585	13.77	1.73
NEW YORK	8	340390003	7	18	4501.31	566.373	21.5854	8382	0.1	1.94353	14.76	1.94
NEW YORK	8	360470071	8	18	4505.07	585.770	5.7954	8477	0.0	2.59192	15.50	2.59
NEW YORK	8	360470076	9	18	4502.63	586.398	8.2625	8604	0.0	0.97424	8.10	0.97
NEW YORK	8	360590005	10	18	4511.20	619.300	33.5748	8446	0.0	1.05445	11.90	1.06
NEW YORK	8	360610010	11	18	4510.12	585.607	0.7607	8308	0.1	1.17512	8.30	1.18

TABLE 1 (continued)

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
NEW YORK	8	360610056	12	18	4512.31	587.250	2.0980	8328	0.0	1.54977	8.10	1.55
NEW YORK	8	360610062	13	18	4508.03	584.059	3.2885	8418	0.4	2.78738	16.60	2.80
NEW YORK	8	360610082	14	18	4512.56	587.140	2.2052	8549	0.0	4.02014	17.00	4.01
NEW YORK	8	360610092	x	18	4510.73	586.340	0.6284	2188	0.1	2.33583	11.30	xxxx
NEW YORK	8	360810054	x	18	4512.08	589.947	4.3910	1344	0.0	1.32143	11.60	xxxx
PHILADELPHIA	9	CENTER	0	18	4422.2305	487.5796	0.0	0	0	0	0	
PHILADELPHIA	9	100032002	1	18	4400.81	453.195	40.5126	7117	0.0	0.97781	10.00	0.99
PHILADELPHIA	9	100033001	2	18	4406.80	461.000	30.7339	7101	0.0	0.75159	7.00	0.75
PHILADELPHIA	9	340051001	3	18	4436.20	512.100	28.2205	8234	0.1	1.32944	12.30	1.33
PHILADELPHIA	9	340070003	4	18	4419.01	491.692	5.2221	8343	0.1	0.71566	10.44	0.72
PHILADELPHIA	9	340071001	5	18	4391.36	512.265	39.5267	8140	0.1	0.42925	2.07	0.43
PHILADELPHIA	9	340210006	6	18	4452.15	520.019	44.1277	8131	0.1	0.94254	6.09	0.95
PHILADELPHIA	9	420170012	7	18	4439.45	510.038	28.3025	8576	0.0	0.62382	13.70	0.62
PHILADELPHIA	9	420910013	8	18	4440.05	473.652	22.6151	8541	0.0	0.75725	7.00	0.76
PHILADELPHIA	9	421010004	9	18	4428.54	491.655	7.5096	8304	0.0	0.90474	10.00	0.91
PHILADELPHIA	9	421010027	10	18	4428.73	487.030	6.5227	7973	0.0	1.65396	11.00	1.64
PHILADELPHIA	9	421010029	11	18	4422.81	485.218	2.4326	8209	0.0	0.87258	8.00	0.88
PHILADELPHIA	9	421010047	12	18	4418.31	485.803	4.3033	8332	0.0	1.15578	9.00	1.15
PHILADELPHIA	9	421010051	13	18	4422.44	487.600	0.2105	8450	0.0	1.65124	14.00	1.65
PHOENIX	10	CENTER	0	12	3705.6689	400.2964	0.0	0	0	0	0	
PHOENIX	10	040130013	1	12	3696.51	400.190	9.1605	8177	0.2	1.26340	9.50	1.26
PHOENIX	10	040130016	2	12	3706.48	395.028	5.3308	6311	0.0	2.64262	12.80	2.64
PHOENIX	10	040130019	3	12	3705.32	393.879	6.4267	8177	0.0	1.69571	10.90	1.69
PHOENIX	10	040131004	4	12	3713.75	401.115	8.1205	8566	0.0	1.56258	11.60	1.56
PHOENIX	10	040132001	5	12	3714.86	389.523	14.1613	7697	0.0	1.14237	8.40	1.14
PHOENIX	10	040132004	6	12	3718.46	413.922	18.6887	8699	0.0	0.93129	5.90	0.93
PHOENIX	10	040133002	7	12	3702.43	403.039	4.2479	8432	0.0	1.57608	13.20	1.60
PHOENIX	10	040133003	8	12	3704.62	414.832	14.5732	8199	0.0	1.31981	11.90	1.33
PHOENIX	10	040131003	x	12	3696.88	419.645	21.2504	5473	0.0	0.81129	5.10	xxxx
PHOENIX	10	040132005	x	12	3729.74	420.718	31.5645	2575	0.1	0.57445	5.50	xxxx
SAN FRANCISCO	11	CENTER	0	10	4155.4	566.3	0.0	0	0	0	0	
SAN FRANCISCO	11	060010003	1	10	4171.38	608.799	45.4030	8672	0.0	0.89795	8.00	0.90
SAN FRANCISCO	11	060010005	2	10	4183.61	564.561	28.2645	8679	0.0	1.05312	8.00	1.05

TABLE 1 (continued)

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
SAN FRANCISCO	11	060011001	3	10	4154.59	591.812	25.5249	8621	0.0	1.00278	8.00	1.01
SAN FRANCISCO	11	060130002	4	10	4199.25	585.701	47.9475	8633	0.0	1.57558	11.00	1.58
SAN FRANCISCO	11	060130003	5	10	4200.26	556.572	45.8997	8590	0.0	1.19709	7.00	1.20
SAN FRANCISCO	11	060750003	6	10	4181.77	552.492	29.7655	8495	0.0	1.79859	12.00	1.80
SAN FRANCISCO	11	060750005	7	10	4178.60	554.004	26.2606	8553	0.0	1.58892	8.00	1.59
SAN FRANCISCO	11	060811001	8	10	4148.34	570.487	8.2082	8462	0.0	1.24273	12.00	1.24
SAN FRANCISCO	11	060850004	9	10	4132.97	598.547	39.2824	8596	0.0	1.80142	17.00	1.80
SAN FRANCISCO	11	060852004	10	10	4131.08	595.148	37.7328	8686	0.0	1.39051	12.00	1.39
SAN FRANCISCO	11	060850004	x	10	4132.97	598.547	39.2824	8432	0.0	1.66556	18.00	xxxx
SPOKANE	12	CENTER	0	11	5278.1953	468.7935	0.0	0	0	0	0	
SPOKANE	12	530630028	1	11	5278.20	468.100	0.6935	8252	0.0	1.72977	12.20	1.74
SPOKANE	12	530630029	2	11	5280.75	469.210	2.5884	7828	0.0	1.58674	12.50	1.60
SPOKANE	12	530630040	3	11	5279.41	470.330	1.9587	8458	0.1	2.57089	21.90	2.56
SPOKANE	12	530630043	4	11	5279.37	469.200	1.2430	8317	0.0	2.31152	14.50	2.31
SPOKANE	12	530630044	5	11	5277.67	468.560	0.5749	7958	0.0	3.13150	19.60	3.14
SPOKANE	12	530630027	x	12	5278.38	469.040	0.3080	2211	0.0	1.45269	9.60	xxxx
SPOKANE	12	530630045	x	12	5278.25	468.080	0.7156	5535	0.0	1.89438	10.60	xxxx
ST. LOUIS	13	CENTER	0	15	4279.125	743.5127	0.0	0	0	0	0	
ST. LOUIS	13	171190017	1	15	4287.36	747.923	9.3452	8380	0.0	1.04426	10.00	1.05
ST. LOUIS	13	171630010	2	15	4277.36	747.251	4.1327	7885	0.0	0.65860	8.90	0.66
ST. LOUIS	13	291890001	3	15	4266.84	731.587	17.1243	7705	0.0	0.73563	4.60	0.73
ST. LOUIS	13	291890006	4	15	4276.70	718.034	25.5940	6403	0.1	0.57884	4.00	0.58
ST. LOUIS	13	291893001	5	15	4280.15	731.007	12.5475	8633	0.0	0.93944	8.90	0.94
ST. LOUIS	13	291895001	6	15	4294.11	735.821	16.8473	8527	0.0	0.69618	4.80	0.70
ST. LOUIS	13	291897001	7	15	4289.59	727.810	18.8709	7651	0.0	1.11270	14.50	1.13
ST. LOUIS	13	295100080	8	15	4284.97	739.503	7.0848	8549	0.0	1.12167	11.23	1.11
ST. LOUIS	13	295100083	x	15	4279.37	744.246	0.7731	5161	0.0	1.22098	11.15	xxxx
WASHINGTON,DC	14	CENTER	0	18	4308.4375	324.0149	0.0	0	0	0	0	
WASHINGTON,DC	14	110010017	1	18	4307.87	322.089	2.0069	8476	0.0	1.23493	12.60	1.24
WASHINGTON,DC	14	110010023	2	18	4307.89	322.548	1.5643	7519	0.0	1.53357	11.90	1.53
WASHINGTON,DC	14	510130020	3	18	4297.77	322.532	10.7701	8618	0.0	0.75032	6.80	0.76
WASHINGTON,DC	14	510590018	4	18	4290.03	319.435	18.9638	8546	0.0	1.05448	11.30	0.91
WASHINGTON,DC	14	510595001	5	18	4311.33	309.390	14.9082	8533	0.2	1.04034	9.90	0.90

TABLE 1 (continued)

STUDY AREA	NUM	SITE	SITENUM	UTMZ	UTMNORTH	UTMEAST	DIST	N	MIN	MEAN	MAX	AQAVG
WASHINGTON,DC	14	515100009	6	18	4302.98	321.357	6.0739	8389	0.0	0.87751	9.30	0.88
WASHINGTON,DC	14	516000005	7	18	4301.89	299.355	25.5143	8427	0.1	0.79294	9.70	0.68
WASHINGTON,DC	14	240311008	x	18	4324.62	316.862	17.6919	5951	0.0	1.25120	8.60	xxxx
WASHINGTON,DC	14	240330001	x	18	4289.65	331.825	20.3471	5815	0.0	1.45070	16.50	xxxx
WASHINGTON,DC	14	510591004	x	18	4304.09	314.073	10.8489	2852	0.0	0.90438	7.20	xxxx

NUM = Study area number (used in ordering the Appendices)
SITENUM = Site number within each study area
UTMZ = Universal Transverse Mercator (UTM) zone number
UTMNORTH = UTM Northing in kilometers
UTMEAST = UTM Easting in kilometers
DIST = Distance of monitor from city center in kilometers
N = Number of valid hourly CO measurements in 1990
MIN = Minimum of all valid hourly CO measurements in 1990
MEAN = Average (mean) of all valid hourly CO measurements in 1990
MAX = Maximum of all valid hourly CO measurements in 1990
AQAVG = Average CO concentration output from AQAVG program (8760 hours)

TABLE 2

Microenvironmental Factors used in model runs.
All additive factors are set to zero.
Lag=0 indicates current ambient value used,
Lag=-1 indicates previous hour's ambient value used.

Microenvironment	Multiplicative Factor	Lag
IN TRANSIT, CAR	0.960	0
IN TRANSIT, BUS	1.650	0
IN TRANSIT, TRUCK	1.140	0
IN TRANSIT, VAN	0.960	0
INDOORS, PUBLIC GARAGE	0.590	-1
OUTDOORS, IN A PARKING LOT	0.570	0
OUTDOORS, ALONG A ROADWAY	0.850	0
IN TRANSIT, MOTORCYCLE	1.030	0
INDOORS, SERVICE STATION	0.870	-1
OUTDOORS, SERVICE STATION	0.000	0
INDOORS, RESIDENTIAL GARAGE	0.780	-1
INDOORS, IN A REPAIR SHOP	1.320	0
INDOORS, HOME	0.380	-1
INDOORS, OFFICE	0.380	-1
INDOORS, STORE	0.410	-1
INDOORS, RESTAURANT	0.710	-1
INDOORS, MFG. FACILITY	0.470	-1
INDOORS, SCHOOL	0.450	-1
INDOORS, CHURCH	0.280	-1
INDOORS, SHOPPING MALL	2.110	-1
INDOORS, AUDITORIUM	0.120	-1
INDOORS, HEALTH CARE FACILITY	0.310	-1
INDOORS, PUBLIC BUILDING	0.350	-1
INDOORS, OTHER LOCATION	0.790	-1
INDOORS, NOT SPECIFIED	0.790	-1
OUTDOORS, CONSTRUCTION	0.960	0
OUTDOORS, RESIDENTIAL GROUNDS	0.550	0
OUTDOORS, SCHOOL GROUNDS	0.960	0
OUTDOORS, SPORTS ARENA	0.280	0
OUTDOORS, PARK/GOLF COURSE	0.280	0
OUTDOORS, OTHER LOCATION	0.600	0
OUTDOORS, NOT SPECIFIED	0.960	0
IN TRANSIT, TRAIN	0.960	0
IN TRANSIT, AIRPLANE	0.000	0

TABLE 3

Summary of Meteorological Data for 1990.

Frequency of day types by quarter.
Summer consists of June, July and August.
Warm in summer means maximum temperature > 83°F.
Warm at other times means maximum temperature > 54°F.

		Non-summer Weekday		Non-summer Weekend		Summer Weekday		Summer Weekend	
STUDY AREA	QUARTER	Warm	Cool	Warm	Cool	Warm	Cool	Warm	Cool
Baltimore	Jan-Mar	29	36	9	16	-	-	-	-
Baltimore	Apr-Jun	39	5	14	3	9	12	4	5
Baltimore	Jul-Sep	20	0	10	0	29	16	13	4
Baltimore	Oct-Dec	44	22	19	7	-	-	-	-
Boston	Jan-Mar	10	55	5	20	-	-	-	-
Boston	Apr-Jun	29	15	9	8	3	18	2	7
Boston	Jul-Sep	20	0	10	0	17	28	6	11
Boston	Oct-Dec	35	31	16	10	-	-	-	-
Chicago	Jan-Mar	9	56	2	23	-	-	-	-
Chicago	Apr-Jun	32	12	15	2	7	14	4	5
Chicago	Jul-Sep	20	0	10	0	13	32	6	11
Chicago	Oct-Dec	27	39	9	17	-	-	-	-
Denver	Jan-Mar	17	48	14	11	-	-	-	-
Denver	Apr-Jun	33	11	16	1	16	5	5	4
Denver	Jul-Sep	20	0	10	0	26	19	6	11
Denver	Oct-Dec	36	30	13	13	-	-	-	-
Houston	Jan-Mar	63	2	22	3	-	-	-	-
Houston	Apr-Jun	44	0	17	0	21	0	9	0
Houston	Jul-Sep	20	0	10	0	44	1	17	0
Houston	Oct-Dec	61	5	24	2	-	-	-	-
Los Angeles	Jan-Mar	63	2	23	2	-	-	-	-
Los Angeles	Apr-Jun	44	0	17	0	21	0	9	0
Los Angeles	Jul-Sep	20	0	10	0	2	43	0	17
Los Angeles	Oct-Dec	65	1	25	1	-	-	-	-
Minneapolis	Jan-Mar	8	57	0	25	-	-	-	-
Minneapolis	Apr-Jun	32	12	12	5	5	16	2	7
Minneapolis	Jul-Sep	20	0	9	1	13	32	3	14
Minneapolis	Oct-Dec	22	44	5	21	-	-	-	-

TABLE 3 (continued)

STUDY AREA	QUARTER	Regular Weekday Warm	Regular Weekday Cool	Regular Weekend Warm	Regular Weekend Cool	Summer Weekday Warm	Summer Weekday Cool	Summer Weekend Warm	Summer Weekend Cool
New York	Jan-Mar	15	50	6	19	-	-	-	-
New York	Apr-Jun	36	8	13	4	5	16	3	6
New York	Jul-Sep	20	0	10	0	24	21	7	10
New York	Oct-Dec	37	29	16	10	-	-	-	-
Philadelphia	Jan-Mar	26	39	8	17	-	-	-	-
Philadelphia	Apr-Jun	37	7	14	3	7	14	4	5
Philadelphia	Jul-Sep	20	0	10	0	30	15	14	3
Philadelphia	Oct-Dec	41	25	16	10	-	-	-	-
Phoenix	Jan-Mar	62	3	25	0	-	-	-	-
Phoenix	Apr-Jun	44	0	17	0	21	0	9	0
Phoenix	Jul-Sep	20	0	10	0	44	1	17	0
Phoenix	Oct-Dec	64	2	24	2	-	-	-	-
San Francisco	Jan-Mar	46	19	18	7	-	-	-	-
San Francisco	Apr-Jun	44	0	17	0	3	18	0	9
San Francisco	Jul-Sep	20	0	10	0	0	45	0	17
San Francisco	Oct-Dec	57	9	20	6	-	-	-	-
Spokane	Jan-Mar	8	57	3	22	-	-	-	-
Spokane	Apr-Jun	35	9	13	4	2	19	1	8
Spokane	Jul-Sep	20	0	10	0	21	24	10	7
Spokane	Oct-Dec	14	52	5	21	-	-	-	-
St. Louis	Jan-Mar	32	33	9	16	-	-	-	-
St. Louis	Apr-Jun	37	7	15	2	14	7	7	2
St. Louis	Jul-Sep	20	0	10	0	30	15	11	6
St. Louis	Oct-Dec	40	26	18	8	-	-	-	-
Washington DC	Jan-Mar	30	35	9	16	-	-	-	-
Washington DC	Apr-Jun	40	4	16	1	11	10	5	4
Washington DC	Jul-Sep	20	0	10	0	34	11	14	3
Washington DC	Oct-Dec	44	22	20	6	-	-	-	-

TABLE 4

Counties, Populations and Exposure from HAPEM runs.
FIPS is the state and county Federal Information Processing Standard code
CENSUS is the total population of the county.
HAPPOP is the population of the part inside the study area.
FLAG is zero if the population criteria are met for the exposure estimate.
EXPOSURE is the county average annual CO exposure for all persons in μg/m³.

STUDY AREA	NUM	COUNTY	FIPS	CENSUS	HAPPOP	FLAG	EXPOSURE
Baltimore	1	Anne Arundel County	24003	427239	411082	1	709
Baltimore	1	Baltimore County	24005	692134	692134	0	631
Baltimore	1	Carroll County	24013	123372	102252	1	692
Baltimore	1	Harford County	24025	182132	164306	1	692
Baltimore	1	Howard County	24027	187328	187328	1	692
Baltimore	1	Queen Anne's County	24035	33953	9167	1	692
Baltimore	1	Baltimore city	24510	736014	736014	0	709
Boston	2	Bristol County	25005	506325	107934	1	617
Boston	2	Essex County	25009	670080	626718	1	613
Boston	2	Middlesex County	25017	1398468	1361314	0	622
Boston	2	Norfolk County	25021	616087	616063	1	626
Boston	2	Plymouth County	25023	435276	319927	1	617
Boston	2	Suffolk County	25025	663906	663906	0	634
Boston	2	Worcester County	25027	709705	86144	1	617
Boston	2	Hillsborough County	33011	335838	9408	1	646
Boston	2	Rockingham County	33015	245845	15204	1	617
Chicago	3	Cook County	17031	5105044	5082451	0	483
Chicago	3	DuPage County	17043	781689	781666	1	493
Chicago	3	Lake County	17097	516418	118106	1	487
Chicago	3	Will County	17197	357313	268052	1	487
Chicago	3	Lake County	18089	475594	419957	0	479
Chicago	3	Porter County	18127	128932	11438	1	487
Denver	4	Adams County	08001	265038	261008	0	553
Denver	4	Arapahoe County	08005	391511	386701	0	524
Denver	4	Boulder County	08013	225339	201015	0	475
Denver	4	Denver County	08031	467610	467610	0	685
Denver	4	Jefferson County	08059	438430	438430	0	561
Denver	4	Weld County	08123	131821	21262	1	568
Houston	5	Brazoria County	48039	191707	79347	1	479
Houston	5	Fort Bend County	48157	225421	203228	1	479
Houston	5	Galveston County	48167	217396	80242	1	479
Houston	5	Harris County	48201	2818101	2814396	0	496
Houston	5	Montgomery County	48339	182201	64628	1	479
Houston	5	Waller County	48473	23389	1227	1	479
Los Angeles	6	Los Angeles County	06037	8863052	8540864	0	925
Los Angeles	6	Orange County	06059	2410668	1471461	0	859
Los Angeles	6	San Bernardino County	06071	1418380	30180	1	1029
Los Angeles	6	Ventura County	06111	669016	45279	1	538

TABLE 4 (continued)

STUDY AREA	NUM	COUNTY	FIPS	CENSUS	HAPPOP	FLAG	EXPOSURE
Minneapolis	7	Anoka County	27003	243641	243641	1	749
Minneapolis	7	Carver County	27019	47915	39178	1	756
Minneapolis	7	Dakota County	27037	275189	269569	1	739
Minneapolis	7	Hennepin County	27053	1032431	1032431	0	807
Minneapolis	7	Ramsey County	27123	485783	485765	0	807
Minneapolis	7	Scott County	27139	57846	53132	1	756
Minneapolis	7	Sherburne County	27141	41945	11201	1	756
Minneapolis	7	Washington County	27163	145896	145896	1	747
Minneapolis	7	Wright County	27171	68710	23536	1	756
Minneapolis	7	Pierce County	55093	32765	5373	1	756
Minneapolis	7	St. Croix County	55109	50251	17789	1	756
New York City	8	Fairfield County	09001	827645	80811	1	933
New York City	8	Bergen County	34003	825380	825380	0	775
New York City	8	Essex County	34013	778206	778206	0	671
New York City	8	Hudson County	34017	553099	553099	0	1166
New York City	8	Middlesex County	34023	671811	520080	0	784
New York City	8	Monmouth County	34025	553093	240818	1	924
New York City	8	Morris County	34027	421361	275218	0	925
New York City	8	Passaic County	34031	453302	432111	0	843
New York City	8	Somerset County	34035	240245	57332	1	933
New York City	8	Bronx County	36005	1203789	1203789	0	971
New York City	8	Kings County	36047	2300664	2300664	0	642
New York City	8	Nassau County	36059	1288563	1288563	0	740
New York City	8	New York County	36061	1487536	1487536	0	1415
New York City	8	Queens County	36081	1951598	1951598	0	916
New York City	8	Richmond County	36085	378977	378977	0	885
New York City	8	Rockland County	36087	265475	235323	1	933
New York City	8	Suffolk County	36103	1321768	85145	1	933
New York City	8	Westchester County	36119	874866	700453	1	934
Philadelphia	9	New Castle County	10003	441946	244573	0	489
Philadelphia	9	Atlantic County	34001	224327	14389	1	218
Philadelphia	9	Burlington County	34005	395066	380258	0	550
Philadelphia	9	Camden County	34007	502824	502824	0	419
Philadelphia	9	Cumberland County	34011	138053	9536	1	510
Philadelphia	9	Gloucester County	34015	230082	230082	1	513
Philadelphia	9	Salem County	34033	65294	52903	1	507
Philadelphia	9	Bucks County	42017	541174	487491	0	441
Philadelphia	9	Chester County	42029	376396	231640	1	504
Philadelphia	9	Delaware County	42045	547651	547651	0	480
Philadelphia	9	Montgomery County	42091	678193	609711	0	498
Philadelphia	9	Philadelphia County	42101	1585577	1585577	0	608
Phoenix	10	Maricopa County	04013	2122101	2096861	0	741
Phoenix	10	Pinal County	04021	116397	20506	1	761
San Francisco	11	Alameda County	06001	1276702	1275586	0	589
San Francisco	11	Contra Costa County	06013	803732	455259	0	709
San Francisco	11	Marin County	06041	230096	71880	1	698
San Francisco	11	San Francisco County	06075	723959	722506	0	882
San Francisco	11	San Mateo County	06081	649623	649623	0	686

TABLE 4 (continued)

STUDY AREA	NUM	COUNTY	FIPS	CENSUS	HAPPOP	FLAG	EXPOSURE
San Francisco	11	Santa Clara County	06085	1497577	1321940	0	772
San Francisco	11	Santa Cruz County	06087	229734	12867	1	698
Spokane	12	Spokane County	53063	361364	361364	0	1172
St. Louis	13	Jefferson County	29099	171380	133376	1	443
St. Louis	13	St. Charles County	29183	212751	175339	1	443
St. Louis	13	St. Louis County	29189	993529	993529	0	420
St. Louis	13	St. Louis city	29510	396685	396685	0	473
Washington DC	14	District of Columbia	11001	606900	606900	0	696
Washington DC	14	Calvert County	24009	51372	16095	1	505
Washington DC	14	Charles County	24017	101154	82668	1	505
Washington DC	14	Montgomery County	24031	757027	757027	1	512
Washington DC	14	Prince George's County	24033	729268	729268	1	506
Washington DC	14	Arlington County	51013	170936	170936	0	491
Washington DC	14	Fairfax County	51059	818584	818584	0	437
Washington DC	14	Loudoun County	51107	86129	54937	1	505
Washington DC	14	Prince William County	51153	215677	192195	1	505

CHAPTER 2

UNCERTAINTY ANALYSIS

Introduction

The fundamental notion of uncertainty that will be discussed in this chapter is that the true value of a quantity in a process is not known with exactitude and can only approximated through estimation. The quantity of interest for which this uncertainty analysis applies is the annual mean exposure of a demographic group to carbon monoxide. The estimate of the annual mean exposure was calculated using the HAPEM-MS3 model. There are no direct measurements of annual mean exposure to carbon monoxide and thus it is not possible to make direct comparisons between measurements of annual mean exposure and estimates derived from HAPEM-MS3. Thus, a statement of the sort "the estimate of annual mean exposure from HAPEM-MS3 is within x% of measured values" cannot be made and statistical quantities such as bias and distributional properties of annual mean exposure to carbon monoxide cannot be fully evaluated.

There are two types of analyses that will be carried out in regard to the uncertainty in exposure estimates to carbon monoxide as carried out using HAPEM-MS3. The first analysis will discuss a comparison of the models HAPEM-MS3 and pNEM/CO as carried out in a report by International Technologies Corporation (IT, 1996). Also, a technical paper that compares short term personal exposure monitor estimates of exposure to carbon monoxide to pNEM/CO (Law, *et al.*, 1997) will be discussed. The second analysis examines the components of HAPEM-MS3 and judges whether or not a component is representative in a quantitative sense of selected parts of the measurement process. The components of HAPEM-MS3 can be divided between those that are derived from empirical measurements and those that are due to modeling assumptions of HAPEM-MS3. The organization of the following material is to identify the components of the HAPEM-MS3 estimates and discuss uncertainty due to the empirical measurements and due to the modeling assumptions together for a given component.

The components of the HAPEM-MS3 estimates that are discussed below are the following:

1) Air quality in terms of the ambient air quality obtained, averaging procedures to obtain quarterly average diurnal data and assignment of air monitoring districts;
2) Population data in terms of data sources;
3) Temperature data in terms of data sources.
4) Activity patterns in terms of the activity database and the selection of activity patterns to compose year long sequences of activity patterns;
5) Microenvironmental factors in terms of sources of data and composition of specific factors.

Comparison of Models to Sample Data

Quantitative comparisons of HAPEM-MS3 annual average exposure estimates to sample data is not possible due to the lack of sample data of annual exposure to carbon monoxide. Some indirect comparisons of the HAPEM-MS3 model estimates to short term estimates of exposure are available through a somewhat circuitous route. Namely, estimates of annual average exposure to CO have been run concurrently (IT, 1996) using the two models HAPEM-MS3 and pNEM/CO. Comparisons of pNEM/CO estimates of 1 hour daily maximum and 8 hour daily maximum exposure to PEM CO measurements have been published (Law, *et al.*, 1997). The results of these two studies are presented in summary form below and inference about uncertainty in HAPEM-MS3 estimates relative to sample measurements is proposed.

The IT report lays out a comparison of HAPEM-MS2 and pNEM/CO exposure estimates for carbon monoxide. The estimates are carried out for Denver, CO for 1988 air quality carbon monoxide data. In the study, estimates for 11 demographic groups for age and work status in 6 home districts were computed. Additionally, analyses were done for the 11 demographic groups and 33 microenvironments. Ten runs of both HAPEM-MS2 and pNEM/CO were carried out and the resulting means and standard errors over the ten runs were compared.

The results indicated that there was no significant bias between the HAPEM-MS2 and pNEM/CO estimates for the exposure means. The investigators examined runs covering six air monitoring districts and eleven demographic groups. They found that the following regression equation best fit the paired results:

HAPEM-MS2 mean = 0.365 + 0.738 (pNEM/CO mean),

with an R^2 value of 0.286. The ratios of (HAPEM-MS2 mean) to (pNEM/CO mean) values were also calculated and had a mean value of 1.028 and a median of 1.0329. This represents a small bias, considering that the individual results (single demographic groups and air district combinations) showed considerable variation in both the means and the ratios.   For the standard errors, the IT report found the relationship

HAPEM-MS2 standard error = 0.010 + 0.169 (pNEM/CO standard error).

The indication is that HAPEM-MS2 standard error is less than the pNEM/CO standard error. This is confirmed by the median value of the ratio (HAPEM-MS3 standard error)/(pNEM/CO standard error) being 0.417, which is significantly smaller than one.

Similar results were found for comparisons among paired runs for various combinations of microenvironments and demographic groups with the derived relationships:

HAPEM-MS2 mean = 0.471 + 0.573 (pNEM/CO mean),
HAPEM-MS2 standard error = 0.0369 + 0.0622 (pNEM/CO standard error).

Finally,   the ratio of the (HAPEM-MS2 mean)/(pNEM/CO mean) had a median value of 0.5405 and the ratio of the (HAPEM-MS2 standard error)/(pNEM/CO standard error) had a median value of 0.1532.

The results indicate that when summarized by district there is little bias evident between the two models. However, for the microenvironmental breakout the HAPEM-MS2 results were usually (but not always) lower than the pNEM/CO results. Since the overall average across districts should equal the overall average across microenvironments, this result seems a little strange at first. The explanation is that the averages must be weighted according to the probability of occurrence of each of the combinations. In practice, HAPEM-MS2 tended to underestimate exposure relative to pNEM/CO in the microenvironments that are infrequently visited, so that a straight (unweighted) mean across microenvironments shows HAPEM-MS2 to underpredict exposure, whereas a weighted (by duration) mean shows little difference between the models. By comparison, the standard error for the HAPEM-MS2 estimates is generally much less than for the pNEM/CO estimates for both of the above sets of runs. This would generally be expected since pNEM/CO uses distributions for several of the factors affecting exposure whereas HAPEM-MS2 uses point estimates. The only distribution modeled in HAPEM-MS2 is the duration due to the stochastic sampling of activity patterns. The general use of point estimates will result in less variability than the use of distributions.

The second study by Law, *et al.* (1997) compared daily maximum 1-hour and 8-hour average carbon monoxide exposure estimates between personal environmental monitor (PEM) results and pNEM/CO exposure estimates. The PEM data was collected during the winter of 1982-1983 in Denver, CO. The authors state in the Summary and Conclusions section that '....pNEM/CO over-predicts the CFD of population exposure at low exposures and under-predicts the CFD at high exposures'. The analysis was limited to four demographic groups. Furthermore, their Table 2 shows that at the median the simulated values are greater than the observed values in two of four cases. Their graphs (shown as Figure A on the next page) show that the measured and modeled Cumulative Probability plots cross each other in all four cases, with small differences near the medians. From this, it can be inferred that the estimates of the mean of the exposure are subject to little bias, but that the dispersion of the simulated values is less than the mean of the observed values.

Combining both of the above analyses, the conclusion would be that there is no evidence for a systematic bias in the means for either the HAPEM-MS2 or pNEM/CO models. However, the pNEM/CO model underestimates the true variance of exposure in the population, and HAPEM-MS2 provides even lower variance estimates than pNEM/CO. Thus, the HAPEM variance estimates do not adequately represent the variance in the population. This was expected, as it has been argued elsewhere that HAPEM attempts to model the average for a whole demographic group, not the individuals who may be at the extremes of the group.

Figure A - not available in electronic copy

Air Quality

The ambient monitored CO data was obtained from the U.S. EPA AIRS database. The data in AIRS is not meant to represent results from a designed air monitoring network. Instead state and local agencies contribute data obtained from local continuous CO monitors. The maps in the appendix B show the geographical locations of the monitors used in the HAPEM-MS3 estimates of CO air quality. The coverage of the HAPEM-MS3 modeling area (within 50 km) of the defined city center) in highly variable from city to city. For example, in Baltimore the CO ambient air monitors are all clustered in a small area. On the other hand, the CO ambient air monitors in Los Angeles are fairly evenly distributed over the modeling area. The spatial distribution of monitors in Los Angeles should result in a more representative CO air quality estimate for the city as a whole than is the case in Baltimore. In either case, uncertainty in ambient air levels of CO in unmonitored locations has not been evaluated for the HAPEM-MS3 estimates.

The treatment of the ambient monitored CO data by HAPEM-MS3 leads to several sources of uncertainty for the air quality estimates used in HAPEM-MS3. The missing values in ambient monitored CO data are first estimated via Fourier analysis and then all hourly average CO values (monitored and estimated) are used to obtain quarterly average diurnal concentrations. The data from each monitor is arbitrarily assigned an influence in a 20 km diameter circle centered on the monitor location. Areas outside the influence of any ambient monitor (air district 19) are assigned a CO concentration that is the average of all other monitors within the modeling area. Some sources of uncertainty that arise from these procedures are the following:

1) CO concentrations cannot be assigned in a spatially continuous way to all geographic areas. The most profound impact of this is that rather extensive geographical areas can be assigned to air district 19 (the average of all ambient monitors). The geometry and population of air district 19 is highly variable from city to city. In Baltimore, for example, air district 19 completely surrounds Baltimore City. In Los Angeles, air district 19 is much less extensive than in

Baltimore but pockets of air district 19 are found between the other air districts. The exposures of people in air district 19 are inherently uncertain.

2) Another spatial effect in HAPEM is that the CO concentration effectively has a discontinuity at each monitor district boundary: people living just on one side of the boundary may have concentrations different by a factor of two or more from people next door on the other side of the district boundary. If it is assumed that the true concentration does not have this discontinuity, then one or both of the estimates on either side of the boundary must be incorrect, and the size of the discontinuity provides a measure of this effect.

3) The current version of HAPEM does not allow for the discrimination of weekday and weekend days for air quality. Generally for CO, the highest concentrations are reported on weekdays due to mobile source contributions during rush hours, which are absent on weekends. If the weekend/weekday distinction were made in the HAPEM air quality programs, then it would result in slightly higher exposure estimates overall (because more people are out driving and getting exposed when the air quality is bad; and of course this is because the air quality is bad whenever more driving occurs.).

.

Population

The population data used for HAPEM-MS3 estimates was obtained from the 1990 U.S. Population Census. The census tract level data is used in HAPEM-MS3 to obtain population estimates for a specified demographic group. There is not any uncertainty introduced in the way the population estimates are incorporated in the HAPEM-MS3 estimates for the non-commuting demographic groups. For the commuters, the location of the workplaces is not known a priori, and is assigned using an iterative algorithm in the ODEST program. In some cases, this is likely to misallocate workers. For example, a large number of workers in a non-central area where few people live (such as a industrial park) might be assigned to tracts closer to the central city by the commuting algorithm. The size of this potential effect has not been estimated. However, relatively few (only five out of 23) of the demographic groups commute, and even in those cases the population data do not affect the exposure of any one of the particular cohorts. The population data is used only to obtain population weighted averages across the cohorts in exposure estimates.

Temperature

The daily maximum temperature is used in HAPEM-MS3 only to determine the number of warm and cool days, used for the selection of daily activity patterns to compose the annual activity patterns. Aside from making this determination in a different way based on temperature in the HAPEM-MS3 estimate, there is no obvious effect on uncertainty in annual average CO exposure due to the way temperature data is treated in HAPEM-MS3.

Activity Patterns

The variance that has been computed for the HAPEM-MS3 estimates of annual average CO exposure arises entirely from variation in duration in the daily activity patterns as represented in the activity data base. In addition to this rather small variance due to variation in time-activity duration, there are several other factors that affect uncertainty due to selection of daily activity patterns and construction of annual activity patterns. These factors include the following:

1) The activity database is constructed from a relatively small number of studies and the studies were not of a consistent quality;
2) The construction of annual activity patterns is carried out in a manner that reduces overall variability in annual average CO exposure.

One of the most obvious sources of uncertainty arises from the construction of the activity database itself. The database was constructed from activity diaries obtained in Denver, CO, Washington, D.C. and Cincinnati, OH in 1982 - 1985. The sampling in Denver and

Washington was carried out only in one winter (November-February) and the Cincinnati sampling was done in March and August of another year. This can lead to several major sources of uncertainty in the HAPEM-MS3 estimates. First, Denver, Washington and Cincinnati are in climatically similar areas. Cities with different climate patterns such as Los Angeles and Houston may not be well represented by the activity patterns in the HAPEM-MS3 activity database. The duration of activity in particular microenvironments is likely to differ at least seasonally between warm climate cities and cold climate cities. In other words, the variation in duration activity is likely to be small in the HAPEM-MS3 activity database relative to variation in duration activity across all cities in the continental United States. Second, seasonal variability in activity is not well represented in the HAPEM-MS3 database. The only summer season data in the database are obtained from the Cincinnati activity diaries. The summer activities in the 14 cities considered in this report are represented only by a survey of activities in Cincinnati in one month of one summer (August 1985). Third, there may be a slow but significant change in activities over time. For example, it is known that the number of vehicle-miles traveled per person is steadily increasing. This means that after several years the activity patterns in the real population may differ substantially from those found in the activity database. All these points could be addressed by examining the impact of using other, more recent activity studies in place of the existing activity database. This would require a substantial effort that is beyond the scope of the current work. Without such comparisons between activity duration estimates, uncertainties in both the level and variation of the HAPEM-MS3 estimates of annual average exposure cannot be meaningfully addressed.

The second major point is that the way in which daily activity duration sequences are selected to represent annual activity duration sequences results in a substantial reduction in variance for the annual exposures. Depending on the state (fixed by weekend\weekday day, summer\winter season, high\low daily maximum temperature) of a particular day of the year, a daily activity duration sequence is selected uniformly from all daily activity sequences in the database for that state and demographic group. A particular year long activity duration sequence does not in general represent the activity duration sequence of any individual in the population of interest. In fact, if a year long activity duration sequence is interpreted as representing the activity duration sequence of a particular individual in the population, this may not conform to a sequence that can ever be realized. This is not a problem when it comes to estimating the mean exposure, since the independent random selection of patterns does not bias the mean, but in general it will result in a great decrease in the variance of the exposure. A sample calculation is given below to examine the effect of altering this method of constructing year-long activity patterns.

Another problem regarding the activity patterns is that the number of daily activity duration sequences for a given demographic group and state may be very small. Table 5 shows the counts of the number of daily activity patterns by demographic group and state. Note that some demographic groups such as ‘Caucasians’ have at least 125 daily activity duration sequences for each of the eight states. In contrast to this, the ‘Outdoor workers’ demographic group is represented by only one sequence for state six. The effect of having a small number of duration activity sequences to select from is estimates of duration are likely to be biased relative to population durations, as the mean for a small handful of individuals may not necessarily be close to the mean for the population. In addition the variability in duration becomes more difficult to estimate. In the extreme case of just one activity pattern, the model exhibits no variance at all.

The way in which daily activity patterns are selected can affect the variance of the annual average exposure. To illustrate this, suppose that instead of selecting a new random pattern every day, a new pattern is selected only the first time a particular combination of quarter and state actually occurs. Thereafter the particular selected daily activity duration sequence is used each time the same combination is encountered. This introduces a correlation in the year long activity duration sequence. It can be shown that the annual average exposure is unaffected by using this new sampling scheme. That is, the annual average exposure is identical as compared to the actual sampling scheme currently used in HAPEM. However, The variance in exposure is greatly increased. The results for this alternative sampling scheme are seen in Table 6, in which the column ‘Annual’ contains the annual average exposure, and the standard deviations are given in the ‘Independent’ and ‘Correlated’ columns. This table presents results for Denver, and shows that the standard deviation for the estimated annual average exposure for the correlated sequence is about five times that of the uncorrelated sequence. However, comparisons of the fractiles (Table 7) and percentiles (Table 8) show only modest changes when comparing the statistics for these sampling schemes. This is because the variance between individuals in the same air district even with the new sampling scheme is still generally less than the variation between different air districts. The conclusion is that the HAPEM-MS3 model underestimates variance, and this variance estimate can be sensitive to the particular algorithms used internally in the model. However, the effect on the mean exposure (if any) is rather small.

Microenvironmental Factors

In general, the discussion of uncertainty in the microenvironmental factors parallels the discussion of activity patterns above. The microenvironmental factors were derived from data collected in the Washington and Denver studies. Thus, the comments on seasonal and annual variation as in the previous comments on activity patterns apply here also. Namely, the Denver and Washington studies were undertaken only in the winter over a four month period. The applicability of the factors to cities with a different climate type also leads to uncertainty of the true range of values (variability) that the factors can assume. Also, the seasonal variation of the factors is not considered (although HAPEM-MS3 allows it), since there is no data available to derive factors for other seasons. The factors were derived for the CO environment in the early 1980's and many changes in the manner in which vehicles and buildings are constructed have taken place since the early 1980's. Thus, three potential sources of biases may be present in the factors: geographical, seasonal, and secular (aperiodic changes over time).

Separate estimates of exposure to carbon monoxide within certain microenvironments can be misleading because of uncertainties in the microenvironmental factors and the durations as derived from time-activity patterns. Any annual exposure estimate for a microenvironment where the total accumulated duration is short is likely to be uncertain due to two causes. First, the amount of data used to derive the microenvironmental factors is very limited, and so the regression equation used to derive the factors may be subject to substantial error (especially for weakly correlated data). Second, the estimate of exposure depends on the duration as obtained from diaries. Any short duration events (i.e. less than ten minutes) tend to be underreported in diaries. Also, for rare events the small sample size for certain demographic groups means that there are large relative errors in both the mean and the variance of duration in the activtity database.

Summary

Overall, most of the discussion on the sources of uncertainty in the estimates of annual average exposure is qualitative in nature. This is for two reasons: a) there are no direct measurements of annual CO exposure to compare the model results with, and b) there are an enormous number of combinations of ways to use alternate modeling assumptions and databases, and only a few of these have been examined in detail. The cited papers used in comparisons between models and comparison with personal monitoring estimates suggest that there is no evident bias in the mean HAPEM exposures, but that the variances (or standard deviations) are substantially underestimated compared to the true variability among individuals. The indirect nature of the comparison of HAPEM to another model (rather than to observed data) makes it difficult to assign a quantitative confidence limit to the exposure estimates. The standard errors that have been calculated for the many tables in the Appendices to this report are very small and represent only the variation arising from the method of selecting activity patterns. The calculation based on an alternative selection scheme presented above illustrates that the model can easily produce larger variances in exposure with only modest changes in the model assumptions. The discussion on the microenvironmental factors points out that the exposures (and the concentrations and durations) in those microenvironments with high accumulated exposure tend to have much smaller relative errors than those with low accumulated exposures.

References

Law P.L., P.J. Lioy, M.P. Zelenka, A.H. Huber, T.R.McCurdy, (1997), “Evaluation of a probabilistic exposure model applied to carbon monoxide (pNEM/CO) using Denver personal exposure monitoring data”, Journal of the Air & Waste Management Association **47**:491-500.

IT (International Technology) Corporation, (1996), “Developmental research for the hazardous air pollutant exposure model (HAPEM) as applied to mobile source pollutants”, IT Project number 453212-3, Report to the U.S. Environmental Protection Agency.

TABLE 5

HAPEM-MS3 Counts of Activity Patterns by Activity Pattern State

State 1 = Cool Non-summer Weekday
State 2 = Cool Non-summer Weekend
State 3 = Warm Non-summer Weekday
State 4 = Warm Non-summer Weekend
State 5 = Cool Summer Weekday
State 6 = Cool Summer Weekend
State 7 = Warm Summer Weekday
State 8 = Warm Summer Weekend

	State								
Demographic Group	1	2	3	4	5	6	7	8	All
Caucasians	316	154	496	162	610	125	209	212	2284
African Americans	14	8	25	10	41	9	15	11	133
Hispanics	16	11	35	11	44	10	17	11	155
Household income lt $10K	19	8	43	6	17	3	15	9	120
Household income $10K-$25K	59	28	69	25	103	21	25	37	367
Household income $25K-$50K	126	69	228	86	241	47	90	90	977
Household income $50K-$75K	73	33	107	30	133	20	41	40	477
Household income gt $75K	12	2	10	6	48	16	20	13	127
Children, 0 to 17	130	54	198	68	225	29	80	80	864
Nonworking men, 18 to 44	13	13	22	5	20	9	2	4	88
Working men, 18 to 44	207	72	135	32	117	29	20	33	645
Nonworking women, 18 to 44	76	35	63	26	52	12	32	28	324
Working women, 18 to 44	219	98	143	47	80	15	32	34	668
Nonworking men, 45 to 64	23	6	7	4	6	1	4	2	53
Working men, 45 to 64	62	25	62	16	36	13	6	12	232
Nonworking women, 45 to 64	67	32	44	15	41	11	19	9	238
Working women, 45 to 64	77	30	57	17	33	5	17	8	244
Men, 65+	21	7	21	3	20	4	7	7	90
Women, 65+	35	9	22	9	26	8	7	6	122
Outdoor workers	11	5	19	4	8	1	2	2	52
Outdoor children	43	14	68	25	112	18	42	46	368
Heart and respiratory	33	14	51	5	42	10	23	12	190
All persons	930	381	774	242	656	136	226	223	3568

TABLE 6

HAPEM-MS3 Average Annual Exposure
Standard Deviations (in μg/m^3) for Independent and Correlated Activity Sequences
Ratio = Correlated/Independent
Denver: Air District=Citywide

Demographic Group	Population	Annual Mean	Independent Std. Dev.	Correlated Std. Dev.	Ratio
Caucasians	1529709	561	1.53	7.76	5.08
African americans	96042	692	3.37	19.34	5.74
Hispanics	225415	586	2.02	11.57	5.73
Household income lt $10K	253323	572	1.20	6.71	5.59
Household income $10K-$25K	517117	582	1.47	7.19	4.91
Household income $25K-$50K	607809	566	1.78	9.14	5.14
Household income $50K-$75K	238244	555	1.44	7.34	5.11
Household income gt $75K	135089	530	1.42	7.84	5.54
Children, 0 to 17	449919	564	1.49	7.34	4.92
Nonworking men, 18 to 44	97037	603	1.41	7.39	5.23
Working men, 18 to 44	318840	604	0.99	5.19	5.24
Nonworking women, 18 to 44	154786	556	1.31	6.47	4.93
Working women, 18 to 44	269914	576	0.71	3.63	5.09
Nonworking men, 45 to 64	37533	557	1.01	4.47	4.41
Working men, 45 to 64	119334	605	0.87	4.76	5.46
Nonworking women, 45 to 64	62442	537	1.00	5.18	5.17
Working women, 45 to 64	100773	564	0.77	3.97	5.16
Men, 65+	79550	590	1.37	7.42	5.40
Women, 65+	85930	561	1.23	6.80	5.51
Outdoor workers	49647	700	0.96	4.97	5.15
Outdoor children	127781	588	1.18	6.04	5.13
Heart and respiratory	236093	544	1.13	5.62	4.97
All persons	1776026	575	1.60	8.35	5.20

TABLE 7

HAPEM-MS3 Percentiles in Micrograms per cubic meter (μg/m^3)
Correlated and Uncorrelated Activity Sequences
Denver: City-wide Annual CO Exposure for 1990

Demographic Group	25th	50th	75th	90th	95th	99th
<u>Correlated Sequences</u>						
Caucasians	507	555	602	819	851	886
African americans	550	761	816	848	865	894
Hispanics	508	564	632	790	821	866
Household income lt $10K	483	539	728	787	804	828
Household income $10K-$25K	507	557	627	826	848	877
Household income $25K-$50K	512	557	600	817	853	892
Household income $50K-$75K	512	562	592	820	857	889
Household income gt $75K	479	552	589	828	858	892
Children, 0 to 17	512	558	600	814	846	879
Nonworking men, 18 to 44	522	579	661	866	888	918
Working men, 18 to 44	556	594	651	843	865	886
Nonworking women, 18 to 44	488	539	606	802	823	850
Working women, 18 to 44	526	566	623	817	831	846
Nonworking men, 45 to 64	485	536	601	801	817	837
Working men, 45 to 64	562	600	661	847	871	892
Nonworking women, 45 to 64	474	522	585	781	799	822
Working women, 45 to 64	519	559	615	805	821	838
Men, 65+	501	558	655	832	851	877
Women, 65+	475	530	633	785	801	824
Outdoor workers	651	681	748	928	952	973
Outdoor children	533	585	623	858	887	916
Heart and respiratory	480	527	589	783	802	825
All persons	511	561	618	832	859	893
<u>Uncorrelated Sequences</u>						
Caucasians	504	561	614	838	847	854
African americans	548	777	817	823	826	831
Hispanics	504	553	610	781	816	827
Household income lt $10K	482	531	748	790	793	798
Household income $10K-$25K	506	558	616	834	839	845
Household income $25K-$50K	510	562	576	838	848	856
Household income $50K-$75K	509	569	576	844	855	862
Household income gt $75K	499	564	573	847	853	859
Children, 0 to 17	510	562	613	817	844	851
Nonworking men, 18 to 44	518	583	644	872	877	883
Working men, 18 to 44	559	597	660	857	861	866
Nonworking women, 18 to 44	484	542	598	810	816	821
Working women, 18 to 44	530	569	629	825	828	831
Nonworking men, 45 to 64	481	540	594	808	812	816
Working men, 45 to 64	558	604	672	864	868	872
Nonworking women, 45 to 64	470	526	581	789	793	798
Working women, 45 to 64	513	563	622	815	818	821
Men, 65+	495	557	656	832	836	841
Women, 65+	467	526	621	785	788	792
Outdoor workers	652	684	752	946	951	955
Outdoor children	531	591	601	860	887	893
Heart and respiratory	474	530	585	791	795	800
All persons	509	565	621	843	850	857

TABLE 8

HAPEM-MS3 Fractiles in Percent (%) of Annual CO Exposure for 1990
Correlated and Uncorrelated Activity Sequences
Denver: City-wide

	μg/m³										
Demographic Group	100	200	300	400	500	600	700	800	900	1000	1100
Correlated Sequences											
Caucasians	0.0	0.0	6.4	12.7	22.0	74.7	85.9	88.1	99.6	100.0	100.0
African americans	0.0	0.0	0.7	1.8	7.5	39.6	42.2	66.7	99.3	100.0	100.0
Hispanics	0.0	0.0	1.4	3.6	21.1	64.6	81.0	91.8	99.9	100.0	100.0
Household income lt $10K	0.0	0.0	1.8	9.1	32.2	68.8	73.0	94.1	100.0	100.0	100.0
Household income $10K-$25K	0.0	0.0	3.0	9.1	21.4	68.4	80.5	84.9	99.9	100.0	100.0
Household income $25K-$50K	0.0	0.0	5.7	10.4	19.7	75.1	86.2	88.4	99.4	100.0	100.0
Household income $50K-$75K	0.0	0.0	9.2	14.6	20.0	78.2	88.0	88.9	99.6	100.0	100.0
Household income gt $75K	0.0	0.0	15.8	23.8	30.3	79.1	86.3	87.5	99.5	100.0	100.0
Children, 0 to 17	0.0	0.0	6.6	10.4	19.1	75.1	86.4	88.5	99.8	100.0	100.0
Nonworking men, 18 to 44	0.0	0.0	3.6	13.4	16.6	60.3	77.9	79.2	97.2	100.0	100.0
Working men, 18 to 44	0.0	0.0	1.1	12.1	12.3	58.2	82.9	86.5	99.8	100.0	100.0
Nonworking women, 18 to 44	0.0	0.0	4.7	12.3	31.5	73.7	80.8	89.6	100.0	100.0	100.0
Working women, 18 to 44	0.0	0.0	4.8	12.0	12.6	73.6	83.4	87.4	100.0	100.0	100.0
Nonworking men, 45 to 64	0.0	0.0	6.1	10.4	33.9	74.8	80.0	89.7	100.0	100.0	100.0
Working men, 45 to 64	0.0	0.0	0.3	13.1	13.3	50.9	83.8	87.2	99.6	100.0	100.0
Nonworking women, 45 to 64	0.0	0.0	7.4	11.6	38.2	78.3	81.7	95.3	100.0	100.0	100.0
Working women, 45 to 64	0.0	0.0	7.9	12.9	15.0	73.4	84.4	89.1	100.0	100.0	100.0
Men, 65+	0.0	0.0	3.8	8.6	24.8	64.3	76.5	81.4	99.9	100.0	100.0
Women, 65+	0.0	0.0	3.8	8.1	36.1	72.2	75.9	94.9	100.0	100.0	100.0
Outdoor workers	0.0	0.0	0.0	2.1	7.8	7.9	67.2	84.2	88.8	100.0	100.0
Outdoor children	0.0	0.0	7.0	11.2	11.7	64.7	86.0	87.2	97.2	100.0	100.0
Heart and respiratory	0.0	0.0	5.5	11.0	34.3	77.7	81.6	94.7	100.0	100.0	100.0
All persons	0.0	0.0	5.7	11.5	19.9	70.7	82.8	85.5	99.4	100.0	100.0
Uncorrelated Sequences											
Caucasians	0.0	0.0	6.4	12.7	21.1	74.4	86.1	86.5	100.0	100.0	100.0
African americans	0.0	0.0	0.7	1.8	6.9	39.8	41.8	56.3	100.0	100.0	100.0
Hispanics	0.0	0.0	1.4	3.6	20.0	60.0	80.7	93.1	100.0	100.0	100.0
Household income lt $10K	0.0	0.0	1.7	9.1	31.4	70.2	72.6	99.8	100.0	100.0	100.0
Household income $10K-$25K	0.0	0.0	3.0	9.1	22.6	65.4	80.6	84.8	100.0	100.0	100.0
Household income $25K-$50K	0.0	0.0	5.7	10.4	17.6	75.6	86.4	86.7	100.0	100.0	100.0
Household income $50K-$75K	0.0	0.0	9.2	14.6	15.3	79.7	88.2	88.2	100.0	100.0	100.0
Household income gt $75K	0.0	0.0	15.8	23.8	25.9	81.0	86.5	86.6	100.0	100.0	100.0
Children, 0 to 17	0.0	0.0	6.6	10.4	18.5	74.1	86.5	87.0	100.0	100.0	100.0
Nonworking men, 18 to 44	0.0	0.0	3.6	13.4	13.4	64.4	78.4	78.4	100.0	100.0	100.0
Working men, 18 to 44	0.0	0.0	0.0	12.3	12.3	65.1	83.4	88.0	100.0	100.0	100.0
Nonworking women, 18 to 44	0.0	0.0	4.7	12.3	35.3	76.9	80.7	86.7	100.0	100.0	100.0
Working women, 18 to 44	0.0	0.0	4.9	12.0	12.0	73.6	83.4	87.1	100.0	100.0	100.0
Nonworking men, 45 to 64	0.0	0.0	6.1	10.4	36.3	77.7	80.0	86.1	100.0	100.0	100.0
Working men, 45 to 64	0.0	0.0	0.0	13.3	13.3	47.0	84.7	88.4	100.0	100.0	100.0
Nonworking women, 45 to 64	0.0	0.0	7.4	11.6	37.1	78.7	81.7	99.8	100.0	100.0	100.0
Working women, 45 to 64	0.0	0.0	8.0	12.9	12.9	73.3	84.4	87.4	100.0	100.0	100.0
Men, 65+	0.0	0.0	3.8	8.6	27.2	59.4	76.5	78.1	100.0	100.0	100.0
Women, 65+	0.0	0.0	3.7	8.1	34.0	72.8	75.6	100.0	100.0	100.0	100.0
Outdoor workers	0.0	0.0	0.0	3.2	7.8	7.8	70.5	84.3	88.9	100.0	100.0
Outdoor children	0.0	0.0	7.0	11.2	11.2	74.9	86.8	87.2	100.0	100.0	100.0
Heart and respiratory	0.0	0.0	5.5	11.0	34.4	79.0	81.6	99.1	100.0	100.0	100.0
All persons	0.0	0.0	5.7	11.5	17.0	70.8	83.0	83.2	100.0	100.0	100.0

CHAPTER 3
QUALITY ASSURANCE

Introduction

In the application of the HAPEM model to the fourteen study areas examined in this report, two main activities were carried out which directly involved quality assurance considerations. These were 1) enhancements to the HAPEM-MS3 model, and 2) obtaining data and running the enhanced model for the 14 study areas. The model enhancements did not change the model assumptions or algorithms, but they improved model performance in several key areas including data management, speed of execution, calculation of mean and variance, the production of tables of exposure broken down by county, and tables broken down by microenvironment. The model runs included the San Francisco study area, which was run previously under HAPEM-MS3 and served as a basis for comparing the model output before and after the enhancements.

Handling of Program Code

The HAPEM-MS3 model is written in Fortran and runs on the EPA's IBM mainframe. Three sets of directories on the mainframe were established to hold program code. Each of these sets has three pieces, a subdirectory for the Fortran source code (.FORT), another for the compiled code (.LOAD), and a third for the compilation programs and the Job Control Language (JCL) used to submit runs (.CNTL). The original code (before enhancements were added) is maintained in the 'EXPO.JLCEXPO.HAPEM3' directory (with .FORT, .LOAD, and .CNTL subdirectories). Modified programs under development and testing are in the 'EXPO.HAPEM3.NEW' directory. Once the modified versions are in final form they are placed in the 'EXPO.HAPEM3.FINAL' directory. Copies of the source code and the JCL are also maintained on local PCs both for backup and for ease of printing. Access to the code (both read and write) on the IBM is open to those individuals who are part of the EXPO group.

Enhancements to HAPEM-MS3

In the following discussion of the program modifications and testing of the HAPEM program, all tables and results report exposure in parts per million (ppm). This is the base unit used by all the HAPEM programs. The output tables for the fourteen study runs report exposure in micrograms per cubic meter (µg/m3). This unit conversion was made during the post-processing phase.

There are 10 main programs in HAPEM-MS3, which can be divided into three groups: those that did not change; those with minor changes; and new programs.

PROGRAM	CHANGES	SUMMARY DESCRIPTION OF CHANGES
TSERIES	none	
AQAVG	none	
DIST90	none	
TVLTIME	none	
ODEST	minor	DIST, ODEST, TTFRA combined (to run faster)
POP90	minor	County level information retained
MECONC	new	Used to be part of MERGE, now separate
DURAVG	new	Part of mean and variance calculations
EXPCODI	new	Creates tables by county and air district
EXPMEHR	new	Creates tables by microenvironment and hour

The change to the ODEST program consisted of simply combining three programs that previously ran separately into one program. This resulted in a considerable improvement in program execution time, in what was the slowest step in the model. The improvement came solely from avoiding the need to write out millions of data records that were only used when read back in again into the next program. The output from the modified ODEST program (a file called HOMEWORK, consisting of a 20 x 20 array of partitioning factors) was compared to the output from the original ODEST. There was general agreement between the files, with occasional differences in the least significant digit. These small differences arise from rounding and do not have an appreciable impact on subsequent analysis.

The POP90 program was modified to retain the information on county of residence that was previously not needed by the model. The output file now has two extra fields (the county name and number) per record. There are also more records, since each county is now done separately. The population numbers on the new output file were checked in two ways: across counties the totals match the totals from the previous version of POP90, and within counties the ALL PERSONS total matches the total county population from the census.

There were four new programs added to HAPEM-MS3 which replaced the MERGE and GRAPH programs. These were DURAVG, MECONC, EXPCODI, and EXPMEHR.

The DURAVG program calculates the average duration in each microenvironment (by hour) for each of the defined demographic groups, by summing all the relevant records in the time-activity database. In the previous version of the model, a random selection of these patterns was made and the resulting time series was used to estimate annual exposure. BY repeated runs, both the mean and variance could be estimated. The direct calculation of the mean by DURAVG is useful in the calculation of the tables with microenvironment and hourly exposure breakdowns, since for these tables no variance estimate is provided. The DURAVG program is not used in the calculation of the quarterly and annual totals, since for those a variance estimate is needed. The DURAVG program was checked by summing the average durations across all microenvironments. These showed a total of 60 minutes of in every hour, as is necessary.

The MECONC program uses the ambient air quality data and a file of microenvironmental factors to calculate pollutant concentration in each of the 37 microenvironments. The resultant output file should have equal concentrations for any pair of microenironments with equal factors. This was found to be the case.

The EXPCODI program calculates exposures summarized two ways: by county, and by air monitor district. Both of these tables can then be summed independently to produce a city-wide average. These were carried out and agree to machine precision (six digits). The program was also tested in another way, by creating a special input file with all the microenvironmental input concentrations set to 1.0 ppm for all air districts. As a result, the exposure for everyone must total 1.0 ppm regardless of activity pattern or location. Table 9 shows the mean exposure by cohort calculated to single precision. The exposure is correct within 3 parts per million, which is the appropriate size for the accumulated roundoff error at single precision. The internal tables used by HAPEM are kept to double precision, and show a mean exposure of 1.0 ppm and a variance of zero to at least seven significant digits.

The EXPMEHR program calculates exposures by hour and demographic group, and also calculates tables of total quarterly and annual exposure by microenvironment. The hourly tables were previously produced by the MERGE program, and the microenvironment tables are new. Both sets of tables were summed to calculate the overall exposure, and were found to be consistent with each other and also consistent with the city-wide average calculated by the EXPCODI program. Tables 10 and 11 show the exposure by hour and by microenvironment. In the latter table, the accumulated exposure is given, which is proportional to the time spent in each microenvironment, so the numbers are not uniform. However, the total across all microenvironments should equal the number of hours in the quarters and in the year (i.e. 2160, 2184, 2208, 2208, and 8760, respectively). These totals come within roundoff error of the correct totals. Table 12 is an extract from the EXPMEHR run on the IBM, showing the mean exposures calculated to single precision.

In general, single precision is adequate for most of the HAPEM-MS3 calculations. There were some exceptions, with the most significant of these being in the accumulation of the sums and the sum of squares of the exposures, used to calculate mean and variance. The reason double precision is needed here is because the variance calculation involves subtracting two nearly equal values, for which single precision is not adequate. The old HAPEM-MS3 programs usually use single precision except in special cases such as the commuting algorithm. The new programs (DURAVG, MECONC, EXPCODI, and EXPMEHR) use double precision as standard.

The new programs were tested on the data used for the 1990 San Francisco study. The results from the earlier version of this study (before program modification) are given in the 1996 report by IT Corporation entitled 'Development and Evaluation of Enhancements to the Hazardous Air Pollutant Exposure Model (HAPEM-MS3)'. These numbers are given in Table 13 under the heading 'IT 1996'. Also provided in the table are the results of a single run of HAPEM-MS3 performed in late 1997 (under the heading 'TEST97'), and the results of the new algorithms for calculating mean and variance of exposure. These algorithms were tested in both SAS and Fortran implementations to ensure that two very different program styles gave similar results (i.e. that the program code actually carried out the algorithm as intended). The SAS versions were not calculated for the five commuting demographic groups since it would have been time-consuming to implement the commuting algorithm in the SAS program. In this table, the populations of the demographic groups are also given as calculated for the present set of model runs. These numbers agree with the population totals for each demographic group given in the 1996 IT report, verifying that the same definitions were used for each of the demographic groups. Note that the exposure of the demographic group 'all persons' was not evaluated in the 1996 IT report and so was not run for these comparative tests.

Data Management and Verification

Before the current enhancements to HAPEM-MS3, the data management problem was significant, especially if 14 study areas were to be run. The main problems were that although HAPEM-MS3 was generally able to handle multiple demographic groups, key programs (e.g. MERGE) only handled a single demographic group at a time. Also, in order to obtain variance estimates, multiple runs (usually ten times) were required per demographic group. This implies more than 200 runs per study area, each producing multiple output files. In the reorganized version, only a single run is needed to calculate mean and variance for all demographic groups, which are sumarized in four output files per study area. On the IBM mainframe, four partitioned data sets (named EXPCO, EXPDI, EXPHR, EMPME) in the 'EXPO.HAPEM3.FINAL' directory. Each of these contains one member for each study area, with a standard three letter code indicating the study area, and a two digit suffix for the year. (In addition, each contains a fifteenth member for the DUMMY run used in program testing). In the sensitivity analysis, several variations of the microenvironmental factors file are being tested. In order to distinguish them, each factors file has a numerical suffix. The same suffix is then added to the HAPEM-MS3 output file names, so as to clearly indicate which data were used for those runs.

Study Area	Code used in file names
Baltimore	BAL
Boston	BOS
Chicago	CHI
Denver	DEN
Houston	HOU
Los Angeles	LAX
Minneapolis	MSP
New York City	NYC
Philadelphia	PHL
Phoenix	PHX
San Francisco	SFB
Spokane	SPO
St. Louis	STL
Washington, D.C.	WDC

The input data to HAPEM-MS3 is of four types. The time-activity data is a single file (‘EXPO.HAPEM3.FINAL.MEDUR.DATA’) used for all the study areas. This file contains 3568 person-days of activity data, the same as the file used for the previous HAPEM-MS3 runs. The data has been reorganized into an array of hour and microenvironment totals per person, whereas the file previously used contained a time sequence of activities. The current form (the array) was summed and compared to the time-sequence file, and was found to agree exactly. This file is read by the DURAVG and MECONC programs, which are new, and which have a logically simpler structure if the data is explicitly in array form.

The second type of input data is meteorological data, consisting of daily mean and maximum temperatures for each day of the study year. This data was uploaded from a PC into the file ‘EXPO.HAPEM3.FINAL.MET’, with each study area saved as a member using the naming convention described above. This data is used to count the frequency of each type of day, for sampling from the activity database. The counts for San Francisco agreed with the counts obtained in the earlier San Francisco runs.

The third type of input data is population counts for demographic groups from the 1990 census. The data for the 14 study areas was provided by the EPA, and was uploaded into 25 partitioned data sets on the IBM, each data set containing 14 members. The structure of creating and maintaining these 25 data sets is somewhat awkward, but this was the form used by the POP90 program. This program was altered slightly to retain county information on its output, but the data input routines were not altered, so the 25 separate files were retained. The population counts at the county level match the known county totals.

The fourth type of input data is air quality data for carbon monoxide (CO) extracted from the AIRS database. This data is summarized in Table 1 in Chapter 1 of this report. The monitor ID, the number of valid measurements, the minimum, the mean, and the maximum at each monitor, were compared to the annual summary statistics provided by AIRS itself, and were found to agree. The data is then input into the TSERIES and AQAVG programs, which were not altered except to remove some unnecessary PRINT statements and comments. The output from TSERIES and AQAVG was checked, as was found to agree with earlier outputs, except in the case where the original AIRS monitor data was missing. In these cases, TSERIES estimates the missing values using a second-order autoregressive algorithm with an additional random term. The random term cannot be reproduced from run to run since it is reinitialized using the computer's internal clock on each run. Therefore, if an exact comparison of two HAPEM runs is desired, then the TSERIES part should not be recalculated (i.e. use the results from the first run directly in the second run).

TABLE 9

Extract of Summary Data from Dummy Run of EXPCODI Program

```
 IEF375I  JOB/NZC     /START 1998043.1459
 IEF376I  JOB/NZC     /STOP  1998043.1501 CPU    1MIN 08.27SEC
SRB    0MIN 00.13SEC
 Compute statistics for cohorts,districts,states,qs
 Read population numbers
 Read commuting fractions
 Compute populations for all cohorts
 Compute states
 Compute statistics for cohorts,districts,quarters
 Compute statistics for cohorts,quarters by county
 Compute statistics for cohorts,quarters by district
 COHORT MEAN=          1  0.999998689
 COHORT MEAN=          2  0.999998033
 COHORT MEAN=          3  0.999998033
 COHORT MEAN=          4  0.999998271
 COHORT MEAN=          5  0.999998868
 COHORT MEAN=          6  0.999999046
 COHORT MEAN=          7  0.999997854
 COHORT MEAN=          8  0.999998093
 COHORT MEAN=          9  0.999998748
 COHORT MEAN=         10  0.999998033
 COHORT MEAN=         11   1.00000000
 COHORT MEAN=         12  0.999997973
 COHORT MEAN=         13   1.00000000
 COHORT MEAN=         14  0.999999046
 COHORT MEAN=         15   1.00000000
 COHORT MEAN=         16  0.999999344
 COHORT MEAN=         17  0.999999344
 COHORT MEAN=         18  0.999998391
 COHORT MEAN=         19  0.999998510
 COHORT MEAN=         20  0.999999881
 COHORT MEAN=         21  0.999998510
 COHORT MEAN=         22  0.999997854
 COHORT MEAN=         23  0.999999106
 COHORT MEAN=         24  0.999999046
 COHORT MEAN=         25  0.999998927
 COHORT MEAN=         26  0.999999166
 COHORT MEAN=         27  0.999998271
```

TABLE 10

Accumulated Exposures by Hour for Dummy Data (Program EXPMEHR)

			QUARTER 1		QUARTER 2		QUARTER 3		QUARTER 4		ANNUAL	
#	HOUR	POP.	EXPOS.	PCT.	EXPOS.	PCT.	EXPOS.	PCT.	EXPOS.	PCT.	EXPOS.	PCT.
1	1	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	2	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	3	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	4	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	5	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	6	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	7	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	8	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	9	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	10	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	11	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	12	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	13	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	14	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	15	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	16	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	17	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	18	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	19	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	20	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	21	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	22	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	23	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17
1	24	1529709.	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17	1.000000	4.17

TABLE 11

Accumulated Exposures (ppm-hr) by Microenvironment for Dummy Data (Program EXPMEHR)

#	MICRO.	POP.	QUARTER 1 EXPOS.	QUARTER 1 PCT.	QUARTER 2 EXPOS.	QUARTER 2 PCT.	QUARTER 3 EXPOS.	QUARTER 3 PCT.	QUARTER 4 EXPOS.	QUARTER 4 PCT.	ANNUAL EXPOS.	ANNUAL PCT.
1	1	1529709.	83.888556	3.88	93.634297	4.29	100.904707	4.57	87.248658	3.95	365.676218	4.17
1	2	1529709.	7.313438	0.34	5.329527	0.24	3.638751	0.16	7.403668	0.34	23.685384	0.27
1	3	1529709.	3.317586	0.15	6.166963	0.28	7.188097	0.33	5.033009	0.23	21.705655	0.25
1	4	1529709.	5.283864	0.24	5.232671	0.24	4.625730	0.21	4.587035	0.21	19.729300	0.23
1	5	1529709.	3.089162	0.14	1.752048	0.08	1.010229	0.05	2.669499	0.12	8.520938	0.10
1	6	1529709.	11.408419	0.53	12.311893	0.56	13.614496	0.62	12.330113	0.56	49.664922	0.57
1	7	1529709.	50.701468	2.35	70.542046	3.23	85.551083	3.87	52.314579	2.37	259.109175	2.96
1	8	1529709.	0.366637	0.02	1.071345	0.05	1.947179	0.09	0.233125	0.01	3.618285	0.04
1	9	1529709.	2.379756	0.11	1.679131	0.08	1.081881	0.05	2.303527	0.10	7.444294	0.08
1	10	1529709.	0.111380	0.01	0.301385	0.01	0.605846	0.03	0.131209	0.01	1.149819	0.01
1	11	1529709.	0.835648	0.04	0.860058	0.04	0.659360	0.03	1.058059	0.05	3.413125	0.04
1	12	1529709.	0.016566	0.00	0.032430	0.00	0.020527	0.00	0.035081	0.00	0.104604	0.00
1	13	1529709.	383.762508	17.77	404.121189	18.50	435.879941	19.74	387.614563	17.56	1611.378202	18.39
1	14	1529709.	294.465560	13.63	274.652999	12.58	252.390261	11.43	295.147155	13.37	1116.655976	12.75
1	15	1529709.	622.674966	28.83	625.144410	28.62	633.391172	28.69	640.340789	29.00	2521.551336	28.78
1	16	1529709.	273.191685	12.65	245.199821	11.23	221.121978	10.01	276.180561	12.51	1015.694045	11.59
1	17	1529709.	76.266908	3.53	79.284855	3.63	82.054748	3.72	80.452520	3.64	318.059031	3.63
1	18	1529709.	48.628480	2.25	46.816278	2.14	50.265892	2.28	48.364887	2.19	194.075537	2.22
1	19	1529709.	28.784362	1.33	28.059512	1.28	29.797676	1.35	29.534727	1.34	116.176277	1.33
1	20	1529709.	25.371086	1.17	23.625065	1.08	23.982064	1.09	24.775413	1.12	97.753628	1.12
1	21	1529709.	132.372796	6.13	99.010908	4.53	62.279231	2.82	133.966412	6.07	427.629347	4.88
1	22	1529709.	14.438729	0.67	13.605791	0.62	11.841169	0.54	14.916047	0.68	54.801735	0.63
1	23	1529709.	2.559327	0.12	2.553705	0.12	2.309947	0.10	3.428247	0.16	10.851226	0.12
1	24	1529709.	6.335354	0.29	8.267466	0.38	8.329330	0.38	6.394365	0.29	29.326515	0.33
1	25	1529709.	22.093345	1.02	19.347466	0.89	15.520672	0.70	23.097307	1.05	80.058790	0.91
1	26	1529709.	5.993680	0.28	8.324292	0.38	9.784408	0.44	6.738874	0.31	30.841254	0.35
1	27	1529709.	21.414750	0.99	24.488207	1.12	27.487000	1.24	23.845650	1.08	97.235608	1.11
1	28	1529709.	0.361528	0.02	0.403156	0.02	0.631482	0.03	0.325029	0.01	1.721195	0.02
1	29	1529709.	1.400590	0.06	2.041705	0.09	2.235540	0.10	2.397774	0.11	8.075610	0.09
1	30	1529709.	16.580161	0.77	39.657468	1.82	55.525212	2.51	18.636365	0.84	130.399206	1.49
1	31	1529709.	6.322506	0.29	6.109973	0.28	5.212024	0.24	7.194568	0.33	24.839072	0.28
1	32	1529709.	0.725828	0.03	4.355367	0.20	7.847709	0.36	1.038109	0.05	13.967013	0.16
1	33	1529709.	4.861221	0.23	15.318178	0.70	23.500621	1.06	4.974094	0.23	48.654114	0.56
1	34	1529709.	2.376354	0.11	14.027173	0.64	24.707418	1.12	3.061130	0.14	44.172075	0.50
1	35	1529709.	0.305798	0.01	0.258884	0.01	0.301025	0.01	0.227852	0.01	1.093558	0.01
1	36	1529709.	0.000000	0.00	0.000000	0.00	0.000000	0.00	0.000000	0.00	0.000000	0.00
1	37	1529709.	0.000000	0.00	0.412334	0.02	0.755595	0.03	0.000000	0.00	1.167929	0.01
1	ALL	1529709.	2160.000002	100.00	2183.999996	100.00	2208.000001	100.00	2208.000000	100.00	8759.999998	100.00

TABLE 12

Extract from Run of EXPMEHR Program Using Dummy Data

```
IEF375I  JOB/NZC     /START 1998043.1516
IEF376I  JOB/NZC     /STOP  1998043.1517 CPU    0MIN 11.68SEC SRB    0MIN 00.09SEC

CALCULATE EXPOSURE BY HOUR
COHORT  1  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  2  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  3  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  4  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  5  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  6  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  7  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  8  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  9  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 10  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 11  MEAN USING HRS =   1.0000070   1.0000070   1.0000070   1.0000070   1.0000070
COHORT 12  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 13  MEAN USING HRS =   1.0000069   1.0000069   1.0000069   1.0000069   1.0000069
COHORT 14  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 15  MEAN USING HRS =   1.0000085   1.0000085   1.0000085   1.0000085   1.0000085
COHORT 16  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 17  MEAN USING HRS =   1.0000066   1.0000066   1.0000066   1.0000066   1.0000066
COHORT 18  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 19  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 20  MEAN USING HRS =   1.0000034   1.0000034   1.0000034   1.0000034   1.0000034
COHORT 21  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 22  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 23  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 24  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 25  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 26  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 27  MEAN USING HRS =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000

CALCULATE EXPOSURE BY MICROENVIRONMENT
```

```
COHORT  1  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  2  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  3  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  4  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  5  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  6  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  7  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  8  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT  9  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 10  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 11  TOT  USING MES =   1.0000070   1.0000070   1.0000070   1.0000070   1.0000070
COHORT 12  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 13  TOT  USING MES =   1.0000069   1.0000069   1.0000069   1.0000069   1.0000069
COHORT 14  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 15  TOT  USING MES =   1.0000085   1.0000085   1.0000085   1.0000085   1.0000085
COHORT 16  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 17  TOT  USING MES =   1.0000066   1.0000066   1.0000066   1.0000066   1.0000066
COHORT 18  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 19  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 20  TOT  USING MES =   1.0000034   1.0000034   1.0000034   1.0000034   1.0000034
COHORT 21  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 22  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 23  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 24  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 25  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 26  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
COHORT 27  TOT  USING MES =   1.0000000   1.0000000   1.0000000   1.0000000   1.0000000
PROGRAM COMPLETED
```

TABLE 13

COMPARISON OF CITYWIDE CO EXPOSURE (PPM) FOR SAN FRANCISCO, 1990

		AVERAGING				IT 1996		TEST97
		SAS		FORTRAN			STANDARD	
DEMOGRAPHIC GROUP	POPULATION	MEAN	VARIANCE	MEAN	VARIANCE	MEAN	DEVIATION	MEAN
CAUCASIANS	2831738	2.27884	.0008992	2.27875	.0008993	2.28	0.026	2.240
AFRICAN AMERICANS	432630	1.34831	.0004823	1.34830	.0004823	1.34	0.021	1.380
HISPANICS	699995	1.67988	.0007567	1.67987	.0007567	1.70	0.030	1.690
HOUSEHOLD INCOME LT $10K	489747	1.42162	.0006038	1.42160	.0006038	1.44	0.028	1.422
HOUSEHOLD INCOME $10K-$25K	887348	1.76267	.0006482	1.76266	.0006482	1.76	0.022	1.687
HOUSEHOLD INCOME $25K-$50K	1378859	2.34770	.0008041	2.34764	.0008041	2.36	0.021	2.341
HOUSEHOLD INCOME $50K-$75K	833376	2.41591	.0007801	2.41589	.0007801	2.42	0.031	2.440
HOUSEHOLD INCOME $75K+	720938	2.07232	.0008495	2.07230	.0008495	2.08	0.030	2.087
CHILDREN, 0 TO 17	962581	2.25152	.0008286	2.25147	.0008286	2.26	0.026	2.276
NONWORKING MEN, 18 TO 44	274000	2.02675	.0005021	2.02674	.0005021	2.02	0.015	2.080
WORKING MEN, 18 TO 44	760144	.	.	2.03289	.0003107	2.04	0.020	2.021
NONWORKING WOMEN, 18 TO 44	420685	2.21329	.0009563	2.21327	.0009563	2.19	0.039	2.216
WORKING WOMEN, 18 TO 44	608473	.	.	1.92241	.0002813	1.91	0.015	1.935
NONWORKING MEN, 45 TO 65	105964	2.30000	.0006274	2.29999	.0006274	2.38	0.023	2.313
WORKING MEN, 45 TO 64	297023	.	.	2.36218	.0002794	2.37	0.015	2.360
NONWORKING WOMEN, 45 TO 64	171842	2.07083	.0011174	2.07081	.0011174	2.09	0.033	2.036
WORKING WOMEN, 45 TO 64	240840	.	.	2.39025	.0002846	2.39	0.011	2.383
MEN,65+	237342	2.28874	.0008037	2.28872	.0008037	2.29	0.035	2.232
WOMEN, 65+	250116	2.30820	.0010813	2.30818	.0010813	2.33	0.032	2.323
OUTDOOR CHILDREN	269922	2.16198	.0006382	2.16197	.0006382	2.17	0.025	2.186
OUTDOOR WORKERS	119781	.	.	2.29645	.0002857	2.31	0.014	2.315
HEART AND RESPIRATORY	609219	1.78158	.0008838	1.78157	.0008838	1.78	0.015	1.770