



# Air Pollutants Exposure Model Documentation (APEX, Version 5.2)

Volume I: User's Guide



EPA-452/R-19-005a  
October 2019

Air Pollutants Exposure Model Documentation (APEX, Version 5.2) Volume I: User's Guide

U.S. Environmental Protection Agency  
Office of Air Quality Planning and Standards  
Health and Environmental Impacts Division  
Research Triangle Park, NC

## **DISCLAIMER**

This document has been prepared at least partially by ICF (under EPA Contract No. EP-W-12-0101 [WA 4-55]). It has been subject to the Agency's review, and has been approved for publication as an EPA document. Mention of trade names or commercial products is not intended to constitute endorsement or recommendation for use.

## **CONTACT**

Questions should be addressed to John E. Langstaff, U.S. Environmental Protection Agency, C504-06, Research Triangle Park, North Carolina 27711 (email: [langstaff.john@epa.gov](mailto:langstaff.john@epa.gov)).

## ACKNOWLEDGEMENTS

The primary author of this document is Graham Glen (ICF). It includes contributions from Melissa Nysewander, Luther Smith, and Casson Stallings (while at Alion Science and Technology, Inc.); Stephen Graham, Kristin Isaacs, Tom McCurdy (retired), and John Langstaff (EPA); and by ICF.

# CONTENTS

CHAPTER 1. INTRODUCTION .....	1
1.1 Overview of the APEX Model.....	1
1.2 Nomenclature.....	3
1.3 Strengths and Limitations of APEX .....	4
1.3.1 Strengths .....	4
1.3.2 Limitations .....	5
1.4 Applicability .....	5
1.5 Brief History of APEX.....	6
1.6 Scope and Organization of This Guide.....	7
CHAPTER 2. SETTING UP AND RUNNING APEX .....	9
2.1 Downloading and Setting Up APEX .....	9
2.2 Setting Up an APEX Simulation.....	11
2.3 Overview of Input and Output Files .....	12
2.3.1 Input Files .....	12
2.3.2 Output Files.....	12
2.4 Overview of Model Settings and Options.....	14
CHAPTER 3. CHARACTERIZING THE STUDY AREA .....	21
3.1 APEX Spatial Units .....	21
3.1.1 Initial Study Area.....	21
3.1.2 Sectors.....	21
3.1.3 Air Quality Districts.....	23
3.1.4 Modeling Commuting.....	24
3.1.5 Meteorological Zones .....	25
3.1.6 The Final Study Area.....	25
CHAPTER 4. APEX INPUT FILES.....	26
4.1 Input File Formats.....	26
4.2 Control Options File .....	29
4.2.1 Input and Output File List Sections of the Control Options File.....	30
4.2.2 Pollutant Parameters Section of the Control Options File.....	32
4.2.3 Job Parameter Settings Section of the Control Options File.....	37
4.3 Population Sector Location File.....	51
4.4 Air District Location File.....	52
4.5 Air Quality Data File .....	53
4.5.1 Air Quality Input Data (Type 1) .....	54
4.5.2 Air Quality Input Defined as Hourly Distributions (Type 2) .....	55
4.6 Meteorology Zone Location File .....	56
4.7 Meteorology Data File .....	56
4.8 Population Data Files.....	58
4.9 Commuting Flow File.....	60
4.10 Commuting Time File.....	61
4.11 Employment Probability File.....	62
4.12 Profile Factors File.....	63
4.13 MET Mapping File .....	65
4.14 MET Distribution File.....	70

4.15	Physiological Parameters File.....	73
4.16	Ventilation File .....	76
4.17	Profile Functions (Distributions) File .....	77
4.17.1	Defining a Profile Function .....	78
4.17.2	Functions for Built-in, User-defined, AQ and Regional APEX Variables.....	80
4.18	Microenvironment Mapping File.....	88
4.19	Diary Questionnaire (DiaryQuest) File.....	91
4.20	Diary Events File .....	92
4.21	Diary Statistics File.....	93
4.22	Diary Occupations File .....	94
4.23	Diary Transitions File .....	95
4.24	Microenvironment Descriptions File .....	95
4.24.1	Microenvironment Descriptions Section .....	96
4.24.2	Parameter Descriptions Section .....	96
4.25	Prevalence File.....	103
CHAPTER 5.    APEX OUTPUT FILES.....		105
5.1	Log File.....	106
5.2	Hourly File.....	107
5.3	Timestep File .....	109
5.4	Daily File .....	111
5.5	Profile Summary (Persons) File.....	114
5.6	Microenvironmental Results File.....	117
5.7	Microenvironmental Summary File.....	120
5.8	Output Tables File.....	121
5.8.1	Exposure Summary Tables .....	122
5.8.2	Dose Summary Tables .....	128
5.8.3	ResponseProb Summary Tables .....	130
5.9	Sites File.....	131
5.10	Events File .....	131
5.11	Multipollutant File .....	133
5.12	Diary Clustering Files .....	134
5.13	Sobol Results File .....	134
REFERENCES		136

## LIST OF TABLES

Table 2-1. APEX Input Files .....	13
Table 2-2. APEX Output Files.....	14
Table 2-3. APEX Settings and Options .....	15
Table 4-1. APEX Input File Descriptions.....	27
Table 4-2. Pollutant-specific Job Parameters.....	32
Table 4-3. Output Parameter Levels in the Output Summary Table .....	34
Table 4-4. Job Parameters in the APEX Control Options File .....	41
Table 4-5. CHAD Activity Codes.....	65
Table 4-6. Available Probability Distributions in APEX .....	71
Table 4-7. Parameters in the Physiological Input File.....	73
Table 4-8. Variables that can be Defined in the Profile Functions File.....	82
Table 4-9. CHAD Location Codes.....	89
Table 4-10. CHAD Occupation Codes .....	92
Table 4-11. CHAD Locations used in Constructing the Outdoor Time and Vehicle Time Diary Statistics Files Supplied with APEX.....	94
Table 4-12. Microenvironment Parameters for the Factors and Massbal Methods.....	96
Table 4-13. Keyword Definitions for the Parameter Descriptions Section of the Microenvironment Descriptions File.....	99
Table 5-1. APEX Output Files.....	106
Table 5-2. APEX Variables Written to the Hourly Output File .....	107
Table 5-3. APEX Variables Written to the Timestep Output File.....	110
Table 5-4. APEX Variables Written to the Daily Output File.....	111
Table 5-5. APEX Variables Written to the Profile Summary File .....	114
Table 5-6. APEX Variables Written to the Microenvironmental Results File .....	118
Table 5-7. Format of the APEX Microenvironmental Summary File .....	120
Table 5-8. Interpretation of the Variables in Exposure Table Type #3 and Other "Person-Days" Based Tables.....	125



## LIST OF EXHIBITS

Exhibit 4-1. Input Files Section of an Example Control Options File.....	31
Exhibit 4-2. Output Files Section of an Example Control Options File .....	31
Exhibit 4-3. Pollutant Parameters Section of an Example Control Options File.....	37
Exhibit 4-4. Job Parameters Sections of an Example Control Options File .....	51
Exhibit 4-5. First Part of the Population Sector Location File (2010 Census) .....	52
Exhibit 4-6. First Part of an Example Air District Location File.....	52
Exhibit 4-7. First Part of an Example Air Quality Data File (Type 1) .....	55
Exhibit 4-8. First Part of an Example Air Quality Data File (Distribution Type).....	56
Exhibit 4-9. First Part of an Example Meteorology Zone Location File .....	56
Exhibit 4-10. Part of an Example Meteorology Data File .....	58
Exhibit 4-11. First Part of a Population Data File (2010 Census, Female Black or African American, Not Hispanic or Latina) .....	60
Exhibit 4-12. First Part of the Commuting Flow File (2010 Census).....	61
Exhibit 4-13. First Part of the Commuting Time File (2010 Census) .....	62
Exhibit 4-14. First Part of the Employment Probability File (2010 Census).....	63
Exhibit 4-15. First Part of an Example Profile Factors File .....	64
Exhibit 4-16. First Part of the MET Mapping File .....	70
Exhibit 4-17. First Part of the Activity-specific MET File.....	73
Exhibit 4-18. An Example of a Portion of the Physiological Parameters File .....	76
Exhibit 4-19. The APEX Ventilation Input File for VEMethod=1 .....	77
Exhibit 4-20. The APEX Ventilation Input File for VEMethod=2 .....	77
Exhibit 4-21. Examples of Profile Functions.....	86
Exhibit 4-22. First Part of an Example Microenvironment Mapping File.....	91
Exhibit 4-23. First Part of the Diary Questionnaire File.....	92
Exhibit 4-24. First Part of the Diary Events File .....	93
Exhibit 4-25. First Part of the Diary Statistics File for Time Spent Outdoors.....	94
Exhibit 4-26. First Part of a Diary Occupations File .....	95
Exhibit 4-27. Example of a Microenvironment Descriptions Section of a Microenvironment Descriptions File.....	96
Exhibit 4-28. Example of a Parameter Descriptions Section of a Microenvironment Descriptions File.....	102
Exhibit 4-29. First Part of an Example Prevalence File.....	104
Exhibit 5-1. First Part of an Example APEX Hourly Output File .....	109
Exhibit 5-2. First Part of an Example APEX Timestep Output File.....	111
Exhibit 5-3. First Part of an Example Daily Output File .....	114
Exhibit 5-4. First Part of an Example Profile Summary File.....	117
Exhibit 5-5. First Part of an Example Microenvironmental Results File .....	120
Exhibit 5-6. First Part of an Example Microenvironmental Summary File.....	121
Exhibit 5-7. Example of Exposure Table Type #1 in the Output Tables File .....	124
Exhibit 5-8. Example of Exposure Table Type #3 in the Output Tables File .....	125
Exhibit 5-9. Example of Exposure Table Type #6 in the Output Tables File .....	127
Exhibit 5-10. Example of Exposure Table Type #11 in the Output Tables File .....	128
Exhibit 5-11. Portion of ResponseProb Table .....	130
Exhibit 5-12. Portion of an Events File .....	133

## LIST OF FIGURES

Figure 3-1. Example of Study Areas, Air Quality Districts, Meteorological Zones, and Sectors	22
Figure 4-1. Relationship between Profile Functions and Microenvironmental Descriptions Files .....	78

# CHAPTER 1. INTRODUCTION

## 1.1 Overview of the APEX Model

The Air Pollutants Exposure model (APEX) is part of EPA's overall Total Risk Integrated Methodology (TRIM) model framework (EPA, 1999). TRIM is a time series modeling system with multimedia capabilities for assessing human health and ecological risks from hazardous and criteria air pollutants; it has been developed to support evaluations with a scientifically sound, flexible, and user-friendly methodology. The TRIM design includes three modules:

- Environmental Fate, Transport, and Ecological Exposure module (TRIM.FaTE);
- Human Inhalation-Dietary-Dermal Exposure module (TRIM.Expo); and
- Risk Characterization module (TRIM.Risk).

APEX is the inhalation exposure component of TRIM.Expo. The APEX model is a multipollutant, population-based, stochastic, microenvironmental model that can be used to estimate human exposure via inhalation for criteria and air toxics pollutants. APEX is designed to estimate human exposure to criteria and air toxic pollutants at the local, urban, and consolidated metropolitan level. The current release of the model is Version 5.2 (October, 2019). Human exposure to a contaminant is defined as “contact at a boundary between a human and the environment at a specific contaminant concentration for a specific interval of time” (National Research Council, 1991). For air pollutants, the contact boundaries are nasal and oral openings in the body. Dose is the amount actually received, or absorbed, in the body, leading to physiological effects. Pollutant exposures are estimated in a microenvironmental model by treating each individual's activities as a sequence of *events*, which are periods with known starting and ending times in particular microenvironments. A *microenvironment* is a defined space with relatively homogeneous air pollution concentration for a simulated individual. “Indoor kitchen,” “outdoor parking lot,” or “in vehicle” are examples of microenvironments. The pollutant concentrations in the air in each microenvironment are estimated from ambient air pollutant concentrations and parameters specific to each microenvironment and each pollutant. A person's inhalation exposures for a time interval are the pollutant concentrations in the microenvironment that person for that interval multiplied by the length of the interval.

The APEX model uses the personal profile approach to generate simulated individuals, for whom exposure time series are calculated. The *profile* is a description of the characteristics of an individual that may affect their activities, their locations, or the concentrations in the microenvironments that they encounter. Each profile is a random sample from the population (or a specified population group) within the study area. Each profile is generated independently of the others, and any number may be run, subject only to computer limitations. Typically, the profile includes demographic variables such as age, gender, and employment, as well as physiological variables such as height and weight, and finally some situational variables such as living in a house with a gas stove or air conditioning. The situational variables are used to help determine the microenvironmental concentrations, and the physiological variables are used in the determination of ventilation rate and dose. The demographic variables are used in the selection of *activity diaries* from EPA's Consolidated Human Activity Database (CHAD, McCurdy et al.,

2000; EPA, 2009) to represent the individual. Note: CHAD is a comprehensive database of human activity studies, which is provided with APEX. However, APEX may utilize other human activity data at the discretion of the user. Throughout this document “CHAD” will be used to denote the human activity diaries, although the reader should note that other data could be used.

APEX calculates the exposure and dose time series for a user-specified number of profiles for any number of pollutants. For modeling the pollutant CO, APEX contains an algorithm for estimating the blood dose (percent carboxyhemoglobin, %COHb). APEX also contains an algorithm for modeling particulate matter (PM) dose. In the case of PM, dose is defined as the rate of mass deposition in the respiratory system. For modeling any other pollutant, APEX calculates dose as Exposure\*Ventilation (see *Volume II* for details of the dose algorithms).

Collectively, the APEX profiles are intended to be a representative random sample of the population in a given study area. To this end, demographic data from the decennial census are used so that appropriate probabilities for any given geographical area can be derived. In APEX, the demographic geographical units are called *sectors*. Using the standard input files provided with the model, each sector is a census tract. Ambient air quality and meteorology data for the study area are also required by the model; the area represented by an air quality monitor (or air quality model grid cell) is called an *air district*, and the area covered by a meteorological monitor (or meteorological model grid cell) is referred to as a *zone*. APEX matches up each sector in the study area with the closest air district and zone to provide the data necessary to simulate exposure and dose for an individual.

For each simulated person (profile), the following general steps are performed:

- select the profile variables to characterize the person;
- construct the event sequence by selecting a sequence of appropriate activity diaries for the person (using demographic and meteorological variables);
- for each pollutant, calculate the concentrations in the microenvironments (using situational variables);
- for each pollutant, calculate the person’s exposure and dose for each event; and
- summarize the results for that profile.

The APEX model reports the results for each profile on various output files (some of which are pollutant-specific), described in detail later in this guide. Once all the profiles have been simulated, the model produces a set of summary tables for each pollutant that indicate the distributions of exposure and dose across all the profiles.

APEX can be thought of as a simulation of a field study that would involve selecting an actual sample of specific individuals who live in (or work and live in) a geographic area and then continuously monitoring their activities and subsequent inhalation exposures to a specific air pollutant during a specific period of time. The main differences between the model and an actual field study are that in the model, the:

- sample of individuals is a “virtual” sample, created by the model according to various demographic variables and census data of relative frequencies, in order to obtain a

- representative sample (to the extent possible) of the actual people in the study area;
- activity patterns of the sampled individuals (e.g., the specification of indoor and other microenvironments, the duration of time spent in each) are assumed by the model to be similar to individuals with similar demographic characteristics, according to activity data such as diaries compiled in EPA’s Consolidated Human Activities Database (CHAD) (EPA, 2009; EPA, 2002; McCurdy et al., 2000);
- pollutant exposure concentrations and doses are estimated by the model using temporally and spatially varying ambient outdoor concentrations, coupled with information on the behavior of the pollutant in various microenvironments; and
- various reductions in ambient air quality levels due to potential emission reductions can be simulated by adjusting air quality concentrations to reflect the scenarios under consideration.

Thus, the model accounts for the most significant factors contributing to inhalation exposure—the temporal and spatial distribution of people and pollutant concentrations throughout the study area and among the microenvironments—while also allowing the flexibility to adjust some of these factors for regulatory assessments and other reasons.

## 1.2 Nomenclature

The terms below are used throughout this guide.

- Diary: a set of events or activities (e.g., cooking, sleeping) for an individual in a given time frame (e.g., a day)
- Air quality district: the geographical area represented by a given set of ambient air quality data (either based on a fixed-site monitor or output from an air quality model)
- Event: an activity (e.g., cooking) with a known starting time, duration, microenvironment, and location (usually home or work)
- Microenvironment: a space in which human contact with an environmental pollutant takes place
- Profile: a set of characteristics that describe the person being simulated (e.g., age, gender, height, weight, employment status, whether an owner of a gas stove or air conditioner)
- Sector: the basic geographical unit for the demographic input to and output from APEX (usually census tracts)
- Study area: the geographical area modeled
- Study area population: total population of persons who live in the study area
- Meteorological zone: the geographical area represented by a given set of meteorological data (either based on a meteorological station or output from a meteorological model)

The labeling conventions below are used in this document.

- *Input and Output File Names* are in italics title case (in some cases, key terms are also introduced in italics, not capitalized, within a paragraph)
- ***Model Variables*** are in bold italics
- ***KEYWORDS***, which are used in the input files to identify variables and settings, are given in uppercase bold italic

- `Input and output file excerpts` are in a box surrounded by a single line
- Courier (fixed space) font is used for folder names, paths, and system commands outside of APEX

### 1.3 Strengths and Limitations of APEX

All models have strengths and limitations, and for each application it is important to carefully select the model that has the desired attributes. With this in mind, it is equally important to understand the strengths and weaknesses of the chosen model. The following sections provide a summary of the strengths and potential limitations of APEX.

#### 1.3.1 Strengths

APEX simulates the movement of individuals through time and space to estimate their exposure to individual or multiple pollutants in indoor, outdoor, and in-vehicle microenvironments. Compared to conducting a field study that would involve identifying, interviewing, and monitoring specific individuals in a study area, APEX provides a vastly less expensive, more expedient, and more flexible approach. The model also allows different air quality data, exposure scenarios, and other inputs and thus is very useful for decision making applications.

An important feature of APEX is its versatility. The model is designed with a great deal of flexibility so that different levels of detail in the input data can be applied for a variety of different applications. The input data sets supplied with APEX contain information for several microenvironments, covering the needs of most applications. The air quality data input to the model can be in the form of monitoring or modeling data. The data can be customized for specific locations—on roadways; within or geo-political units such as counties; census units such as tracts; the locations of air dispersion model receptors; or the grid cells of Eulerian model output. Criteria and hazardous air pollutants can be modeled by APEX.

A key strength of APEX is the way it incorporates stochastic processes representing the natural variability of personal profile characteristics, activity patterns, and microenvironment parameters. In this way, APEX is able to represent much of the variability in the exposure estimates resulting from the variability of the factors effecting human exposure.

Secondly, APEX has the ability to estimate exposures and doses on different timescales (e.g., 5-minute, hourly, daily, or annual) for all simulated individuals in the sample population from the study area. This ability allows for powerful statistical analysis of a number of exposure characteristics (e.g., acute and chronic exposure, correlations with activities and demographics), many of which are provided automatically by APEX in formatted output tables.

APEX also estimates the exposures of workers in the areas where they work, in addition to the areas where they live. The pollutant concentrations in these respective locations may be substantially different from each other.

The use of APEX has been facilitated by the availability of model-ready input files which have been developed from the databases discussed above: national population demographics and

commuting information from either the 2000 or 2010 U.S. Census; CHAD activity data; and microenvironment definitions.

### **1.3.2 Limitations**

The limitations detailed below have been identified with regard to APEX.

- The population activity pattern data supplied with APEX (CHAD activity data) are compiled from a number of studies in different areas and for different seasons and years. Therefore, the combined data set may not constitute a representative sample. Nevertheless, much of CHAD is from random-sample studies of national scope, which could be extracted by the user if desired to create a representative sample.
- The commuting data address only home-to-work travel; travel between sectors for other purposes is not modeled directly. APEX can model time spent in travel; however, based on the model settings, using the ambient air quality for one or more study area sectors, or by using a special roadway algorithm.
- APEX creates seasonal or year-long sequences of activities for a simulated individual by sampling human activity data from more than one subject in CHAD. Thus, uncertainty exists about season-long exposure event sequences. This approach can tend to underestimate the variability from person to person because each simulated person essentially becomes a composite or an “average” of several actual people in the underlying activity data (which tends to dampen the variability). At the same time, this approach overestimates the day-to-day variability for any individual if each simulated person is represented by a sequence of potentially dissimilar activities from different people rather than more similar activities from one person. These uncertainties have been reduced with the implementation in APEX of algorithms for combining diaries which address these limitations to some extent.
- The model currently does not capture certain correlations among human activities that can impact microenvironmental concentrations (e.g., cigarette smoking leading to an individual opening a window, which in turn, affects the amount of outdoor air penetrating the residence).
- Certain aspects of the personal profiles are held constant, though in reality they change every year (e.g., age). This is only an issue for simulations spanning several years.
- At this point in time, no interactions between pollutants are modeled.

Other data and model limitations exist besides those identified above, including physiological data and algorithms, meteorological data, and the data and algorithms associated with estimating concentrations in microenvironments. EPA continues to refine the model and data to reduce these limitations to the extent possible. The uncertainties which result from these limitations of APEX have been characterized for an ozone assessment (Langstaff, 2007).

## **1.4 Applicability**

APEX is an advanced air inhalation exposure model which can be used for a range of applications. APEX can be employed to model episodic "high-end" inhalation exposures that result from highly localized pollutant concentrations (e.g., residual risk assessments). APEX can also provide detailed probabilistic estimates of exposure for urban and greater metropolitan areas

(e.g., for regulatory analyses supporting national decisions such as NAAQS reviews). APEX is appropriate for assessing both long-term chronic and short-term acute inhalation exposures of the general population or of specific segments of the population. The model is designed to look at the range of inhalation exposures among different groups of people across a population, for a range of averaging times, in a single simulation. The current version of APEX produces results for flexible averaging times. By default, APEX produces results for 1 hour, 8 hours, 24 hours, and annual time periods (or the length of a simulation, if shorter than one year). However, APEX can optionally model results for timesteps on a much smaller scale (e.g., 5 minutes) by setting appropriate run parameters and providing air quality data on the appropriate time scale.

Although it is outside the original scope of APEX, the model has been successfully applied to a set of specific individuals by arranging to have exactly one person per sector and just one matching set of activity diaries that may be assigned to each person. However, many properties of these people remain randomly assigned as APEX is basically a stochastic model.

Due to the computational demands (run time and disk space) of running APEX, it is not appropriate for national-level assessments of population exposure. However, this is not an inherent limitation in the model code or algorithms.

## **1.5 Brief History of APEX**

APEX was originally derived from the probabilistic National Ambient Air Quality Standards Exposure Model (pNEM). The NEM series was developed to estimate exposure to the criteria pollutants (e.g., CO, ozone). In 1979, EPA began to develop NEM by assembling a database of human activity patterns that could be used to estimate exposures to outdoor pollutants (Roddin et al., 1979). The data were then combined with measured outdoor concentrations in NEM to estimate exposures to CO (Biller et al., 1981; Johnson and Paul, 1983). In 1988, OAQPS began to incorporate probabilistic elements into the NEM methodology using activity pattern data based on various human activity diary studies in an early version of probabilistic NEM for ozone (pNEM/O<sub>3</sub>). In 1991, a probabilistic version of NEM was developed for CO (pNEM/CO) that included a one-compartment mass-balance model to estimate CO concentrations in indoor microenvironments (Johnson et al., 1992). A newer version of pNEM/O<sub>3</sub> was developed in the 1990s and applied to nine urban areas for the general population, outdoor children, and outdoor workers (Johnson et al., 1996a, b, c). During 1999-2001, an updated version of pNEM/CO (versions 2) was developed that relied on activity diary data from CHAD and enhanced algorithms for simulating gas stove usage, estimating alveolar ventilation rate (a measure of human respiration), and modeling home-to-work commuting patterns.

APEX evolved from pNEM to provide greater applicability, flexibility, and accuracy. The APEX model was substantially different than pNEM, particularly in the use of a personal profile approach rather than a cohort simulation approach. APEX introduced a number of new features including automatic site selection from large (e.g., national) databases; a series of new output tables providing summary statistics; and a thoroughly reorganized method of describing microenvironments and their parameters. Most of the spatial and temporal constraints were removed or relaxed in APEX. Several major improvements to APEX have been introduced in the most recent version, APEX5. Specifically, APEX5 includes the improvements listed below.



- Multipollutant capability
- Algorithms for the assembly of multi-day (longitudinal) activity diaries that model intra-individual variance, inter-individual variance, and day-to-day autocorrelation in diary properties.
- Methods for adjusting diary-based energy expenditures for fatigue and excess post-exercise oxygen consumption
- New equations for estimation of ventilation
- The ability to model commuters leaving the study area
- The ability to model air quality and exposure on different time scales
- The ability to model person-to-person variability in air quality within an air district
- New output files containing diary event-level, timestep level, and hourly-level exposure, dose, and ventilation data, and hourly-level microenvironmental data
- The ability to model the prevalence of disease states such as asthma
- New output exposure tables that report exposure statistics for subpopulations such as children and active people under different ventilation levels.
- The ability to model inhaled dose for pollutants
- The inclusion of commuting data from the 2000 or 2010 Census
- Expanded options for modeling microenvironments
- Expanded location options for sampling ambient air quality
- The option of running Sobol sensitivity analysis

Model enhancements and other changes are occasionally made to APEX, and thus users are encouraged to revisit the APEX website <https://www.epa.gov/fera/human-exposure-modeling-air-pollutants-exposure-model> for notices of these changes. Comments and suggestions for improvements to the model or the input data provided with the model should be addressed to: John Langstaff, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711 (email: [langstaff.john@epa.gov](mailto:langstaff.john@epa.gov)).

## 1.6 Scope and Organization of This Guide

The documentation of the APEX model is currently divided into two volumes. *Volume I: User's Guide* [this document] is designed to be a hands-on guide to the model. It focuses on how to run the APEX computer model, develop the appropriate input files, and understand the model output files.

*Volume II: Technical Support Document* describes the scientific basis of the APEX model and provides scientific background for the model algorithms. It covers the methods implemented in APEX for calculating microenvironmental concentrations, modeling ventilation, estimating dose, and assembling activity diaries.

The remainder of *Volume I* have been organized into the chapters listed below.

- Chapter 2, Setting Up and Running APEX: Provides instructions for setting up APEX and running single or multiple APEX simulations
- Chapter 3, Characterizing the Study Area: Described the procedure for characterizing the area to be modeled, in terms of APEX input files

- Chapter 4, Input Files: Provides a description of the format, data, and options for each of the APEX input files
- Chapter 5, Output Files: Provides a description of the format and data associated with each of the APEX output files

*Volume II* has the chapters listed below:

- Chapter 1, Introduction
- Chapter 2, Overview of Model Design and Algorithms
- Chapter 3, Using Probability Distributions in APEX
- Chapter 4, Characterizing the Study Area
- Chapter 5, Generating Simulated Individuals (Profiles)
- Chapter 6, Constructing a Sequence of Diary Events
- Chapter 7, Estimating Energy Expenditures and Ventilation
- Chapter 8, Calculating Pollutant Concentrations in Microenvironments
- Chapter 9, Calculating Exposures
- Chapter 10, Calculating Dose
- Chapter 11, Sobol Sensitivity Analysis

## CHAPTER 2. SETTING UP AND RUNNING APEX

APEX is written in Fortran using only standard routines and conventions to allow portability to different operating systems and compilers. APEX has been tested on Windows 10, 7, Vista, XP, 2000, NT, and 98 operating systems, as well as Linux, using Intel Fortran. Other compilers may produce warnings and/or errors, and may require some small code changes for compatibility. APEX currently uses ASCII coding internally.

In addition to providing flexibility in modeling options, the APEX code is specifically designed for fast execution time and reasonable memory requirements.

We recommend running APEX on a computer with at least:

- 2 GB of RAM;
- 600 MHz processor; and
- 1000 MB of available hard drive space.

The input files supplied with APEX will require 500 MB of hard drive space, and the additional input files created by the user may take up another 1–10 MB of space, or more, depending mainly on the size of the air quality input files.

APEX run time on a PC with a 2.8 GHz Core i7 8 CPU and 8 GB of RAM, running Windows 7, is about 1.8 hours for a one-year single-pollutant simulation of 20,000 individuals in a large metropolitan area. The combined size of the output files from this simulation is small (~50 MB), unless detailed daily and hourly data are requested, in which case the output files can take up more than 5,000 MB.

This chapter, which describes the steps involved in setting up and running an APEX simulation, is organized as shown below.

- Section 2.1, Downloading and Setting Up APEX
- Section 2.2, Setting up an APEX Simulation
- Section 2.3, Overview of APEX Input and Output Files
- Section 2.4, Overview of Model Settings and Options

### 2.1 Downloading and Setting Up APEX

To install APEX, download the APEX installer from <http://www.epa.gov/fera>. This will install an APEX folder containing all of the necessary files to run the model, as well as supporting documentation and supplementary input files. It is suggested that users install to the directory “C:\APEX” (this is the default directory specified in the installer) as it will make running the example APEX run easier. The file “unins000.exe” can be used to uninstall APEX from a user’s computer. This will remove the installed files but not files created by the user. See *An Introduction to APEX* (EPA, 2017) for more details about installing APEX.

Note that if APEX has previously been installed on a user’s computer, users can overwrite the previous installation of APEX to that same directory or choose to install APEX to a different

location on their computer. If users choose to overwrite the previous installation of APEX, the installer will warn them that “The Folder: [directory] already exists. Would you like to install to that folder anyway?” Selecting “Yes” will re-download APEX to the location it was previously downloaded, and all of the files that exist in the APEX directories with names that match the file names of the incoming installation (i.e., all the files that have their original names) will be overwritten. If, however, users have generated new files with different names in the existing folders or additional folders in the “APEX” directory, these files and folders will not be overwritten with the re-installation of APEX.

To run APEX, you should open a DOS window in the directory your APEX run batch file is, and then type the batch file name. (If the DOS link puts you in the wrong directory, right click on it, select properties, and put the directory you want in the "Start in" box.) APEX should run at this point and its progress will be indicated as it runs. If it fails for some reason, an error message will appear. You should also look in the log file to check that the run was successful.

A batch file is a text file with the extension .bat. For a single APEX run a batch file (e.g., "runAPEX.bat") will contain one line, naming the APEX executable followed by the *Control Options* file (COF), for example:

```
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF.txt
```

Then if you type ‘runAPEX’ in the DOS window, the APEX simulation governed by the COF is run. You can run multiple APEX simulations using one batch file, for example:

```
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF1.txt  
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF2.txt  
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF3.txt  
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF4.txt  
C:\APEX\EXEs\APEX.exe C:\APEX\Input\COF5.txt
```

Then if you type ‘runAPEX’ in the DOS window, the 5 APEX simulations governed by the COFs are run, one after the other. After the initialization of the run, APEX will begin progressing through the simulated profiles. When the model run ends successfully, APEX will stop with the message “Finished APEX model run.”

Note that each of the *Control Options* files used should contain unique names for the model output files to avoid overwriting the output from the previous run. As the model run starts and then progresses, normal status messages will be printed to the screen, in addition to any error or warning messages that may arise from incomplete or incorrect model settings.

Even if an APEX simulation runs to completion, the user should examine the APEX *Log* output file to confirm that the model behaved as expected. The *Log* file contains information on the model settings, input parameter values, and input and output file names. The file also contains a great deal of detailed information about the model run including, but not limited to, summaries of: 1) the modeled profiles; 2) the final study area (including the final sectors, air quality districts, and meteorology zones); and 3) the simulated microenvironments. The *Log* file, which is discussed in Section 5.1, will also contain a listing of any warning or error messages that

resulted from the run. Some of the common error messages are explained in the companion document *An Introduction to APEX* (EPA, 2017).

## 2.2 Setting Up an APEX Simulation

This section describes the steps involved in performing an APEX simulation.

### 1. Select Model Options

After identifying the scope of the analysis, the user must decide which options to select. To determine the appropriate options for the application, the user must answer questions such as those listed below.

- How many pollutants do I want to model in a single run?
- Do I want to model worker commuting?
- How many profiles do I want to model?
- How many microenvironments do I want to model?
- How should I define my microenvironments?
- How should the activity diaries be constructed (i.e., randomly select a new diary every day for each simulated individual, or construct longitudinal diaries based on diary properties?)
- Which other model settings should I select?
- What types of outputs do I want from the model?
- What time resolution do I want to use?

These options and others are described in greater detail in Section 2.4.

### 2. Prepare Input Files

After deciding which model options to use and how to configure them, the next step in configuring an APEX simulation is to set up the input data files with the necessary data. One of these files, the *Control Options* file, is used to specify input and output file names and locations and the simulation settings. The remaining files contain the input data necessary to run APEX. The data contained in these remaining files varies depending on the configuration of the scenario to be modeled and the number of pollutants used. The input files are described in CHAPTER 4.

### 3. Configure the Simulation Settings

The final step in preparing an APEX simulation is to create the *Control Options* input file for the desired simulation settings. This file contains four sections:

- input file names and locations;
- output file names and locations;
- pollutant parameters (including output table specifications); and
- job parameters.

A detailed description of the data for each of the sections of the *Control Options* file is provided in CHAPTER 4.

## 4. Running APEX

To perform an APEX simulation, the user can run the model as described above in Section 2.1.

### 2.3 Overview of Input and Output Files

This section provides a brief overview of the input and output files associated with APEX. For more detailed descriptions of the input and output files, refer to CHAPTER 4 and CHAPTER 5, respectively. All of the input and output files used by APEX are ASCII text files; they can be read and/or modified by the user using a text editor or other software. Note, however, that certain files, such as the commuting flow input file and the hourly and events output files, may be very large (in some cases several gigabytes), and difficult for some text editors to handle.

#### 2.3.1 Input Files

There are several types of APEX input data files. Most of these files are required; however, the *Diary Statistics*, *Diary Occupation*, *Profile Factor* and *Prevalence* files are optional in some cases. With the exceptions of the *Population Data* and the *Air Quality Data* files, only one file of each type is required for a simulation. The input file paths and names are designated in the *Control Options* file using a “keyword” approach. APEX processes the file and searches for particular keywords followed by an equal sign and one or more values for the keyword. Table 2-1 lists each file type and the keyword that must be used to identify it. CHAPTER 4 provides a detailed description of the input files and their syntax.

#### 2.3.2 Output Files

APEX utilizes a total of 16 possible output file types. These files contain such information as: 1) a summary of the properties of the simulated persons; 2) hourly or event-level exposures, doses, and breathing data for the simulated profiles; 3) hourly or daily values of microenvironmental parameters and pollutant concentrations; 4) dose and exposure summary tables for the modeled population; and 5) exposure statistics for the modeled microenvironments. The creation of some of the output files is dependent on settings in the *Control Options* file, which also contains the path and file name for each output file.

Table 2-1 lists each of the input data files and their corresponding keywords, while Table 2-2 lists the output data files. If an output file is specified with the same name and location as an existing file, the old files are overwritten. Therefore, if the user wishes to conduct a series of model runs, the output files for each run should be named differently or written to a different directory, or the output from a prior run should be moved elsewhere before the next model run is submitted.

The APEX output files are covered in detail in CHAPTER 5.

**Table 2-1. APEX Input Files**

<b>Input Files</b>	<b>Pollutant Specific?<sup>a</sup></b>	<b>Control Options File Keyword</b>
<i>Control Options</i>		-
<i>Commuting Time</i>		<b>COMMTIME</b>
<i>Commuting Flow</i>		<b>COMMUT</b>
<i>ME Mapping File for Clustering Diary Assembly</i>		<b>DIARYCLUS</b>
<i>Diary Events (DiaryEvents)</i>		<b>DIARYEVE</b>
<i>Diary Occupation (DiaryOcc)</i>		<b>DIARYOCC</b>
<i>Diary Questionnaire (DiaryQuest)</i>		<b>DIARYQUEST</b>
<i>Diary Statistics (DiaryStat)</i>		<b>DIARYSTA</b>
<i>Diary Raw Transitions</i>		<b>DIARYTRANS</b>
<i>MET Distribution</i>		<b>DISTRIB</b>
<i>Air District Location</i>		<b>DISTRICT</b>
<i>Employment Probability</i>		<b>EMPLOY</b>
<i>Profile Functions</i>		<b>FUNCTIONS</b>
<i>Microenvironment Mapping (MEMap)</i>		<b>MEMAP</b>
<i>Meteorology Data</i>		<b>METEOR</b>
<i>MET Mapping</i>		<b>METMAP</b>
<i>Microenvironment Descriptions</i>		<b>MICROENV</b>
<i>Physiological Parameters</i>		<b>PHYSIOL</b>
<i>Population Data</i>		<b>POP</b>
<i>Prevalence</i>		<b>PREVAL</b>
<i>Profile Factors</i>		<b>PROFILE</b>
<i>Air Quality Data</i>	YES	<b>QUALITY</b>
<i>Population Sector Location</i>		<b>SECTOR</b>
<i>Seed offsets and Sobol grouping</i>		<b>SEED</b>
<i>Ventilation</i>		<b>VENTIL</b>
<i>Meteorology Zone Location</i>		<b>ZONE</b>

<sup>a</sup> if yes, then a separate file is required for each pollutant modeled.

**Table 2-2. APEX Output Files**

<b>Output Files</b>	<b>Pollutant Specific?<sup>a</sup></b>	<b>Control Options File Keyword</b>
<i>Daily</i>		<b><i>DAILY</i></b>
<i>Events</i>		<b><i>EVENT</i></b>
<i>Hourly</i>		<b><i>HOUR</i></b>
<i>Log</i>		<b><i>LOG</i></b>
<i>Microenvironmental Results</i>	YES	<b><i>MICRORES</i></b>
<i>Microenvironmental Summary</i>	YES	<b><i>MICROSUM</i></b>
<i>Profile Summary (Persons)</i>		<b><i>PERSON</i></b>
<i>Sites</i>		<b><i>SITE</i></b>
<i>Sobol</i>		<b><i>SOBOL</i></b>
<i>Output Tables</i>	YES	<b><i>TABLE</i></b>
<i>Timestep</i>		<b><i>TIMESTEP</i></b>
<i>MultiPollutant</i>	YES <sup>b</sup>	<b><i>MULTIPOL</i></b>
<i>Diary Clustering</i>		<b><i>CLUSTER</i></b>
<i>Diary Clustering Weights</i>		<b><i>CWEIGHT</i></b>
<i>Diary Clustered Transition</i>		<b><i>TRANSITIONS</i></b>

<sup>a</sup> if yes, then a separate file is written for each pollutant modeled.

<sup>b</sup> the output is specific to the first two pollutants in a multipollutant run.

## 2.4 Overview of Model Settings and Options

This section briefly describes the primary settings and options available in APEX. These are specified by the user in the *Control Options* file or other input files. There are six general categories of settings and options in APEX:

- general model settings and options;
- study area location;
- pollutants;
- profiles;
- microenvironments; and
- outputs.

Table 2-3 describes the settings and options in each of these categories, how they are selected or changed, and the impact of changing a setting or option on the other input files. See CHAPTER 4 for additional details of input files and their content, how to edit or create them, and how they interact with other files.

“YES” or “NO” settings within the *Control Options* file are not case-sensitive and may be abbreviated to a single letter; thus, “Y” or “y” means “YES,” and “N” or “n” means “NO.” However, writing these out in full may provide more clarity.



**Table 2-3. APEX Settings and Options**

<i>Setting/Option</i>	<i>How Option is Selected</i>	<i>Impact</i>
<b>GENERAL MODEL SETTINGS AND OPTIONS</b>		
Simulation start/end dates	Specified in YYYYMMDD format (e.g., 19960704 is July 4, 1996) using the <i>START_DATE</i> and <i>END_DATE</i> keywords in the <i>Control Options</i> file. The user must define the appropriate start and end dates for an application.	The indicated start and end dates must be included in the date ranges specified on the <i>Air District Location</i> , <i>Meteorology Zone Location</i> , <i>Meteorology Data</i> , and <i>Air Quality Data</i> files. These files may contain additional data before and/or after the start and end dates, but must contain data for the entire period between the specified start and end dates.
Adjust for Daylight Saving Time (DST)	Specified using the <i>DSTADJUST</i> keyword in the <i>Control Options</i> file. If this parameter is set to “YES,” then the <i>Air Quality Data</i> file will be adjusted for Daylight Saving Time (DST) in the summer; if it is set to “NO,” no adjustment is made. This keyword should be set to “YES” if the <i>Air Quality Data</i> file is based on Standard Time yet the activity data are based on DST. The default is “YES.”	Changing this setting means that the <i>Air Quality Data</i> file is based on DST (which is not usual) or that the activity data are based on Standard Time (CHAD data are based on current local time, which is typically DST in summer). Regardless of this setting, the output (hourly exposure and dose) for all simulated days will contain exactly 24 hours, and all input activity diaries must contain exactly 24 hours.
Model worker commuting	Specified using the <i>COMMUTING</i> keyword in the <i>Control Options</i> file. If this keyword is set to “YES,” commuting to work is allowed and the user must provide a <i>Commuting Flow</i> file in the appropriate format and employment data must be specified in the <i>Employment Probabilities</i> file; if it is set to “NO,” all workers are assumed to work at home and the user is not required to provide the <i>Commuting Flow</i> or <i>Commuting Time</i> files. The <i>Commuting Flow</i> file accompanying APEX contains commuting flows between all census tracts from the 2010 Census. These commuting data are sufficient for all applications within the United States in which the sectors are defined as census tracts. The <i>Commuting Time</i> file specifies the duration of commutes for each census tract. The default is “NO.”	If commuting is modeled, the <i>Diary Questionnaire</i> file must have an additional column that lists the total commuting time on each diary. Also, if the user chooses to define sectors as something other census tracts, new <i>Commuting Flow</i> and <i>Commuting Time</i> files (in addition to a number of other input files) must be created corresponding to the new sectors.
Air quality rollback adjustment (for estimating exposure in hypothetical scenarios)	Specified using the <i>ROLLBACK</i> keyword in the <i>Control Options</i> file. If this keyword is set to “YES,” the user must specify appropriate values for the <i>RBTARGET</i> , <i>RBACKGND</i> , and <i>RBMAX</i> keywords in the <i>Control Options</i> file; if it is set to “NO,” values are not required for these keywords (and any present will be ignored).	If the <i>ROLLBACK</i> keyword is changed to “YES” in the <i>Control Options</i> file accompanying APEX, the <i>RBTARGET</i> , <i>RBACKGND</i> , and <i>RBMAX</i> keywords must be set to appropriate values.
Time resolution (length of APEX timestep)	Specified using the <i>TIMESTEPSPERDAY</i> keyword in the <i>Control Options</i> file, which must be an integer. The timestep can be either smaller or larger than an hour. However if the timestep is larger than an hour, it must be an integer multiple of an hour. If it is smaller than an hour, there must be an integer number of timesteps in an hour. The default APEX timestep is one hour. If <i>TIMESTEPSPERDAY</i> is not set, then APEX uses the default, which is 24.	The timestep dictates the required time resolution of the following air quality input. The time resolution of the <i>Air Quality Data</i> file must match that indicated by <i>TIMESTEPSPERDAY</i> .
Random number seed	Specified using the <i>RANDOMSEED</i> keyword in the <i>Control Options</i> file. If <i>RANDOMSEED</i> = 0, the program uses the clock to determine the initial seed. Otherwise, the seed may range from 1 to 2147483646.	If all input data and settings are unchanged in two model runs, along with <i>RANDOMSEED</i> , then the output will be identical.

<b>Setting/Option</b>	<b>How Option is Selected</b>	<b>Impact</b>
Run Sobol analysis	Specified using the <b>SOBOLRUN</b> keyword in the <i>Control Options</i> file. If YES, then a Sobol sensitivity analysis is performed (which takes considerably longer). If “NO” (the default), a regular model run is performed.	If <b>SOBOLRUN</b> = YES, the SobolVarList keyword is needed. A Sobol run should have fewer profiles than a regular run, because the program performs multiple passes through the profiles and creates additional arrays.
<b>STUDY AREA LOCATION</b>		
Center of study area	Specified as the latitude and longitude of the center of the study area in decimal degrees using the <b>LATITUDE</b> and <b>LONGITUDE</b> keywords in the <i>Control Options</i> file. The user must always define the appropriate study area center for an application. Latitude and longitude are in decimal degrees, without a compass direction. American locations (west of the prime meridian) have negative longitude and positive latitude.	If the study area is changed, the user must ensure that the following files contain appropriate data for the new location: <i>Population Sector Location</i> file (unless the included file is used), <i>Air District Location</i> file, <i>Meteorology Zone Location</i> file, <i>Meteorology Data</i> file, and <i>Air Quality Data</i> file.
Radius of study area	Specified as the distance (in km) from the center to the edge of the study area using the <b>CITYRADIUS</b> keyword in the <i>Control Options</i> file. The user must always define the appropriate study area radius for an application.	If the study area is changed, the user must ensure that the following files contain appropriate data for the new location: <i>Population Sector Location</i> file (unless the included file is used), <i>Air District Location</i> file, <i>Meteorology Zone Location</i> file, <i>Meteorology Data</i> file, and <i>Air Quality Data</i> file.
Restrict study area to selected counties	Specified using the <b>COUNTYLIST</b> keyword in the <i>Control Options</i> file. The default is “NO.” If set to “YES,” the “county” is identified by the first 5 characters of the sector ID, which for the default files is the state and county FIPS code. Use the <b>COUNTY</b> keyword once for each county to be included. This option operates in conjunction with the study area center and radius.	None, normally. However, if the user does not use the included <i>Population Sector Location</i> file, they must ensure that the new <i>Population Sector Location</i> file provides suitable “county IDs,” which consist of the first 5 characters of the sector IDs.
Restrict study area to selected census tracts	Specified using the <b>TRACTLIST</b> keyword in the <i>Control Options</i> file. The default is “NO.” If set to “YES,” the user must list the sector (tract) ID for the tracts to which the study area will be restricted using the <b>Tract</b> keyword in the <i>Control Options</i> file. The sector IDs for all census tracts in the 2010 Census are included in the <i>Population Sector Location</i> file accompanying APEX.	This is similar to <b>COUNTYLIST</b> , except that “tracts” are identified using the first 11 characters of the sector name (the standard for U.S. census tracts). If the user is not using actual census tracts, the <b>TRACTLIST</b> option will still select all sectors from the <i>Population Sector Location</i> file that match the first 11 characters (and are within <b>CITYRADIUS</b> of the study area center).
Locations of sectors (filename)	Specified as sector IDs and locations (latitude and longitude) in the <i>Population Sector Location</i> file. The <i>Population Sector Location</i> file accompanying APEX use the census tracts from the 2010 Census as sectors. This file also specifies the county associated with each sector (via the first five characters of the sector ID, which are the county FIPS codes in the supplied data), which allows the user to restrict the study area to selected counties. The default files provide sectors with the necessary population and commuting data for the entire United States.	Sectors in <i>Population Sector Location</i> file must match the sectors in <i>Population Data</i> files. The default file contains all U.S. census tracts from the 2010 Census. If commuting is modeled, the sectors on the two commuting files must match those on the <i>Population Sector Location</i> file and the <i>Population Data</i> files; if other sector definitions are used, the user must provide compatible <i>Commuting</i> files.
Locations of air districts (filename)	Specified in the <i>Air District Location</i> file. The user must always define the appropriate air districts for an application.	The locations of the air districts must be selected such that they can provide reasonable estimates of air quality for the sectors and study period included in the analysis. Data for each AQ monitor for each pollutant in the simulation must be provided in the <i>Air Quality Data</i> files (one for each pollutant).

<b>Setting/Option</b>	<b>How Option is Selected</b>	<b>Impact</b>
Radius of air district	Using either a single the <i>AirRadius</i> keyword in the <i>Control Options</i> file or multiple ones via the <i>Air District Location</i> file, the user can specify the maximum distance (in km) that a sector can be from the nearest air district to remain in the study. If a sector has no district centers within <i>AIRRADIUS</i> of its center, that sector is removed from the study area and is not modeled. Users must always define an appropriate value for this radius based on their application.	The radius of the air districts must be selected such that they will include the sectors the user would like to include in the analysis.
Type of air quality data	The keyword <i>MODELAQVAR</i> specifies the type of air quality data to be used in the simulation. The air quality may be entered as values for each timestep in the simulation (the default, <i>MODELAQVAR</i> = N) or as distributions for each hour ( <i>MODELAQVAR</i> = Y).	The value of <i>MODELAQVAR</i> dictates the expected format of the <i>Air Quality Data</i> file. See Section 4.5 for details.
Using roadway concentrations	The user has the option to model roadway concentrations if they set <i>ROADWAY</i> = Y. The <i>ROADWAY</i> parameter can be set for each pollutant, so that in a multipollutant run, it is possible to selectively use roadway concentrations for each pollutant. If <i>ROADWAY</i> = NO, then all MEs in the microenvironmental mapping file that are mapped to R will instead be mapped to O. If the user specifies using roadway locations based on the last home/work location ( <i>ROADLAST</i> = YES), then roadway locations will be chosen from either the home or work location, whichever most recently occurred.	Additional concentrations must be specified in the <i>Air Quality Data</i> and <i>Air District Location</i> files.
Location of meteorological data stations (filename)	Specified as zone IDs and locations (latitude and longitude) in the <i>Meteorology Zone Location</i> file. The user must always define the locations of meteorological stations for an application.	Data for each meteorological data station specified in the <i>Meteorology Zone Location</i> file must be provided in the <i>Meteorology Data</i> file.
Radius of meteorological station coverage	Using the <i>ZONERADIUS</i> keyword in the <i>Control Options</i> file, the user can specify the maximum allowed distance (in km) from a sector to the nearest meteorological station. If all zone centers are further than <i>ZONERADIUS</i> from the sector center, the sector is removed from the study area and is not modeled.	The radius of the zones must be selected such that they will include the sectors the user would like to include in the analysis.
<b>POLLUTANTS</b>		
Number of pollutants	The number of different pollutants to be modeled must be specified using the <i>#Pollutants</i> keyword	
Pollutant Names	The user must specify each pollutant with the keyword <i>POLLUTANT</i> . The pollutant name may contain only alphanumeric characters and the underscore (“ ”) character, as it is used to generate filenames.	Must be followed in the <i>Control Options</i> file by the pollutant-specific parameters and output table levels. Must match the pollutant names in the specification of the air quality data input files.
Model dose for pollutant	Specified using the <i>DODOSE</i> keyword in the <i>Control Options</i> file. Pollutant-specific. If this keyword is set to “YES,” APEX will calculate dose for the pollutant; if it is set to “NO,” the dose calculations will be suppressed.	If <i>DODOSE</i> is set to “YES” and CO is being modeled, the user must specify the correct values for the <i>ALTITUDE</i> and <i>COHBFACT</i> keywords in the <i>Control Options</i> file.
<b>PROFILES</b>		
Number of profiles	Set to a positive integer using the <i>#PROFILES</i> keyword in the <i>Control Options</i> file. Users must determine an appropriate value based on the application.	None.

<b>Setting/Option</b>	<b>How Option is Selected</b>	<b>Impact</b>
Modeled populations (filenames)	Specified in the <i>Control Options</i> file following the specification of the file names. A population file is required for each combination of gender and race to be modeled. All gender/race combinations without specified population files are assumed to have zero populations and not considered further. Users can select from the sets of available <i>Population Data</i> files accompanying APEX, or generate their own.	These files must have the same sectors (in the same order) as the <i>Population Sector Locations</i> file.
Profile function options (filename)	Specified in the <i>Profile Functions</i> file. The user must develop data relevant to a particular application prior to performing an APEX simulation. The <i>Profile Functions</i> file is required.	None.
Employment status (filename)	Specified in the <i>Employment Probability</i> file for implementation of commuting and occupation. The file accompanying APEX should cover all applications where sectors are defined as census tracts. The <i>Employment Probability</i> file is required.	None.
Occupation status	Occupational status can be modeled by using the <i>Profile Factor</i> file and specifying that it is an employment-related parameter. The user can match by occupational status and it can also affect work-related MET values.	
Minimum and maximum ages for simulated profiles	Specified using the <i>AGEMIN</i> and <i>AGEMAX</i> keywords in the <i>Control Options</i> file.	None.
Modeled age groups	Specified in the <i>Population Data</i> files. The files that accompany APEX define the age groups as single years up to 99, and are sufficient for all applications where sectors are defined as census tracts.	None.
Size of age window	The <i>AGECUTPCT</i> and <i>AGE2PROBAB</i> keywords in the <i>Control Options</i> file are used to specify the window around the assigned age of a profile from which activity data can be selected.	None.
Probabilities for selecting diaries with missing characteristics	Using the <i>MISSGENDER</i> , <i>MISSEMPL</i> , <i>MISSAGE</i> , and <i>MISSOCC</i> keywords in the <i>Control Options</i> file, the user can specify the probability that activity diary data with missing gender, employment status, age, or occupation will be selected.	None.
Type of diary assembly	Determined by the <i>LONGITDIARY</i> and <i>CLUSTERDIARY</i> keywords. If <i>LONGITDIARY</i> = YES, longitudinal dairy assembly will be performed based on the statistic in the <i>Diary Statistics</i> file. If <i>CLUSTERDIARY</i> = YES, the clustering algorithm will be performed on the entire CHAD database.	If <i>LONGITDIARY</i> is YES, then the <i>Diary Statistics</i> file must be designated in the <i>Control Options</i> file, and the <i>DIARYD</i> and <i>DIARYAUTO</i> keywords must be set. If <i>CLUSTERDIARY</i> = YES then three output file switches <i>CLUSTEROUT</i> , <i>TRANSOUT</i> , and <i>CWEIGHTOUT</i> are active, and two clustering parameters <i>CLUSTERAGES</i> and <i>USEADJACENT</i> may be set.
Physiological parameters for the simulated population (filename)	Specified in the <i>Physiological Parameters</i> file. The default values in this file are suitable for most APEX applications. This file is required.	None.

<b>Setting/Option</b>	<b>How Option is Selected</b>	<b>Impact</b>
Activity-specific energy expenditures for the simulated population (filenames)	Specified in the <i>MET Mapping</i> and <i>MET Distribution</i> input files. The default values in these files are suitable for most APEX applications. These files are required.	None.
Method of selecting height and weight	Determined by the <b>HTWTMETHOD</b> keyword. If omitted (or set to 2), the method implemented in 2017 is used, in which an explicit correlation is maintained (so that for a given age and gender, the taller people are heavier, based on empirical data from NHANES). <b>HTWTMETHOD</b> = 1 uses the older method, which was the only option previously.	
Method of determining resting metabolic rate (RMR)	Determined by the <b>RMRMETHOD</b> keyword. If omitted (or set to 2), the new (2017) method is used. Set <b>RMRMETHOD</b> = 1 to use the older method.	
Method of determining breathing ventilation rate VE	Determined by the <b>VEMETHOD</b> keyword. If omitted (or set to 2), the method implemented in 2017 is used, in which VE is a function of both VO <sub>2</sub> and VO <sub>2</sub> max. Set <b>VEMETHOD</b> = 1 to use the older method.	
Apply a maximum to VE?	Determined by the <b>VEMAX</b> keyword. If omitted (or set to YES), then VE is limited to 150 L/min for events up to 5 minutes, or to 100 L/min for longer events. This was automatically applied prior to the implementation of the new VE method. The maximum is not as relevant when using the new method because it does not generate an excess of high VE values.	
Modeling of disease prevalence (option and filename)	Determined by the <b>DISEASE</b> keyword. If <b>DISEASE</b> is given a value (a string of maximum length 12 characters containing the condition name, spaces allowed) in the input file, APEX will then use data in the <i>ePrevalence</i> file to assign a YES/NO value to the physiological profile variable <b>III</b> , and produce output tables for the subpopulation of modeled persons with <b>III</b> = YES. If disease is not defined, this file is not needed.	If <b>DISEASE</b> is given a value (a string of maximum length 12 characters containing the condition name, spaces allowed) in the input file, then APEX requires that a <i>Prevalence</i> file be designated as well.
<b>MICROENVIRONMENTS</b>		
Maximum number of microenvironments	Set to an integer using the <b>#MICRO</b> keyword in the <i>Control Options</i> file; must not exceed 127.	Number of APEX microenvironments in the <i>Microenvironment Mapping</i> and <i>Microenvironment Descriptions</i> files must not exceed the specified value in the <i>Control Options</i> file.
Microenvironment definitions (filename)	Specified in the <i>Microenvironment Descriptions</i> file. The user must develop data relevant to a particular application prior to performing an APEX simulation. This file is required.	Each location referenced in the activity database (e.g., CHAD) must be mapped to one of the microenvironments specified in the <i>Microenvironment Descriptions</i> file using the <i>Microenvironment Mapping</i> file. The user may choose to define custom microenvironmental parameter definitions that depend on conditional variables. If so, these variables must be defined on the <i>Profile Functions</i> file.
<b>OUTPUTS</b>		
Produce hourly outputs	Specified using the <b>HOURLYOUT</b> keyword in the <i>Control Options</i> file. If this keyword is set to "YES," the hourly output file named in the Output File section is created; if it is set to "NO," the file is not created (even if it is named). The variables to be written are listed using <i>HOURLYLIST</i> .	None.

<b>Setting/Option</b>	<b>How Option is Selected</b>	<b>Impact</b>
Produce daily outputs	Specified using the <b>DAILYOUT</b> keyword in the <i>Control Options</i> file. If this keyword is set to “YES,” the daily output file is created; if it is set to “NO,” the file is not created. The variables to be written are listed using <b>DAILYLIST</b> .	None.
Produce microenvironmental output	Specified using the <b>MSUMOUT</b> and <b>MRESOUT</b> keywords in the <i>Control Options</i> file. If these keywords are set to “YES,” the <i>Microenvironmental Summary</i> and/or <i>Microenvironmental Results</i> output files are created; if they set to “NO,” these files are not created.	None.
Produce event output	Specified using the <b>EVENTSOUT</b> keyword in the <i>Control Options</i> file. If this keyword is set to “YES,” the events output file is created; if it is set to “NO,” this file is not created. This refers to activity diary events, which are of variable duration between one minute and the timestep size.	None.
Produce timestep output	Specified using the <b>TIMESTEPOUT</b> keyword in the <i>Control Options</i> file. If this keyword is set to “YES” and the timestep is not one hour, the timestep output file is created; if it is set to “NO,” this file is not created. If the timestep is one hour, then this file would be identical to the Hourly output file, so it is not written. The optional variables to be written are specified using <b>TIMESTEPLIST</b> . The “timestep” is the spacing in time of the air quality data.	None
Produce multipollutant table output	Specified using the <b>MULTIPOL</b> keyword on the <i>Control Options</i> file. If this keyword is set to “YES” and at least two pollutants are modeled, then the user must also specify level cut-points for each pollutant using the keywords <b>Multi1</b> and <b>Multi2</b> . The cut-points are written on one line in a comma-separated list. If there are N cut-points, then there will be (N+1) bins, because the lowest below means “below the first cut-point”, and the last means “equal to or above the last cut-point”. Values exactly on a cut-point go into the higher bin.	

## CHAPTER 3. CHARACTERIZING THE STUDY AREA

An initial study area in an APEX analysis consists of a set of basic geographic units called *sectors*, typically defined as census tracts (see Section 1.2). The user provides the geographic center (latitude/longitude) and radius of the study area. APEX contains a database with the central locations of every census tract in the U.S. Each tract is assigned a central point as its location, as determined by the U.S. Census Bureau. APEX calculates the distances to the center of the study area of all the sectors included in the sector location database, and then selects the sectors within the radius of the study area. One can also provide a list of counties or census tracts as part of the specification of the initial study area. APEX then maps the user-provided *air quality district* and *meteorological zone* data to the selected sectors. The sectors identified as having acceptable air and meteorological data within the radius of the study area are selected to comprise a final study area for the APEX simulation analysis. This final study area determines the population make-up of the simulated persons (profiles) to be modeled.

All geographical data in APEX use latitude and longitude to locate places. The convention is that both are specified in decimal degrees (do not use degrees, minutes, and seconds), and longitudes in the western hemisphere are negative, as are latitudes in the southern hemisphere. Longitude “wraps around” with -180 degrees equal to +180 degrees. The following sections describe in more detail how a final study area is determined in an APEX simulation analysis.

### 3.1 APEX Spatial Units

#### 3.1.1 Initial Study Area

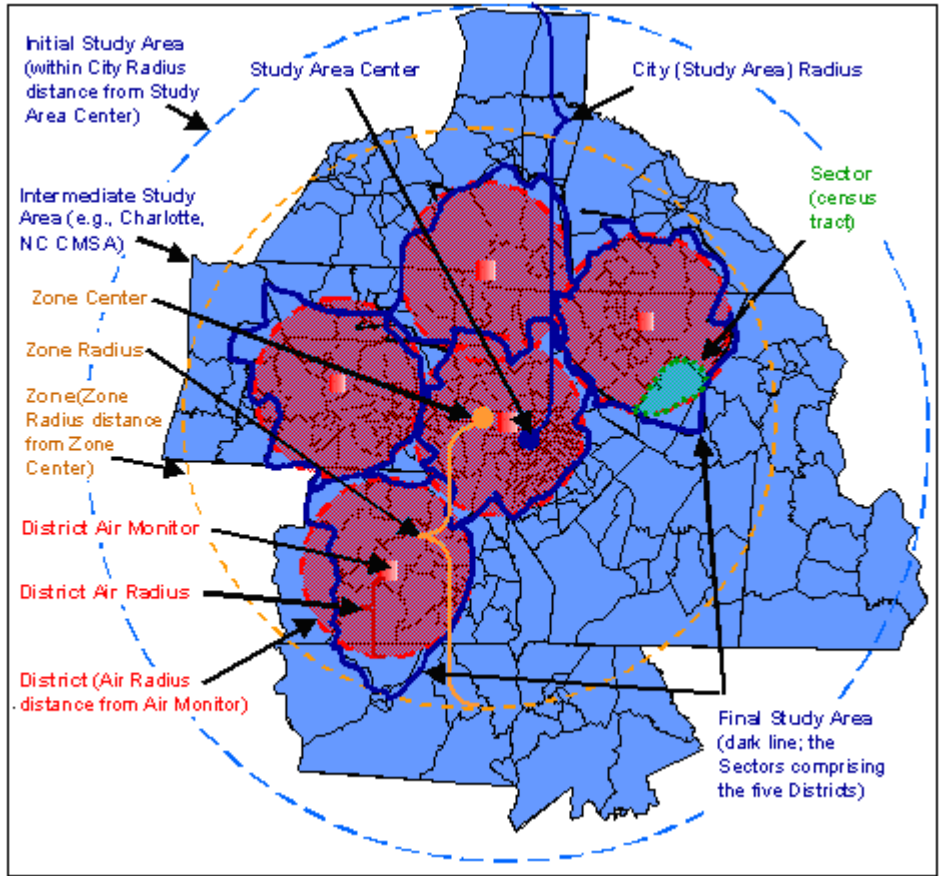
The APEX study area has typically been the neighborhood around an emission source, a city, or a larger metropolitan area. Larger study areas are possible to simulate, depending on computing capabilities, available data, and the desired precision of the run.

The user defines an initial study area by specifying the latitude and longitude of a central point (referred to here as the *study area central location*), together with a radius. The user also has the option of providing a list of counties or census tracts to be modeled. If present, this list further restricts the area to be modeled to the counties or tracts to be modeled which are within the specified study area radius. The final study area is a function of the availability of the user-supplied demographic data, pollutant concentration data, and the meteorological data within the initial study area, as determined respectively by population sectors, air quality districts, and meteorological zones. Figure 3-1 and the subsections below provide additional details about these geographical units.

#### 3.1.2 Sectors

The demographic data used by the model to create personal profiles is provided at the sector level. For each sector the user must provide demographic information allowing the determination of age, gender, race, and work status. This is most commonly done by equating sectors with census tracts and providing input files with counts at the tract level for each age, gender, and race combination. The current release of APEX includes input files that already contain this demographic and location data for all census tracts in the 50 states and Washington,

D.C., based on the 2010 Census. Data files based on the 2000 Census are also available. One of the APEX input files, named *Population Sector Location* file in this guide, lists the sector ID and location for all sectors that have associated population data. The supplied *Population Sector Location* file has been prepared listing all the census tracts in the 2010 U.S. Census. Corresponding *Population Data* files have been supplied as well. This allows the user to model any desired study area in the country without having to make any changes to these input files. The 2000 Census files have also been used with APEX. However, one cannot mix 2000 and 2010 files because the lists of tracts are not the same.



**Figure 3-1. Example of Study Areas, Air Quality Districts, Meteorological Zones, and Sectors**

If available, finer scales such as census block groups could be used instead. Also, data could be aggregated to larger regions such as counties if fewer sectors were desired. Regardless of the specific meaning for sectors in APEX, the shape of sectors is irrelevant in the sense that the model only uses the central location each sector, determined by the latitude and longitude for some representative point. Sector names in APEX are 40 characters or fewer.

In the *Control Options* file for an APEX run, the user specifies the area to be modeled by specifying the latitude and longitude of a central location for the study area, along with a radius (the *CITYRADIUS* parameter). Optionally, the user may also provide a list of counties or



sectors (typically census tracts, but can be census blocks or other geographic units) to be modeled. If present, this list further restricts the study area.

For each model run, APEX selects the sectors that meet the study area conditions in the following way. First, the sector location must be within the specified distance (radius) of the designated center of the study area. Second, if the user provided a specific list of counties identified by their FIPS codes (using *COUNTYLIST* and *COUNTY* in the *Control Options* file), then the sector must belong to one of these counties. This is achieved using the first five characters of the sector ID, which should contain the FIPS code for the county in which the sector is located. The user may also provide a list of census tracts (sectors) using the *TRACTLIST* and *TRACT* variables, which match the first 11 characters of the sector ID. If census tract names are not being used as the sector names, the *COUNTYLIST* (or *TRACTLIST*) option will still select those sectors matching the first 5 (or 11) characters with any names on the designated list.

If no county or tract list is provided, the initial study area is roughly circular, consisting of all sectors with sector locations within the specified radius. Sectors are never divided; each one is either entirely inside or entirely outside the study area.

One way to exert greater control over the selection of sectors is to edit the census input files to eliminate any undesired sectors. However, the *Population Sector Location* file and the various *Population Data* files must all have the same set of sectors in them, so consistent editing is necessary. The *Population Sector Location* and *Population Data* files provided with the current release of APEX contains data for every census tract in the 50 states and Washington, D.C. from the 2010 Census.

### 3.1.3 Air Quality Districts

The spatial units for ambient air quality data are called districts. Ambient air quality data are provided as time series at specific locations. Previous versions of APEX required hourly time series, but now the user may specify other timesteps. All the air quality data in a single APEX run must use the same timestep. The locations could be monitoring sites, political units such as counties, census units such as tracts, or receptor locations or grid points as used by some air quality models. As with sectors, each air quality district has a nominal central location indicated by its latitude and longitude. The air quality district locations are stored in the *Air Quality District Location* input file. The user designates the maximum representative radius of the air quality districts as a modeling parameter called *AIRRADIUS*. A single *AIRRADIUS* specified on the *Control Options* file applies to all air districts. Otherwise, the user can specify a different *AIRRADIUS* for each district on the *Air Quality District Location* file. The same set of air quality districts must be used for each pollutant. Users can also indicate air data to be used specifically for roadway concentrations in the *Air Quality District Location* file by including “Road” as part of the district name. These data will be used for microenvironments set to “R (Road)” or “RW (Road Work)” in the *Microenvironment Mapping* file instead of using the ambient AQ data for the person’s Home or Work district.

The model checks each air quality district listed in the *Air Quality Data* input files (one for each pollutant) to determine if the district has data covering the entire simulation period, as indicated

by the start and end dates. Districts without complete data are dropped. Each air quality district may have a different period of operation (i.e., different start and/or end dates). When the air data are read, there can be no gaps (missing data) between the simulation start and end dates, or else APEX stops, reporting error #4 in the ReadAirQuality module. If the user is supplying monitoring data as inputs, for example, then any missing values within the simulation period must be filled in before running APEX.

APEX calculates the distance from each sector location to each air quality district center, and then assigns the sector to the nearest air quality district, provided it is within the maximum representative air quality radius. If there is no air quality district within range (that is, all air quality district centers are further than *AIRRADIUS* from the sector center), the sector is deleted from the study area and is not modeled. It is possible, and perhaps even likely, that some air quality districts in the *Air Quality Data* input files will have no sectors assigned to them. Such air quality districts are not used. This feature allows the user to prepare an input file in the simplest manner, perhaps containing more air quality districts than are necessary. For example, one might prepare a single *Air Quality Data* input file for a pollutant for all air quality monitors in the state of Texas. This same input file could then be run on a study area around Houston, Dallas, or some other location in Texas, without having to alter the input file.

By default, APEX will assign each person within a sector the corresponding appropriate ambient values from the sector's matching air district. Thus, for each timestep in the simulation, all persons within the same district will have the same outdoor air quality value. However, APEX can optionally model person-to-person variation in air quality within an air district. In this case, an optional form of the *Air Quality Data* file is provided, which lists air quality distributions for each hour for each air district. Each person in the district is assigned a randomly-sampled value from the appropriate hourly distribution. Note that this option can only be used when the APEX timestep is equal to 1 hour (the APEX default).

### 3.1.4 Modeling Commuting

APEX models commuting by assigning a work sector to each employed individual based on commuting data for that individual's home sector. Two commuting data files are required. These are files consisting of: 1) commuting flow data (the *Commuting Flow* file), and 2) commuting time data (the *Commuting Time* file). Nationwide files are supplied with APEX.

The nationwide input files use census tracts as the sectors. The *Population Data* files and the two commuting files must refer to the same census (for example, the 2010 Census), as the list of tracts must match in all three files. APEX extracts the flows for the selected home sectors from the *Commuting Flow* file and derives profile level commuting times from the *Commuting Time* file.

In APEX5, a new option has been added which allows the user to use the national *Commuting Flow* file with finer sectors, such as census blocks or block groups. The reason is that this is the only input file that provides a sector-to-sector mapping. An exhaustive listing of sector pairs may become unmanageable.

To use the national *Commuting Flow* file with finer sectors, it is necessary that each sector name be the same length, and that the first 11 characters of the name indicate the proper census tract. If APEX detects that the commuting flow place names are shorter than the sector names, then it automatically performs this mapping. The same destination probabilities are applied to each sector within a given commuting tract. Once a destination tract is chosen, each sector within that tract (and also within the study area) is equally likely to be selected as the work sector.

This capability has not been extended to other input files that list sectors. All the other such files are linear in the number of sectors, and even with very fine sectors, these files are still of reasonable length.

### 3.1.5 Meteorological Zones

Another spatial unit in APEX is the *meteorological zone*, which is the equivalent to air quality districts but for meteorological data. Most of the rules that apply to air quality districts also apply to meteorological zones. Each meteorological zone is associated with a central location (specified by latitude and longitude), a maximum representative radius given by **ZONERADIUS**, a **START\_DATE**, and an **END\_DATE**. The start and end dates may differ for each meteorological zone and must encompass the entire simulation period, otherwise the meteorological zone is deleted. If there are missing data between the start and end dates of the simulation, APEX stops, reporting error #4 in the ReadMeteorologyData module.

APEX calculates the distance from each sector location to each meteorological zone center and assigns each sector to the nearest meteorological zone if within range (**ZONERADIUS**); otherwise the sector is deleted.

### 3.1.6 The Final Study Area

The final study area consists of all the sectors within **CITYRADIUS** of the study area central location, restricted to the listed counties or tracts (if provided), that have both an air quality district and a meteorological zone within specified ranges. If both tracts and counties are listed, then the resulting study area is the union of the two lists. Sectors for which a valid air quality district (that is, within a distance **AIRRADIUS**) or a valid meteorological zone (within **ZONERADIUS**) cannot be found are discarded from the final study area. The study area population is the total population in the input *Population Data* files that reside in these sectors.

## CHAPTER 4. APEX INPUT FILES

This chapter provides the details necessary for creating and modifying the APEX input files. The first section describes the general format and properties that pertain to all of the APEX input files, while the remaining sections cover each input file in detail.

### 4.1 Input File Formats

The APEX input files and their descriptions are given in Table 4-1. Some input files are not required if certain features of the model are turned off. For example, the *Diary Statistics* file is not needed if the D&A longitudinal diary assembly is not being used, the *Prevalence* file is not needed if disease is not defined, and the *Commuting Flow* and *Commuting Time* files are not needed if commuting is not considered. All input files are ASCII text files that can be edited using a text editor. Each input line of these files is categorized into one of four types, as noted below.

1. Keyword (or variable or parameter) line: Keywords are used in the *Control Options* file to indicate to APEX where the input files are located and what values should be assigned to certain variables. A keyword line always contains an “=” sign. The part of the line to the left of “=” is called the “keyword” and the part to the right is called the “value.” The keyword must start with a letter and must match the spelling sought by the program code, after which the keyword may contain additional letters, blanks, or commas. Keywords are not case sensitive. APEX uses the keyword to identify and set the input values. The values may be character, logical, or numeric values, or file names.
2. Numeric line: Any line beginning with a digit (0–9) is recognized as a numerical data line by APEX. Non-digits may appear later in a numeric line.
3. Character line: A line that begins with a character but does not contain an “=” sign is recognized as a character data line.
4. Comment line: Any blank lines and any lines beginning with “!” generally are regarded as comment lines by APEX and are used by the user to help document the file. However, comment lines should not be inserted in the middle of a block of data. That is, if the computer code is expecting to read a long series of numbers without a break, then comments may break the flow.

The keywords and input values are not case sensitive, except as noted. Additionally, each line on an input file is processed independently by APEX. Continuation of data values across multiple lines is not permitted unless specifically noted for a particular file. APEX uses “list” (or “free”) format for all input values. This means that the values or data do not have to be fixed in specific positions on an input line. Multiple items on an input line can generally be separated by either a blank or comma, although certain items (as noted below) insist on one or the other. Words on numeric and character input lines should not contain internal blanks, as these will be interpreted as delimiters between input fields. This does not apply to keyword lines, as those lines have only two fields (separated by the “=” sign), so either or both sides may contain internal blanks. Keywords can have additional characters after the required characters given in the table (e.g., “commut” can be “commuting”).

**Table 4-1. APEX Input File Descriptions**

<b>Input File</b>	<b>Keyword</b>	<b>Description</b>
<i>Control Options</i>	-	(Required) Specifies the overall settings (or parameters) for an APEX simulation, including input and output file names, job settings and switches, and output table levels.
<i>Commuting Time</i>	<b>COMMTIME</b>	(Optional) Provides a distribution of recorded commuting times for workers in each home census tract. This file is only required when work commuting is modeled ( <b>COMMUTING</b> = Y).
<i>Commuting Flow</i>	<b>COMMUT</b>	(Optional) Provides probabilities of a worker commuting to various destination census tracts from any given home tract. This file is only required when commuting is modeled ( <b>COMMUTING</b> = Y).
<i>Diary Cluster Mapping</i>	<b>DIARYCLUS</b>	(Optional) Provides the mapping from activity location codes in the <i>Diary Events</i> file (e.g., from CHAD) to user-defined microenvironments for the purposes of calculating transition probabilities used in the clustering method of diary assembly (needed only when <b>CLUSTERDIARY</b> =Y).
<i>Diary Events</i>	<b>DIARYEVE</b>	(Required) Provides the 24 hour event descriptions (i.e., start time, duration, activity, and location) for all the diary days in the original activity database. This file contains the same list of diary IDs as the <i>Diary Questionnaire</i> file, in the same order, but it has multiple records (one per event) for each ID.
<i>Diary Occupations</i>	<b>DIARYOCC</b>	(Optional) Contains an occupation for all CHAD activity diaries. If provided, these occupations will overwrite the default CHAD occupations listed on each diary.
<i>Diary Questionnaire</i>	<b>DIARYQUEST</b>	(Required) Provides personal and other information (e.g., day type, gender, age, race, occupation) relating to each 24 hour activity record from the original CHAD activity database. If commuting is used, this file must also contain commuting times.
<i>Diary Statistics</i>	<b>DIARYSTA</b>	(Optional) Contains the value of the key statistic for all CHAD activity diaries, used in the D&A method of longitudinal diary assembly. Statistics files are included with APEX for outdoor time and time spent in vehicles. Users could construct other statistics files from CHAD. (Required if <b>LONGITDIARY</b> = Y).
<i>Diary Transitions</i>	<b>DIARYTRANS</b>	(Optional) Contains a list of diary IDs that belong to the distinct persons. Multiple diaries from the same person are used to calculate empirical transition probabilities when the <b>CLUSTERDIARY</b> = YES option is active.
<i>MET Distribution</i>	<b>DISTRIB</b>	(Required) Provides distribution types and parameters for calculating the metabolic equivalent of task (MET) value for each distribution number in the <i>MET Mapping</i> file. A MET value is a dimensionless ratio of the activity-dependent energy expenditure rate to the basal or resting energy expenditure rate.
<i>Air District Location</i>	<b>DISTRICT</b>	(Required) Provides the site IDs and locations (degrees latitude and longitude) of air quality monitoring or modeling locations. The file is used along with the user-defined <b>AIRRADIUS</b> to define the geographical area covered by the air quality data. Sectors within the study area that are close enough to an air district that has data between the simulation start and end dates are retained for modeling.
<i>Employment Probability</i>	<b>EMPLOY</b>	(Required) Contains employment probabilities by age group, gender, and study sector. The default file is based on the tracts from the 2010 U.S. Census. For other definitions of sectors, the user would have to supply a new employment file.
<i>Profile Functions</i>	<b>FUNCTIONS</b>	(Required) Contains user-defined functions for several model variables, which in turn can be used by the model for a variety of purposes, including calculating microenvironmental concentrations.

<b>Input File</b>	<b>Keyword</b>	<b>Description</b>
<i>Microenvironment Mapping</i>	<b>MEMAP</b>	(Required) Provides the mapping from activity location codes in the <i>Diary Events</i> file (e.g., from CHAD) to user-defined microenvironments in the <i>Microenvironment Descriptions</i> file.
<i>Meteorology Data</i>	<b>METEOR</b>	(Required) May contain temperature, wind, humidity, and precipitation data for the meteorological stations and dates indicated in the <i>Meteorology Zone Location</i> file. These data can be used to determine window positions, group activity pattern pools, and microenvironmental parameters in APEX.
<i>MET Mapping</i>	<b>METMAP</b>	(Required) Maps each activity codes present in the <i>Diary Events</i> file to an MET distribution. (A MET is a dimensionless ratio of the activity-dependent energy expenditure rate to the basal or resting metabolic rate). The MET distributions are defined by number in the <i>MET Distribution File</i> .
<i>Microenvironment Descriptions</i>	<b>MICROENV</b>	(Required) Contains the definitions of the microenvironments and the microenvironment parameters used to determine the exposure concentrations in microenvironments. The data on this file will change greatly from one pollutant to another, so no default version of this file is provided.
<i>Physiological Parameters</i>	<b>PHYSIOL</b>	(Required) Contains tables of age- and gender-specific physiological parameters.
<i>Population Data</i>	<b>POP</b>	(Required) Contains information on the population (by age group) in each study sector. Each race/gender combination being modeled has its own file.
<i>Prevalence</i>	<b>PREVAL</b>	(Optional) Contains prevalence rates (probabilities) for disease (or any other condition) for different age/gender cohorts. This file is optional only if the <i>Control Options</i> file variable <b>DISEASE</b> is not defined.
<i>Profile Factors</i>	<b>PROFILE</b>	(Optional) Contains details of a user-specified profile factor that can vary by age group, gender, and study sector, and can be used to specify microenvironmental parameters.
<i>Air Quality Data</i>	<b>QUALITY</b>	(Required) Provides the air quality data for each air monitoring/modeling location listed in the <i>Air District Location</i> file. The time resolution on this file depends on the <i>Control Options</i> file setting <b>TIMESTEPSPERDAY</b> . There is one <i>Air Quality Data</i> file per pollutant. Optionally, the file may include distributions for hourly air quality values, see Section 4.5 for details.
<i>Population Sector Location</i>	<b>SECTOR</b>	(Required) Provides the IDs and locations (in degrees latitude and longitude) of sectors (e.g., census tracts). The file is used along with the user-defined <b>CITYRADIUS</b> and other data to select the sectors within the modeled area.
<i>Seed offsets and Sobol grouping</i>	<b>SEED</b>	(Required) Contains variable-specific offsets for random number seeds. This file also contains the Sobol grouping numbers for each random variable, but these are used only if <b>SOBOLRUN</b> = Yes. If not a Sobol run, these numbers must still be present for file formatting reasons, but their values do not matter, and may all be zero (for example).
<i>Ventilation</i>	<b>VENTIL</b>	(Required) This file contains regression parameters used to estimate total ventilation (VE) from MET. The format of the ventilation file is different for <b>VEMETHOD</b> = 2 versus <b>VEMETHOD</b> = 1.
<i>Meteorology Zone Location</i>	<b>ZONE</b>	(Required) Provides the site IDs and locations (degrees latitude and longitude) of the meteorological stations. The file is used along with the user-defined <b>ZONERADIUS</b> to determine the area covered by the meteorological data. Start and end dates indicate the dates during which the data for a particular location are valid.

The following sections discuss the details of APEX input files and provide several examples. Note that these example files in this and the next chapter are provided for illustration purposes only. These are provided for the purpose of highlighting various aspects and options of APEX. Most of these examples are only portions of the necessary input files. Thus, these example files will not work as an actual set of input files. Users are encouraged to view the example input files (which can be downloaded separately) for a complete set of input files.

## 4.2 Control Options File

The *Control Options* file is APEX's master file. The *Control Options* file names all the other input and output files, sets model parameters, and controls formats of output files. APEX only processes keyword lines in this file. Any other types of input lines are ignored. However, the very first line of this file (even if it is a comment beginning with !) is saved to be used as part of the header that is written to each output file for audit trail purposes. Therefore, it is helpful for this line to include information that describes or identifies the simulation.

When creating the *Control Options* file, the rules listed below should be used.

- The very first line of the file should identify the specific simulation (up to 224 characters in length) (it does not need to start with !)
- Keywords (or parameter or variable names) are placed to the left of the equal sign in a keyword line, and are not case sensitive
- Parameter values are to the right of the equal sign
- Yes/no parameters are not case sensitive and may be abbreviated to "Y" or "N"
- Lines may appear in any order after the first line, with the exceptions listed below
  - Lines using the **COUNTY** keyword, which must immediately follow the line with the **COUNTYLIST** keyword
  - Lines using the **Tract** keyword, which must immediately follow the line with the **TRACTLIST** keyword
  - Lines using the **Block** keyword, which must immediately follow the line with the **BLOCKLIST** keyword
  - Lines using pollutant-specific parameters or table levels, which must immediately follow the line with their corresponding **POLLUTANT** keyword
- Lines for specific keywords may be omitted if defaults are allowed and are acceptable
- Only one equal sign is allowed per keyword line
- Anything after an exclamation mark in a line is treated as a comment and ignored
- Any unexpected line without an equal sign treated as a comment and is ignored
- The information on the *Control Options* file is not case-sensitive, except that quoted strings are left as they are. On Windows systems, file names are not case sensitive.

It is useful to keep a copy of the *Control Options* file associated with each simulation to provide a record of the input and output files and model settings associated with the simulation and to make it easier to run the model again, either with or without modifications.

We describe the *Control Options* file in terms of four sections of the file: input files, output files, pollutant parameters (including output table levels), and job parameters. Organizing the *Control Options* file in this manner is not required. The details of each section are discussed below.

#### 4.2.1 Input and Output File List Sections of the Control Options File

In the Input Files section of the *Control Options* file (Exhibit 4-1), the user needs to specify the names and path names of all of the input files. The details on the content and format of these input files are provided in the subsequent sections of this chapter.

The keywords for these files were given in Table 2-1. The keyword may be extended, as long as the listed keyword is contained within the text. For example, “Employ” could become “Employment,” but “MEMap” cannot be “MicroenvironmentMap.” The keyword **FILE** must appear (with a blank space before it) right after each of the file keywords and before the “=”. Full paths including the drive letter are required. If any of these files are not found at the specified locations, then APEX will print an error listing the file that is missing.

The *Air Quality Data* files are the only input files that are pollutant-specific with one file for each pollutant modeled. Each **AIR QUALITY FILE** keyword must be followed by a comma and the name of its corresponding pollutant (i.e., the pollutant names must match the names given by the **POLLUTANT** keyword in the *Control Options* file; see Section 4.2.2). Exhibit 4-1 provides an example of designating *Air Quality Data* files for a two-pollutant scenario (CO and ozone).

The example in Exhibit 4-1 has 10 population data files. The number of population files could change, depending on how the user classifies the population. For example, the user could provide two population files, one for all females and one for all males.

For the population input files, the keywords **POP** and **FILE** must appear at the beginning of the keyword part of the keyword input line in the *Control Options* file, followed by a comma and **GENDER** and another comma and **RACE**. **GENDER** currently must be either male or female and it can be shortened to M or F. If the population files provided with APEX are to be used, the **RACE** must be White, Black, Asian, NatAm, or Other. If the user provides different population files, **RACE** can be customized, however, the first 5 characters of each race must be unique. It is necessary for **RACE** to match the designation in the header of the population files, or an error will result. Further information on population files is given in Section 4.8. It is not necessary to specify all genders and race combinations for APEX to run. However, the model assumes that any missing gender/race combinations have zero population. A warning message is returned if one gender for a race is present but the other is missing.

In the Output File section of the *Control Options* file, illustrated in Exhibit 4-2, the user needs to specify the keywords (see Table 2-2), names, and paths for the output files. If the user turns off the hourly file creation, event file creation, or microenvironmental summary file creation, the corresponding output files will not be generated, and file names do not need to be specified. The *Microenvironmental Summary*, *Microenvironmental Results*, and *Tables* files are pollutant-specific; one of each of these files will be created for each pollutant. However, only one filename for each type has to be defined in the *Control Options* file – output filenames for each pollutant are constructed by appending the pollutant name (as defined using the *Control Options*



file **POLLUTANT** keyword) to the end of the filename base. Further details on output files are given in CHAPTER 5.

! INPUT FILES		
Zones file	=	C:\APEX\Input\METsites.txt
Air Quality file, Ozone	=	C:\APEX\Input\AirQuality_ozone.txt
Air Quality file, CO	=	C:\APEX\Input\AirQuality_co.txt
Districts file	=	C:\APEX\Input\AQdistricts.txt
Meteorology file	=	C:\APEX\Input\METdata_h.txt
Functions file	=	C:\APEX\Input\ProfileFunctions.txt
Microenvironment file	=	C:\APEX\Input\MicroDescriptions.txt
MEMap file	=	C:\APEX\Input\ME_Mapping.txt
DiaryEvent file	=	C:\APEX\Input\CHADEvents.txt
DiaryQuest file	=	C:\APEX\Input\CHADQuest.txt
METMap file	=	C:\APEX\Input\CHADMap.txt
MET Distribution file	=	C:\APEX\Input\MetsDists.txt
DiaryStat file	=	C:\APEX\Input\CHADSTATSoutdoor.txt
Physiology file	=	C:\APEX\Input\Physiology.txt
Ventilation file	=	C:\APEX\Input\Ventilation.txt
Diarycluster file	=	C:\APEX\Input\Cluster_ME_Mapping.txt
Prevalence file	=	C:\APEX\Input\Asthma.txt
Seed offset file	=	C:\APEX\Input\Groupings.txt
!		
! POPULATION INPUT FILES		
Pop file, Female, Asian	=	C:\APEX\Input\pop_fa.txt
Pop file, Female, Black	=	C:\APEX\Input\pop_fb.txt
Pop file, Female, Natam	=	C:\APEX\Input\pop_fn.txt
Pop file, Female, Other	=	C:\APEX\Input\pop_fo.txt
Pop file, Female, White	=	C:\APEX\Input\pop_fw.txt
Pop file, Male, Asian	=	C:\APEX\Input\pop_ma.txt
Pop file, Male, Black	=	C:\APEX\Input\pop_mb.txt
Pop file, Male, NatAm	=	C:\APEX\Input\pop_mn.txt
Pop file, Male, Other	=	C:\APEX\Input\pop_mo.txt
Pop file, Male, White	=	C:\APEX\Input\pop_mw.txt
Sectors file	=	C:\APEX\Input\pop_geo.txt
Employment file	=	C:\APEX\Input\Employment.txt
Commuting file	=	C:\APEX\Input\Commuting2000.txt
CommTime file	=	C:\APEX\Input\CommutingTimes2000.txt

**Exhibit 4-1. Input Files Section of an Example Control Options File**

! OUTPUT FILES		
log file	=	C:\APEX\Output\log.txt
hourly file	=	C:\APEX\Output\hours.txt
daily file	=	C:\APEX\Output\days.txt
events file	=	C:\APEX\Output\events.txt
persons file	=	C:\APEX\Output\psum.txt
microsum file	=	C:\APEX\Output\msum.txt
microres file	=	C:\APEX\Output\mres.txt
tables file	=	C:\APEX\Output\tables.txt
site file	=	C:\APEX\Output\sites.txt
Sobol file	=	C:\APEX\Output\Sobol.txt

**Exhibit 4-2. Output Files Section of an Example Control Options File**

If the user is intending to run APEX multiple times, then the output files from each run should be given different names; otherwise, the later runs will overwrite the earlier ones. It may be convenient to create a folder for each run. However, the inputs often refer to the same file in multiple runs. This avoids having a separate copy of a large file (like CHAD, or the commuting databases) for each run.

#### 4.2.2 Pollutant Parameters Section of the Control Options File

Several *Control Options* file keywords described in Table 4-2 are pollutant-specific.

**#POLLUTANTS** determines the number of different pollutants being modeled, and must precede all of the pollutant-specific keywords. The **POLLUTANT** keyword is used to: 1) assign a name to each pollutant being modeled, and 2) designate the start of a keyword list for another pollutant. The pollutant name may be up to 40 alphanumeric characters long and may contain underscores “\_”. When modeling particulate matter (PM), each discrete size of PM must be modeled as a separate pollutant. All PM pollutants must start with the characters “PM.” The **SIZE** and **DENSITY** parameters must be defined for these pollutants.

**Table 4-2. Pollutant-specific Job Parameters**

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>#POLLUTANTS</b>	Integer	(Optional) The number of pollutants in the simulation. Any number of pollutants may be modeled - the maximum is limited only by the available system memory. Default = 1
<b>POLLUTANT</b>	Char(40)	(Required) Pollutant name. There must be one <b>POLLUTANT</b> statement for each pollutant modeled, which must be immediately followed by the other pollutant-specific job parameters and table levels. If the pollutant is a PM size category, then the pollutant name must start with the characters “PM.”
<b>INPUTUNITS</b>	Char(40)	(Optional) Pollutant concentration units used for the input data for the pollutant (ppm, ppb, or $\mu\text{g}/\text{m}^3$ ). The last of these is indicated by “UGM3”. Default is ppm (parts per million).
<b>OUTPUTUNITS</b>	Char(40)	(Optional) Pollutant concentration units used for the output data for the pollutant (ppm, ppb, or $\mu\text{g}/\text{m}^3$ ). The last of these is indicated by “UGM3”. Default is ppm (parts per million).
<b>#SOURCES</b>	Integer	(Optional) Largest number of sources in any one microenvironment for the pollutant. Any number of sources may be modeled - the maximum is limited only by the available system memory. Default = 0.
<b>PPMFACTOR</b>	Real	(Optional) Units conversion factor, the number of $\mu\text{g}/\text{m}^3$ that equal 1 ppm. For example, when modeling CO, set PPMFactor = 1145 (because 1 ppm = 1145 $\mu\text{g}/\text{m}^3$ for CO at typical indoor temperature and pressure). PPMFactor would be the same for a given chemical regardless of the choice of <b>INPUTUNITS</b> or <b>OUTPUTUNITS</b> , although in principle it would be reduced when modeling a high-altitude study area. See discussion below.
<b>DODOSE</b>	Char(1)	(Optional) Y = perform dose calculations, N = don't perform dose calculations. Like all Y/N flags, the words YES and NO may be spelled out. <b>DODOSE</b> = N may save some job execution time if doses are not of interest. Default = No.

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>SIZE</b>	Real	(Only used for modeling particulates) Aerodynamic diameter (particle size) in micrometers for a particle pollutant. This parameter is not required for gaseous pollutants.
<b>DENSITY</b>	Real	(Only used for modeling particulates) Density (in g/cm <sup>3</sup> ) of a particle pollutant. This parameter is not required for gaseous pollutants.
<b>ALERTTHRESH</b>	Real	(Optional, but default is zero) Timestep exposure threshold for alerting the user that a simulated person has high exposure. The number of exceedances and the time over the threshold are reported, for each person that goes over the limit. If this is not set then every person will exceed the threshold.
<b>ROADWAY</b>	Char(1)	(Optional) If “Y”, then APEX will use roadway concentrations while in roadway microenvironments. Default = No.
<b>ROADLAST</b>	Char(1)	(Optional) If “Y”, then APEX will use roadway concentrations based on the last event occurring in either the Home or Work location. Default = No.
<b>HOURLYFEVE1</b>	Char(1)	(Optional) If “Y”, then the constant intra-individual variability term used by the % $\Delta$ FEV1 model will be sampled hourly, otherwise it will be sampled daily. Default = No. See Section 7.4 in <i>Volume II</i> for details.
<b>HOURLYFEVE2</b>	Char(1)	(Optional) If “Y”, then the ozone-dependent intra-individual variability term used by the % $\Delta$ FEV1 model will be sampled hourly, otherwise it will be sampled daily. Default = No.

In addition to the parameters listed above, the output table level specifications discussed below are pollutant specific.

The role of **PPMFACTOR** was restricted starting with APEX version 4.69, so earlier *Control Options* files might need to be altered to run correctly with the newer code. It now always represents the conversion factor for gaseous pollutants between molar fraction and mass density at a typical indoor temperature and pressure. It is the number of micrograms per cubic meter equivalent to 1 part per million volume, and is independent of the choice of **INPUTUNITS** or **OUTPUTUNITS**. It is required in two cases: 1) if one (but not the other) of **INPUTUNITS** and **OutputUnits** are “UGM3”, or 2) if some microenvironments contain emission source (**ESUM**) terms and **INPUTUNITS** are not “UGM3”. In other cases it is not required and not used. For particulate pollutants, both **INPUTUNITS** and **OUTPUTUNITS** must be “UGM3” as molar volume is not well defined, and **PPMFACTOR** is not relevant.

In the Pollutant Parameters section, the user specifies the levels of each of the parameters used in the creation of the output summary tables for each pollutant. These specification parameters include: **PERCENTILES**, **DAVGEXP**, **DM1HEXP**, **DM8HEXP**, **DMTSEXP**, **SAVGEXP**, **TIMEEXP**, **TSEXP**, **DAVGDOSE**, **DM1HDOSE**, **DM8HDOSE**, **DMEHDOSE**, **DMTSDOSE**, **H\_EHDOSE**, **SAVGDOSE**, **TIMEDOSE**, **TSDOSE**, **TSMULTI**, and **RESPONSEPROB**. The table specifications for each pollutant must come after the corresponding **POLLUTANT** keyword (and prior to the next occurrence of “Pollutant” in a multi-pollutant run). Each parameter is identified by a single keyword, and the values are a list of numbers ordered from smallest to largest and separated by commas. All the values are read as real numbers, although

the decimal points are optional if the values happen to be integers. Items in each list must be separated by commas. Except for the **PERCENTILES**, all of these parameters are used to bin exposures or doses into categories in order to create output tables. Note that there is always one more bin than there are number of values in the list, since the first bin is less than the first value in the list and the last bin is greater than or equal to the last number in the list. The specific meanings of the parameters are explained in Table 4-3. These parameters, with the exception of **PERCENTILES**, are optional; if they are omitted, then the corresponding table is simply not written in the output file. See CHAPTER 5 for more information on the APEX output tables.

**Table 4-3. Output Parameter Levels in the Output Summary Table**

<b>Table Parameter</b>	<b>Keyword</b>	<b>Data Type</b>	<b>Description</b>
Percentiles	<b>PERCENTILES</b>	Real	(Required) This does not refer to a separate table, but specifies the population percentiles to be analyzed in the other exposure and dose tables. Each requested percentile occupies one row in each table. Values can include up to one digit beyond the decimal point (e.g. the 99.5 or 99.9 percentile).
Daily Average Exposure Cut points	<b>DAVGEXP</b>	Real	(Optional) This parameter specifies cut-points for daily average exposure for binning all the person-days in the simulation period.
Daily Max 1-Hour Exposure Cut points	<b>DMIHEXP</b>	Real	(Optional) This parameter specifies the daily maximum 1-hour exposure cut-points for binning all the person-days in the simulation period. (Note: 1-hour tables are not generated with the APEX timestep is greater than one hour.)
Daily Max 8-Hour Exposure Cut points	<b>DM8HEXP</b>	Real	(Optional) This parameter specifies daily maximum 8-hour average exposure cut-points for binning all the person days in the simulation period. It is similar to <b>DMIHEXP</b> except for the longer averaging time. (Note: 8-hour tables are not generated with the APEX timestep is greater than one hour.)
Daily Max Timestep Exposure Cut points	<b>DMTSEXP</b>	Real	(Optional) This parameter specifies daily maximum timestep exposure cut points for binning all the person days in the simulation period. It is equivalent to <b>DMIHEXP</b> , but for timesteps. (Note: If using the default timestep of one hour, then only the hour tables are generated - the timestep tables are not.)
Simulation Average Exposure Cut points	<b>SAVGEXP</b>	Real	(Optional) This parameter specifies cut-points for average exposure over the simulation period. The cut points are used to bin all simulated persons created in a run.
Exposure Cut points	<b>TIMEEXP</b>	Real	(Optional) This parameter specifies the exposure cut points for summing time spent at various exposure levels. The time is expressed in minutes and is summed across all profiles. <b>TIMEEXP</b> is used in two tables. (Exposure Tables Type 1 and 2; see discussion of <i>Tables</i> file in CHAPTER 5)
Timestep Exposure Cut points	<b>TSEXP</b>	Real	(Optional) This parameter specifies timestep exposure cut-points for counting multiple exceedances of timestep levels over the simulation (Exposure Table Type #9; see discussion of <i>Tables</i> file in CHAPTER 5)

Table Parameter	Keyword	Data Type	Description
Daily Average Dose Cut points	<i>DAVGDOSE</i>	Real	(Optional) This parameter specifies cut points for the daily average dose. The cut-points are used to bin all the person/days in the simulation period.
Daily Max 1-Hour Dose Cut points	<i>DMIHDOSE</i>	Real	(Optional) This parameter specifies cut points for daily maximum 1-hour dose. The cut points are used to bin all the person-days in the simulation period.
Daily Max 8-Hour Dose Cut points	<i>DM8HDOSE</i>	Real	(Optional) This parameter specifies cut points for daily maximum 8-hour dose. The cut points are used to bin all the person-days in the simulation period.
Daily Max End-of-hour Dose Cut points	<i>DMEHDOSE</i>	Real	(Optional) This parameter specifies cut points for daily maximum end-of-hour dose. The cut-points are used to bin all the person/days in the simulation period. Note that <i>DMEHDOSE</i> uses the instantaneous level at the end of each hour, whereas <i>DMIHDOSE</i> uses the time-averaged level over each hour. For CO, these two statistics track each other fairly closely. For other pollutants, the end-of-hour dose is just the dose on the last event of the hour.
Daily Max Timestep Exposure Cut points	<i>DMTSEXP</i>	Real	(Optional) This parameter specifies daily maximum timestep dose cut points for binning all the person days in the simulation period. It is similar to <i>DMIHDOSE</i> except that that the time period considered is a timestep rather than an hour. (Note: If using the default timestep of one hour, then only the hour tables are generated - the timestep tables are not.)
Hourly End-of-hour Dose	<i>H_EHDOSE</i>	Real	(Optional) Similar to <i>DMEHDOSE</i> , except that instead of using just the highest single end-of-hour dose on each day, it collects results for all 24 end-of-hour doses on each day. As with the other keywords, the values specified here refer to the cut points used for tabulating the dose results.
Simulation Average Dose Cut points	<i>SAVGDOSE</i>	Real	(Optional) This parameter specifies cut points in dose for the Average Dose over the entire simulation. The cut-points are used to bin all the persons (or profiles) created in the APEX run.
Dose Cut points	<i>TIMEDOSE</i>	Real	(Optional) This parameter specifies cut-points in dose for summing time spent at various dose levels. Apart from the statistic, the tables resemble the <i>TIMEEXP</i> tables.
Timestep Dose Cut points	<i>TSDOSE</i>	Real	(Optional) This parameter specifies timestep dose cut-points for counting multiple exceedances of timestep levels over the simulation (Exposure Table Type #9; see discussion of <i>Tables</i> file in CHAPTER 5)
Time Step Multiple Exceedance Cut-points	<i>TSMULTI</i>	Real	(Optional) This parameter lists the number of exceedances to use as cut-points in Exposure Table Type #9 and Dose Table Type #5 (multiple exposure or dose exceedances of timestep values of the simulation; see discussion of <i>Tables</i> file in CHAPTER 5). For example, if the user may want to track the number of persons that have 1, 10, 50, and 100 exceedances of the levels indicated by <i>TSEXP</i> and <i>TSDOSE</i> over the course of the simulation.

Table Parameter	Keyword	Data Type	Description
Risk Levels	<b>RESPONSEPROB</b>	Real	(Optional) These are lists of risk probabilities, with one probability for each exposure bin. Multiple lists may be provided. The first 20 characters after the “=” is a label, followed by comma-separated risk values starting 22 places after “=”. Each risk probability is multiplied by the number of persons in the corresponding exposure bin (see <b>RESPONSEVAR</b> ). A useful practice is to provide a final “ResponseProb=” line labeled “Bin counts” with a probability of 1.0 for each bin, to provide a summary line with the number of persons in each exposure bin.
Risk Exposure Metric	<b>RESPONSEVAR</b>	Char(4)	(Optional) One of “DM1H”, “DM8H”, or “DMTS”. APEX will use the exposure levels of the corresponding exposure table to calculate the number of persons expected to have a positive response (based on the <b>RESPONSEPROB</b> probabilities), for each exposure bin.
Multi-pollutant summary	<b>MULTIPOL</b>	Real	(Optional) This parameter specifies that a file be prepared and output containing information on the simultaneous exposure levels for two pollutants. Can only be used in a run with 2 or more pollutants. If used, <b>MULTI1</b> and <b>MULTI2</b> are also required.
Multipollutant #1 levels	<b>MULTI1</b>	Real	(Optional) A comma-separated list of exposure cut-points for the first named pollutant in the run, similar in format to the <b>TIMEEXP</b> list. The values may differ from those in <b>TIMEEXP</b> .
Multipollutant #2 levels	<b>MULTI2</b>	Real	(Optional) A comma-separated list of exposure cut-points for the second pollutant in the run, similar in format to the <b>TIMEEXP</b> list. The values may differ from those in <b>TIMEEXP</b> .
Hours in Running Averages	<b>RUNHOURS</b>	Integer	(Optional) The number of hours used in computing the running averages for EVR, Exposure, Dose, and HomeAmb. Default value is 8. Can use only if <b>TIMESTEPSPERDAY</b> >= 24. Note that the keywords for affected variables such as <b>DM8HEXP</b> and <b>DM8HDOSE</b> still contain an ‘8’ even if <b>RUNHOURS</b> is set to a different value.

The following example using a *Control Options* file excerpt (Exhibit 4-3) provides an illustration of a pollutant parameters section for the simulation of two pollutants: ozone and CO. Note that while the input units for ozone are ppb, the output is in ppm, and therefore the table cut-points are also in ppm. **PPMFACTOR** is not relevant for ozone because there are no microenvironment-specific sources and can be omitted. The user is alerted by a note on the log file for each simulated person who encounters ozone levels over 0.16 ppm. The **RESPONSEPROB** levels are reported for ozone exposure. This example has seven risk probabilities because the **DM8HEXP** table for ozone has six cut-points (and therefore seven bins). There are two sets of **RESPONSEPROB** probabilities, and a “Bin counts” line.

```

! POLLUTANT PARAMETERS
#Pollutants = 2

Pollutant = Ozone
DoDose = NO
RoadWay = YES
RoadLast = NO
InputUnits = ppb
OutputUnits = ppm
PPMFactor = 1.
#Sources = 0
AlertThresh = 0.16
Percentiles = 10, 25, 50, 75, 90, 95, 99
TimeExp = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08
DM1HExp = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08
DM8HExp = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08
DAvgExp = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08
SAvgExp = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08
ResponseVar = DM8H
ResponseProb = LO10 : 0,0.08,0.15,0.25,0.40,0.55,0.75
ResponseProb = LO15 : 0,0.02,0.04,0.07,0.10,0.15,0.25
ResponseProb = Bin Counts : 1,1,1,1,1,1,1

Pollutant = CO
DoDose = YES
RoadWay = NO
RoadLast = NO
InputUnits = ppb
OutputUnits = ppm
PPMFactor = 1145.0
#Sources = 1
Percentiles = 10, 25, 50, 75, 90, 95, 99
TimeExp = 2, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60
DM1HExp = 5, 10, 20, 30, 40, 50, 75
DM8HExp = 3, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 20, 25
DAvgExp = 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 16, 18, 20
SAvgExp = 0.5, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4, 5, 6, 8, 10
DM1HDose = 0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0
DM8HDose = 0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0
DMEHDose = 0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0
H_EHDose = 0.5, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0
DAvgDose = 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75
SAvgDose = 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8
AlertThresh = 100.

```

**Exhibit 4-3. Pollutant Parameters Section of an Example Control Options File**

### 4.2.3 Job Parameter Settings Section of the Control Options File

In the Job Parameter Settings section of the *Control Options* file, the user can specify a number of different job parameters for APEX runs. Table 4-4 provides a description of the keyword, data type, and uses of these job parameters. As with Input and Output Files, the keyword is the part of the Parameters input line that is necessary to allow APEX to identify the parameter. Data type must be integer, real, or character. Each character variable has a specified length; input values longer than allowed will be truncated to this length, and values shorter than allowed are

simply padded with blanks. In all cases in this section, with the exception of **COUNTY** or **TRACT**, if the same keyword appears more than once, then the last occurrence overwrites the others. Exhibit 4-4 shows an example of this section of the *Control Options* file.

When APEX runs, the values of all the job settings (including the default settings for parameters not explicitly set), will be printed to the *Log* file.

**#PROFILES:** This is the number of simulated persons to randomly generate and assess. Each one is independent of the others, and is followed for the duration of the simulation, from the **START\_DATE** to the **END\_DATE** (see below). The APEX execution time is essentially linear in the number of profiles, until a limit is reached when the data arrays must be stored in virtual memory, which slows it down substantially. If more profiles than this are desired, it is best to run APEX multiple times with the same inputs except for the **RANDOMSEED**. Then the output can be combined into what is essentially a single large run.

**RANDOMSEED:** APEX uses many random number seeds internally, but all are derived from this one. The Fortran random number generator is known to produce correlated streams of numbers in certain cases for simple offsets, so a second level of randomization was added to APEX to eliminate this. Effectively, this extra randomization amounts to a total scrambling of the list of 2,147,483,646 possible seeds, using a different random number generator. For this reason, simply incrementing **RANDOMSEED** by one, between two runs, will change the seeds in an unpredictable way, resulting in virtually no chance for correlations between runs.

Two APEX runs with the same inputs and same **RANDOMSEED** will produce identical output. If only **RANDOMSEED** is changed between runs, the two runs may have their output combined into a larger sample. Using the same seed with slightly different inputs in multiple runs allows a sensitivity analysis of the importance of the differing inputs. If different seeds were used in this case, then stochastic differences between the runs may obscure the effects of the changed inputs.

**START\_DATE** and **END\_DATE:** Dates in APEX are specified using an 8-digit string, with the year occupying the first 4 digits, the next 2 digits for the month, and the last 2 for the day. Months and days less than 10 require a leading zero as a place holder. This format is shown symbolically as: YYYYMMDD. While month/day/year or day/month/year formats are more common, the APEX order is used so that numerical sorting results in chronological sorting. The end date may not be less than the start date. If the two are the same, the simulation consists of a single day. Often, APEX is run for a calendar year (e.g., from 20140101 to 20141231). APEX may be run over longer periods, but the profiles will remain at their initial ages throughout the simulation. APEX knows which years contain February 29. Dated inputs such as air quality data must include the entire simulation period, although these files can contain data that extend beyond the start and end dates.

**TIMESTEPSPERDAY:** APEX uses three levels of time resolution within a day: the hour, the timestep, and the diary event. The air quality data is input as time series data with a fixed timestep, controlled by the parameter **TIMESTEPSPERDAY**. The default is 24, representing hourly data. The meteorological data must be hourly, regardless of **TIMESTEPSPERDAY**. If the timestep is less than one hour, then there must be a whole number timesteps in one hour to permit the determination of hourly output data. If the timestep is longer than one hour, then



there must be a whole number of hours in one timestep, and a whole number of timesteps in a day. Thus, if *TIMESTEPSPERDAY* is less than 24, it must be one of 12, 8, 6, 4, 3, 2, or 1.

The activity diary events may range from 1 to 60 minutes in length, and always break at the top of each hour. With hourly air quality data, which also change at the top of each hour, each diary event is associated with one (constant) air concentration, which simplifies the exposure calculations. In this case, multiple diary events within the same hour use the same ambient air concentration, although the microenvironmental concentration may change. When using the default hourly timestep, there is no point in requesting APEX to output both hourly files and timestep files, since the information on them would be the same.

If the timestep is less than one hour, any diary events that cross a timestep boundary are split. If the timestep is greater than one hour, then by insisting it must be a whole number of hours ensures that no diary event will cross a timestep boundary, and the calculation of (say) average temperature over the timestep is simple, given hourly temperature data. There must be a whole number of timesteps in one day to permit aggregation to daily totals.

***DSTADJUST***: Every day in an APEX simulation is 24 hours long, as is every activity diary. If a particular location uses daylight saving time (DST), one spring day is 23 hours long and one fall day is 25 hours long. Air quality data are reported as a time series, typically in Standard Time throughout the year. Suppose a particular pollutant spikes at 8 a.m. every day, due to traffic. The air quality data will show regular spikes every 24 hours, except for a 23 hour spacing on the first day of DST, and a 25 hour spacing on the last day, because the traffic follows human behavior (local time) patterns. When matched with APEX diaries, the spikes will appear to occur at 7 a.m. in the summer. This may affect the exposure calculations, since there may be fewer commuters available to be exposed at 7 a.m. If *DSTADJUST* = Yes, then APEX duplicates one hour of air quality data at the start of DST in the spring, and deletes one hour at the end of DST in the fall, effectively shifting the summertime air quality data so the spikes always occur at 8 a.m. APEX time; that is, 24 hours apart every day of the year. The start and end dates for DST are coded into APEX and change each year. U.S. states may opt out of DST, but if they use it, they must follow the national dates. At present, only Arizona and Hawaii do not use DST.

***COUNTYLIST***, ***TRACTLIST***, and ***CITYRADIUS***: The user may specify the study area in three ways: using counties, tracts, and a target circle. The basic geographical unit in APEX is the “sector”, which in a standard run is the same as a census tract. The APEX sector names (or ID) may be up to 40 characters in length, but census tract names are just 11 characters long. A “county” is defined by the first five characters of the sector ID. For example, “06037” is Los Angeles County when using census tract names. If the user supplies a list of their own sector names, APEX will still match the first five characters if *COUNTYLIST* = Yes in the *Control Options* file. If *COUNTYLIST* = Yes, then each county in the study area is listed on a separate line as “County = xxxxx”. For example, County = 06037 selects Los Angeles County. All counties must be listed on consecutive lines; once a different keyword is encountered, the county list is assumed to have ended.

If the county resolution is too coarse, the option *TRACTLIST* = Yes may be used. This operates just like *COUNTYLIST*, except that the first 11 characters of the sector name are matched. If

both *COUNTYLIST* and *TRACTLIST* are used, then sectors that match either one are included in the study area (the union of the lists is used).

The study area is also defined by a specific latitude and longitude and must lie within a given radius (in km) of this point (*CITYRADIUS*). This defines a circular area. All sectors in the study area must lie within this circle (that is, the intersection of this circle and the county/tract list is used). If the user wants to effectively use just the county/tract list, then specify a very large radius. If a small radius is used, it may select only some of the sectors in a listed county. If no sectors remain, an error occurs.

**KEEPLAVERS:** The default target population in APEX is all persons who live inside the study area. If commuting is modeled, some people who live inside the study area will work outside of it. The air quality at these work places is unknown to APEX, and is assumed to be related to the average concentration over all air districts in the study area at the same point in time. Calling this average  $C_{avg}$ , the ambient concentration  $C$  for work outside the study area is:  $C = LEAVERMULT * C_{avg} + LEAVERADD$ . For **KEEPLAVERS** = No, those persons are still modeled, but are excluded from all output tables. This ensures that the random numbers assigned to each individual are not affected by the **KEEPLAVERS** status, if comparisons are being made to earlier runs using the same random number seed but altered job settings. The point is that sometimes the corresponding person in the two runs may stay inside the study area in one run but leave in the other, depending on which job settings were changed. For **KEEPLAVERS** = Yes, the target population is all persons who live in the study area. For **KEEPLAVERS** = No, the target population for the output tables is all persons who both live and work inside the study area.

**NEARBYRADIUS:** APEX assigns a home sector (and also a work sector if employed) for each simulated person. Certain diary events like shopping or restaurants may be assigned to “other” sectors. Every sector has a list of nearby sectors, where “nearby” means within a distance of **NEARBYRADIUS**, as measured (in km) between the nominal centers of each sector. The default is 20 km. If there are multiple nearby sectors to the current location, one is selected at random for diary events in “other” locations. The air quality at the “other” sector might (or might not) be different from the “base” sector, depending on the mapping of sectors to air quality districts.

**COMCUT1, COMCUT2, COMPROBAB1, COMPROBAB2:** Each simulated individual who works is assigned a commuting time. APEX uses this parameter to match the commute time on each diary day to the target time of the person. For example, if a person’s commute time is 30 minutes and **COMCUT1** = 10 minutes, then diaries with a commute time between 20 and 40 minutes inclusive will be given full (100%) weight. Diaries outside of this window, but inside **COMCUT2** of the target time, receive a weight of **COMPROBAB1**. All other diaries in the appropriate pool receive a weight of **COMPROBAB2**. These weights are used when randomly selecting diaries. Do not set all four variables to zero, or APEX may stop prematurely due to the lack of any possible matches.

**Table 4-4. Job Parameters in the APEX Control Options File**

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>SIMULATION PARAMETERS</b>		
<b>#PROFILES</b>	Integer	(Required) Number of profiles to simulate.
<b>RANDOMSEED</b>	Integer	(Optional) <b>RANDOMSEED</b> > 0 is specified by the user, <b>RANDOMSEED</b> = 0 gets seed from clock. <b>RANDOMSEED</b> may be any natural number up to 2147483646. Default = 0.
<b>END_DATE</b>	Integer	(Required) Simulation end date in YYYYMMDD format.
<b>FIRSTPROFILE</b>	Integer	(Optional) First profile number to simulate. For example, this can be used for skipping to a particular person's profile when performing repeated runs using a controlled <b>RANDOMSEED</b> . Default = 1.
<b>START_DATE</b>	Integer	(Required) Simulation start date in YYYYMMDD format (e.g., 19960704 for July 4, 1996).
<b>TIMESTEPSPERDAY</b>	Integer	(Optional) Number of timesteps in a day. This setting dictates the required time resolution of the air quality input data, as well as the resolution of calculated exposures and doses. Default = 24.
<b>OCCFACTOR</b>	Char(1)	(Optional) This parameter tells APEX that the <i>Profile Factor</i> file contains a parameter that is related to occupation. If this parameter is set to "Y", then the profile factor will only be applied to employed individuals, and an extra factor group will be made for all unemployed profiles. Default = No.
<b>STUDY AREA PARAMETERS</b>		
<b>LOCATION</b>	Char(40)	(Optional) Study area location (for output labeling only; not used internally). Default = empty character string.
<b>LATITUDE</b>	Real	(Required) Latitude in decimal degrees for the center of the study area. Note that latitude south of the equator is negative.
<b>LONGITUDE</b>	Real	(Required) Longitude in decimal degrees for the center of study area. Note that longitude between 0 and 180 degrees west of the prime meridian is negative (e.g., in the United States).
<b>ALTITUDE</b>	Real	(Optional) Altitude of study area in feet. The altitude is assumed constant for the study area. It is used in the Coburn-Forster-Kane (CFK) equation for determining blood COHb concentration. Only necessary when simulating CO dose. Default = 0.
<b>DSTADJUST</b>	Char(1)	(Optional) Y = adjust air quality data for Daylight Saving Time (DST), N = do not use DST. Default = Yes.
<b>CITYRADIUS</b>	Real	(Optional) Radius of study area in km. The population sectors (e.g., census tracts) with centers (or representative locations) within this radius will be automatically selected for modeling. If the default is used, then <b>COUNTYLIST</b> should be specified. Default = 99999.
<b>AIRRADIUS</b>	Real	(Optional) Maximum representative radius (km) of air quality data. Each sector uses the nearest air quality site, as long as it is within a distance <b>AIRRADIUS</b> . This term can be specified via the <i>Air District Location</i> File; however, if <b>AIRRADIUS</b> is in the <i>Control Options</i> file, it will overwrite those found in the <i>Air District Location</i> File. Default = 99999.
<b>MODELAQVAR</b>	Char(1)	(Optional) Dictates the expected format of the <i>Air Quality Data</i> file. The default is ModelAQVar = N, when APEX expects AQ values for each timestep. If <b>MODELAQVAR</b> = Y, then APEX expects AQ distributions for each hour of the simulation. Default = No.
<b>ZONERADIUS</b>	Real	(Optional) Maximum representative radius (km) of temperature data collected at a weather station. Default = 99999.

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>COUNTY</b>	Char(5)	(Optional) FIPS code for listed county (or other relevant portion of the sector ID if the supplied sector files are not used). <b>COUNTY</b> is used only if <b>COUNTYLIST</b> = Y. Repeat this line for each additional county code.
<b>COUNTYLIST</b>	Char(1)	(Optional) <b>COUNTYLIST</b> = Y means that the study area is composed of sectors in the listed counties (previous variable) and within <b>CITYRADIUS</b> ; otherwise the study area is restricted to sectors within the specified <b>CITYRADIUS</b> only or defined with the <b>TRACTLIST</b> . (May be used in conjunction with <b>TRACTLIST</b> ; final study area is union of tracts and counties listed). Default = No.
<b>TRACT</b>	String	(Optional) Sector ID for a listed sector (usually census tract). <b>Tract</b> is used only if <b>TRACTLIST</b> = Y. Repeat this line for each additional sector to be used.
<b>TRACTLIST</b>	Char(1)	(Optional) <b>TRACTLIST</b> = Y means that the study area is composed of the sectors (usually census tracts) listed using the <b>TRACT</b> keyword which are within <b>CITYRADIUS</b> ; N = the study area is restricted to sectors within the specified <b>CITYRADIUS</b> only or defined with the <b>COUNTYLIST</b> . (May be used in conjunction with <b>COUNTYLIST</b> ; final study area is union of tracts and counties listed). Default = No.
<b>SCENARIO</b>	Char(40)	(Optional) Scenario description (for output labeling only; not used internally). Default = blank character string.
<b>NEARBYRADIUS</b>	Real	(Optional) Radius (km) for use when randomly selecting nearby tracts to the home or work location (For use with the NW, NH, L locations). Default = 20.
<b>RESAMPLEN</b>	Char(1)	(Optional) This allows the user to select either a new random nearby tract for each day, or use a single tract for the whole simulation. Default = No.
<b>MICROENVIRONMENTAL PARAMETERS</b>		
<b>#MICROS</b>	Integer	(Required) Number of microenvironments defined in the <i>Microenvironment Mapping</i> file and on the <i>Microenvironment Descriptions</i> file.
<b>COMMUTING PARAMETERS</b>		
<b>COMMUTING</b>	Char(1)	(Optional) <b>COMMUTING</b> = Y allows a simulated person to commute to a work sector. <b>COMMUTING</b> = N means any work occurs in the home sector. If "Y", a work sector is randomly selected for each simulated profile based on the probabilities in the commuting database. Default = No.
<b>KEEPLAVERS</b>	Char(1)	(Optional) <b>KEEPLAVERS</b> = Y means that people who live in the study area but work outside it area still modeled. Default = No.
<b>LEAVERADD</b>	Real	(Optional) Additive concentration term applied when working outside study area (only used if <b>KEEPLAVERS</b> = yes). Default = 0.
<b>LEAVERMULT</b>	Real	(Optional) Multiplicative factor for city-wide average concentration, applied when working outside study area (only used if <b>KEEPLAVERS</b> = yes). Default = 0.
<b>COMCUT1</b>	Real	(Optional) The width (in minutes) of the window of commuting times within which all times will be weighted by 100%. Default = 0.
<b>COMCUT2</b>	Real	(Optional) The width (in minutes) of the second commuting time window. This parameter works similar to the previous commuting time window, <b>COMCUT1</b> . For example, if a person's commute time target is 60 minutes, <b>COMCUT1</b> = 10 min, <b>COMCUT2</b> = 20 min, and a diary has a commute time of 77 minutes (within ±20 min but not within 10 min), then that diary will be given a weight of <b>COMPROBABI</b> . Default = 0.

Keyword	Type (length)	Description
<b>COMPROBAB1</b>	Real	(Optional) The weight given to diaries with commuting times in the window between <b>COMCUT1</b> and <b>COMCUT2</b> on either side of the target commuting time assigned to the current profile. Default = 0.
<b>COMPROBAB2</b>	Real	(Optional) The weight given to all diaries that fall outside the <b>COMCUT1</b> and <b>COMCUT2</b> windows. For example, with a target of 60 minutes and <b>COMCUT2</b> = 20 min, this probability is assigned to all diaries with <40 minutes or >80 minutes of commuting time. Default = 0.
<b>DIARY SELECTION PARAMETERS</b>		
<b>AGE2PROB</b>	Real	(Optional) Diary probability factor for “shoulder” ages. This parameter allows an optional shoulder window of ages outside the primary age window. The shoulders have the same width in years as the main age window, so in the example under <b>AGECUTPCT</b> the shoulders are ages 20–29 and 51–60. The <b>AGE2PROBAB</b> parameter operates like <b>MISSAGE</b> , by suppressing the selection probability in the shoulders. If <b>AGE2PROBAB</b> = 0 then shoulder ages are never selected. Default = 0.
<b>AGECUTPCT</b>	Real	(Optional) Width of main age window (%). Each simulated profile (person) is assigned a specific year of age, but the activity diaries assigned to this person do not need to match this age exactly. A window is created around this target age, of size equal to <b>AGECUTPCT</b> percent of the target age. If the profile age is 40 and <b>AGECUTPCT</b> = 25, then the age window is ten years wide (25% of 40) and diaries for persons from 30 to 50 years of age inclusive are permitted to be selected. The age window is always at least 1 year wide, even when using the default. Default = 0.
<b>AGEMAX</b>	Integer	Maximum age for simulated profiles (persons). Each profile is assigned a specific age, used for selecting diaries and for physiological variables. This age does not change over the simulation, even if it is one year or longer.
<b>AGEMIN</b>	Integer	Minimum age for simulated profiles (persons).
<b>MISSAGE</b>	Real	(Optional) Diary probability factor for missing age. Some of the supplied CHAD diaries are for persons of unknown age. This factor operates just like <b>MISSGENDER</b> and <b>MISSEMPL</b> to lower the selection probability for such diaries. Default = 0.
<b>MISSEMPL</b>	Real	(Optional) Diary probability factor for missing employment. Some of the supplied CHAD diaries are for persons of unknown employment status. Like <b>MISSGENDER</b> , this factor lowers the selection probability for such diaries. If <b>MISSEMPL</b> = 0, then such diaries will never be selected. Default = 0.
<b>MISSGENDER</b>	Real	(Optional) Diary probability factor for missing gender. Some of the supplied CHAD diaries are for persons of unknown gender. All APEX profiles are assigned gender, however, and the CHAD diaries are selected from those of the same gender or from the unknowns. <b>MISSGENDER</b> is used as a multiplicative factor to reduce the probability of selecting diaries of unknown gender. If <b>MISSGENDER</b> = 0, then diaries with missing gender will never be selected. If <b>MISSGENDER</b> = 1, then such diaries are equally likely to be selected as diaries of the correct gender. <b>MISSGENDER</b> can also be set to values between zero and one. Allowing small but nonzero values for <b>MISSGENDER</b> and the other <b>MISS</b> parameters may prevent empty diary pools. Default = 0.

Keyword	Type (length)	Description
<b>MISSOCC</b>	Real	(Optional) Diary probability factor for missing occupation. This parameter will only be used if diaries are to be weighted via occupation, which requires the inclusion of a <i>Profile Factors</i> file, and <b>OCCFACTOR</b> = Y and <b>USEOCCGROUPS</b> = Y. Most diaries in CHAD have missing occupation; <b>MISSOCC</b> provides a weight for these diaries. If <b>MISSOCC</b> = 0, then no diaries other than those matching the occupation of the individual will be selected. Age, gender, employment and occupation each produce a probability, which are then multiplied together to determine the overall diary selection probability for that profile. Default = 0.
<b>USEOCCGROUPS</b>	Char(1)	(Optional) This parameter tells APEX to match diaries based on occupation group. To use this option, occupations must be specified in the <i>Profile Factors</i> file. Default = No.
<b>DOSE PARAMETERS</b>		
<b>COHBFACTOR</b>	Real	(Only needed when modeling CO) Convergence parameter for COHb algorithm. This is a safety factor that limits the permitted error in determining the solution to the CFK equation. Larger factors mean greater accuracy but slower evaluation. Numerical tests indicate that factors in the range of 2–3 are optimal for most purposes. Only necessary when simulating CO dose.
<b>LOCATION PARAMETERS</b>		
<b>#OTHER</b>	Integer	(Optional) Number of other districts to use in calculating the air quality for diary events with location = O (“Other”) when <b>SAMPLEOTHERLOCS</b> is used. (The probability of the person’s home district being one of these districts is given by <b>HOMEPROBAB</b> .) Default = 1.
<b>CUSTOMWORK</b>	Comma-delimited list	(Optional) List of CHAD activity codes that will be assigned to location = W (Work). Default is all CHAD activity codes < 11000.
<b>HOMEPROBAB</b>	Real	(Optional) Probability (0–1) of a person’s home district being one of the districts used to calculate the air quality for diary events with location = O (“Other”) when <b>SAMPLEOTHERLOCS</b> is used. Default = 0.
<b>SAMPLEOTHERLOCS</b>	Char(1)	(Optional) If <b>SAMPLEOTHERLOCS</b> = Y, a random list of air districts will be selected for each person for calculating the air quality for diary events with location = O (“Other”). The number of districts selected for each person is given by <b>#OTHER</b> , and the probability of the person’s home district being the list is given by <b>HOMEPROBAB</b> . Default = No.
<b>ROLLBACK PARAMETERS</b>		
<b>RBACK</b>	Real	(Optional) Rollback background concentration. Use same units as <b>INPUTUNITS</b> . The background is the part of the air quality that is not affected by Rollback controls. Default = 0.
<b>RBMAX</b>	Real	(Optional) Rollback maximum concentration. Use same units as <b>INPUTUNITS</b> . Default = 0.
<b>RBTARGET</b>	Real	(Optional) Rollback target concentration. Use same units as <b>INPUTUNITS</b> . The target must be less than the maximum for any air quality improvement to occur. If <b>RBTARGET</b> = <b>RBMAX</b> , the air quality is unchanged by rollback. If <b>RBTARGET</b> = 0, the air quality always stays at <b>RBACK</b> . Default = 0.
<b>ROLLBACK</b>	Char(1)	(Optional) <b>ROLLBACK</b> = Y uses air quality rollback adjustments. The default is <b>ROLLBACK</b> = N. <b>ROLLBACK</b> adjusts the ambient air quality data before the exposure calculations occur. The purpose is to determine exposure in hypothetical scenarios where the ambient concentrations have been reduced by various controls. Default = No.

Keyword	Type (length)	Description
<b>DIAGNOSTIC PARAMETERS</b>		
<b>DEBUGLEVEL</b>	Integer	(Optional) A value > 0 results in more information being written to the log file than for a value of zero. For production runs with a large number of profiles, use <b>DEBUGLEVEL</b> = 0, or else the log file will become too large. <b>DEBUGLEVEL</b> may also be 1 or 2. Default = 0.
<b>LOG FILE SWITCHES</b>		
<b>LOGDISTRICT</b>	Char(1)	(Optional) Y = the name and location of each of the air districts will be written to the <i>Log</i> file. Both a preliminary list (all the air districts in the <i>Air District Location</i> file that are within the study area and have data for the entire simulation period) and a final list (those required to simulate the final list of study sectors) are printed. Default = Yes.
<b>LOGPOPULATION</b>	Char(1)	(Optional) Y = the following population information will be written to the <i>Log</i> file for each study area sector: The total population of the sector (TotalPop); the base population for the study (StudyPop), which will be smaller than TotalPop if only certain age ranges are being considered; the total population of workers in the sector (Workers); the sector population of workers who work inside the study area (WorkInside); and the final population (FinalPop) for the simulation, which may be smaller than StudyPop if the workers who leave the sector are excluded (if <b>KEEPLEAVERS</b> = NO). Default = Yes.
<b>LOGPROFILES</b>	Char(1)	(Optional) Y = the following population information will be written to the <i>Log</i> file for each study area sector: The total population of the sector (TotalPop); the base population for the study (StudyPop), which will be smaller than TotalPop if only certain age ranges are modeled; the total number of workers in the modeled age range who live in the sector (Workers); the population of these workers who work inside the study area (WorkInside); and the final study population of the sector (FinalPop), which may be smaller than StudyPop if the commuters who leave the study area are not modeled (if <b>KEEPLEAVERS</b> = NO). Default = Yes.
<b>LOGSECTORS</b>	Char(1)	(Optional) Y = the name and location of each study sector will be written to the <i>Log</i> file. Both a preliminary list (all the sectors geographically within the study area) and a final list (those sectors within the study area having available air quality and temperature data) are printed. Default = No.
<b>LOGTABLES</b>	Char(1)	(Optional) Y = all the tables that are written to the <i>Tables</i> file are also written to the <i>Log</i> file. Default = No.
<b>LOGZONES</b>	Char(1)	(Optional) Y = the name and location of each of the meteorological zones will be written to the <i>Log</i> file. Both a preliminary list (all the air districts in the <i>Meteorology Zone Location</i> file that are within the study area and have data for the entire simulation period) and a final list (those required to simulate the final list of study sectors) are printed. Default = Yes.
<b>OUTPUT FILE SWITCHES AND KEYWORDS</b>		
<b>CLUSTEROUT</b>	Char(1)	(Optional) Y = Write the cluster assignments for all diaries in the database to an output file. Default = No.

Keyword	Type (length)	Description
<b>CUSTOMSAMPLE</b>	Comma-separated list of integers	(Optional) The profiles designated by <b>CUSTOMSAMPLE</b> are written in addition to the profiles specified by the <b>EVENTSAMPLE</b> variable. If both <b>EVENTSSAMPLE</b> and <b>CUSTOMSAMPLE</b> are set, then all the <b>EVENTSSAMPLE</b> events are written as before and any additional <b>CUSTOMSAMPLE</b> events are written in the appropriate place in the numerical profile order. Writing of <b>CUSTOMSAMPLE</b> events is dictated by the value of the <b>EVENTSOUT</b> variable, so no events will be written if <b>EVENTSOUT=N</b> , even if a <b>CUSTOMSAMPLE</b> is specified. If neither <b>CUSTOMSAMPLE</b> nor <b>EVENTSAMPLE</b> is set, then events are written as dictated by the default <b>EVENTSAMPLE</b> value (if <b>EVENTSOUT = Y</b> ). If the user wishes to write only the <b>CUSTOMSAMPLE</b> events, then <b>EVENTSAMPLE</b> should be set to 0. Default = empty list.
<b>CWEIGHTOUT</b>	Char(1)	(Optional) Y= Write the cluster weights for all diary pool-cluster combinations to an output file. Default = No.
<b>DAILYLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Daily</i> output file. See Section 5.4 for details. Only relevant if <b>DAILYOUT = Yes</b> .
<b>DAILYOUT</b>	Char(1)	(Optional) Y= the <i>Daily</i> output file containing values of daily parameters (exposures, doses, etc.) is created. Otherwise it is not written. Default = No.
<b>EVENTSAMPLE</b>	Integer	(Optional) Dictates which profiles have their event data written to the <i>Events</i> file. If <b>EVENTSAMPLE = K</b> , then K diaries are written out, spaced evenly through the run. For example, <b>#PROFILES = 10,000</b> and <b>EVENTSAMPLE = 10</b> prints profile #1000, 2000, etc. This only has effect if <b>EVENTSOUT = Yes</b> . Default = 10.
<b>EVENTSLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Events</i> output file. See Section 5.10 for details. Only relevant if <b>EVENTSOUT= Yes</b> .
<b>EVENTSOUT</b>	Char(1)	(Optional) Y = the output file containing the event-level model outputs for each simulated individuals is written. Otherwise, the file is not written. Default = No.
<b>HOURLYLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Hourly</i> output file. See Section 5.2 for details. Only relevant if <b>HOURLYOUT = Yes</b> .
<b>HOURLYOUT</b>	Char(1)	(Optional) Y= the <i>Hourly</i> output file containing values of hourly parameters (exposures, doses, etc.) is created. Otherwise it is not written. Default = No.
<b>MRESHOME</b>	Char(1)	(Optional) If = Y, then only values associated with “home” locations will be written to the <i>Microenvironmental Results</i> file. Otherwise, values will be written for each of the locations specified in the <i>Microenvironment Mapping</i> file. Default = No.
<b>MRESLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Microenvironmental Results</i> output file. See Section 5.6 for details. Only relevant if <b>MRESOUT = Yes</b> .
<b>MRESMICROS</b>	comma-separated list of integers	(Optional) A comma-separated list of integers that indicate the microenvironments for which data will be written to the <i>Microenvironmental Results</i> file. Only relevant if <b>HOURLYOUT = Yes</b> .



<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>MRESOUT</b>	Char(1)	(Optional) Y = the <i>Microenvironmental Results</i> file will be created. Otherwise, the file is not written. Default = No.
<b>MSUMOUT</b>	Char(1)	(Optional) Y = the <i>Microenvironmental Summary</i> file will be created. Otherwise, the file is not written. Default = No.
<b>PSUMLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Profile Summary</i> output file. See Section 5.5 for details. By default, only 12 variables are printed, but the user may add up to 59 others using this option.
<b>SOBOLRUN</b>	Char(1)	(Optional) Y = perform Sobol sensitivity analysis, using the input variable groupings on the <i>Seed</i> file. See CHAPTER 11 in <i>Volume II</i> for details. Default = No.
<b>SOBOLVAR</b>	comma or space-separated strings	(Optional) List of APEX output variables subject to Sobol analysis. See CHAPTER 11 in <i>Volume II</i> for details. Default = none.
<b>TABLESLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Tables</i> output file. See Section 5.8 for details. Default = none.
<b>TIMESTEPOUT</b>	Char(1)	(Optional) Y= the <i>Timestep</i> file will be created. Otherwise, the file is not written. If the default timestep (1 hour) is used, then this file will not be written because it will be identical to the <i>Hourly</i> file. Default = No.
<b>TIMESTEPLIST</b>	comma or space-separated strings	(Optional) List of keywords indicating which variables are to be written to the <i>Timestep</i> output file. See Section 5.3 for details. Only relevant if <b>TIMESTEPOUT</b> = Yes.
<b>TRANSOUT</b>	Char(1)	(Optional) Y= Write the pool-cluster transitions for all diaries in the database to an output file. Default = No.
<b>TABLE PARAMETERS</b>		
<b>ACTIVE</b>	Real	(Optional) Threshold median daily PAI (MET) value for defining active persons. Simulated individuals having median PAI equal to or greater than this value over the simulation period will be included in the "active persons" subgroup in the output exposure tables. Default = 0 (this will catch all person-days).
<b>CHILDMAX</b>	Integer	(Optional) Maximum age for inclusion in the "child" and "active child" population subgroups in the output exposure tables. Default = 0.
<b>CHILDMIN</b>	Integer	(Optional) Minimum age for inclusion in the "child" and "active child" population subgroups in the output exposure tables. Default = 0.
<b>HEAVYEVRI</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining one-hour heavy exertion. It is used in generating the APEX output tables for one-hour exposures under heavy exertion. Default = 0.
<b>HEAVYEVRI8</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining eight-hour heavy exertion. It is used in generating the APEX output tables for eight-hour exposures under heavy exertion. Default = 0.
<b>HEAVYEVRTS</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining timestep-level heavy exertion. It is used in generating the APEX output tables for timestep exposures under heavy exertion. Thus, this value should be dependent on the length of timestep used. Default = 0.

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>MODEVR1</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining one-hour moderate exertion. It is used in generating the APEX output tables for one-hour exposures under moderate exertion. Default = 0.
<b>MODEVR8</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining eight-hour moderate exertion. It is used in generating the APEX output tables for eight-hour exposures under moderate exertion. Default = 0.
<b>MODEVRTS</b>	Real	(Optional) This parameter sets the threshold for equivalent ventilation rate defining timestep-level moderate exertion. It is used in generating the APEX output tables for timestep exposures under moderate exertion. Thus, this value should be dependent on the length of timestep used. Default = 0.
<b>LONGITUDINAL DIARY SELECTION PARAMETERS</b>		
<b>DIARYAUTOC</b>	Real	(Optional) Lag-1 autocorrelation statistic for the D&A longitudinal diary assembly algorithm. Provides a target for the autocorrelation in the key diary statistic. Default = 0.
<b>DIARYD</b>	Real	(Optional) Provides a target D statistic for the D&A longitudinal diary assembly algorithm. The D statistic reflects the relative importance of within person variance and between person variance in the key diary statistic. Default = 0.
<b>LONGITDIARY</b>	Char(1)	(Optional) Y = APEX will use the D&A longitudinal diary assembly algorithm to construct the activity diaries for the simulated persons, based on the statistics in the <i>DiaryStat</i> file. In this case, <b>DIARYAUTOC</b> , <b>DIARYD</b> , and the name of the diary statistics file must all be designated in the <i>Control Options</i> file. Default = No.
<b>CLUSTERDIARY</b>	Char(1)	(Optional) Y = APEX will use the clustering algorithm for longitudinal diary assembly based on the transitional probabilities calculated by analyzing the CHAD input files. Default = No
<b>CLUSTERAGES</b>	comma-separated list of integers	(Optional) Age cut-points for binning the empirical transition probabilities. A list of N cut-points gives N+1 bins. The cut-point age goes into the higher bin. Default = 0 (all ages in one bin).
<b>USEADJACENT</b>	Integer	(Optional) Minimum number of diaries in a pool age bin combination for restricting the empirical transition probabilities to consecutive calendar dates (no gaps allowed). Groups with fewer examples also use non-adjacent cases.
<b>PHYSIOLOGY PARAMETERS</b>		
<b>DISEASE</b>	Char(12)	(Optional) Provides the name of a condition or disease. If set, then APEX expects the <i>Prevalence</i> file to be defined as well, and a subpopulation of persons with the condition will be modeled, resulting in exposure summary tables corresponding to the subpopulation. The tables will be labeled using this variable; spaces are allowed. Default = none.
<b>HTWTMETHOD</b>	Integer	(Optional) Selects the method of setting height and weight for each profile. <b>HTWTMETHOD</b> = 1 is the method used by APEX prior to 2017. <b>HTWTMETHOD</b> = 2 has distributions with explicit correlation between height and weight, with parameters matched to NHANES data. Default = 2.
<b>MODEVRMETHOD</b>	Integer	(Optional) Selects the method of determining the cutoff for moderation exertion (8-hour average only) for output tables. <b>MODEVRMETHOD</b> = 1 applies the same cutoff (given by ModEVR8) to everyone, while <b>MODEVRMETHOD</b> = 2 assigns a value for each profile using a distribution from the physiology input file.

<b>Keyword</b>	<b>Type (length)</b>	<b>Description</b>
<b>VEMAX</b>	Char(1)	(Optional) Y = apply an upper limit to VE, which is 100 L/min for events over 5 minutes, and 150 L/min for shorter events. N = calculate VE without explicit truncation. Default = Y.
<b>VEMETHOD</b>	Integer	(Optional) Selects the method of determining VE on each diary event. <b>VEMETHOD</b> = 1 is the method used in APEX prior to 2017. <b>VEMETHOD</b> = 2 is a new method that uses both VO2 and VO2max. Default = 2.

```

! ----- PARAMETER SETTINGS -----
! SIMULATION PARAMETERS
#Profiles = 40000
RandomSeed = 0
Start_date = 20040401
End_date = 20040930
!
! STUDY AREA PARAMETERS
Location = Description of Location of the Study Area
Latitude = 33.7629
Longitude = -84.4004
Altitude = 150.
DSTadjust = YES
CityRadius = 100.
AirRadius = 25.
ZoneRadius = 100.
CountyList = YES
TractList = NO
NearbyRadius= 5.
ResampleN = NO
County = 01017
County = 13013
County = 13015
!
! MICROENVIRONMENT PARAMETERS
#Micros = 12
!
! COMMUTING PARAMETERS
Commuting = YES
KeepLeavers = YES
LeaverMult = 0.0
LeaverAdd = 0.0
ComCut1 = 10.0
ComCut2 = 20.0
ComProbab1 = 0.20
ComProbab2 = 0.05
!
! DIARY SELECTION PARAMETERS
AgeMin = 0
AgeMax = 99
ChildMin = 5
ChildMax = 18
MissGender = 0.0
MissEmpl = 0.0
MissAge = 0.0

```

```

MissOcc      = 0.5
AgeCutPct   = 20.0
Age2Probab  = 0.05
UseOccGroups= YES
OccFactor   = YES
!
! DOSE PARAMETERS
COHbFact    = 2.5
!
! LOCATION PARAMETERS
CustomWork  =
SampleOtherLocs = YES
#OtherDistricts = 2
HomeProbab  = 0
!
! ROLLBACK PARAMETERS
Rollback    = NO
RBtarget    = 5.0
RBbackgnd   = 0.0
RBmax       = 10.0
!
! DIAGNOSTICS PARAMETERS
DebugLevel  = 0
!
! LOG FILE SWITCHES
LogDistrict = NO
LogPopulate = NO
LogProfiles = NO
LogSectors  = NO
LogTables   = NO
LogZones    = NO
!
! OUTPUT FILE SWITCHES AND KEYWORDS
EventsOut   = YES
EventSample = 2
CustomSample = 3092
MResOut     = NO
MSumOut     = NO
HourlyOut   = NO
DailyOut    = YES
PSumList    = AVGEXP,MAXEXP,AVGEXP,MAXEXP
HourlyList  = CONC1, AMB, EXP, EVR, VE, VA, EE, MET, EF, DFEV1
DailyList   = MAX1DOSE MAX8DOSE MAX1FDOSE AVGDOSE
MResList    = VOL, AER, RR, PRX, PEN, CSUM, AMB
MResHome    = YES
MResMicros  = 1,2,8,12
!
! TABLES PARAMETERS
HeavyEVR1   = 30
HeavyEVR8   = 99
ModEVR1     = 16
ModEVR8     = 13
ActivePAI   = 1.76
!
! LONGITUDINAL DIARY PARAMETERS
LongitDiary = YES
DiaryAutoC  = 0.2

```

```
DiaryD      = 0.5
!  
! CLUSTERING DIARY PARAMETERS  
ClustDiaryA = NO  
ReRunClus  = NO  
ClustDiaryB = NO
```

**Exhibit 4-4. Job Parameters Sections of an Example Control Options File**

### 4.3 Population Sector Location File

The *Population Sector Location* file provides the latitude and longitude of a representative location such as the geographic center of all the sectors (e.g., census tracts) to be included in the population data files. Each line includes a **Sector Name**, **Latitude**, and **Longitude**. The sector name (also called “ID”) may be any string, numeric or character, and is stored as a character string (up to length 40). The string may contain any characters except “!” or embedded spaces. When census tracts are used as the sectors, the names are composed entirely of numerical digits, and the term “ID” is common in such cases. The sector names must match those in the *Commuting Flow* and *Commuting Time* files (if worker commuting is being modeled), although APEX now optionally accepts a coarser level of spatial resolution for the former. The names are case-sensitive, so the values in the two files must match exactly. Tract-level 2010 Census demographic files covering the U.S. are provided with APEX.

The population sector location file is used along with the user-specified **CITYRADIUS** to automatically select population sectors within the study area (after also addressing an optional county test and ensuring suitable air district and meteorology zone data). APEX calculates the distance between the location of a sector and the center of the study area and then compares it with the **CITYRADIUS**. Sectors with a distance from the study area center greater than the city radius will not be included in the exposure assessment.

The tract-level population sector location file supplied with APEX contains the 11-character names and latitudes and longitudes of the corresponding year U.S. census tracts. APEX expects that the left-most five characters of a sector name will be the state and county FIPS code, or the county-level code used in the **COUNTY** list (if the study area is limited in that way). APEX reads counties and tracts in character format, so for example Los Angeles County is 06037, with the leading zero retained.

If not modeling actual United States counties and/or tracts, the user may create custom names for their sectors that follow the same rules. The “county” refers to the first 5 characters of the sector name, and the “tract” refers to the first 11 characters (i.e., the county name plus six more). The **COUNTYLIST** and **TRACTLIST** options can be used for subsetting customized sector files to smaller study areas.

The latitude and longitude should be in decimal degrees. At least three digits should be provided after the decimal point to prevent significant rounding error. Note that the longitude west of the prime meridian (e.g., United States locations) should be negative. Exhibit 4-5 provides an example of the first few records of this input file. These tracts are all in county “01001”.

```

e! TractID   Latitude Longitude
01001020100 32.47711 -86.4903033
01001020200 32.47576 -86.4724678
01001020300 32.47402 -86.4597033
01001020400 32.47108 -86.4446805
01001020500 32.45892 -86.4218165
01001020600 32.44735 -86.4768023
01001020700 32.43052 -86.4369107
01001020801 32.41172 -86.531683
01001020802 32.54713 -86.531596
...

```

**Exhibit 4-5. First Part of the Population Sector Location File (2010 Census)**

#### 4.4 Air District Location File

The *Air District Location* file provides the *Site Name*, *Latitude*, *Longitude*, air data *Start Date*, air data *End Date*, and optionally, *AirRadius*, for all air quality (modeling or monitoring) sites included in the *Air Quality Data* file (Section 4.5). As for sector names, the site name (or ID) may be any string, numeric or character, and is stored as a character string (up to length 40), but must not contain an ! character or embedded spaces. Latitude and longitude are in decimal degrees. The start and end dates are in YYYYMMDD format (for example, 19951231 is December 31, 1995). If the user wishes to define unique air district radii for each district, instead of supplying a single one using the *AIRRADIUS* parameter in the *Control Options* File, then the user can supply a sixth column in this file, which is the *AirRadius* in km. If no *AirRadius* values are set either in *Control Options* File or on the *Air District Location*, each air monitor will have an unlimited effective radius, and each sector will use the nearest monitor.

It is a good practice to insert comments on the first few lines of each input file to indicate the source or type of data used. See Exhibit 4-6 for an example of the first few records of an *Air District Location* file. The variables are site name, latitude, longitude, start date, end date, and effective radius. The designated start and end dates for the simulation must be entirely covered by the date range indicated on this file, or else the monitor will be discarded.

```

! Hourly ozone air quality districts for an example metropolitan area
! This file contains the locations of 105 air quality districts
! Created on November 4, 2005
Roadway1   33.500000 -85.300000 20040301 20041031 25
Roadway2   34.500000 -85.300000 20040301 20041031 25
0000100010 34.371470 -85.461103 20040301 20041031 30
0000100009 34.194947 -85.461103 20040301 20041031 30
0000100008 34.018423 -85.461103 20040301 20041031 35
0000100007 33.841899 -85.461103 20040301 20041031 35
0000100006 33.665375 -85.461103 20040301 20041031 30
0000100005 33.488851 -85.461103 20040301 20041031 30
0000100004 33.312327 -85.461103 20040301 20041031 25
0000100003 33.135804 -85.461103 20040301 20041031 25
0000200011 34.547994 -85.239577 20040301 20041031 10
...

```

**Exhibit 4-6. First Part of an Example Air District Location File**

APEX uses the *Air District Location* file to determine the “air district” or geographical area represented by the ambient air quality data for a specified location. All pollutants must use the same air districts and thus there is only one file of this type. APEX first compares the start and end dates for each air quality site with the start and end dates for the APEX exposure simulation. Only the sites with air quality data covering the entire simulation period are accepted. If a site is encountered with incomplete data (gaps) between the listed start and end dates, APEX prints a warning to the log file and stops execution. For example, the warning message is “*No appropriate districts found*” if the start date of the simulation is before the first date on the concentration input file. Air quality data in the file for dates before or after the simulation period are simply ignored. If the user wishes to specify a roadway air district, then “road” must appear within the district name. Consequently, “road” cannot appear in the name of a regular AQ district.

APEX then calculates the distance of an air district location from the study area center and compares it with the sum of *CityRadius* and *AirRadius*. This allows air quality data to be used from sites a little outside the study area, in case they happen to be the nearest to some population sectors. Only the sites with a distance less than this sum are retained for further calculations.

APEX then calculates the distances of each site from the sector locations. Sectors within a distance *AirRadius* of an air site are included in the final study area, and use the nearest site for their ambient air data. Each sector is assigned to only one air district. Sectors within the study area that lack a matching air district are not included in the simulation. If no sectors remain, the model stops during initialization and does not simulate anyone.

Not all air districts on the air quality input file need sectors assigned to them. Such air districts are simply not included in the modeling. This feature allows the user to prepare an input file in the simplest manner, perhaps containing more air districts than are necessary. For example, a single input file could be prepared for all air districts in a given state. This same input file could then be run on several study areas in the state without having to alter the air quality input file.

Internally, APEX refers to air quality districts by a sequential index (district #1, #2, etc.) that is assigned when the district-sector mapping is established. The *Log* file for the model run reports the names and locations for each air quality district number (these are also in the *Sites* output file). Note that district #1 for a particular study area might not always mean the same location on the ground for all model runs. For example, if a series of runs for different years in Denver were performed, different monitors might be online during different years, in which case district #1 might change meaning from year to year. This can be avoided by preparing an *Air Quality Data* input file (see next section) that has complete data for all air quality districts for all years being modeled, in which case the mappings should remain the same from year to year.

## 4.5 Air Quality Data File

The *Air Quality Data* file provides air concentration data for air sites listed in the *Air District Location* file for a given pollutant; there is one file of this type for each pollutant in the simulation. Only keyword or numeric input lines are processed; other types of input lines are ignored in this file with the exception of the first line, which (even if it is a comment) is always

echoed to the header in each output file. Therefore, the first line should contain information describing the content of the file.

There are two different types of air quality (AQ) data files that may be used in APEX. The first type of file simply contains values of the AQ data for each air district for each timestep (e.g., hour) in the simulation. The second type of AQ data file contains distributions that allow for person-to-person variability in the AQ data for each hour of the simulation. This type of file may only be used when the APEX timestep is 1 hour. APEX expects type 1 by default; if a type 2 file is to be used, the user must set the *Control Options* file flag **MODELAQDATA** = YES.

The formats of the two types of files are described in detail below.

#### 4.5.1 Air Quality Input Data (Type 1)

Type 1 is the APEX default. Within this file, the data for each site begins with a header section containing the site **Name** (see Exhibit 4-7). Recall that these site names must match those in the *Air District Location* file exactly; the names are case sensitive and must not contain an ! or embedded spaces. The sites can be in any order in this file. APEX locates the air data set by matching a site name in the *Air District Location* file with the site name in this file. There can be no missing data within the simulation period. If monitoring data contains gaps or missing values, then these should be filled in by the user prior to running APEX.

Each of the subsequent numeric records includes a list of **Timestep Average Air Concentrations** followed by a **Date**. The date should be in YYYYMMDD format (e.g., 20010507 is May 7, 2001). Air quality data should be in the units specified in the *Control Options* file for the pollutant. The data values can be either comma or space delimited. Note that the length of each data line in an air quality file should not exceed 5000 characters. For example, if the APEX timestep is one hour, then each numeric record will list 24 hourly average concentration values followed by a date. Thus, if the APEX timestep is 5 minutes, then each line will have 288 5-minute averages followed by the date. An example of the beginning portion of this type of file is depicted in Exhibit 4-7. The ellipsis on each line indicates data that were removed in this exhibit so the final date entry on each line could be seen.



```

! Ozone air quality data for an example metropolitan area
! For 105 air quality districts, for the period 03/01/04 to 10/31/04
! Created on November 4, 2005
Name = Site0000100003
0.01553 0.01825 0.02621 0.02989 0.02975 0.02650 0.02310 ... 0.03891 20040301
0.03822 0.03738 0.03749 0.03754 0.03687 0.03550 0.03240 ... 0.00948 20040302
0.00577 0.00570 0.00528 0.00477 0.00394 0.00453 0.00430 ... 0.01169 20040303
0.01456 0.01828 0.01916 0.01810 0.01547 0.00925 0.00591 ... 0.03326 20040304
0.03354 0.03244 0.02412 0.01705 0.01293 0.01076 0.01066 ... 0.02849 20040305
...

```

#### Exhibit 4-7. First Part of an Example Air Quality Data File (Type 1)

#### 4.5.2 Air Quality Input Defined as Hourly Distributions (Type 2)

AQ input data defined as hourly distributions can be used to model person-to-person variability within an hour within an AQ district. Consequently, data of this type can only be used if the APEX timestep is equal to 1 hour (that is, *TIMESTEPSPERDAY* = 24, the APEX default).

Within this input file, the data for each site begins with a header section containing the site ID or *Name* (see Exhibit 4-7). Recall that these site IDs must match those in the *Air District Location* file exactly; the IDs are case sensitive and must not contain an ! or embedded spaces. The sites can be in any order in this file. APEX locates the air data set by matching a site name in the *Air District Location* file with the site name in this file. There can be no missing data within the simulation period.

In this type of AQ file, each numerical record begins with a date and an hour number, followed by any APEX distribution definitions. See Table 2-1 and *Volume II* for a discussion of available probability distributions in APEX. If this type of input is to be used, the *Control Options* file flag *MODELAQVAR* must be set to Y, otherwise an APEX error will result. An example of the first part of an AQ data file (distribution type) is shown below in Exhibit 4-8. In this example, the AQ value for each hour is defined by a normal distribution. The ambient AQ value for the hour for will be sampled from this distribution for each person in the air quality district.

```

! Hourly ozone air quality distributions for an example metropolitan area
! This file contains data for 127 air quality districts, for the period 01/01/04 to 12/30/04
! Created on February 26, 2008.
! Format is Date Hour DistributionDef
! Where DistributionDef is any standard APEX distribution definition
Name =0000200006
! Date Hr Distribution
20040101 1 Normal 0.01066 .00005 . . 0 0.10 Y
20040101 2 Normal 0.01121 .00005 . . 0 0.10 Y
20040101 3 Normal 0.01184 .00005 . . 0 0.10 Y
20040101 4 Normal 0.01067 .00005 . . 0 0.10 Y
20040101 5 Normal 0.01231 .00005 . . 0 0.10 Y
20040101 6 Normal 0.01515 .00005 . . 0 0.10 Y
20040101 7 Normal 0.01537 .00005 . . 0 0.10 Y
...

```

**Exhibit 4-8. First Part of an Example Air Quality Data File (Distribution Type)**

## 4.6 Meteorology Zone Location File

The format and use of the *Meteorology Zone Location* file is analogous to the *Air District Location* file. Each record represents one site, and contains five values: **Site ID**, **Latitude**, **Longitude**, **Start\_Date**, and **End\_Date**. In the same way, the **Site ID** may be any string up to 40 characters long; it cannot contain an ! or embedded spaces. The IDs must match those in the *Meteorology Data* file exactly; the IDs are case sensitive. The site selection process is also analogous to that described above for the *Air District Location* file. The file is used to map the set of meteorology data collected at a weather station to sectors within its zone radius for exposure calculations. An example file is provided in Exhibit 4-9. Similar to air districts, zones within the sum of **CITYRADIUS** and **ZONERADIUS** are used. Study area sectors for which no meteorology data are available are not included in the simulation.

APEX makes an internal list of meteorological zones that have sectors assigned to them and assigns them sequential numbers for convenience. This mapping is reported in the *Log* file and *Sites* file, which are output from each model run.

```

! Example APEX Meteorological Station Locations (Zones) File
! Created 11/4/05
03812 35.4333 -82.5333 20040101 20041231
03813 32.7000 -83.6500 20040101 20041231
03816 37.0667 -88.7667 20040101 20041231
03820 33.3667 -81.9667 20040101 20041231
03856 34.6500 -86.7667 20040101 20041231
03870 34.9000 -82.2167 20040101 20041231
03937 30.1167 -93.2167 20040101 20041231
...

```

**Exhibit 4-9. First Part of an Example Meteorology Zone Location File**

## 4.7 Meteorology Data File

The *Meteorology Data* file provides hourly temperature and meteorological data for the sites listed in the *Meteorology Zone Location* file. Only numeric input lines or lines containing the keyword “name” followed by an equal sign are processed. All other types of input lines are ignored.

The meteorology sites may be in any order in this file. The section of data for each site must begin with the “name” keyword input line. An example is shown in Exhibit 4-10 (some times have been removed for display, as indicated by ellipses). The site names (site IDs) must match those in the *Meteorology Zone Location* file exactly; the IDs are case sensitive and must not contain an ! or embedded spaces.

APEX matches a site name in the *Meteorology Zone Location* file with the data set site name to locate its data in this file. If desired, the user can add more comment lines in the header section of a data set.

Temperatures can be used to assign activity diaries to days (via the profile function *DiaryPools*, see Section 4.17), and any meteorological variable present in the file may be used as conditional variables for microenvironment parameters (see Section 4.17.2 and *Volume II*).

The “site name” input line is followed by the meteorological data. Each data line may contain the data listed below.

- Date (YYYYMMDD)
- Hour (1–24)
- Temperature (degrees Fahrenheit)
- Relative Humidity (percent)
- Precipitation (character code, 1 or 2 characters)
- Wind speed (km per hour)
- Wind Direction (degrees clockwise from north)

The data are not required to be in fixed columns, but must be separated by whitespace only.

The numerical data may be integer or real (decimal)—they are truncated to integers when the file is read.

The precipitation code may be any string of 1 or 2 characters. The codes used for precipitation must match those used in the *Profile Functions* file (see Section 4.17).

Only date, hour, and temperature are required. To use an optional variable such as wind speed, then precipitation and humidity must also be present. This does not imply that the user must make use of the precipitation data in the model run (e.g., to set microparameter distributions, see Section 4.17). Therefore, a dummy code could be entered for precipitation in this case.

Each data set should cover the entire exposure simulation period, but it may extend further. Thus, the user may prepare a file with a full year or many years of data for each site and then use the same meteorology file for a series of different simulation periods. There can be no missing data within the simulation period.

```

!Hourly Meteorological Data
!Date      Hr  Temp  Humidity Prec Windspeed  Direction
name=03812
20040101   1   64.0   30.0    RA   12.0    180.0
20040101   2   64.1   30.1    CL   12.3    182.5
20040101   3   64.7   31.8    RA   15.5    215.3
20040101   4   65.3   33.3    CL   18.7    246.7
20040101   5   65.9   34.8    CL   21.6    276.3
20040101   6   65.7   34.2    CL   20.4    263.8
...
name=03813
20040101   1   66.1   30.4    CL   12.7    187.4
20040101   2   67.6   34.1    CL   20.1    261.1
20040101   3   67.7   34.2    CL   20.4    264.2
20040101   4   67.0   32.4    CL   16.9    228.6
20040101   5   66.4   31.0    CL   14.0    199.8
20040101   6   66.9   32.3    CL   16.7    227.0
...

```

**Exhibit 4-10. Part of an Example Meteorology Data File**

## 4.8 Population Data Files

Each *Population Data* file contains sector-level data for a single gender/race combination. Ten gender/race specific population data files for all year 2010 census tracts have been prepared and provided with the APEX release. However, user-defined population data files may be constructed, if the format given below is followed.

The population files contain the population counts for each sector contained in the *Sector Location* file. In general, each population file is for a single race/gender combination, although composite files containing more than one gender or race can be used. The population counts are given by age group. The age groups are designated in the first part of the file (i.e., the descriptor records).

Four descriptor records must appear in each population file. These records must appear immediately after any header comment records (which start with “!”) and before the population data records (i.e., the actual population counts). The data on these four records are read starting to the right of the “=” sign, if present. Text descriptors to the left of the “=” signs are optional. The contents of these four records must be as shown below.

- Descriptor record 1: **Gender** (“Female” “Male” or “All”), **Race** (5 characters), **#Age** (number of age groups)
- Descriptor record 2: **Race** description (may contain blanks, up to 200 characters)
- Descriptor record 3: **Age Group Minimum**
- Descriptor record 4: **Age Group Maximum**

The fields in descriptor records 1, 3, and 4 are space-delimited. **Gender** must be “Female” “Male” or “All.” The 5-character label for race also appears as a column header on the *Profile Summary* output file. If the population files provided with APEX are to be used, the **Race** must be White, Black, Asian, NatAm, or Other, which may be shortened to W, B, A, N, or O. If the user provides the population files, **Race** could be different. For example, if one file each is given

for all males and all females, **Race** could be specified as *All*. However, it is necessary for **Race** to match the designation in the *Control Options* file, or a fatal error will result.

The race description is not used, but is echoed in the log file for the benefit of the user. Only the shorter 5-character race label that is given on the first line is written to the other output files, to save space.

The next two records specify the minimum and maximum ages for the age groups. The ages must be delimited by a single space. Note that all the population data files must contain the same number of population groups, and furthermore, all the group age limits (minima and maxima) must match as well, or APEX will exit with a fatal error. The *Employment Probability* file, *Prevalence* file, and *Profile Factors* file can have different age groupings. The *Population Data* files provided with APEX contain single-year age groups.

The actual population data follows the descriptors records. Each population record has the **Sector ID**, (which must match the IDs in the *Sector Locations* file exactly, and thus can be any alphanumeric string of 40 or fewer characters without embedded spaces or an !) followed by a **Count** for each population age group (youngest first). The counts are the number of people in a given age group living in the sector; they must be delimited by a single space. Each *Population Data* file used in a model run must have a record for each sector listed in the population *Sector Location* file or a fatal error will result. The sectors do not necessarily have to be in the same order in every population file in order for APEX to run, however, a warning message will result if APEX finds that the order of the sectors in any population file differs from the order of the sector list. A single error message will be written for each population file having out-of-order sectors, no matter how many differences are found. APEX will exit with a fatal error message if a sector in the final list of study area sectors cannot be found in a population file.

Exhibit 4-11 provides an example of a portion of a *Population Data* file. The ellipsis on each line indicates data that were removed for brevity in this exhibit.

```

! Population file by census tract, extracted from 2010 census.
! File prepared by Redhorse Corporation and ICF International, January 2014.
! The 99 to 99 age group includes people older than 99.
Gender,Race,#Age = Female, Black, 100
Race description = Black or African American Alone, Not Hispanic or Latina
Age group minimum = 0 1 2 3 4 5 6 7 8 9 10 11 12 ... 98 99
Age group maximum = 0 1 2 3 4 5 6 7 8 9 10 11 12 ... 98 99
!
01001020100          1 1 0 2 2 0 1 2 1 1 5 1 2 ... 0 0
01001020200          4 11 5 6 4 10 5 10 13 12 7 11 2 ... 0 0
01001020300          3 7 2 5 4 3 2 3 6 12 3 8 4 ... 0 0
01001020400          2 2 0 0 2 2 1 2 3 2 1 1 1 ... 0 0
01001020500          9 9 14 7 15 13 11 11 14 17 16 9 15 ... 0 1
...

```

**Exhibit 4-11. First Part of a Population Data File (2010 Census, Female Black or African American, Not Hispanic or Latina)**

### 4.9 Commuting Flow File

The *Commuting Flow* file provides cumulative fractions of the population in a home sector that commute to different work sectors. An example portion of this file is provided in Exhibit 4-12. After the header lines, this file is composed of sections, each starting with a home place and an arbitrary number of work places. The names of these places must be matched to those contained in the *Sector Location* file, as discussed below. The first record of each section lists the **Home Sector ID** followed by two values of -1. These have no meaning; it is simply used by APEX to recognize the beginning of a new data section (i.e., a new home). After the home record, each of the work places for that home is listed. Each work record contains the **Work Place ID**, a **Cumulative Fraction** of the home sector population commuting to this work place, and the **Distance (km)** between the home and work places. The cumulative fraction for the last work place in each group should always be equal to 1. APEX uses this file to determine which work sector a simulated individual may commute to by using the cumulative fractions as commuting probabilities.

The commuting place names must all be of the same length. These either match the population sector names, or can be shorter. If shorter, APEX assumes that each one applies to all population sectors that start with the commuting place name. In this case, the destination place is determined using the probabilities on the commuting flow file, then a single population sector is selected at random from all that map to the same commuting place.

If the population sectors are U.S. census tracts, or subdivisions of tracts, then the national database of *Commuting Flow* data may be used. Otherwise, the user must prepare a replacement file using the same format, unless commuting is not modeled.

The default *Commuting Flow* database contains all the U.S. census tracts and their associated work tracts. For the 2000 Census files, the mean number of associated work tracts per home tract was 79, with a minimum of 1 and a maximum of 413. The 2010 database uses the same format, but the list of home and work tracts is compatible with the 2010 population files (and is different from those used in 2000).

```

! APEX U.S. Tract-Level Commuting File for use with 2010 Census
! Prepared by Alion Science and Technology, March 2014
! ID          CumulFraction  Distance(km)
01001020100 -1.00000      -1.0
01001020200  0.19075       1.8
01001020400  0.27168       4.4
01101000200  0.34104      21.0
01101003302  0.40462      28.3
01001020500  0.46243       6.9
01001020100  0.50867       0.0
01001020700  0.55491       7.6
...

```

**Exhibit 4-12. First Part of the Commuting Flow File (2010 Census)**

### 4.10 Commuting Time File

The *Commuting Time* file provides the distribution of the duration of one-way commuting times for all workers in all tracts. An example portion of this file is shown in Exhibit 4-13. The first column lists all census tracts (*Home Sector ID*), and the following columns show the number of people in each of fifteen bins, which for the 2000 Census were:

1. Total: Workers 16 years and over
2. Total: Did not work at home
3. Total: Did not work at home: Less than 5 minutes
4. Total: Did not work at home: 5 to 9 minutes
5. Total: Did not work at home: 10 to 14 minutes
6. Total: Did not work at home: 15 to 19 minutes
7. Total: Did not work at home: 20 to 24 minutes
8. Total: Did not work at home: 25 to 29 minutes
9. Total: Did not work at home: 30 to 34 minutes
10. Total: Did not work at home: 35 to 39 minutes
11. Total: Did not work at home: 40 to 44 minutes
12. Total: Did not work at home: 45 to 59 minutes
13. Total: Did not work at home: 60 to 89 minutes
14. Total: Did not work at home: 90-120 minutes
15. Total: Worked at home (0 commuting time)

APEX uses these bins to create a cumulative probability distribution of one-way commuting time, which it uses in conjunction with commuting distance information, to assign a profile-level commuting time variable to each employed person in the population. These data are from the 2000 or 2010 Census, and as such, include all census tracts included for that year. Take care to note that the census data for the bins shown here are one-way commuting times, while the times in the *Diary Questionnaire* file should indicate the *total* daily commuting time. APEX accounts for the difference internally, as the time from the census commuting bin is doubled before matching to activity diary times.

```

! APEX U.S. Tract-Level Commuting File from 2010 American Community Survey
! Prepared by WGG at Alion Science and Technology, March 24 2014
! Variables: Tract, allworkers, allnonhome, timebins 1-9, work-at home
! New headers in March 2014: first is #bins (do not include work-at-home)
! Second header has bin boundaries (lowest time in each bin)
! The final bin may be open-ended (if so, APEX assumes 30 min width)
number of bins = 9
boundaries =
01001020100 905 890 40 200 80 135 335 60 10 25 0 15
01001020200 810 810 10 110 115 210 265 90 0 0 10 0
01001020300 1540 1520 45 270 260 505 355 55 15 0 15 25
01001020400 2190 2155 55 700 215 585 495 45 30 4 15 40
...

```

**Exhibit 4-13. First Part of the Commuting Time File (2010 Census)**

The 2010 *Commuting Time* file uses different time bins. There are only 12 numbers per tract, as follows:

1. Total: Workers 16 years and over
2. Total: Did not work at home
3. Total: Did not work at home: Less than 5 minutes
4. Total: Did not work at home: 5 to 14 minutes
5. Total: Did not work at home: 15 to 19 minutes
6. Total: Did not work at home: 20 to 29 minutes
7. Total: Did not work at home: 30 to 44 minutes
8. Total: Did not work at home: 45 to 59 minutes
9. Total: Did not work at home: 60 to 74 minutes
10. Total: Did not work at home: 75 to 89 minutes
11. Total: Did not work at home: 90-120 minutes
12. Total: Worked at home (0 commuting time)

#### 4.11 Employment Probability File

A nationwide *Employment Probability* file was prepared for ages 16 and above, covering all of the tracts from the 2010 Census. Each record (tract) contains 26 probabilities—13 age groups each for males and females. The age groups in the provided file are comprised of the following: 16–19, 20–21, 22–24, 25–29, 30–34, 35–44, 45–54, 55–59, 60–61, 62–64, 65–69, 70–74, and 75 and older.

The employment probability age groups do *not* have to match the population file age groups, providing increased flexibility in the demographic inputs to APEX. As a result, users may create their own employment files, as long as the file format is followed. The ages in the *Employment Probability* file may extend beyond those in the population files, but APEX will never generate a profile outside of the ages in the *Population Data* files.

An example portion of the *Employment Probability* file is provided in Exhibit 4-14. The file contains optional header lines, followed by three required lines. The first required line reports the **Gender** for each column of data, the second line reports the age group minimum (**MinAge**),



and the third line reports the age group maximum (*MaxAge*). Below that, each line starts with the *Sector ID*, followed by a vector of decimal probabilities (one per column). The first item on each line below the header lines is the *Sector ID*, followed by the eight employment probabilities for that sector. Each probability in the national file is calculated by dividing the number of employed persons by the total sector population for the specified age range and gender. Whenever the total sector population for a particular age range and gender is zero, then obviously, the employed persons must also be zero. These data are reported as zero probabilities in the file. It should not matter what values are assigned, since no simulated persons of that type should ever be generated by the model. Users should be aware that a custom *Employment Probability* file must be created if custom *Population Data* files are used. The sectors in the *Employment Probability* file must include all those in the study area. If the file includes extra sectors not on the population files, the extras must be outside the study area. The ellipsis on each line indicates data that were removed for brevity in this exhibit.

Note that any ages not covered by one of the employment age groups will automatically have an employment probability of zero. In the example below this would apply to persons younger than age 16.

```
! Employment probability fractions by gender and age group from 2010 census.
! File prepared by Redhorse Corporation and ICF International, January 2014.
Gender=      M      M      M      M      ... F      F
MinAge=     16     20     22     25     ... 70     75
MaxAge=     19     21     24     29     ... 74     200
01001020100 0.00000 0.56098 1.00000 0.73913 ... 0.00000 0.00000
01001020200 0.00000 1.00000 0.85714 0.78947 ... 0.00000 0.00000
01001020300 0.37903 0.35000 1.00000 1.00000 ... 0.11538 0.00000
01001020400 0.26866 0.76923 0.68293 0.82178 ... 0.10577 0.02459
01001020500 0.51103 1.00000 0.87903 0.84733 ... 0.00000 0.00000
01001020600 0.06667 1.00000 0.68750 1.00000 ... 0.00000 0.00000
01001020700 0.00000 1.00000 0.00000 1.00000 ... 0.30769 0.06977
01001020801 0.13830 1.00000 0.00000 0.56250 ... 0.00000 0.00000
...
```

**Exhibit 4-14. First Part of the Employment Probability File (2010 Census)**

## 4.12 Profile Factors File

The *Profile Factors* file allows the user to specify particular profile factors by age, gender, and sector. These profile factors can then be used to select microenvironmental scaling factors for each group. The groups can be applied to all individuals, or to employed individuals only.

The *Profile Factors* file is indicated by the keyword **PROFILE** in the *Control Options* File. If a *Profile Factors* file is supplied, APEX will automatically use it. Groups are assigned in a very similar manner as in the *Employment Probability* file. Apart from comment lines, the *Profile Factors* file starts with the following lines. The first line uses the syntax “*Level = Number*” to indicate how many factor groups are included in the file. In the example below (Exhibit 4-15), five groups are specified. The next three lines define the age-gender categories. The second line indicates the gender, and the next two lines indicate the minimum and maximum ages for each category. All modeled ages must be defined. The data section for each group begins with the keyword (*Name = Name*). Each sector ID starts a new line, followed by a list of probabilities—

one for each age/gender category. Finally, an “END” must be inserted after every sector/probability block to inform APEX that this block is complete.

If the profile factors are tied to employment, then groups will only be applied to employed individuals. Unemployed individuals will automatically be assigned to Group 1, while Group 2 will be defined by the first block listed in the *Profile Factors* file, Group 3, the next block, and so on. To specify employment groups, a new keyword, **OCCFACTOR**, was added to the *Control Options* file. Set **OCCFACTOR** = YES to use this option. When matching diaries based on occupation, it is important to be careful to match the name of the group to the name of the diary occupations.

APEX uses these probabilities to select a group for each individual. These probabilities do not have to equal one—APEX will automatically scale these values before using them to randomly assign a group to each individual. Groups are stored as a profile variable that may be printed in the *Profile Summary* file by using the keyword **FGROUP**. The **FGROUP** number corresponds to the order of the groups listed in the *Profile Factors* File. The name of each profile group can also be printed using the **GROUPNAME** keyword.

Finally, these groups can be used as a conditional in the *Microenvironmental Descriptions* file. To specify profile factor groups, the user should use the keyword **FACTORGROUP** as the *Conditional* name. Microenvironmental parameters must be defined for all groups, even for unemployed individuals, when using employment-related groups.

```

! Profile Factors fractions by gender and age group
! Prepared by Alion Science and Technology, Inc. for EPA in November 2011
Levels          = 5
Gender=         M      M      M      F      F      F
MinAge=         0      46     78     0      46     78
MaxAge=         45     77     99     45     77     99
Name = 0-0.5km
36025991300 0.71  0.74  0.11  0.00  0.38  0.56
36025991400 0.62  0.70  0.12  0.82  0.34  0.42
36105950200 0.76  0.57  0.04  0.81  0.56  0.95
36105950300 0.49  0.14  0.11  0.11  0.96  0.39
36105950400 0.66  0.86  0.37  0.27  0.95  0.96
END

Name = 0.5-2km
36025991300 0.45  0.22  0.26  0.55  0.75  0.53
36025991400 0.71  0.93  0.50  0.09  0.95  0.53
...

```

**Exhibit 4-15. First Part of an Example Profile Factors File**

The *Profile Factors* file can also be used to match diaries by occupation. If occupations are defined in the file (**OCCFACTOR** = Y), and diary matching is selected (**USEOCCGROUPS** = Y), then APEX will weight diaries by occupation. Each diary is already assigned an occupation in the standard diary input file, but the user can choose to define a different list of occupations in the *Profile Factors* file. If the pre-assigned diary occupation is not on the *Profile Factors* list, the occupation for that diary is changed to be missing (“X”).

APEX gives full selection probability weight (100%) to diaries that have the same occupation as the individual's occupation. APEX will not select diaries with a different occupation and will give a user-specified weight (*MISSOCC*) to diaries that have missing occupation. *MISSOCC* has a default value of zero, so if this is not changed, then exact occupation matching will be forced. Setting *MISSOCC* to a non-zero value may avoid situations of diary pools with no possible matches for some occupation, which would cause APEX to stop without completing the run. Occupation has no influence on the selection of diaries for individuals who are not considered to be workers.

Different occupations have different MET distributions for work activities. Under normal diary matching (i.e., not using *OCCFACTOR* and *USEOCCGROUPS*), these occupations are the pre-assigned ones on each diary. Hence, the occupation (and the MET values while at work) for a simulated individual may change from day to day, whenever the selected diary changes. When matching diaries by occupation, each profile is assigned one occupation, which determines the work-time MET values on all simulation days.

### 4.13 MET Mapping File

The *MET Mapping* file maps each CHAD (or other database) activity code to an internal APEX distribution number in order to calculate the energy expended by a simulated person for each diary event. Energy expenditures are used for estimating activity level and ventilation for each simulated person. These quantities are used for used in creating tables of exposures at different exertion levels and for estimating pollutant dose. A MET value is a dimensionless ratio of the activity-dependent energy expenditure rate to the basal or resting energy expenditure (metabolic) rate, and the CHAD activity code is an identifier associated with each diary event that indicates the type of activity being performed. The current CHAD activity codes are listed in Table 4-5.

**Table 4-5. CHAD Activity Codes**

<b>Code</b>	<b>Activity Description</b>	<b>Code</b>	<b>Activity Description</b>
10000	Work and other income producing activities, general	17121	Passive, sitting
10100	Work, General	17140	Create art, music, participate in hobbies
10110	Work, general, for organizational activities	17141	Participate in hobbies
10111	Work for professional/union organizations	17142	Create domestic crafts
10112	Work for special interest identity organizations	17143	Create art
10113	Work for political party and civic participation	17144	Perform music / drama / dance
10114	Work for volunteer/helping organizations	17150	Play, unspecified, general
10115	Work of/for religious groups	17151	Play, unspecified, low level
10116	Work for fraternal organizations	17152	Play, unspecified, moderate level
10117	Work for child/youth/family organizations	17160	Use of computers
10118	Work for other organizations	17170	Participate in recess and physical education
10120	Work, income-related only	17180	Other sports and active leisure, general
10130	Work, secondary (income-related)	17200	Passive leisure, general
10200	Unemployment	17201	Indoor passive leisure
10300	Breaks	17210	Watch

<b>Code</b>	<b>Activity Description</b>	<b>Code</b>	<b>Activity Description</b>
11000	Household activities, general	17211	Watch adult at work
11001	Other household	17212	Watch someone provide childcare
11100	Prepare food, general	17213	Watch personal care
11101	Washing	17214	Watch education
11110	Prepare and clean-up food	17215	Watch organizational activities
11120	Cooking	17216	Watch recreation
11121	Baking	17220	Listen to radio/listen to recorded music/ watch T.V.
11122	Fry, grill, sauté	17221	Listen to radio
11130	Simple food preparation	17222	Listen to recorded music
11131	Cutting/chopping	17223	Watch TV
11200	Indoor chores, general	17230	Read, general
11210	Move things	17231	Read books
11220	Put things away	17232	Read magazines / not ascertained
11230	Straighten up	17233	Read newspaper
11231	Make bed	17240	Converse / write
11240	Clean-up food	17241	Converse
11300	Outdoor chores, general	17242	Write for leisure / pleasure / paperwork
11301	Shoveling	17250	Think and relax
11310	Clean outdoors	17260	Other passive leisure
11320	Chop wood	17300	Other leisure
11330	Garden	17400	Walk, bike, or jog (not in transit), general
11331	Harvest	17410	Bike, general
11332	Watering	17411	Bike
11333	Weeding	17412	Cycles, other
11340	Lawn/grass	17413	Tricycle
11341	Lawn watering	17420	Run or jog, general
11342	Lawn weeding	17421	Run around, casual
11343	Mowing	17422	Running, vigorous/sustained
11344	Pruning	17430	Walk, general
11345	Use rake/leaf blower	17431	Crawl
11350	Load/unload	17432	Use of walker
11360	Mechanical chores	17433	Walk dog
11370	Move objects	17434	Walk for chores
11400	Care of clothes, general	17435	Walk inside
11401	Fold/sort	17500	Participate in sports, general
11410	Wash clothes	17501	Archery
11411	Hand wash/dry	17502	Equestrian sports
11412	Laundry	17503	Frisbee
114201	Maintain clothes	17504	Gymnastics
11421	Mend/sew	17505	Skateboarding
11422	Press/iron/steam	17506	Skating
11500	Build a fire	17507	Track
11600	Repair, general	17510	Combat sports

<b>Code</b>	<b>Activity Description</b>	<b>Code</b>	<b>Activity Description</b>
11610	Repair of boat	17511	Boxing
11620	Paint home / room	17512	Fencing
11630	Repair / maintain car	17513	Martial arts
11640	Home repairs, general	17514	Wrestling
11641	Home improvement/ construction, moderate level	17520	Racquet sports
11642	Home maintenance, low level	17521	Badminton
11650	Other repairs	17522	Racquetball
11700	Care of plants, general	17523	Squash
11710	Care of plants, low level	17524	Tennis
11720	Care of plants, moderate level	17530	Team sports
11800	Care for pets/animals	17531	Baseball
11900	Clean house, general	17532	Basketball
11901	Collect/empty trash	17533	Cheerleading
11902	Mop	17534	Dodgeball
11903	Sweep	17535	Football
11904	Vacuum	17536	Hockey
11910	Clean house, heavy	17537	Kickball
11911	Clean bathroom	17538	Lacrosse
11912	Clean carpet	17539	Rugby
11913	Clean floors	17541	Soccer
11914	Clean kitchen	17542	Softball
11915	Wash windows	17543	Volleyball
11920	Clean house, light	17550	Water sports
11921	Dust	17551	Surfing
11922	Use aerosol cleaner/freshener	17552	Swimming
12000	Child care, general	17560	Winter sports
12100	Care of baby	17600	Play games, general
12200	Care of child	17610	Active games
12300	Help / teach	17611	Fighting
12400	Talk /read	17612	Running games
12500	Play indoors	17613	Trampoline
12600	Play outdoors	17620	Board games/card games
12700	Medical care-child	17621	Board games
12800	Other child care	17622	Bingo
13000	Obtain goods and services, general	17623	Card games
13100	Dry clean	17630	Dress-up/make believe
13200	Shop/run errands, general	17640	Low energy games
13201	Errands for children or pets	17641	Arts and crafts
13202	Shopping, general	17642	Play with books
13203	Shop at mall or superstore	17643	Writing/drawing
13210	Shop for food, general	17650	Outdoor play
13211	Grocery shopping	17651	Playground/swings
13212	Shop for meals/snacks	17660	Play with animals

<b>Code</b>	<b>Activity Description</b>	<b>Code</b>	<b>Activity Description</b>
13220	Shop for clothes or household goods	17670	Puzzles
13230	Run errands	17671	Jigsaw puzzle
13300	Obtain personal care service	17672	Word puzzle
13400	Obtain medical service	17680	Toys
13500	Obtain government / financial services	17681	Toy balls
13600	Obtain car services	17690	Video games
13700	Other repairs	17691	Active video games
13800	Other services	17692	Computer games
14000	Personal needs and care, general	17700	Active leisure, general
14001	Wake up	17701	Camping
14100	Shower, bathe, personal hygiene	17702	Caving/rock climbing
14110	Shower, bathe, general	17703	Climb trees/structures
14111	Bathe	17710	Dance
14112	Shower	17720	Hiking
14120	Personal hygiene, general	17730	Horseback riding
14121	Use restroom	17740	Water recreation
14122	Wash hands/teeth/face	17741	Boating
14200	Medical care, general	17742	Recreational swim
14201	Use nebulizer/oxygen machine	17743	Scuba diving
14210	Feel sick	17800	Exercise, general
14300	Help and care	17810	Cardiovascular exercise
14400	Eat	17811	Aerobics
14500	Sleep or nap	17812	Bike for exercise
14600	Dress, groom	17813	Run or jog for exercise
14700	Other personal needs	17814	Swim for exercise
15000	General education and professional training	17815	Walk for exercise
15100	Attend full-time school	17820	Strength/stretching
15110	Attend day-care	17821	Lift weights
15120	Attend K-12	17822	Physical therapy
15130	Attend college or trade school	17823	Stretching
15140	Attend adult education and special training	18000	Travel, general
15200	Attend other classes	18010	Travel by bus, general
15300	Do homework	18020	Travel by foot, general
15400	Use library	18030	Travel by motor vehicle, general
15500	Other education	18031	Drive a motor vehicle, general
16000	Social activities, general	18032	Ride in a motor vehicle, general
16001	Gamble	18040	Wait, general
16002	Go to park or festival	18100	Travel during work, general
16100	Attend sports events	18110	Travel during work by bus
16200	Participate in social, political, or religious activities	18120	Travel during work by foot
16210	Practice religion	18130	Travel during work by motor vehicle
16300	Watch movie	18131	Travel during work, drive a motor vehicle

Code	Activity Description	Code	Activity Description
16400	Attend theater	18132	Travel during work, ride in a motor vehicle
16500	Visit museums	18140	Travel during work, wait
16600	Visit	18200	Travel to/from work, general
16700	Attend a party, general	18210	Travel to/from work by bus
16701	Attend a party, dance	18220	Travel to/from work by foot
16702	Attend a party, eat/drink	18230	Travel to/from work by motor vehicle
16703	Attend a party, sit/stand	18231	Travel to/from work, drive a motor vehicle
16704	Attend a party, talk	18232	Travel to/from work, ride in a motor vehicle
16705	Attend a party, walk	18240	Travel to/from work, wait
16800	Go to bar / lounge	18300	Travel for education, general
16900	Other entertainment / social events	18310	Travel for education by bus
17000	Leisure, general	18320	Travel for education by foot
17010	Indoor leisure	18330	Travel for education by motor vehicle
17111	Hunting, fishing, hiking	18331	Travel for education, drive a motor vehicle
17112	Golf	18332	Travel for education, ride in a motor vehicle
17113	Bowling / pool / ping pong / pinball	18340	Travel for education, wait
17114	Yoga	U	Uncertain
17120	Participate in outdoor leisure	X	Missing

Each of the CHAD codes is mapped to an internal APEX distribution number; activities that have identical energy expenditure associated with them map to the same distribution. The distributions themselves are defined by number in the *MET Distribution* File (Section 4.14). Each line of the *MET Mapping* file contains the variables described below.

- **Activity** code: This maps the CHAD activity to the internal APEX distribution number.
- **Age** category: Some MET distributions differ for persons of different ages. This variable maps the age groups to the correct distribution number. The age category given in this file is a label representing the age group. APEX will assign distributions as shown below.
  - **Age** is “0”: APEX will use for persons of all ages
  - **Age** is “20”: APEX will use for persons age 0 to 25
  - **Age** is “30”: APEX will use for persons age 26 to 39
  - **Age** is “40”: APEX will use for persons age 40 and older
- **Occ**: The MET distributions for the “Work” CHAD activity differ based on the occupation of the profile. This variable maps the different occupations to the correct distribution number. If the user has defined their own occupational groups for each profile, then profile occupations will be used instead of diary occupations.
- **APEX Dist #**: This is an internal index used by APEX to access the distribution. These values range from 1 to a maximum of 256.
- **Notes**: Description of the activity being modeled by the MET distribution. This is for the convenience of the user and is not used internally by APEX.

An example portion of the *MET Mapping* file is given in Exhibit 4-16.

```

! CHAD MET Distribution Mapping file for input to APEX
! Created 9-28-2015 by ICF International, with updates to many activity codes and !
distributions
Activity Age Occ. APEX Dist # Notes
10000 0 ADMIN 7 Work and other income producing activities, general
10000 0 ADMSUP 7 Work and other income producing activities, general
10000 0 FARM 23 Work and other income producing activities, general
10000 0 HSHLD 13 Work and other income producing activities, general
10000 0 LABOR 85 Work and other income producing activities, general
10000 0 MACH 176 Work and other income producing activities, general
10000 0 PREC 82 Work and other income producing activities, general
10000 0 PROF 67 Work and other income producing activities, general
10000 0 PROTECT 67 Work and other income producing activities, general
10000 0 SALE 67 Work and other income producing activities, general
10000 0 SERV 75 Work and other income producing activities, general
10000 0 TECH 81 Work and other income producing activities, general
10000 0 TRANS 11 Work and other income producing activities, general
10000 0 X 66 Work and other income producing activities, general
10100 0 X 66 Work, general
10110 0 X 66 Work, general, for organizational activities
10111 0 X 66 Work for professional/union organizations
10112 0 X 66 Work for special interest identity organizations
10113 0 X 66 Work for political party and civic participation
10114 0 X 66 Work for volunteer/ helping organizations
10115 0 X 66 Work of/for religious groups
10116 0 X 66 Work for fraternal organizations
10117 0 X 66 Work for child / youth / family organizations
10118 0 X 66 Work for other organizations
10120 0 X 66 Work, income-related only
10130 0 X 66 Work, secondary (income-related)
10200 0 X 66 Unemployment
10300 0 Any 93 Breaks
11000 0 Any 14 Household activities, general
11001 0 Any 1 Other household
11100 0 Any 125 Prepare food, general
11101 0 Any 52 Washing
11110 0 Any 4 Prepare and clean-up food
11120 0 Any 125 Cooking
...

```

**Exhibit 4-16. First Part of the MET Mapping File**

The user should not change the *MET Mapping* file unless the user has developed their own activity codes, or if they have defined their own occupation groups. User-defined occupation groupings must match the occupations listed for activity code 10000.

#### **4.14 MET Distribution File**

The *MET Distribution* file provides the actual distributions for calculating the MET value for each diary event (activity). The distributions are defined by APEX distribution numbers as given in the *MET Mapping* file. A MET value is a dimensionless ratio of the activity-dependent energy expenditure rate to the basal or resting energy expenditure (metabolic) rate, and the CHAD activity code is an identifier associated with each diary event that indicates the type of activity being performed. In general, the user should not change the distributions in this file, as these data were developed from extensive experimental data on human energy expenditures.



The distribution definitions use the standard APEX distribution format (i.e., a distribution shape, followed by 4 distribution parameters, upper and lower truncation bounds, and a resampling flag). The 4 parameters used are dependent on the shape of the distribution. Each data line in this file provides the information shown below in list format.

- **Dist#**: This is an internal index used by APEX to access the distribution. These values range from 1 to a maximum of 256. This matches the distribution numbers used in the *MET Mapping* file.
- **Shape**: This variable gives the type of the MET distribution.
- **Par1**: Parameter 1 of the MET distribution.
- **Par2**: Parameter 2 of the MET distribution.
- **Par3**: Parameter 3 of the MET distribution.
- **Par4**: Parameter 4 of the MET distribution.
- **LTrunc**: Lower truncation point of the MET distribution.
- **UTrunc**: Upper truncation point of the MET distribution.
- **ResampOut**: Distribution resampling flag.
- **General Use**: Text description of the general use of the particular distribution in APEX. (Optional—not used by the model code).

*Volume II* provides complete details for defining probability distributions in APEX; a summary of the available distributions is presented in Table 4-6.

**ResampOut** is a Yes/No variable. If it is set to “Y,” then generated random values outside the truncation limits are effectively thrown out and replaced by new samples, repeated until falling within the bounds. If **ResampOut** = N, such values are shifted to the appropriate truncation point. This may result in a “spike” of probability at **LTrunc** and/or **UTrunc**. Both methods ensure that all returned values equal or exceed **LTrunc** and are less than or equal to **UTrunc**. Both methods, however, may alter the statistical properties of the returned values. For example, truncation will always reduce the variance and may in some cases alter the mean as well. The default is **ResampOut** = Yes.

**Table 4-6. Available Probability Distributions in APEX**

Distribution	Shape	Par1	Par2	Par3	Par4	LTrunc (optional)	UTrunc (optional)	ResampOut (optional)
Beta	<i>Beta</i>	Minimum	Maximum	Shape1 (s1) > 0	Shape2 (s2) > 0	Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Burr	<i>Burr</i>	Scale(b) > 0	Shape1(s1) > 0	Shape2 (s2) > 0	Shift(a)	Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Cauchy	<i>Cauchy</i>	Median	Scale (b) > 0			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Discrete	<i>Discrete</i>	This type of distribution has no parameters, rather the keyword is simply followed by a list of up to 100 specific values. One of these values is selected at random, with equal probability for each. Duplicate values are acceptable.						
Exponential	<i>Exponential</i>	Decay constant, k > 0	Shift (a)			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Extreme Value	<i>Evalue</i>	Scale (b) > 0	Shift (a)			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)

Distribution	Shape	Par1	Par2	Par3	Par4	LTrunc (optional)	UTrunc (optional)	ResampOut (optional)
Gamma	<b>Gamma</b>	Shape (s) > 0	Scale (b) > 0	Shift (a)		Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Logistic	<b>Lgt</b>	Mean	Scale (b) > 0			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Lognormal	<b>Lognormal</b>	Geometric mean (gm) of unshifted distribution	Geometric standard deviation (gsd) > 1	Shift (a)		Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Loguniform	<b>LUniform</b>	Minimum > 0	Maximum > 0			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Normal	<b>Normal</b>	Mean	Standard deviation			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
OffOn	<b>OffOn</b>	Probability of being 0 (0-1)						
Pareto	<b>Pareto</b>	Shape (s) > 0	Scale (b) > 0	Shift (a)		Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Point	<b>Point</b>	Point Value						
Triangle	<b>Triangle</b>	Minimum	Maximum	Peak		Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Uniform	<b>Uniform</b>	Minimum	Maximum			Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)
Weibull	<b>Weibull</b>	Shape (s) > 0	Scale (b) > 0	Shift		Lower truncation limit	Upper truncation limit	Resample outside truncation? (Y/N)

Periods (“.”) must be used as placeholders in the file if a parameter is not needed for a particular distribution (except for the *Discrete* distribution).

See *Volume II* for further information about the use of MET probability distributions in APEX. A portion of this file is shown in Exhibit 4-17. For brevity in this exhibit, the *General Use* descriptions of the corresponding activities have been truncated with ellipses.

```

! APEX MET Distribution File
! Created 9-28-2015 by ICF International, with updates to many activity codes and
! distributions
! For use with CHAD mapping file
!Dist# Shape      Par1 Par2 Par3 Par4 LTrunc UTrunc ResampOut General Use
1      Exponential 0.3  3    .    .    1     9     Y     Other household
2      Exponential 0.4  2.6  .    .    1     6     Y     Clean outdoors
3      Exponential 0.7  3.5  .    .    1     6     Y     Paint home/room
4      Exponential 1.1  1.9  .    .    1     4     Y     Prepare and clean-up...
5      LogNormal   0.9  1.1  0    .    0.8   1.1   Y     Sleep or nap
6      LogNormal   1.2  1.4  0    .    0.9   2.3   Y     Listen to radio/...
7      LogNormal   1.7  1.2  0    .    1.4   2.7   Y     Work and other...
8      LogNormal   1.8  1.4  0    .    1.4   4     Y     General education...
9      LogNormal   2.6  1.7  0    .    1.1   6     Y     Play games, general
10     LogNormal   3     1.2  0    .    2.5   5     Y     Help and care
11     LogNormal   3     1.5  0    .    1.3   8.4   Y     Work and other...
12     LogNormal   3     1.5  0    .    1.5   8     Y     Go to bar/lounge
13     LogNormal   3.5  1.2  0    .    2.5   6     Y     Work and other...
14     LogNormal   3.6  1.5  0    .    1.8   7.3   Y     Household activities...
15     LogNormal   3.9  1.4  0    .    2     9     Y     Participate in...
16     LogNormal   4     1.6  0    .    1.9   8.4   Y     Walk, general/Crawl...
...

```

**Exhibit 4-17. First Part of the Activity-specific MET File**

## 4.15 Physiological Parameters File

The *Physiological Parameters* file provides age- and gender-specific distributions for a number of physiological parameters (see Exhibit 4-18). The parameters are listed in Table 4-7. See *Volume II* for details of these parameters and the equations in which they are used in APEX.

The variables in Table 4-7 that have RMR = 1 in their descriptions are required when setting **RMRMETHOD** = 1 on the *Control Options* file. The ones with RMR = 2 are required when using **RMRMETHOD** = 2, which is the default method. Similarly, the variables with HtWt = 1 are required when **HTWTMETHOD** = 1 is specified on the *Control Options* file, and those with HtWt = 2 are required when using the default method (which is **HTWTMETHOD** = 2).

**Table 4-7. Parameters in the Physiological Input File**

Variable	Description	Units
<b>NVO2Max</b>	Normalized maximum oxygen uptake	ml-O <sub>2</sub> /(min·kg) (Note: while the APEX inputs for <b>NVO2Max</b> are in ml-O <sub>2</sub> /(min·kg), APEX outputs VO2Max in the <i>Profile Summary</i> file in L-O <sub>2</sub> /min)
<b>Bm</b>	Body mass (HtWt=1)	kg
<b>LogBM</b>	Natural logarithm of body mass (kg) (HtWt=2)	-
<b>Height</b>	Height (HtWt=2) (Note: units different from other HtWt method)	cm
<b>HtWtCorr</b>	Pearson correlation between height and logBM (HtWt=2)	-
<b>RMRInt</b>	Intercept of resting metabolic rate regression (both RMR=1 and RMR2, but with different values)	MJ/day (Note: while the APEX inputs for RMR are in MJ/day, APEX outputs RMR in the <i>Profile Summary</i> file in kcal/min).
<b>RMRSp</b>	Slope of resting metabolic rate regression (RMR=1)	MJ/(day·kg)

<b>Variable</b>	<b>Description</b>	<b>Units</b>
<i>RMRErr</i>	Standard deviation for resting metabolic rate regression (both RMR=1 and RMR=2)	MJ/day
<i>RMR BM</i>	RMR regression coefficient for BM (RMR=2)	-
<i>RMR_LBM</i>	RMR regression coefficient for Log(1+BM) (RMR=2)	-
<i>RMR Age</i>	RMR regression coefficient for Age (RMR=2)	-
<i>RMR_LAge</i>	RMR regression coefficient for Log(1+Age) (RMR=2)	-
<i>RMR Ht</i>	RMR regression coefficient for Height (RMR=2)	-
<i>RMR_LHt</i>	RMR regression coefficient for Log(1+Height) (RMR=2)	-
<i>Hmg</i>	Blood hemoglobin density	g/dl
<i>BSAExp1</i>	Exponent 1 for calculating body surface area	-
<i>BSAExp2</i>	Exponent 2 for calculating body surface area	-
<i>MaxOxD</i>	Maximum oxygen deficit	ml/kg
<i>BldFac1</i>	Blood volume factor 1	ml/lb
<i>BldFac2</i>	Blood volume factor 2	ml/inches <sup>3</sup>
<i>HeightInt</i>	Intercept of height regression (HtWt=1)	inches
<i>HeightSlp</i>	Slope of height regression (HtWt=1)	children under 18: inches/(year of age) adults: inches/lb (lbs body weight)
<i>HeightErr</i>	Standard deviation of height regression (HtWt=1)	inches
<i>ECF</i>	Energy conversion factor	L-O <sub>2</sub> /kcal
<i>RecTime</i>	Time required to recover maximum oxygen deficit	hours
<i>Endgn1</i>	Endogenous CO production rate 1	ml/min
<i>Endgn2</i>	Endogenous CO production rate 2 (used for women in 2 <sup>nd</sup> half of menstrual cycle)	ml/min
<i>B1-B9</i>	Model parameters for calculation of % $\Delta$ FEV1 (note:B7 not currently used)	-
<i>FEVU</i>	Uncertainty term used in calculation of % $\Delta$ FEV1	-
<i>FEVE1</i>	1 <sup>st</sup> error term used in calculation of % $\Delta$ FEV1	-
<i>FEVE2</i>	2 <sup>nd</sup> error term used in calculation of % $\Delta$ FEV1	-
<i>FEVBMI</i>	Cut point for body mass index in regression	-
<i>FEVSlp</i>	Slope of age fit in % $\Delta$ FEV1 calculation	-
<i>FEVInt</i>	y-intercept of age fit in % $\Delta$ FEV1 calculation	-
<i>ModEVR</i>	Distribution of moderate 8-hour average exertion cutoff	L/(min-m <sup>2</sup> )

Distributions for the above parameters are assigned to persons of every age and gender combination in the *Physiology* file. The distributions are defined in the APEX distribution format (i.e., a distribution shape, followed by 4 distribution parameters, upper and lower truncation bounds, and a resampling flag—see *Volume II*). Thus, each data line contains the information listed below.

- **Variable** keyword
- **AgeMin** (minimum age) for the current parameter distribution definition
- **AgeMax** (maximum age) for the current parameter distribution definition
- **Gender** for the current parameter distribution
- **Shape** (type of distribution)
- **Par1** (parameter 1) of the distribution, which depends on shape

- **Par2** (parameter 2) of the distribution, which depends on shape
- **Par3** (parameter 3) of the distribution, which depends on shape
- **Par4** (parameter 4) of the distribution, which depends on shape
- **LTrunc** (lower truncation point of the distribution)
- **UTrunc** (upper truncation point of the distribution)
- **ResampOut** (distribution resampling flag)

Thus, each line of the *Physiology* file can define the distribution for a range of ages, but only a single gender. The physiological parameters must be defined for both genders for all ages 0–100 years, with the exception of **ENGN2**, which need only be defined for females. An APEX fatal error will result if not all data are provided. In general, the distributions in this file should not be changed from their default values, as they were derived from available physiological data.

See Table 4-6 for the available distribution shapes and required parameters. Periods (“.”) must be used as placeholders if a parameter is not needed for a particular distribution.

```

! APEX Physiology Data Input File
! Reference: Isaacs, K. and L. Smith, Alion Science and Technology.
! "New Values for Physiological Parameters for the Exposure Model Input File Physiology.txt,"
! December 20, 2005. Memorandum to T. McCurdy, US EPA
! Updated 09/29/2010 (Data reorganized, some added; existing data unchanged from 05/04/06
! version)
! Updated 02/04/2011 with dFEV1 model parameters from McDonnell, Stewart & Smith, 2010
! "Predictions of ozone-induced lung function..."
! Updated 4/11/11 to correct the Slope for RMR regression for males ages 0-2 to 0.249 from
! 0.244, per Table 1 in Layton (1993).
! Updated 8/22/12 with the B9 threshold parameter for dFEV1, MSS model parameters from MSS 2012
! Table 2 Threshold model
! Updated 12/31/12 with age term extensions for the MSS 2012 model
! Updated 4/8/14 to add the FEVCUT parameter for APEX v463 (note that the definition of FEVE2
! changes with v463, per the McDonnell et al. 2013 model)
! Height and weight data updated on 11/15/2016 to include correlation, using data from Jonathan
! Cohen at ICF
!
!Variable AgeMin AgeMax Gen Shape Par1 Par2 Par3 Par4 LTrunc UTrunc ResampOut
!NVO2Max
NVO2MAX 0 0 M Normal 48.3 1.7 . . 44.3 52.2 N
NVO2MAX 1 1 M Normal 48.6 2.0 . . 43.8 53.3 N
NVO2MAX 2 2 M Normal 48.9 2.4 . . 43.4 54.4 N
NVO2MAX 3 3 M Normal 49.2 2.7 . . 43.0 55.4 N
NVO2MAX 4 4 M Normal 49.5 3.0 . . 42.5 56.5 N
NVO2MAX 5 5 M Normal 49.8 3.3 . . 42.1 57.6 N
NVO2MAX 6 6 M Normal 50.1 3.7 . . 41.6 58.6 N
NVO2MAX 7 7 M Normal 50.4 4.0 . . 41.2 59.7 N
NVO2MAX 8 8 M Normal 50.8 4.3 . . 40.8 60.8 N
NVO2MAX 9 9 M Normal 51.1 4.6 . . 40.3 61.8 N
...
! Body Mass
BM 0 0 M Lognormal 7.8 1.301 0 . 3.6 11.8 N
BM 1 1 M Lognormal 11.4 1.143 0 . 8.2 16.1 N
BM 2 2 M Lognormal 13.9 1.146 0 . 9.8 20.9 N
BM 3 3 M Lognormal 16.0 1.154 0 . 11.7 23.7 N
BM 4 4 M Lognormal 18.5 1.165 0 . 11.1 28.1 N
BM 5 5 M Lognormal 21.6 1.234 0 . 13.7 42.4 N
BM 6 6 M Lognormal 23.1 1.213 0 . 16.1 41.1 N
BM 7 7 M Lognormal 27.1 1.216 0 . 19.3 46.8 N
BM 8 8 M Lognormal 31.7 1.302 0 . 19.1 66.2 N
BM 9 9 M Lognormal 34.7 1.265 0 . 24.0 69.9 N
BM 10 10 M Lognormal 38.3 1.280 0 . 24.3 72.9 N
BM 11 11 M Lognormal 44.1 1.308 0 . 26.2 83.8 N
BM 12 12 M Lognormal 48.0 1.315 0 . 27.7 94.8 N
...
! Intercept for RMR regression for RMRMethod=1
RMRINT 0 2 M Point -0.127 . . . . .
RMRINT 3 9 M Point 2.110 . . . . .
RMRINT 10 17 M Point 2.754 . . . . .
RMRINT 18 29 M Point 2.896 . . . . .
...

```

**Exhibit 4-18. An Example of a Portion of the Physiological Parameters File**

**4.16 Ventilation File**

The *Ventilation* file contains a set of regression parameters used by APEX to estimate ventilation quantities from the event MET. This is a small file containing the parameters for each of four age groups (Exhibit 4-19). This file should not be edited except by advanced users who understand the APEX ventilation algorithm.

For more information on the ventilation algorithm and the derivation of the values in this file, see *Volume II* and Graham and McCurdy (2005).

! APEX Ventilation Data File												
! MinAge	MaxAge	b0	seb0	b1	seb1	b2	seb2	b3	seb3	eb	ew	R2
0	19	4.4329	0.0579	1.0864	0.0097	-0.2829	0.0124	0.0513	0.0045	0.0955	0.1117	0.925
20	33	3.5718	0.0792	1.1702	0.0067	0.1138	0.0243	0.045	0.0031	0.1217	0.1296	0.8927
34	60	3.1876	0.1271	1.1224	0.012	0.1762	0.0335	0.0415	0.0095	0.1260	0.1152	0.8922
61	100	2.4487	0.3646	1.0437	0.0195	0.2681	0.0834	-0.0298	0.0100	0.1064	0.0676	0.8932

**Exhibit 4-19. The APEX Ventilation Input File for VEMethod=1**

The newer *VEMETHOD* = 2 uses an altered format, shown in Exhibit 4-20. In this case, the same regression applies to everyone.

! Ventilation file - ventilation regression parameters for method 2							
! Reference: "Analysis of the VE/VO2 relationship in the Adams data set"							
! Graham Glen, ICF, September 2016							
! MinAge	MaxAge	Intercept	LOG (VO2)	(VO2/VO2max)**4	eb	ew	
0	100	3.300	0.8128	0.5126	0.09866	0.07852	

**Exhibit 4-20. The APEX Ventilation Input File for VEMethod=2**

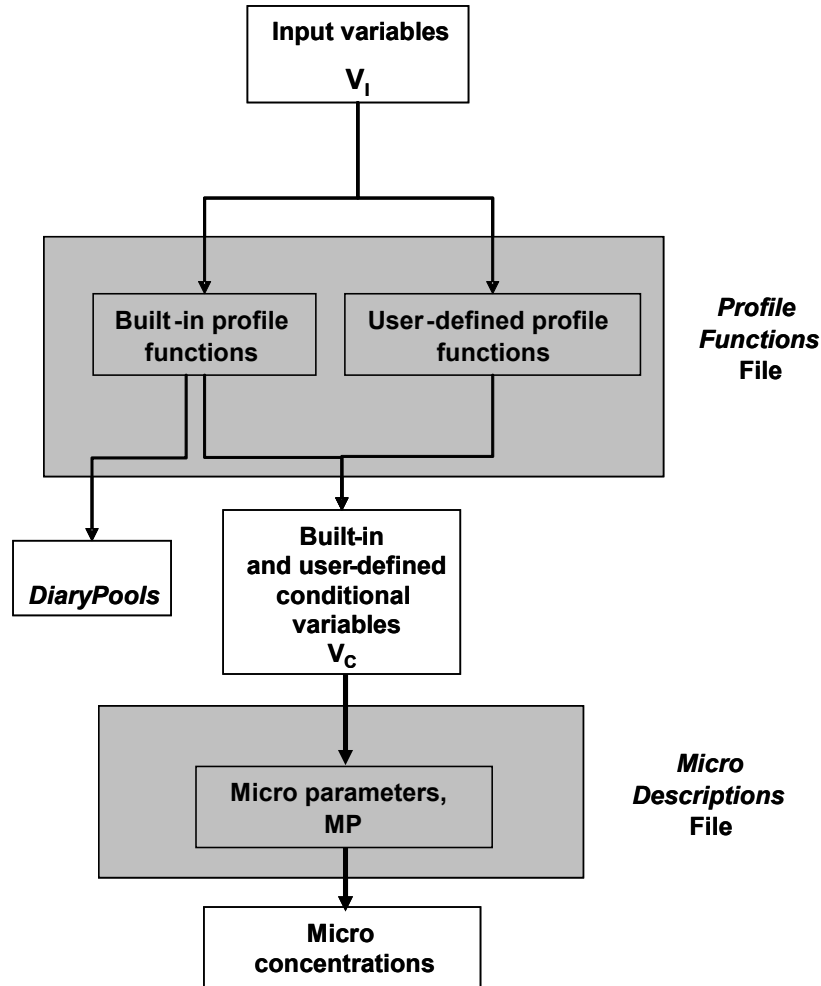
#### 4.17 Profile Functions (Distributions) File

The *Profile Functions* input file defines functions for variables associated with each simulated profile. There are four types of functions that can be defined, as indicated below.

- Functions for built-in APEX variables: These are variables that are predefined in APEX, and whose values under different circumstances can be customized by the functions defined in this file. Most of these variables are also “conditional variables” since microenvironmental parameters can depend on their values.
- Functions for creating user-defined APEX conditional variables: These are generic variables that the user may define and then use in calculating microenvironmental parameters. The names of these variables have no set intrinsic meaning in APEX; they can be used to represent whatever the user wishes. Up to eight of these variables may be defined in APEX.
- Functions for creating user-defined APEX conditional variables that vary by region: These are generic variables that the user can define and then use for calculating microenvironmental parameters. These conditional variables can vary by region (either county or sector) and thus may be evaluated differently for individuals who reside in these different regions. Up to five of these functions may be defined.
- Functions for creating user-defined APEX conditional variables that vary with ambient air quality: These are also generic variables that the user can define and then use for calculating microenvironmental parameters. These variables vary by ambient air quality, and thus are recalculated during each timestep. Up to five of these functions may be defined.

The relationships among the different functions that can be defined in the *Profile Functions* file and the microenvironmental descriptions are shown in Figure 4-1. The built-in and user-defined functions are used to define a set of conditional variables,  $V_C$ , which are functions of input

APEX variables ( $V_I$ ). These conditional variables are used in determining microenvironmental parameters. Microenvironment parameters are quantities that appear in the equations for the microenvironmental concentrations. The relationship between the conditional variables and the microenvironment parameters are described in the *Microenvironmental Descriptions* file (see Section 4.24).



**Figure 4-1. Relationship between Profile Functions and Microenvironmental Descriptions Files**

#### 4.17.1 Defining a Profile Function

The general procedure for defining a profile function is as shown below.

1. A function definition begins with its name on the first input line.
2. The user may add as many comment lines as necessary to describe the profile function or units of the involved parameters.
3. If the function is of type regional (*RegionalConditional1* through *RegionalConditional5*), then a statement is required to define how the regions are defined, either by county or sector, and how many different regions are being modeled.



4. The number of subsequent input lines varies with the number of input variables required to define the function. At least one (and usually two) input lines are needed for each input variable of the function. In addition, at least two lines are also needed for the function result. For each input variable (table dimension), the first line starts with the keyword ***Input***, followed by the indexing number of the variable in the function, the ***Type of Input Variable***, and the ***Number of Values (Nvals)*** allowed for the input variable. At the end of this input line, the user may add comments in double quotes to explain input variables. The lines directly following define the input variable data—specifically, they define how the input variable is grouped into integer categories for indexing the table of results. The ***Type Of Input Variable*** must be one of the following:

- ***Probability***,
- ***Realrange***,
- ***Intrange***,
- ***Intvalue***,
- ***Intindex***
- ***Conditional***, or
- ***Regionindex***

***Probability*** means the result is randomly determined, using fixed probabilities for each outcome. The input variable data for ***Probability*** is a list of the Nvals fixed probabilities. The sum of the probabilities must equal 1. ***RealRange*** represents a set of discrete categories, each consisting of a range of real numbers. In this case, the categories are defined by (Nvals-1) cut-points. If the input variable falls exactly on a cut point, it falls into the higher bin. ***IntRange*** is similar, except that each category consists of a range of integers. ***IntValue*** denotes that each possible value that the input variable may take on is listed on the data line. ***IntIndex*** signifies that the input variable is an integer and is to be used to index the table of results directly (e.g., a value of 3 means that it uses the third cell of a table dimension). Thus, this type of input variable does not require a second line. ***Conditional*** refers to conditional probabilities that depend on the values of other input variables. Only one ***Conditional*** input variable is allowed in a function, and it comes last in the function specification. A table of probabilities follows. The number of entries in the probability table must be equal to the product of the number of category combinations for the other inputs and the number of possible function results. ***RegionIndex*** is an option for the ***RegionalConditional*** only. This option uses two columns of input data—a region and an index—and maps regions to the index to be used directly in the *Microenvironmental Results* file. The index must be between 1 and the total number of regions used.

The examples provided in the sections that follow illustrate the appropriate use of these input variable types. Exhibit 4-21 contains examples of ***Probability***, ***IntRange***, ***IntValue***, and ***Conditional***. An example of ***IntIndex*** occurs in the *DiaryPools* definition just below Exhibit 4-21. ***RealRange*** works exactly like ***IntRange***, except the cut-points are real numbers (that is, containing decimal points) instead of integers.

5. After all the input variables, except for ***RegionIndex***, are specified, the next line must contain the keyword ***Result***, followed by a type (***Integer***, ***Real***, or ***Histogram***) and the

- number of possible results (Nresults). **RegionIndex** variables must be an integer.
6. The table results are then listed in order in subsequent lines. If the result type is designated as **Integer**, the results must be a list of integers of length Nresults. If the type is **Real**, then the list of results must contain Nresults real numbers. If the result type is **Histogram**, the results are a series of (Nresults+1) cut points that define Nresults bins. If there are multiple inputs, the indexing follows the Fortran convention of increasing the leftmost (first defined) variable over its range, then the second variable, and so on.
  7. The profile function ends with a new line that has a # sign.

The types of profile functions are discussed in detail below along with examples. Note that when preparing or editing a profile functions file, be careful not to use Tab to separate the items on a line. APEX explicitly searches for blanks (spaces) as delimiters, and does not recognize Tabs as such.

#### 4.17.2 Functions for Built-in, User-defined, AQ and Regional APEX Variables

The built-in APEX variables for which functions can be assigned are given in Table 4-8. All of these variables are conditional variables which can be used to define microenvironment parameters, with the exception of the variable **DiaryPools**. Note that a few other APEX variables, such as gender, can be used as conditional variables (see Section 4.24.2). **DiaryPools** is the only function that APEX requires to be defined, as it is used in the selection of appropriate CHAD diaries for different days in the simulation. The input variables required for each of these functions are hard coded; the required inputs for each variable are listed in the table. Some conditional variables in this file must be defined if others that require them are defined. For example, **AC\_Home** must be present if **WindowRes** is defined.

Several generic user-defined conditional variables are listed in the table as well. These functions take a single input variable, which must be defined by fixed probabilities for each of the function results (categories). The four types are distinguished by the sampling properties of their input variable: **DailyConditional** requires a variable that is sampled daily; **ProfileConditional** requires a variables assigned once per profile, such as age or gender; **RegionalConditional** requires a region-dependent variable, where regions are defined as groups of home sectors; and finally, **AQConditional** variables depend on the ambient air quality (and so can change their status over time, although not necessarily daily).

Each of the functions in the table returns an integer category for each combination of input parameters. For the conditional variables, these category numbers can be used in defining the microenvironment parameters in the *Microenvironment Descriptions* file (see Section 4.24.2). Note that one conditional based on the *Profile Factors* file, **FactorGroup**, can be used in the *Microenvironment Descriptions* file, but is defined in its own file, rather than in the *Profile Functions* file.

Several examples are shown in Exhibit 4-21. The first is the definition for a function for **AvgTempCat**. It returns an integer category number for the average temperature, which will be used in the definition of one or more microenvironment parameters. The first and only input variable defines the integer ranges (via **IntRange**) for the three categories of average temperature. In this case, the ranges are < 50 degrees, 50–77 degrees, and ≥ 78 degrees. The

function reads the daily average temperature and determines which category it falls in. The resulting categories are 1, 2, and 3. If the average temperature were 69 degrees, then the *AvgTempCat* function would return a value of 2.

The second example is a definition for *WindowRes*. The first input variable is *AC\_Home*, and the categories for it are defined by its two possible integer values (via *IntValue*), as 1 (has AC) or 2 (does not have AC). The second input variable is the maximum daily temperature; the categories for it are defined via *IntRange* in a manner similar to that demonstrated in the first example. The third input variable for this function as indicated in Table 4-8 is the average daily temperature. In this case, *AvgTemp* is not used to determine *WindowRes*, and leaving it out of the function definition means that by default means that all possible values for *AvgTemp* are assigned to the first bin. The fourth and final input variable is a *Conditional* probability. Each row lists the probabilities for each possible output, for one combination of inputs. The first probability in each row is the chance of YES (which equals 1 in APEX), while the second probability is the chance of NO (which equals 2). Some functions might have more than two possible outcomes, so there could be more than two probabilities on each line. Thus, the first row of the table contains probabilities for *WindowRes* = 1 (YES/OPEN) and *WindowRes* = 2 (NO/CLOSED) for *AC\_Home* = 1 (YES), *MaxTemp*<56, and any *AvgTemp* value. The second line has the same probabilities for *AC\_Home* = 2 (NO), *MaxTemp*<56, and any *AvgTemp* value. The last line has the probabilities for *AC\_Home* = 2 (NO), *MaxTemp*>78, and any *AvgTemp* value. In this example, when it is hot outside, the window status is just 10% likely to be open when *AC\_Home* = YES, but 90% likely when *AC\_Home* = NO.

**Table 4-8. Variables that can be Defined in the Profile Functions File**

<b>Conditional Variable</b>	<b>Purpose</b>	<b>Input Variables</b>	<b>Number of Categories</b>	<b>Function Reevaluated</b>
<i>MaxTempCat</i>	Binning hourly maximum temperatures into categories	INPUT1: Maximum temperature on hour of simulation	any number	hourly
<i>AvgTempCat</i>	Binning hourly average temperatures into categories	INPUT1: Average temperature on hour of simulation	any number	hourly
<i>HumidCat</i>	Binning hourly humidities into categories	INPUT1: Humidity on hour of simulation	any number	hourly
<i>WindCat</i>	Binning hourly wind speeds into categories	INPUT1: Wind speed on hour of simulation	any number	hourly
<i>DirCat</i>	Binning hourly wind directions into categories	INPUT1: Wind direction on hour of simulation	any number	hourly
<i>PrecipCat</i>	Assigning precipitation codes to categories	INPUT1: Precipitation code on hour of simulation	any number (equal to or less than the number of precipitation codes in the <i>Meteorology Data</i> file)	hourly
<i>MaxTempCat</i>	Binning daily maximum temperatures into categories	INPUT1: Temperature on hour of simulation	any number	daily
<i>AvgTempCat</i>	Binning daily average temperatures into categories	INPUT1: 24-hour average temperature on day of simulation ( <i>AvgTemp</i> )	any number	daily
<i>DiaryPools</i> (required)	Assigning diary pools	INPUT1: Maximum temperature on simulated day ( <i>MaxTemp</i> ) INPUT2: Average temperature on simulated day ( <i>AvgTemp</i> ) INPUT3: Day of the week	any number	daily
<i>PoolTrans</i>	Grouping of diary pools for empirical cluster transition probabilities	INPUT1: Pool number on “to” day INPUT2: Pool number on “from” day	(#pools) <sup>2</sup>	once per run

<b>Conditional Variable</b>	<b>Purpose</b>	<b>Input Variables</b>	<b>Number of Categories</b>	<b>Function Reevaluated</b>
<i>HasGasStove</i>	Probability of having a gas stove	INPUT1: Probabilities for the 2 results	2 (Y/N)	once per profile
<i>HasGasPilot</i>	Probability of having a pilot light, conditional on <i>Hasgasstove</i>	INPUT1: Has Gas Stove (Y/N)? ( <i>HasGasStove</i> ) INPUT2: Conditional probabilities for the result categories for both <i>HasGasStove</i> = Y and <i>HasGasStove</i> = N	2 (Y/N)	once per profile
<i>AC_Home</i>	Probability of having different types of home air conditioning or ventilation	INPUT1: Fixed probabilities for the types of air conditioning/ventilation (the number of types is user-defined)	any number	once per profile
<i>AC_Car</i>	Probability of having A/C in car	INPUT1: Probabilities for the 2 results	2 (Y/N)	once per profile
<i>WindowRes</i>	Probability of residence windows being open or closed, conditional on <i>AC_Home</i> , <i>MaxTempCat</i> , and <i>AvgTempCat</i>	INPUT1: Type of home A/C ( <i>AC_Home</i> ) INPUT2: Max. temperature on day of simulation ( <i>MaxTemp</i> ) INPUT3: Average temperature on day of simulation ( <i>Avgtemp</i> ) INPUT4: Conditional probabilities for the result categories for every combination of input1-input3 categories	2 (Y/N)	daily
<i>WindowCar</i>	Probability of car windows being open or closed, conditional on <i>AC_Car</i> , <i>MaxTempCat</i> , and <i>AvgTempCat</i>	INPUT1: Has car A/C ( <i>AC_Car</i> ) INPUT2: Max. temperature on day of simulation ( <i>MaxTemp</i> ) INPUT3: Average temperature on day of simulation ( <i>AvgTemp</i> ) INPUT4: Conditional probabilities for the result categories for every combination of input1-input3 categories	2 (Y/N)	daily
<i>SpeedCat</i>	Probability of average speed categories for vehicles	INPUT1: Fixed probabilities for the result categories	any number	daily
<i>DailyConditional1</i>	Generic daily conditional variable #1	INPUT1: Fixed probabilities for the result categories	any number	daily
<i>DailyConditional2</i>	Generic daily conditional variable #2	INPUT1: Fixed probabilities for the result categories	any number	daily

<b>Conditional Variable</b>	<b>Purpose</b>	<b>Input Variables</b>	<b>Number of Categories</b>	<b>Function Reevaluated</b>
<i>DailyConditional3</i>	Generic daily conditional variable #3	INPUT1: Fixed probabilities for the result categories	any number	daily
<i>ProfileConditional1</i>	Generic profile conditional variable #1	INPUT1: Fixed probabilities for the result categories	any number	once per profile
<i>ProfileConditional2</i>	Generic profile conditional variable #2	INPUT1: Fixed probabilities for the result categories	any number	once per profile
<i>ProfileConditional3</i>	Generic profile conditional variable #3	INPUT1: Fixed probabilities for the result categories	any number	once per profile
<i>ProfileConditional4</i>	Generic profile conditional variable #4	INPUT1: Fixed probabilities for the result categories	any number	once per profile
<i>ProfileConditional5</i>	Generic profile conditional variable #5	INPUT1: Fixed probabilities for the result categories	any number	once per profile
<i>RegionalConditional1</i>	Generic regional conditional variable #1	INPUT1: Fixed probabilities or region index for the result categories, defined for each region (sector or county) modeled	any number	once per profile, based on profile's home sector
<i>RegionalConditional2</i>	Generic regional conditional variable #2	INPUT1: Fixed probabilities or region index for the result categories, defined for each region (sector or county) modeled	any number	once per profile, based on profile's home sector
<i>RegionalConditional3</i>	Generic regional conditional variable #3	INPUT1: Fixed probabilities or region index for the result categories, defined for each region (sector or county) modeled	any number	once per profile, based on profile's home sector
<i>RegionalConditional4</i>	Generic regional conditional variable #4	INPUT1: Fixed probabilities or region index for the result categories, defined for each region (sector or county) modeled	any number	once per profile, based on profile's home sector
<i>RegionalConditional5</i>	Generic regional conditional variable #5	INPUT1: Fixed probabilities or region index for the result categories, defined for each region (sector or county) modeled	any number	once per profile, based on profile's home sector

<b>Conditional Variable</b>	<b>Purpose</b>	<b>Input Variables</b>	<b>Number of Categories</b>	<b>Function Reevaluated</b>
<i>AQConditional1</i>	Generic air quality conditional variable #1	INPUT1: Fixed probabilities for the result categories, defined for each air quality bin	any number	once per time step based on current location
<i>AQConditional2</i>	Generic air quality conditional variable #2	INPUT1: Fixed probabilities for the result categories, defined for each air quality bin	any number	once per time step based on current location
<i>AQConditional3</i>	Generic air quality conditional variable #3	INPUT1: Fixed probabilities for the result categories, defined for each air quality bin	any number	once per time step based on current location
<i>AQConditional4</i>	Generic air quality conditional variable #4	INPUT1: Fixed probabilities for the result categories, defined for each air quality bin	any number	once per time step based on current location
<i>AQConditional5</i>	Generic air quality conditional variable #5	INPUT1: Fixed probabilities for the result categories, defined for each air quality bin	any number	once per time step based on current location

```

AvgTempCat
! Temperature ranges (categories) in Fahrenheit
INPUT1 INTRANGE 3      "AvgTemp"
50 78
RESULT INTEGER 3      "TempCatA"
1 2 3
#
WindowRes
! Home windows open(1) or closed
INPUT1 INTVALUE 2     "AC_Home"
1 2
INPUT2 INTRANGE 3     "MaxTemp"
56 80
INPUT4 CONDITIONAL 12
0.2 0.8
0.2 0.8
0.5 0.5
0.7 0.3
0.1 0.9
0.9 0.1
RESULT INTEGER 2
1 2
#
DailyConditional3
! DailyConditional3 - Penetration values for vehicle micro
INPUT1 PROBABILITY 4
0.2 0.5 0.2 0.1
RESULT INTEGER 4
1 2 3 4
#
RegionalConditionall
! Has attached garage
BY Sector 14
INPUT1 PROBABILITY 2
01017953800 0.05 0.95
01017953900 0.05 0.95
01017954000 0.05 0.95
01017954200 0.05 0.95
01017954300 0.05 0.95
01017954400 0.05 0.95
01017954500 0.05 0.95
13013180101 0.8 0.2
13013180102 0.8 0.2
13013180201 0.8 0.2
13013180202 0.8 0.2
13013180300 0.8 0.2
13013180400 0.8 0.2
13013180500 0.8 0.2
RESULT INTEGER 2
1 2
#

```

#### Exhibit 4-21. Examples of Profile Functions

Note: There may be spaces between the word, "INPUT," and the following number.

The third example is a definition for a user-defined conditional variable *DailyConditional3*. In this case, the user wants four categories of a variable (penetration) for a given microenvironment



and assigns each category a probability of being selected on a given day. All user-defined conditional variables are designated in an analogous manner. User-defined conditional variables require a **Probability** input. Note that the probabilities for the four categories in the example sum to one. The resulting category number is saved to the profile on each day and can be used to determine the microenvironment parameters (Section 4.24.2).

The final example is a function definition for a regional conditional variable, **RegionalConditional1**. In this example, it is being used to describe the differences in housing conditions (i.e., presence of an attached garage) in different sectors of the study area. The BY statement indicates how the prevalence vary regionally, either by county or by sector (in this case, sector). This line must additionally contain the number of regions (i.e., counties or sectors) that will be used (in this case, 14). After this BY line, the probability input for each sector or county is listed. APEX matches these regions to the appropriate study area sector (or sectors, in the case of a county), and uses them when assigning the value of **RegionalConditional1** through **RegionalConditional5** to each profile. The “By county” option matches the first 5 characters of the sector name. An APEX warning will result if a listed region does not match up with any study area sector, and APEX will give an error message and stop if there exists a study area sector for which there is no corresponding region. The result of this function is that profiles in each sector will be assigned an attached garage (the RESULT,1 = yes, 2 = no) based on their sector’s listed probabilities.

The **DiaryPools** mapping is required. A typical example is the following:

```
DiaryPools
! Group activity diaries into pools
TABLE
INPUT1 INTRANGE 3          "MaxTemp"
55 84
INPUT2 INTRANGE 1          "AvgTemp"
INPUT3 INTINDEX 7         "DayOfWeek"
RESULT INTEGER 21         "Pool number"
4 5 6
1 2 3
1 2 3
1 2 3
1 2 3
1 2 3
1 2 3
4 5 6
#
```

In the above example, there are three bins for **MaxTemp**, one for **AvgTemp**, and seven for **DayOfWeek**. The 21 results could all be placed on one line, but this layout is clearer. The first row is Sunday, for low, medium, and high **MaxTemp**. If there were instead (say) two **AvgTemp** bins, then there would be six numbers for Sunday, the first three for the low **AvgTemp** bin and the last three for the second bin. Although there are 21 “results”, there are only six different diary pools. Later on, these pools will be subset by gender and age before choosing diaries for particular individuals.

The minimum number of categories for all the variables defined in the *Profile Functions* file is one, in which case all profiles will have the same value for the variable. In the case of all functions EXCEPT **DiaryPools**, having one category is the default case and can be implemented

by simply omitting the function definition from the *Profile Functions* file. **DiaryPools**, however, is required to be defined in the file. Therefore, if one wishes to define only a single diary pool, this must be done explicitly by setting all the RESULT values for the function equal to one. For example:

```
DiaryPools
! Group all activity diaries into one pool
TABLE
INPUT1 INTRANGE 1      "MaxTemp"
INPUT2 INTRANGE 1      "AvgTemp"
INPUT3 INTINDEX 7      "DayOfWeek"
RESULT INTEGER 7       "Pool number"
1 1 1 1 1 1 1
```

There is no explicit upper limit on the number of categories, and in practice, it is only limited by what is convenient. See Section 6.1 in *Volume II* for more information on diary pools.

## 4.18 Microenvironment Mapping File

The *Microenvironment Mapping* file provides the mapping of the **Location Codes** (e.g., for CHAD) to **Microenvironments** defined in APEX. The current CHAD location codes are given in Table 4-9, and an example portion of a *Microenvironment Mapping* file is provided in Exhibit 4-22. This file only allows comment lines and keyword input lines, except for the first two header lines. Each keyword input line begins with a location code followed by a short description; an “=,” an integer that designates a microenvironment defined in the *Microenvironment Descriptions* file; and a character variable that assigns the location code as belonging to a “Home,” “Work,” “Other,” “Road,” “Road Work,” “Near Work,” “Near Home,” “Last,” or “Unknown” location (the codes are H/W/O/R/RW/NW/NH/L/U, respectively). CHAPTER 8 in *Volume II* explains how APEX calculates air concentrations. The location determines the choice of ambient air data, and the microenvironment determines the rules for modifying that ambient data (see Section 4.24).

“Road” specifies that air concentrations are drawn from the roadway district assigned to the home location, while “Road Work” draws air concentrations from roadways near the work location. If multiple pollutants are used, not all of them are required to use roadway concentrations. This option can be set for each pollutant. However, because APEX requires all pollutants to use the same district information, AQ data are required in all districts for all pollutants. If the user chooses to not use these data, all “Road” designations will then be replaced by “Other.”

The “Near Home” and “Near Work” locations are randomly sampled within a certain radius of the home and work tract. These can be used to simulate movements outside the standard home and work locations. These tracts can either be sampled at each time step, or once for each person by using the **RESAMPLEN** keyword in the *Control Options* file. By default, APEX does not resample. The **NEARBYRADIUS** keyword controls how large of a radius is used when sampling nearby tracts. By default, this value is set to 20 km. If the user selects the “Last” location (L), then APEX will use either the NH or NW location, based on the location last visited. Since the Last location depends on the events of the individual, this location is treated differently than the others, and is not included in the *Microenvironmental Results* file. In the

events file, locations mapped to the Last location are designated by “99.” If the simulation includes people who leave the study area, all near work locations will use the ambient concentrations of the work location.

The supplied file contains microenvironment assignments for the CHAD location codes. The user must assign each location code to microenvironments defined in the *Microenvironment Descriptions* file by specifying the microenvironment number in the *APEX Micro* column. The file must contain assignments for all CHAD location codes, or APEX will exit with a fatal error.

A zero in the *APEX Micro* column will result in no exposure in that particular CHAD microenvironment location. A value of -1 means that APEX will use whichever microenvironment was previously in use in the composite diary time series for an individual (typically used for CHAD locations “U” and “X”).

The *DiaryCluster* file has the same layout as the *Microenvironment Mapping* file, except it maps to the axes used by the clustering method for longitudinal diary assembly. This file is required only if the clustering algorithm is being used (i.e., if *ClusterDiary* = Y). This file is used for creating the transitional probabilities of individuals moving throughout different microenvironments. The microenvironments defined in the *DiaryCluster* file need not be the same as in the *Microenvironment Mapping* file. The method currently uses up to five axes for clustering; that is, the diary times spent in those five “cluster axes” locate each diary as a point in a five dimensional space, and then clusters of diaries are defined, one per axis. For purposes of defining air concentrations, an APEX run that uses this diary clustering file may use a *Microenvironment Mapping* file with more (or fewer) than five microenvironments.

**Table 4-9. CHAD Location Codes**

<b>Loc. Code</b>	<b>Description</b>	<b>Loc. Code</b>	<b>Description</b>
X	No data	31210	Walk
U	Uncertain of correct code	31230	In stroller or carried by adult
30000	Residence- general	31300	Waiting for travel
30010	Your residence	31310	... bus or train stop
30020	Other residence	31320	... indoors
30100	Residence- indoor	31900	Travel- other
30120	Your residence- indoor	31910	... other vehicle
30121	... kitchen	32000	Non-residence indoor- general
30122	... living room or family room	32100	Office building/ bank/ post office
30123	... dining room	32200	Industrial/ factory/ warehouse
30124	... bathroom	32300	Grocery store/ convenience store
30125	... bedroom	32400	Shopping mall/ non-grocery store
30126	... study or office	32500	Bar/ night club/ bowling alley
30127	... basement	32510	Bar or night club
30128	... utility or laundry room	32520	Bowling alley
30129	... other indoor	32600	Repair shop
30130	Other residence- indoor	32610	Auto repair shop/ gas station
30131	... kitchen	32620	Other repair shop
30132	... living room or family room	32700	Indoor gym /health club
30133	... dining room	32800	Childcare facility
30134	... bathroom	32810	... house
30135	... bedroom	32820	... commercial

<b>Loc. Code</b>	<b>Description</b>	<b>Loc. Code</b>	<b>Description</b>
30136	... study or office	32900	Large public building
30137	... basement	32910	Auditorium/ arena/ concert hall
30138	... utility or laundry room	32920	Library/ courtroom/ museum/ theater
30139	... other indoor	33100	Laundromat
30200	Residence- outdoor	33200	Hospital/ medical care facility
30210	Your residence- outdoor	33300	Barber/ hair dresser/ beauty parlor
30211	... pool or spa	33400	Indoors- moving among locations
30219	... other outdoor	33500	School
30220	Other residence- outdoor	33600	Restaurant
30221	... pool or spa	33700	Church
30229	... other outdoor	33800	Hotel/ motel
30300	Residential garage or carport	33900	Dry cleaners
30310	... indoor	34100	Indoor parking garage
30320	... outdoor	34200	Laboratory
30330	Your garage or carport	34300	Indoor- none of the above
30331	... indoor	35000	Non-residence outdoor- general
30332	... outdoor	35100	Sidewalk- street
30340	Other residential garage or carport	35110	Within 10 yards of street
30341	... indoor	35200	Outdoor public parking lot /garage
30342	... outdoor	35210	... public garage
30400	Residence- none of the above	35220	... parking lot
31000	Travel- general	35300	Service station/ gas station
31100	Motorized travel	35400	Construction site
31110	Car	35500	Amusement park
31120	Truck	35600	Playground
31121	Truck (pickup truck or van)	35610	... school grounds
31122	Truck (not pickup truck or van)	35620	... public or park
31130	Motorcycle or moped	35700	Stadium or amphitheater
31140	Bus	35800	Park/ golf course
31150	Train or subway	35810	Park
31160	Airplane	35820	Golf course
31170	Boat	35900	Pool/ river/ lake
31171	Boat- motorized	36100	Outdoor restaurant/ picnic
31172	Boat- other	36200	Farm
31200	Non-motorized travel	36300	Outdoor- none of the above

! APEX Microenvironment Mapping File					
! Mapping of CHAD activity locations to APEX microenvironments					
!CHAD Loc.	Description		Micro #	Location	Notes
-----					
U	Uncertain of correct code	=	-1	U	Unknown
X	No data	=	-1	U	Unknown
30000	Residence, general	=	1	H	Home
30010	Your residence	=	1	H	Home
30020	Other residence	=	1	H	Home
30100	Residence, indoor	=	1	H	Home
30120	Your residence, indoor	=	1	H	Home
30121	..., kitchen	=	1	H	Home
30122	..., living room or family room	=	1	H	Home
...					

**Exhibit 4-22. First Part of an Example Microenvironment Mapping File**

#### 4.19 Diary Questionnaire (DiaryQuest) File

The *Diary Questionnaire* file provides the personal information component of each 24-hour activity diary (Exhibit 4-23). Each record contains values for the variables listed below.

- **CHADID** (9 characters)
- **Weekday** (day of week: MON, TUE, ..., SUN, Missing (X))
- **Sex** (Male (M), Female (F), Missing (X))
- **Race** (White (W), Black (B), Asian (A), Hispanic (H), Other (O), not available (X))
- **Employed** (Yes (Y), No (N), Missing (X))
- **Age** (integer years)
- **MaxTemp** (maximum hourly temperature for this diary day; degrees F)
- **AvgTemp** (daily mean temperature for this diary day; degrees F)
- **Occupation** (code; see Table 4-10)
- **QCMiss** (missing time: the total number of minutes associated with events in the *Diary Events* file for which the activity and/or location codes are missing for this diary day)
- **RecCount** (record count: the number of records in the *Diary Events* file corresponding to this diary day)
- **Commute** (commuting time; only required if commuting is modeled: the total time in minutes spent commuting on this diary day)

The user should not change this input file unless the CHAD database has changed or other activity data are to be used instead. In the latter case, the input file format restrictions must be met, the CHAD coding conventions must be used, and the other CHAD files must be modified to be consistent with this file. Note that this file has one record per **CHADID**, whereas the CHAD *Diary Events* file has **RecCount** of records per **CHADID**. The **Commute** column is only required when commuting is modeled. If a diary is missing the average or maximum temperature, weekday, or is missing time greater or equal to three hours, then the diary will be dropped.

```

! CHAD Questionnaire File
! based on the 110116 CHAD master files
! Created      July 26, 2017 by JEL
! chadid, weekday, sex, race, employed, age, maxtemp, avgtemp, occupation, qcmiss, reccount, commute
BAL97001A,TUE,    F,  W,  N,    77, 43,   34,   X,    45,  29,   0
BAL97001B,WED,    F,  W,  N,    77, 51,   41,   X,   135,  28,   0
BAL97001C,THU,    F,  W,  N,    77, 57,   47,   X,    15,  30,   0
BAL97001D,FRI,    F,  W,  N,    77, 45,   36,   X,     0,  28,   0
BAL97001E,TUE,    F,  W,  N,    77, 47,   39,   X,     0,  27,   0
BAL97001F,WED,    F,  W,  N,    77, 36,   29,   X,     0,  28,   0
BAL97001G,THU,    F,  W,  N,    77, 38,   29,   X,     0,  26,   0
BAL97001H,FRI,    F,  W,  N,    77, 43,   36,   X,     0,  28,   0
BAL97001I,TUE,    F,  W,  N,    77, 41,   39,   X,    15,  28,   0
BAL97001J,WED,    F,  W,  N,    77, 54,   44,   X,    15,  28,   0
BAL97001K,THU,    F,  W,  N,    77, 48,   40,   X,     0,  30,   0
BAL97001L,FRI,    F,  W,  N,    77, 42,   36,   X,    30,  30,   0
BAL97006A,WED,    M,  W,  N,    80, 51,   41,   X,     0,  31,   0
BAL97006B,THU,    M,  W,  N,    80, 57,   47,   X,    60,  36,   0
BAL97006C,FRI,    M,  W,  N,    80, 45,   36,   X,    75,  31,   0
...

```

### Exhibit 4-23. First Part of the Diary Questionnaire File

Note: In Exhibit 4-23, spaces (not tabs) may be used as delimiters.

**Table 4-10. CHAD Occupation Codes**

Code	Description
ADMIN	Executive, administrative, and managerial
PROF	Professional
TECH	Technicians
SALE	Sales
ADMSUP	Administrative support
HSHLD	Private household
PROTECT	Protective services
SERV	Service
FARM	Farming, forestry, and fishing
PREC	Precision production, craft, and repair
MACH	Machine operators, assemblers, and inspectors
TRANS	Transportation and material moving
LABOR	Handling, equipment cleaners, helpers, and laborers
X	Missing

## 4.20 Diary Events File

The *Diary Events* file provides descriptions of events in each day for all the diary days in the CHAD database. Events may last from one minute to one hour in duration. Each record includes the variables listed below.

- **CHADID**
- **StartTime** (the time the event began; HHMM, with 0000 representing midnight)
- **Duration** (the duration of the event, in minutes)
- **Act** (activity code; see Table 4-5)
- **Loc** (location code; see Table 4-9)

The diary events file should be generated from the CHAD database at the same time as the *Diary Questionnaire* file to ensure that the CHAD IDs are in the same order. Each diary day begins and ends at midnight, and there should be exactly 24 hours of data per diary. See Exhibit 4-24

for an example of a portion of this file. See the previous section on the *Diary Questionnaire* file if user-supplied data are to be provided.

```

! CHAD Events File
! based on the 110116 CHAD master files
! Created      July 26, 2017 by JEL
! chadid, starttime,duration,act,  loc
BAL97001A,0000,    60,    14500,30125
BAL97001A,0100,    60,    14500,30125
BAL97001A,0200,    60,    14500,30125
BAL97001A,0300,    60,    14500,30125
BAL97001A,0400,    60,    14500,30125
BAL97001A,0500,    60,    14500,30125
BAL97001A,0600,    60,    14500,30125
BAL97001A,0700,    30,    14500,30125
BAL97001A,0730,    30,    14400,30121
BAL97001A,0800,    60,    16000,30122
BAL97001A,0900,    60,    14500,30125
BAL97001A,1000,    30,    14500,30125
BAL97001A,1030,    30,     X,     X
BAL97001A,1100,    45,    14500,30125
BAL97001A,1145,    15,     X,     X
BAL97001A,1200,    60,    14500,30125
BAL97001A,1300,    60,    14500,30125
...

```

**Exhibit 4-24. First Part of the Diary Events File**

Note: In Exhibit 4-24, spaces (not tabs) may be used as delimiters.

## 4.21 Diary Statistics File

The *Diary Statistics* file contains a diary statistic for each diary in the CHAD database. This file is used in constructing multi-day (longitudinal) diaries in APEX from the CHAD one-day diaries using the D&A method. Refer to *Volume II* for information on how to construct this file.

APEX has three options for assembling simulation-length diaries. The first method is to randomly pick a new day-long diary from CHAD for each day in the simulation. For the second option, APEX has a longitudinal diary assembly algorithm for selecting diaries based on some key statistic of each CHAD diary. This algorithm requires the selection of a diary based on some key diary statistic relevant to the pollutant being studied. For example, the statistic may be time spent outdoors or time spent in a vehicle. The third method creates longitudinal diaries via transitional probabilities calculated for clusters derived from the CHAD inputs files. Details of both longitudinal diary algorithms are provided in *Volume II*.

The *Diary Statistics* file must contain the CHAD ID for each diary and the value of this statistic (ID and statistic separated by a comma or a space, one diary per row). The order of the CHAD IDs in this file must be the same as on the *Diary Questionnaire* file, or an error will result.

Two *Diary Statistics* files have been generated from CHAD and are included starting in the APEX Version 4 release. These files are for time spent outdoors and time spent in vehicles. The files were constructed by summing up the time spent in locations considered “outdoors” or “in vehicle” in each CHAD diary. Table 4-11 gives the CHAD location codes that were used to

generate these files. Users may construct other files from the CHAD database. An example portion of a diary statistic file is shown in Exhibit 4-25.

The use of the second longitudinal algorithm is invoked by setting the *Control Options* file keyword **LONGITDIARY** = YES. If **LONGITDIARY** = NO, the *Diary Statistics* file is not needed, and need not be specified in the *Control Options* file.

**Table 4-11. CHAD Locations used in Constructing the Outdoor Time and Vehicle Time Diary Statistics Files Supplied with APEX**

CHAD Location IDs Considered “Outdoors”	CHAD Location IDs Considered “In Vehicle”
30332, 30342, 30320, 30200, 31310, 35000–36300	31000–31172

<pre>! CHAD Longitudinal Activity Statistics File for Time Outdoors ! Prepared by EPA for EPA ! Created      July 26, 2017 by JEL ! CHAD ID, time spent outdoors (minutes) BAL97001A, 0 BAL97001B, 165 BAL97001C, 0 BAL97001D, 0 BAL97001E, 0 BAL97001F, 0 BAL97001G, 0 BAL97001H, 0 BAL97001I, 0 BAL97001J, 0 BAL97001K, 0 BAL97001L, 0 BAL97006A, 75 BAL97006B, 270 BAL97006C, 90 BAL97006D, 60 BAL97006E, 30 BAL97006G, 135 BAL97006H, 135 BAL97006I, 60 BAL97006J, 90 ...</pre>
---

**Exhibit 4-25. First Part of the Diary Statistics File for Time Spent Outdoors**

## 4.22 Diary Occupations File

The *Diary Occupations* file contains a diary occupation for each diary in the CHAD database. This file can be used at any time, but is most useful when defining an occupation-related *Profile Factor*. Using this file will overwrite all occupations found in the *Diary Questionnaire* file. If occupation is a factor, and if an occupation is found in a diary that is not used in the *Profile Factors* file, then that diary occupation will be set to missing (X).

The *Diary Occupations* file must contain the CHAD ID for each diary and the name of the occupation (ID and occupation separated by a comma or a space, one diary per row). The order



of the CHAD IDs in this file must be the same as on the *Diary Questionnaire* file. An example portion of a *Diary Occupations* file is shown in Exhibit 4-26.

```
! CHAD Optional Occupation file
! Created 11/10/2011
BAL97001A, SALE
BAL97001B, ADMIN
BAL97001C, TRANS
BAL97001D, FARM
BAL97001E, TECH
BAL97001F, PROTECT
BAL97001G, PROF
BAL97001H, OTHER
BAL97001I, LABOR
BAL97001J, OTHER
BAL97001K, LABOR
BAL97001L, ADMIN
BAL97006A, TRANS
...
```

**Exhibit 4-26. First Part of a Diary Occupations File**

### 4.23 Diary Transitions File

This file is used only with the *CLUSTERDIARY* = YES option. It has the same records as the *Diary Questionnaire* file, one for each CHAD ID. The fields are: *CHADID*, *Study*, *Base*, *Year*, *Month*, *Day*, *Wkday*, *DayNum*, and *Sasdate*. The most important data relate each *CHADID* to a *Base* and a *SASDate*. The *Base* is unique to each distinct person, so multiple diaries from one person share the same *Base*. The *SASDate* increases by one each time a day passes, which makes it easy to decide which diaries from the same person are on consecutive days.

### 4.24 Microenvironment Descriptions File

The *Microenvironment Descriptions* input file serves two purposes. First, it lists the names and numbers assigned to each microenvironment, and defines the methods by which pollutant concentrations are calculated. Second, this file tells APEX how to define the parameters that are required to calculate these concentrations. These are called “microenvironment parameters,” abbreviated as “MP.” With the exceptions of air exchange rate and microenvironment volume, the MPs are specific to each combination of pollutant and microenvironment. The layout of the *Microenvironment Descriptions* file typically makes a clear distinction between the two sections using headers lines which start with !, which are therefore not processed by APEX. An example of the first section is shown in Exhibit 4-27, while an example of a “Parameter Description” section is shown in Exhibit 4-28. The examples shown in these figures will be discussed in detail below.

APEX is a microenvironmental human exposure model, and the definitions and properties of the microenvironments are crucial to an APEX run. The user may spend more time preparing this input file than on any other file.

#### 4.24.1 Microenvironment Descriptions Section

In the “Microenvironment Descriptions” section of the *Microenvironment Descriptions* file, the user specifies a **Micro** (microenvironment number), a **Name**, and a calculation **Method** for each microenvironment, as shown in Exhibit 4-27. The **Micro** cannot exceed the number of microenvironments specified in the *Control Options* file, nor can it exceed 127. It also has to correspond with each of the microenvironment numbers in the *Microenvironment Mapping* file. A microenvironment **Name** may be a word up to 40 characters (without spaces). The calculation **Method** could be either **MassBal** or **Factors**. In the **MassBal** method, the concentration in a microenvironment is calculated using a mass balance approach, while in the **Factors** method the microenvironment concentration is assumed to be a linear function of ambient concentration. See *Volume II* for further description of the **MassBal** and **Factors** methods.

Micro	Name	Method
1	Indoor-Residence	MASSBAL
2	Indoor-Other	MASSBAL
3	Outdoor	FACTORS
4	Near-road	FACTORS
5	Vehicle	FACTORS

**Exhibit 4-27. Example of a Microenvironment Descriptions Section of a Microenvironment Descriptions File**

#### 4.24.2 Parameter Descriptions Section

The “Parameter Descriptions” section of the *Microenvironment Descriptions* file consists of the specification of probability distributions for the MP that are required for calculating pollutant concentrations in the microenvironments. See *Volume II* for further information on the **MassBal** and **Factors** concentration calculation methods. Three types of MP can be defined for the **Factors** method and eight for the **MassBal** method. In each method, some of the MP may be left at default values, and do not be explicitly defined if the defaults are acceptable to the user. The parameters and their default values (if present) are given in Table 4-12. Air exchange rate and volume are not pollutant-specific, so they are only defined once. Otherwise, there must be one definition for each MP for each pollutant for each microenvironment, with the exception of the two pollutant source types, concentration (CSource) and emissions (ESource), which permit multiple sources in the same microenvironment.

**Table 4-12. Microenvironment Parameters for the Factors and Massbal Methods**

Calculation Method	Parameter Type	Parameter Code	Parameter Units	Parameter Default Value
<b>Factors</b>	Proximity	<b>PR</b>	None	1
	Penetration	<b>PE</b>	None	1
	Csource	<b>CS</b>	ppm, ppb, or $\mu\text{g}/\text{m}^3$ (same as <b>INPUTUNITS</b> )	0
<b>MassBal</b>	Proximity	<b>PR</b>	None	1
	Penetration	<b>PE</b>	None	1
	Decay Rate	<b>DE</b>	1/hr	0

Calculation Method	Parameter Type	Parameter Code	Parameter Units	Parameter Default Value
	Air Exchange Rate	<i>AE</i>	1/hr	none
	Volume	<i>VO</i>	m <sup>3</sup>	None (only needed if ESource is present)
	MeanR	<i>MR</i>	1/hr	AE+DE
	Csource	<i>CS</i>	ppm, ppb, or μg/m <sup>3</sup> (same as <i>INPUTUNITS</i> )	0
	ESource	<i>ES</i>	μg/hr	0

As mentioned above, not all of the parameters must be explicitly defined for each microenvironment. If the default values in Table 4-12 are acceptable for a microenvironment, then a given parameter definition may be omitted from the input file. For **Factors**, default values exist for all the parameters. If no parameters are defined for microenvironments using the **Factors** method, then the microenvironment concentration is always equal to the current ambient concentration. For a **MassBal** microenvironment, the air exchange rate parameter must always be defined as it has no default value. The volume parameter does not have a default either, but it is only used if ESource terms exist for that microenvironment and may be omitted otherwise. All other parameters are optional.

The proximity and penetration factors are used to model the ambient pollutant concentrations immediately outside and inside a microenvironment. The air exchange rate and volume variables define the air flow rate in and out of the microenvironment as well as the microenvironment air volume. The decay rate defines the rate of removal of pollutant from the microenvironment via various means. The parameter MeanR is a factor that describes the removal of pollutant by both air flow and decay. The CSource and Esource terms are concentration and emission pollutant sources, respectively. See *Volume II* for a detailed description of these parameters and the microenvironmental concentration equations.

As part of the estimation of microenvironment concentrations, each MP for each pollutant is given a value for each hour of the simulation, for each profile generated. This value may or may not be different from the values at other hours, depending on how it is defined. Some MPs, such as house volume, should logically remain constant throughout the simulation, while others may change seasonally, daily, or hourly. Values may recur in patterns, such as the same set of 24 hourly values for some parameter may recur each Saturday during the winter season. These patterns are determined from the four mapping options and the three resampling options specified in each microenvironmental parameter definition.

The definitions for the microenvironment parameters may appear in any order in the *Microenvironment Descriptions* file. Therefore, the user (for example) may choose to group definitions by microenvironment or by pollutant. Each definition should be separated from the next either by blank lines or by comment lines (starting with an exclamation point) to aid in

clarity. A parameter description consists of keywords and distribution definitions, described in the following sections.

## Keywords

The first part of a microenvironment parameter description is a list of settings, each described by a keyword. The first keyword for a given MP should be *MP#*. The various keywords serve a number of purposes, including specifying:

- Which microenvironment is being considered;
- Which pollutant is being considered (not needed for air exchange rate or volume);
- Which parameter is being defined for that microenvironment (the parameter code);
- The source number for the current parameter (if it is ESource or CSource);
- How that parameter varies over hours in the day, days of the week, or months of the year;
- Whether the parameter depends on any conditional variables;
- Whether the parameter is correlated with any other parameter (by being sampled using the same random numbers);
- Which group the MP belongs to (for Sobol analysis); and
- Whether or not a new value of parameter is generated for each hour, for each day, and for the workplace

The keywords and their descriptions are provided in Table 4-13.

The conditional variable keywords must be either one of the conditional variables listed in Table 4-8; i.e., *TempCat*, *HumidCat*, *WindCat*, *DirCat*, *PrecipCat*, *MaxTempCat*, *AvgTempCat*, *HasGasStove*, *HasGasPilot*, *AC\_Home*, *AC\_Car*, *WindowRes*, *WindowCar*, *SpeedCat*, *DailyConditional1–DailyConditional3*, *ProfileConditional1–ProfileConditional5*, *RegionalConditional1–RegionalConditional5*, *AQConditional1–AQConditional5*, *Gender*, *Employed*, *FactorGroup*, Or *PopCat*. All variables, with the exception of the last four, must be defined in the *Profile Functions* file in order to be used as a conditional variable in a microenvironmental parameter description. *PopCat* is the “population category,” or gender/race combination (e.g., “white males” is a population category). Therefore, *Gender* and *PopCat* should not both be used as conditional variables for the same MP.

In APEX, the user has the option of correlating the random samples for microenvironmental parameters. Such correlation would make sense, for example, when the value of the parameter is assumed to be mainly a function of the properties of a simulated individual’s home and the pollutants have similar properties (for example, if the pollutants are all particles). In addition, in some cases it may be that the same parameters may be correlated in different microenvironments. APEX uses a simple method of correlating MPs—by sampling them using the same random numbers. This results in values being selected for correlated parameters at the same percentile from the appropriate distributions. The percentiles will correspond each hour as long as the 2 (or more) parameters use the same conditional variables, time and area mappings, and resampling rates and thus have the same number of required distributions and samples. Otherwise, the samples get out of phase and any correlation is lost. APEX checks that the conditionals, mappings, and resampling are the same when correlating parameters, and writes a

warning if they are not. APEX will still run, but the user should be aware that the correlation is lost.

Correlation is handled by an optional keyword in the MP definition, *CorrNum*. Each subset of MPs that the user desires to be correlated (sampled at the same percentile each hour) are assigned a unique integer 1–N, where N is the total number of correlated subsets.

All of the keywords for the MP come at the beginning of the microenvironmental parameter definition. Except for the *Conditional* and *Resamp* variables, APEX examines only the first 5 characters of each keyword to decide what it is; therefore, the user may extend them any way they like. After the definition of all the keywords, the next line should be the header line for the data section; that is, the section that contains the actual distribution definitions for the MP. The header line must begin with the word *Block*, as APEX recognizes this word as indicating the end of the keyword section (see Exhibit 4-28 for an example of an appropriate header).

**Table 4-13. Keyword Definitions for the Parameter Descriptions Section of the Microenvironment Descriptions File**

<b>Keyword</b>	<b>Abbrev.</b>	<b>Description</b>
<i>MP#</i>	<i>MP#</i>	This number is used for tracking this particular MP. Every MP must be assigned a distinct number, but they do not have to be consecutive.
<i>Microenvironment Number</i>	<i>Micro</i>	These numbers must match the microenvironment numbers in the “Microenvironment Descriptions” section.
<i>Pollutant</i>	<i>Pollu</i>	Integer corresponding to the pollutant being considered. (Number corresponds to the order of the pollutant definition in the <i>Control Options</i> file). The default is 1, so this is not needed in a single-pollutant APEX run. Not used for <i>AER</i> and <i>Volume</i> definitions (ignored if defined).
<i>Parameter Code</i>	<i>Param</i>	A parameter code such as <i>PR</i> (proximity) or <i>PE</i> (penetration) provided in Table 4-12, used to specify the parameter type.
<i>Correlation Number</i>	<i>Corrn</i>	Integer number corresponding to correlation subset. Each subset of MPs that the user desires be correlated (sampled at the same percentile each hour) are assigned a unique integer 1–N, where N is the total number of correlated subsets. If using this option with Sobol analysis, be certain that the correlated MPs are assigned to the same Sobol group.
<i>Source Number</i>	<i>Sourc</i>	Numbers multiple sources in the same microenvironment. Not needed if there is only one source present (or none at all).
<i>Hours - Block</i>	<i>Hours</i>	This variable is used to map hours of a day to different time blocks. A “time block” is a group of hours for which the same MP distribution(s) will be used. The input line always contains a list of 24 integers, representing 24 hours a day. The first hour is midnight to 1 a.m. and the 24th is 11 p.m. to midnight. The position of an integer in the input line represents the hour in a day. The integer represents the number of a time block that an hour belongs to. The hours in a time block do not need to be consecutive, nor does a time block have to have the same number of hours. If this line is missing, the default value is that all 24 hours are in a single time block–block #1.
<i>Weekday - Daytype</i>	<i>Weekd</i>	This variable is used to map days in a week to different day types. A “day type” is a set of days for which the same MP distribution(s) will be used. Seven integers must be given in this input line. The position of an integer in the input line represents a day, beginning on Sunday and ending on Saturday. The integer represents the day type a day belongs to. If this variable is not defined, all days of a week will belong to day type #1.

<b>Keyword</b>	<b>Abbrev.</b>	<b>Description</b>
<b><i>Month - Season</i></b>	<b><i>Month</i></b>	This variable is used to map months of a year to different seasons. A “season” is a set of months for which the same MP distribution(s) will be used. Twelve integers must be given in this input line. The position of an integer represents a month of a year, beginning in January and ending in December. The integer represents the season that a month belongs to. If this line is missing, all 12 months belong to season #1.
<b><i>District - Area</i></b>	<b><i>Distr</i></b>	This variable is used to map air districts to larger areas. The number of integers in this line must match the number of air districts in the study area. This variable is a holdover from APEX2 and should not be used unless really necessary. The user could delete this line or place the same number of 1 in this line as the number of air districts.
<b><i>Condition # 1</i></b>	<b><i>Condi #1</i></b>	Choice for the first conditional variable. A conditional variable is a variable whose value affects the choice of MP distribution(s). If not used, this line may either be omitted or the value set to zero.
<b><i>Condition # 2</i></b>	<b><i>Condi #2</i></b>	Choice for the second conditional variable.
<b><i>Condition # 3</i></b>	<b><i>Condi #3</i></b>	Choice for the third conditional variable.
<b><i>ResampHours</i></b>	<b><i>Resamph</i></b>	Either “YES” or “NO”. If “YES”, a random value is selected from distribution for a parameter in each hour within a time block. If “NO”, a random value is selected for a parameter for a time block and used for every hour within the time block. The default value is “NO”.
<b><i>ResampDays</i></b>	<b><i>Resampd</i></b>	Either “YES” or “NO”. If “YES”, a random value is selected from a distribution for a parameter for each day within a day type. If “NO”, a random value is selected for a day type and used for every day within the same day type. The default is “NO”.
<b><i>ResampWork</i></b>	<b><i>Resampw</i></b>	Either “YES” or “NO”. If “YES”, a separate set of random values is selected from a distribution for the workplace. If “NO”, the same set of random values are used (for the same day and hour) both for home and at work. The default is “NO”.
<b><i>ResampTime Step</i></b>	<b><i>Resampt</i></b>	Either “YES” or “NO”. If “YES”, a new random value is generated each time step, if longer than one hour. For timesteps equal to or less than one hour, <b><i>Resampt</i></b> is equivalent to <b><i>Resamph</i></b> . The default is “NO”.
<b><i>Sobol Group</i></b>	<b><i>Sobol</i></b>	Group number of this MP for Sobol analysis. MPs in the same group are still sampled independently; this does NOT correlate the inputs. Grouping the inputs reduces the number of passes through the code needed to obtain sensitivity indices. Two indices (main and total) are generated for each group. No default exists. This only needs to be specified if the Sobol method is being used.

## Distribution Definitions

The last part of an MP definition lists the probability distributions for the MP at different times or under different circumstances during the simulation. Sets of distribution data may exist for all possible combinations of the user-specified cases of the seven indexing variables defined below.

- ***Block***: time block (as described by ***Hours - Block***)
- ***DayType***: day type (as described by ***Weekday - DayType***)
- ***Season***: season of the year (as described by ***Month - Season***)
- ***Area***: air quality area (as described by ***District - Area***)
- ***C1***: conditional variable # 1
- ***C2***: conditional variable # 2
- ***C3***: conditional variable # 3

The above labels are listed in the header line for the data section (which starts with ‘Block’). Each subsequent line lists seven indices (which reference combinations of the above), followed by a distribution. Each possible combination of indices requires one line. For most MPs, most or even all of the indices have just a single value (which is 1).

The number of cases for the indexing variables *Block*, *DayType*, and *Season* are specified by mappings in the keyword section, and is 1 by default. For example, the number of time blocks would be the highest integer indicated in the *Time - Block* mapping. For the conditional variables, *MaxTempCat*, *AvgTempCat*, *HasGasStove*, *HasGasPilot*, *AC\_Home*, *AC\_Car*, *WindowRes*, *WindowCar*, *SpeedCat*, *DailyConditional1–DailyConditional3*, *ProfileConditional1–ProfileConditional5*, *RegionalConditional1–RegionalConditional5*, or *AQConditional1–AQConditional5*, the number of cases is determined by the number of *Results* indicated on the *Profile Functions* file (Section 4.17). If fewer than 3 conditional variables are used for an MP, the ones not used are given an index of 1. For *Gender*, there are always 2 cases. For *PopCat*, the number of cases is indicated by the number of population groups (population files) defined on the *Control Options* file (Section 4.2), and the groups are indexed in the order they appear in the file (for example, if the population file for white females happened to be defined first in the *Control Options* file, then that group would correspond to the case *PopCat* = 1). For *FactorGroup*, the number is determined by the number of groups in the corresponding *Profile Factors* file. If the groups are tied to employment, then an additional group is defined to be the unemployed. For example, if your *Profile Factors* group defines probabilities for 8 occupations, 9 cases will need to be defined here; the additional one will be applied to unemployed people.

The user specifies the MP distribution using the standard APEX distribution format (a distribution shape, followed by 4 distribution parameters, upper and lower truncation bounds, and a resampling flag). The 4 parameters used are dependent on the shape of the distribution. See *Volume II* for a complete discussion of the use of probability distributions in APEX. Thus the following data must be present in each specification.

- **Shape**: This variable gives the type of the distribution
- **Par1**: Parameter 1 of the MP distribution. Depends on type.
- **Par2**: Parameter 2 of the MP distribution. Depends on type.
- **Par3**: Parameter 3 of the MP distribution. Depends on type.
- **Par4**: Parameter 4 of the MP distribution. Depends on type.
- **LTrunc**: Lower truncation point of the distribution
- **UTrunc**: Upper truncation point of the distribution
- **ResampOut**: Distribution resampling flag

See Table 4-6 for the available distribution types and required parameters. The parameters that are not used for specifying a distribution should be marked with a period (“.”) as a place holder.

### Examples of Parameter Descriptions

Two examples of parameter descriptions are shown in Exhibit 4-28. These examples should provide the user with a good idea of how the keywords and distribution definitions work.

In the first example, the *MP#* is 1. The air exchange rate (code *AE*) is defined for microenvironment #1. The pollutant number is not used with air exchange rates, so this keyword is missing. This MP is assigned to Sobol group 51. In this case, the parameter distribution is only a function of two conditional variables, *AvgTempCat*, and *AC\_Home*. The parameter is not resampled from the distribution every hour (*ResampHours* = NO) nor each day (*ResampDays* = NO), although the parameter is resampled if the simulated person moves between home and work (*ResampWork* = YES). Consequently, the conditional variable *AvgTempCat* has five possible values (1–5) and *AC\_Home* has two possible values (1–2); these variables and their values were defined in the *Profile Functions* file. Thus, after the line starting with ‘Block’ there must be 5x2 = 10 lines, one for each combination of the two conditional variables. The ten distributions are lognormal in shape (although they have different parameters), and are listed in order—first looping over the values of *AvgTempCat* and then *AC\_Home*.

```

MP# = 1
Micro number = 1
Parameter Type = AE
Condition # 1 = AvgTempCat
Condition # 2 = ACHome
ResampHours = NO
ResampDays = NO
ResampWork = YES
Sobol group = 51
Block DType Season Area C1 C2 C3 Shape Par1 Par2 Par3 Par4 LTrunc UTrunc ResampOut
1 1 1 1 1 1 1 Lognormal 0.95 1.7 0 . 0.111 10.0 Y
1 1 1 1 2 1 1 Lognormal 0.65 1.7 0 . 0.111 10.0 Y
1 1 1 1 3 1 1 Lognormal 0.35 1.7 0 . 0.111 10.0 Y
1 1 1 1 4 1 1 Lognormal 0.33 1.9 0 . 0.111 10.0 Y
1 1 1 1 5 1 1 Lognormal 0.33 1.9 0 . 0.111 10.0 Y
1 1 1 1 1 2 1 Lognormal 0.50 2.0 0 . 0.111 10.0 Y
1 1 1 1 2 2 1 Lognormal 0.50 2.0 0 . 0.111 10.0 Y
1 1 1 1 3 2 1 Lognormal 0.60 2.0 0 . 0.111 10.0 Y
1 1 1 1 4 2 1 Lognormal 0.80 2.0 0 . 0.111 10.0 Y
1 1 1 1 5 2 1 Lognormal 1.00 2.0 0 . 0.111 10.0 Y

MP# = 25
Micro number = 12
Pollutant = 3
Parameter Type = PE
Hours - Block = 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1
Weekday-DayType = 1 2 2 2 2 2 1
Month-Season = 1 1 2 2 2 3 3 3 4 4 4 1
Sobol group = 51
Block DType Season Area C1 C2 C3 Shape Par1 Par2 Par3 Par4 LTrunc UTrunc ResampOut
1 1 1 1 1 1 1 Point 1.0 . . . . .
2 1 1 1 1 1 1 Point 0.5 . . . . .
1 2 1 1 1 1 1 Point 0.9 . . . . .
2 2 1 1 1 1 1 Point 0.4 . . . . .
1 1 2 1 1 1 1 Point 0.8 . . . . .
2 1 2 1 1 1 1 Point 0.3 . . . . .
1 2 2 1 1 1 1 Point 1.0 . . . . .
2 2 2 1 1 1 1 Point 0.9 . . . . .
1 1 3 1 1 1 1 Point 0.8 . . . . .
2 1 3 1 1 1 1 Point 0.7 . . . . .
1 2 3 1 1 1 1 Point 0.6 . . . . .
2 2 3 1 1 1 1 Point 0.5 . . . . .
1 1 4 1 1 1 1 Point 0.5 . . . . .
2 1 4 1 1 1 1 Point 0.3 . . . . .
1 2 4 1 1 1 1 Point 0.2 . . . . .
2 2 4 1 1 1 1 Point 0.1 . . . . .

```

**Exhibit 4-28. Example of a Parameter Descriptions Section of a Microenvironment Descriptions File**



In the second example, penetration factor (*PE*) is defined for Microenvironment #12. Here, the distributions are not a function of any conditional variable, but rather different time blocks, day types, and seasons. Distributions for *PE* must be defined for all possible combinations of these time variables. The *Hour - Block* keyword line indicates a mapping of the hours of the day into two different time blocks (1 and 2) roughly defining night and day. Thus a different parameter distribution for *PE* will be used for these two time blocks. Similarly, the *Weekday - DayType* mapping keyword line defines two different day types, “1” for Saturday and Sunday, and “2” for the rest of the days of the week. Finally, the *Month - Season* mapping keyword line defines four seasons, labeled 1-4, corresponding to winter, spring, summer, and autumn. The distributions follow (again looping first over block, then day type, then season), and in this example the parameter is defined as a single point value in all cases. Hence  $2 \times 2 \times 4 = 16$  distributions are needed.

It is clear that these methods allow the user a great deal of flexibility in defining different distributions for the microenvironmental parameters. In most cases, many of the features of these descriptions will not be used, but in some cases the user may wish to define a large number of distributions for a single parameter. There is no limit in APEX on the number of distributions that can be defined for a microenvironment parameter.

Control of the resampling is distinct from the number of distributions. APEX produces random values in two steps. First, a uniform random number between zero and one is produced for every hour of the simulation (separately for each person and each variable). The same uniform sample may be used for multiple hours, depending on the resampling options. If all resampling is “NO” (the default), then the same random sample applies to all hours of the simulation, for the given person. The distributions determine how these samples are transformed. The uniform value indicates the percentile of the CDF of the final distribution that is to be assigned. When several distributions use the same uniform value, the same percentile is drawn from each. The value for an MP may change whenever the underlying uniform sample changes, or the distribution changes, or both.

## 4.25 Prevalence File

The *Prevalence* file is an optional APEX input file for modeling a subpopulation of persons with a particular disease or condition. The *Prevalence* file is only required when the setting *DISEASE* is set in the *Control Options* file. APEX uses the prevalence rates to assign a YES/NO value to a physiological profile variable, *ILL*, and to produce output exposure summary tables for persons with *ILL* = YES. If *DISEASE* is not set, then the *Prevalence* file is not required and no summary tables for ill persons will be printed.

The *Prevalence* file must contain prevalence rates (probabilities) for all age and gender cohorts from ages 0–99. The *Gender*, *MinAge*, and *MaxAge* lines define the bins. Each line of the prevalence file contains a sector ID, followed by the values for each of the bins. The values in the *Prevalence* file may be separated by one or more spaces. A portion of an example *Prevalence* file is shown in Exhibit 4-29.

```
! asthma prevalence
!  
Gender=      M          M          M          F          F          F  
MinAge=      0          20         60         0          20         60  
MaxAge=      19         59         99         19         59         99  
06071001110 0.09744  0.22000  0.12258  0.13636  0.18462  0.14524  
06071001210 0.05056  0.12857  0.09230  0.15200  0.24553  0.19099  
06071001220 0.14921  0.29310  0.13249  0.11818  0.11339  0.05856  
06071001300 0.17143  0.18889  0.21798  0.16503  0.06861  0.14479  
...
```

**Exhibit 4-29. First Part of an Example Prevalence File**

## CHAPTER 5. APEX OUTPUT FILES

APEX can produce the output files listed below.

- *Log*
- *Hourly*
- *Timestep*
- *Daily*
- *Profile Summary (Persons)*
- *Microenvironmental Summary*
- *Microenvironmental Results*
- *Output Tables*
- *Sites*
- *Events*
- *MultiPollutant*
- *Diary Clustering*
- *Sobol Results*

All of these output files are ASCII files which can be opened and reviewed using a text editor or other software (e.g., spreadsheet, database, statistical analysis, or graphics). A brief summary of these files is given in Table 5-1. Details of each file are provided in Sections 5.1 to 5.13 below. If the clustering method of longitudinal diary assembly is used, then APEX may record an additional output file, *ChadClust.dat*. This file is created for APEX internal use only and is not intended as output for the user (see *Volume II* for details).

All output files (except for *ChadClust.dat*, which is a binary file) contain the same set of header records, allowing files generated from the same run to be identified, and for audit trail requirements. This header section consists of six lines followed by a blank line, as indicated below.

- Line 1: Type of output file
- Line 2: APEX version, date and time of start of run
- Line 3: Location description (from *Control Options* file)
- Line 4: Scenario description (from *Control Options* file)
- Line 5: Echoes first line of *Control Options* file
- Line 6: List of the Pollutants (as given in *Control Options* file)
- Next N lines: Echo the first line of the each of the *Air Quality Data* files for the N pollutants in the simulation. If the output file is pollutant-specific, then only the line from its corresponding *Air Quality Data* file is echoed.

The Location, Pollutant, and Scenario descriptions echo what the user provided for those keywords in the *Control Options* file. In the first line of the *Control Options* file the user typically gives general identifying information for the simulation. Similarly, the first lines of the *Air Quality Data* files can identify the contents of the files.

**Table 5-1. APEX Output Files**

<b>Output File</b>	<b>Description</b>
<i>Log</i>	Contains the record of the APEX model simulation as it progresses. If the simulation completes successfully, the log file indicates the input files and parameter settings used for the simulation and reports on a number of different factors. If the simulation ends prematurely, the log file contains error messages describing the critical errors that caused the simulation to end.
<i>Hourly</i>	Provides an hour-by-hour time series of exposures, doses, and other variables for each modeled profile.
<i>Daily</i>	Provides a day-by-day time series of exposures, doses, and other variables for each modeled profile.
<i>Profile Summary</i>	Provides a summary of each profile modeled in the simulation. Each line lists the person's age, gender and race, in addition to a number of other personal profile variables that the model uses to simulate exposure.
<i>Microenvironmental Summary</i>	Provides a summary of the time and exposure by microenvironment for each profile modeled in the simulation.
<i>Microenvironmental Results</i>	Provides an hour-by-hour time series of microenvironment concentrations and parameters for a pollutant for each modeled profile for each location (e.g., "Home," "Work," and "Other"). A <i>Microenvironmental Results</i> file is generated for each pollutant.
<i>Output Tables</i>	Contains a series of tables summarizing the exposure (and dose, if calculated) results of the simulation for a pollutant. The percentiles and exposure/dose cut-off points used in these tables are defined in the <i>Control Options</i> file. A <i>Tables</i> file is generated for each pollutant.
<i>Sites</i>	Lists the sectors, air districts, and zones in the study area, and identifies the mapping between them.
<i>Events</i>	Contains event-level information (including MET, exposure, ventilation, and dose) for individuals in the simulation. Settings in the <i>Control Options</i> file allow the user to write this information for all persons, every Nth person, or for a set of specified profile IDs.
<i>Timestep</i>	Has the same format as the Hourly file, except that it reports variables on every timestep.
<i>Multipollutant</i>	Contains one row for each combination of micro, clock hour (0–23), and level for each pollutant.
<i>Sobol</i>	Is produced only if the user requests a Sobol sensitivity analysis run. In that case, most of the other output files are suppressed. The output consists of main and total indices for both the average day and maximum day, for selected exposure metrics.

## 5.1 Log File

The *Log* file records the information listed below as a model run progresses.

- Input files used
- Settings for job control options
- Number of diaries in total, and the number in each diary pool
- Model execution time
- Sectors in the study area
- Air districts in the study area
- Meteorology zones in the study area
- Mappings of sectors to air districts and meteorology zones
- Diary activities and counts of MET distributions
- Descriptions of the microenvironment-specific parameters

- Statistical summaries of the simulated profiles
- Output summary tables

If a model run stops abnormally, an error message will be written to the *Log* file. The user should review the *Log* file after a model run to ensure that the simulation executed and terminated normally and that the output results are valid. Output summary tables in this file (if the **LOGTABLES** option is set to YES) are exactly the same as the tables in the *Output Table* file. The level of detail of the information written to the *Log* file is controlled by the *Control Options* file setting, **DEBUGLEVEL**. **DEBUGLEVEL** can have a value of 0, 1, 2, or 3; the higher the level, the more information is written to the log. The *Control Options* file settings, **LOGDISTRICT**, **LOGPOPULATION**, **LOGPROFILES**, **LOGSECTORS**, **LOGTABLES**, and **LOGZONES** also control the writing of information to the *Log* file. See Table 4-4 for more information on these settings.

If there are 40 or fewer profiles in the run, the person-level demographic variables are written to the *Log* file for each person, as are the number of days that each person equals or exceeds the exposure cut-points for the **DMIHEXP** (daily maximum 1-hour) table.

## 5.2 Hourly File

The *Hourly* file contains hourly time series of a number of APEX variables including concentrations and doses for each simulated person or profile. Note: if the APEX timestep is greater than 1 hour (**TIMESTEPSPERDAY** < 24), the *Hourly* file will not be written. In this case, the *Timestep* file (see next section) provides the best summary of the exposure and dose time series. The user can control which variables are written to the *Hourly* file via a list of keywords using the *Control Options* file keyword **HOURLYLIST**. The variables and their corresponding keywords are provided in Table 5-2.

**Table 5-2. APEX Variables Written to the Hourly Output File**

Variable	Description	Units	Control Options File Keyword	Optional
<b>Person</b>	Simulated profile number	-	-	N
<b>Hour</b>	Hour # of the simulation	-	-	N
<b>Ve</b>	Ventilation	ml/min	<b>VE</b>	Y
<b>Va</b>	Alveolar ventilation	ml/min	<b>VA</b>	Y
<b>EVR</b>	Equivalent ventilation rate, VE divided by body surface area	L/(min-m <sup>2</sup> )	<b>EVR</b>	Y
<b>MET</b>	Metabolic equivalent. Time-averaged multiple of basal energy expenditure for the hour.	-	<b>MET</b>	Y
<b>EE</b>	Energy expenditure	kcal/min	<b>EE</b>	Y
<b>FEV E1 term</b>	Ozone-dependent variability term in the %ΔFEV1 model	% <sup>2</sup>	<b>FEVE1</b>	Y
<b>FEV E2 term</b>	Non-ozone dependent variability in %ΔFEV1	% <sup>2</sup>	<b>FEVE2</b>	Y
<b>Hourly Max %ΔFEV1</b>	Hourly maximum of the event-level %ΔFEV1 values	%	<b>DFEV1</b>	Y
<b>Micro Time</b>	Time spent in microenvironment N	min	<b>TIME1 - TIMEN</b>	Y

<b>Variable</b>	<b>Description</b>	<b>Units</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>Micro Exposure</i>	Exposure in microenvironment N	<i>OUTPUTUNITS</i>	<i>EXPI – EXPN</i>	Y
<i>Ambient Concentration</i>	Ambient pollutant concentration, time averaged over events	<i>INPUTUNITS</i>	<i>AMB</i>	Y
<i>Ambient Concentration (Home)</i>	Ambient pollutant concentration in the home district, time averaged over events	<i>INPUTUNITS</i>	<i>HOMEAMB</i>	Y
<i>Exposure</i>	Time-averaged exposure for the hour	<i>OUTPUTUNITS</i>	<i>EXP</i>	Y
<i>Dose</i>	Time-averaged dose for the hour. Units of dose depend on pollutant, see <i>Volume II</i> .	<pollutant-specific>	<i>DOSE</i>	Y
<i>Intake Dose</i>	PM pollutants only. Average mass inhaled per minute (includes mass not deposited) during the hour	micrograms /minute (µg/min)	<i>INTAKEDOSE</i>	Y
<i>Deposited Dose</i>	PM pollutants only. Total mass deposited in the respiratory system during the hour.	micrograms (µg)	<i>DEPDOSE</i>	Y
<i>Exposure Factor</i>	The ratio of the hourly exposure to the hourly ambient concentration	-	<i>EF</i>	Y
<i>Exposure Factor (Home)</i>	The ratio of the hourly exposure to the hourly ambient concentration in the home district	-	<i>HOMEEF</i>	Y

See *Volume II* for a description of the APEX ventilation algorithms and further information on *VE*, *VA*, *EVR*, and *EE*. The variables *VE*, *VA*, *EVR*, *MET*, and *EE*, and the variables for exposure, dose, and ambient concentration, are the time-weighted averages of the event values for these variables. The ambient concentration is time-averaged over the events because the simulated individual may move between home/work/other locations (and thus possibly between air districts) in the course of an hour. Thus, the hourly ambient concentration may not be equal to the home district AQ data for that hour.

The hourly exposure in microenvironment N is the portion of the total exposure for the hour occurring in microenvironment N, equal to:

$$ExpN = \sum \frac{ConcN * Duration}{60}$$

where ConcN is the concentration in microenvironment N for the event and Duration is the event duration in minutes. A weighted average is used because it is possible for concentrations in a given microenvironment to vary as the person moves between home/work/other locations during the hour. ExpN is the total hourly exposure. The hourly exposure factor, *EF*, is simply the ratio of the hourly exposure to the hourly ambient air concentration.

The variables may be listed in any order in the *Control Options* file using the keyword **HOURLYLIST**; however, they are printed in the output file in the order they appear in the above table. The list can be on a single line or multiple lines, and may be comma or space-delimited. The **EXP**, **DOSE**, **EXPN**, **AMB**, and **EF** keywords control the writing of that variable for all pollutants in the simulation; the file headers for these variables will contain the pollutant name. The dose variables will not be written for a pollutant if it has **DODOSE** = NO in the *Control Options* file, even if a dose keyword is included in the **HOURLYLIST**.

An example regarding the use of the **HOURLYLIST** keyword would be:

```
HourlyList      = EVR VE VA EE MET FEVE1 FEVE2 DFEV
HourlyList      = AMB HOMEAMB EXP EF HOMEEF DOSE INTAKEDOSE DEPDOSE
```

An illustration of a portion of the resulting *Hourly* file for an example two-pollutant run (Pol1 and Pol2) is shown in Exhibit 5-1.

APEX Hourly File									
APEX Version 4.0 (dated February 19, 2007) Run Date = 20070321 Time = 101001									
Location = Description of Location of the Study Area									
Scenario = APEX Sensitivity Simulation									
Simulation = ! APEX Sensitivity Simulation									
Pollutant = Pol1 Pol2									
Air Quality = ! Hourly Pol1 air quality data for an example metropolitan area									
Air Quality = ! Hourly Pol2 air quality data for an example metropolitan area									
P	Hour	Time_4	Time_12	Amb-Poll	Exp-Poll	EF-Poll	Amb-Pol2	Exp-Pol2	EF-Pol2
1	1	60	0	1.039E-02	5.842E-03	0.562	1.132E-02	6.351E-03	0.561
1	2	60	0	9.000E-03	3.295E-03	0.366	9.000E-03	3.351E-03	0.372
1	3	60	0	7.000E-03	2.710E-03	0.387	7.000E-03	2.756E-03	0.394
1	4	60	0	2.000E-03	9.189E-04	0.459	2.000E-03	9.379E-04	0.469
1	5	25	0	7.000E-03	2.297E-03	0.328	7.000E-03	2.333E-03	0.333
1	6	26	0	2.300E-02	8.118E-03	0.353	2.300E-02	8.241E-03	0.358
1	7	60	0	2.100E-02	7.262E-03	0.346	2.100E-02	7.392E-03	0.352
1	8	60	0	1.800E-02	6.180E-03	0.343	1.800E-02	6.291E-03	0.349
1	9	30	0	1.800E-02	7.378E-03	0.410	1.800E-02	7.491E-03	0.416
1	10	12	30	1.900E-02	6.683E-03	0.352	1.900E-02	6.799E-03	0.358
1	11	60	0	2.026E-02	8.227E-03	0.406	2.044E-02	8.427E-03	0.412
...									

**Exhibit 5-1. First Part of an Example APEX Hourly Output File**

Note that the hourly file could be very large if a large number of profiles are simulated. The hourly file is only written if the **HOURLYOUT** parameter is to “YES” in the *Control Options* file.

### 5.3 Timestep File

The *Timestep* file contains the timestep-level time series of a number of APEX variables, including exposure and doses, for each simulated person or profile. The user can control which variables are written to the *Timestep* file via a list of keywords using the *Control Options* file keyword **TIMESTEPLIST**. The variables and their corresponding keywords are listed in Table 5-3.

**Table 5-3. APEX Variables Written to the Timestep Output File**

<b>Variable</b>	<b>Description</b>	<b>Units</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>Person</i>	Simulated profile number	-	-	N
<i>Hour</i>	Hour # of the simulation	-	-	N
<i>Timestep</i>	Timestep # of the simulation	-	-	N
<i>Ve</i>	Ventilation	ml/min	<i>VE</i>	Y
<i>Va</i>	Alveolar ventilation	ml/min	<i>VA</i>	Y
<i>EVR</i>	Equivalent ventilation rate, $V_e$ , divided by body surface area	L/(min-m <sup>2</sup> )	<i>EVR</i>	Y
<i>MET</i>	Metabolic equivalents. Time-averaged multiple of basal energy expenditure for the timestep.	-	<i>MET</i>	Y
<i>EE</i>	Energy expenditure	kcal/min	<i>EE</i>	Y
<i>Ambient Concentration</i>	Ambient pollutant concentration, time-averaged over timestep	<i>INPUTUNITS</i>	<i>AMB</i>	Y
<i>Exposure</i>	Exposure, time-averaged over events in the timestep	<i>OUTPUTUNITS</i>	<i>EXP</i>	Y
<i>Dose</i>	Time-averaged dose for the hour. Units of dose depend on pollutant, see <i>Volume II</i> .	<pollutant-specific>	<i>DOSE</i>	Y
<i>Intake Dose</i>	PM pollutants only. Average mass inhalation rate (includes mass not deposited) over timestep	µg/min	<i>INTAKEDOSE</i>	Y
<i>Deposited Dose</i>	PM pollutants only. Total mass deposited in the respiratory system during the timestep	µg	<i>DEPDOSE</i>	Y
<i>Exposure Factor</i>	The ratio of the timestep exposure to the timestep ambient concentration	-	<i>EF</i>	Y

See *Volume II* for a description of the APEX ventilation algorithms and further information on *VE*, *VA*, *EVR*, and *EE*. The variables *VE*, *VA*, *EVR*, *MET*, and *EE*, and the variables for exposure, dose, and ambient concentration, are the time-weighted averages of the event values for these variables. The ambient concentration is time-averaged over the events because the simulated individual may move between home/work/other locations (and thus possibly between air districts) in the course of a timestep. Thus, the timestep ambient concentration may not be equal to the home district AQ data for that timestep.

The variables may be listed in any order in the *Control Options* file using the keyword *TIMESTEPLIST*, but they are printed in the output file in the order they appear in the above table. The list can be on single or multiple lines and may be comma or space-delimited. The *EXP*, *DOSE*, *AMB*, and *EF* keywords control the writing of that variable for all pollutants in the simulation; the file headers for these variables will contain the pollutant name. However, the dose variables will not be written for a pollutant if it has *DODOSE* = NO in the *Control Options* file, even if a dose keyword is included in the *TIMESTEPLIST*.



An example of the use of the **TIMESTEPLIST** keyword would be:

```
TimestepList      = VE AMB EXP
```

An example of a portion of the resulting *Timestep* file for a one-pollutant run (ozone) is depicted in Exhibit 5-2.

```
APEX Timestep File
APEX Version 4.0 (dated February 21, 2008)  Run Date = 20080227  Time = 111351
Location      = Description of Location of the Study Area
Scenario      = APEX Sensitivity Simulation
Simulation    = ! APEX Sensitivity Simulation
Pollutant     = ozone
Air Quality   = Name =0000200006

      P   Hour Timestep  Ve      Amb-ozone  Exp-ozone
      1   1     1        4858.  3.760E-03  3.760E-03
      1   1     2        5951.  1.027E-02  1.027E-02
      1   1     3        4156.  3.570E-03  3.570E-03
      1   1     4        4949.  8.480E-03  8.480E-03
      1   1     5        5060.  3.680E-03  3.680E-03
      ...
```

**Exhibit 5-2. First Part of an Example APEX Timestep Output File**

Note that the timestep file could be very large if a large number of profiles are simulated or if the APEX timestep is very small. The timestep file is only written if the **TIMESTEPOUT** parameter is to “YES” in the *Control Options* file. Also note that if the APEX timestep is equal to the default (1 hour, or **TIMESTEPSPERDAY**=24), then the *Timestep* file in general would contain the same information as the *Hourly* file, and thus in this case it is not written.

## 5.4 Daily File

The *Daily Exposure* file contains a daily time series of a large number of APEX variables for each simulated person or profile. Writing of the file is controlled by the *Control Options* file variable **DAILYOUT**. The user can control which variables are written to the file via a list of keywords using the *Control Options* file keyword **DAILYLIST**. The variables and their corresponding keywords are listed in Table 5-4.

**Table 5-4. APEX Variables Written to the Daily Output File**

Variable	Description	Units	Control Options File Keyword	Optional
<b>Person</b>	Simulated profile number	-	-	N
<b>Day</b>	Day number of the simulation	-	-	N
<b>Diary ID</b>	ID of CHAD diary selected for the current day for the profile	-	<b>CHADID</b>	Y
<b>Diary Age</b>	Age associated with the selected CHAD diary (may be different from the age of the simulated profile)	years	<b>CHADAGE</b>	Y
<b>Diary Employment</b>	Employment status associated with the selected CHAD diary	-	<b>CHADEMP</b>	Y
<b>Diary Occupation</b>	The occupation of the CHAD diary selected for that day.	-	<b>CHADOCC</b>	Y

<b>Variable</b>	<b>Description</b>	<b>Units</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>Diary pool</i>	Index of the APEX diary pool for the current day (as determined by profile functions file)	-	<b>DIARYPOOL</b>	Y
<i>PAI</i>	Physical activity index, the time-averaged MET over the day for the simulated person	-	<b>PAI</b>	Y
<i>Key Diary Variable</i>	Daily value of the key diary variable (statistic) used for the D&A longitudinal diary assembly for the simulated day for the profile (such as time spent outdoor or in vehicles)	<depends on statistic>	<b>KEYVAR</b>	Y
<i>WindowRes</i>	Conditional variable value indicating whether residence windows are open or closed (as determined by profile functions file)	-	<b>WINDOWRES</b>	Y
<i>WindowCar</i>	Conditional variable value indicating whether car windows are open or closed (as determined by profile functions file)	-	<b>WINDOWCAR</b>	Y
<i>SpeedCat</i>	Conditional variable value indicating the speed at which a vehicle is traveling (as determined by profile functions file)	-	<b>SPEEDCAT</b>	Y
<i>DailyCond1</i>	Value of daily conditional variable 1 (as determined by profile functions file)	-	<b>DCOND1</b>	Y
<i>DailyCond2</i>	Value of daily conditional variable 2 (as determined by profile functions file)	-	<b>DCOND2</b>	Y
<i>DailyCond3</i>	Value of daily conditional variable 3 (as determined by profile functions file)	-	<b>DCOND3</b>	Y
<i>MaxTempCat</i>	Conditional variable giving the category for the maximum temperature for the day (as determined by profile functions file)	-	<b>MAXTEMPCAT</b>	Y
<i>AvgTempCat</i>	Conditional variable giving the category for the average temperature for the day (as determined by profile functions file)	-	<b>AVGTEMPCAT</b>	Y
<i>Maximum Temperature</i>	Maximum hourly temperature for the current day	Fahrenheit	<b>MAXTEMP</b>	Y
<i>Average Temperature</i>	Average of the hourly temperatures for the current day	Fahrenheit	<b>AVGTEMP</b>	Y
<i>Daily Max %<math>\Delta</math>FEV1</i>	Daily maximum of the event-level % $\Delta$ FEV1 calculations	%	<b>DFEV1</b>	Y
<i>Average Exposure</i>	Time-averaged pollutant exposure for the day.	<b>OUTPUTUNITS</b>	<b>AVGEXP</b>	Y
<i>Max 1 Hour Exposure</i>	Maximum 1-hour exposure on the given day; each hourly exposure time-averaged over events.	<b>OUTPUTUNITS</b>	<b>MAX1EXP</b>	Y
<i>Max 8 Hour Exposure</i>	Maximum 8-hour exposure on the given day; each 8-hour exposure time-averaged over events.	<b>OUTPUTUNITS</b>	<b>MAX8EXP</b>	Y

Variable	Description	Units	Control Options File Keyword	Optional
<b>Max 8 Hour Exposure Factor (Home)</b>	The ratio of the maximum 8-hour exposure to the corresponding average ambient concentration in the home microenvironment	-	<b>HOME8MAXEF</b>	Y
<b>Average Dose</b>	Time-averaged pollutant dose for the day. Units of dose depend on pollutant, see <i>Volume II</i> .	<pollutant-specific>	<b>AVGDOSE</b>	Y
<b>Intake Dose</b>	PM pollutants only. Average mass inhaled per minute (includes mass not deposited) during the day	µg/min	<b>INTAKEDOSE</b>	Y
<b>Deposited Dose</b>	PM pollutants only. Total mass deposited in the respiratory system during the day	µg	<b>DEPDOSE</b>	Y
<b>Max 1 Hour Dose</b>	Maximum 1-hour dose on the given day; each hourly dose time-averaged over events	<pollutant-specific>	<b>MAX1DOSE</b>	Y
<b>Max 8 hour Dose</b>	Maximum 8-hour dose on the given day; each 8-hour dose time-averaged over events.	<pollutant-specific>	<b>MAX8DOSE</b>	Y
<b>Max End-of-Hour Dose</b>	Maximum dose as calculated at the end of each hour of the day.	<pollutant-specific>	<b>MAX1FDOSE</b>	Y

See *Volume II* for further information on the diary selection variables and conditional variables on this list. The exposure and dose keywords will control printing for all pollutants in the simulation; the file headers for these variables will contain the pollutant name.

Note that the *Daily* file could be very large if a large number of profiles or pollutants are simulated. The daily file is only written if the **DAILYOUT** parameter is “YES” in the *Control Options* file.

The keywords may be separated by either spaces or commas. An example of a **DAILYLIST** would be:

```
DailyList = CHADID CHADAGE CHADEMP CHADOCC DIARYPOOL PAI KEYVAR WINDOWRES WINDOWCAR AVGEXP
```

An example portion of a *Daily* file created with the **DAILYLIST** example above for a theoretical two-pollutant run (Pol1 and Pol2) is shown in Exhibit 5-3. Note that in the daily file, the values may not fall directly under the corresponding label in the file header (in order to minimize file size).

```

APEX Daily File
APEX Version 4.0 (dated February 19, 2007) Run Date = 20070321 Time = 101001
Location = Description of Location of the Study Area
Scenario = APEX Sensitivity Simulation
Simulation = ! APEX Sensitivity Simulation
Pollutant = Poll Pol2
Air Quality = ! Hourly Poll1 air quality data for an example metropolitan area
Air Quality = ! Hourly Poll2 air quality data for an example metropolitan area

P Day CHADID CHADAge CHADEmp DiaryPool PAI KeyVar WindowRes WindowCar AvgExp-Poll1 AvgExp-Poll2
1 1 NHW19167A 24 Works 2 2.20 540.00 0 0 7.833E-03 1.436E-02
1 2 CIN02759A 21 Works 2 2.41 49.00 0 0 7.748E-03 5.456E-02
1 3 NHW10859A 20 Works 5 2.60 490.00 0 0 9.537E-03 7.764E-02
1 4 NHA16047A 27 Works 5 1.95 0.00 0 0 8.256E-03 8.379E-02
1 5 NHW13255A 24 Works 1 2.43 525.00 0 0 4.343E-03 5.747E-02
1 6 NHW15968A 21 Works 1 1.58 475.00 0 0 6.938E-03 7.345E-02
1 7 NHW12055A 20 Works 1 1.82 600.00 0 0 6.196E-03 3.253E-02
1 8 WAS96832A 25 Works 1 2.24 15.00 0 0 4.372E-03 6.744E-02
1 9 DEN34716B 22 Works 1 3.63 11.00 0 0 6.306E-03 9.222E-02
1 10 CIN80040B 21 Works 4 2.70 390.00 0 0 5.712E-03 2.543E-02
1 11 CIN00339B 24 Works 4 2.46 457.00 0 0 6.089E-03 4.334E-02
1 12 WAS63046A 24 Works 1 2.14 0.00 0 0 6.366E-03 6.435E-02
1 13 CIN61737C 26 Works 2 2.95 91.00 0 0 4.539E-03 6.765E-02
1 14 CAA06251A 21 Works 2 2.37 230.00 0 0 2.629E-03 6.279E-02
...

```

**Exhibit 5-3. First Part of an Example Daily Output File**

## 5.5 Profile Summary (Persons) File

The *Profile Summary (Persons)* file provides a summary of profile characteristics and exposure/dose for each simulated person. Each record contains values for a number of variables for each simulated individual. A small set of variables are written by default to the file, and additional variables are only written if designated by the user in the *Control Options* file. The variables are defined using the *PSUMLIST* keyword, followed by an equals (=) sign and a list of variable-specific keywords. The available variables and their corresponding keywords are provided in Table 5-5.

**Table 5-5. APEX Variables Written to the Profile Summary File**

Variable	Description	Control Options File Keyword	Optional
<i>Person</i>	Sequential index number for simulated individual	-	N
<i>Home Sector</i>	Sector in which the person lives (home)	-	N
<i>Work Sector</i>	Sector in which the person works (=home sector for non-workers, or = 0 for those working outside the study area)	-	N
<i>Home District</i>	Air district for the home sector	-	N
<i>Work District</i>	Air district for the work sector	-	N
<i>Zone</i>	Meteorology zone for the home sector	-	N
<i>Age</i>	Age of the simulated profile (years)	-	N
<i>Gender</i>	Male or female	-	N
<i>Race</i>	e.g., White, Black, Asian, Native American (NatAm), Other (depending on pop. files)	-	N
<i>Employment</i>	Indicates employment outside the home	-	N
<i>Height</i>	Person height (inches)	-	N
<i>Weight</i>	Body mass (pounds)	-	N
<i>Number of Diaries</i>	Number of diaries used	<b>#DIARIES</b>	Y
<i>Group Number</i>	Number of the Profile Factor subgroup	<b>FGROUP</b>	Y
<i>Group Name</i>	Name of the Profile Factor subgroup	<b>GROUPNAME</b>	Y

<b>Variable</b>	<b>Description</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>Roadway District</i>	Air district for the home roadway sector	<b>ROADDIST</b>	Y
<i>Roadway Work District</i>	Air district for the work roadway sector	<b>RWDIST</b>	Y
<i>Commuting Distance</i>	The distance (in km) from the home to the work sector.	<b>COMMDIST</b>	Y
<i>Commuting Time</i>	The estimated time (in minutes) it takes to travel from the home to work sector	<b>COMMTIME</b>	Y
<i>Car AC type</i>	Type of air conditioning in the car (depends of <i>Profile Functions</i> file)	<b>ACCAR</b>	Y
<i>Home AC type</i>	Type of air conditioning in the residence (depends of <i>Profile Functions</i> file)	<b>ACHOM</b>	Y
<i>Disease status</i>	Whether or not a profile is ill (depends on <i>Prevalence</i> file)	<b>DISEASE</b>	Y
<i>Gas Pilot</i>	Indicates the presence of a gas pilot light in the home (depends of <i>Profile Functions</i> file)	<b>PILOT</b>	Y
<i>Gas Stove</i>	Indicates the presence of a gas stove in the home (depends of <i>Profile Functions</i> file)	<b>STOVE</b>	Y
<i>ProfileConditional1</i>	Value of profile conditional variable # 1 for the person	<b>PCOND1</b>	Y
<i>ProfileConditional2</i>	Value of profile conditional variable # 2 for the person	<b>PCOND2</b>	Y
<i>ProfileConditional3</i>	Value of profile conditional variable # 3 for the person	<b>PCOND3</b>	Y
<i>ProfileConditional4</i>	Value of profile conditional variable # 4 for the person	<b>PCOND4</b>	Y
<i>ProfileConditional5</i>	Value of profile conditional variable # 5 for the person	<b>PCOND5</b>	Y
<i>RegionalConditional1</i>	Value of regional conditional variable # 1 for the person	<b>RCOND1</b>	Y
<i>RegionalConditional2</i>	Value of regional conditional variable # 2 for the person	<b>RCOND2</b>	Y
<i>RegionalConditional3</i>	Value of regional conditional variable # 3 for the person	<b>RCOND3</b>	Y
<i>RegionalConditional4</i>	Value of regional conditional variable # 4 for the person	<b>RCOND4</b>	Y
<i>RegionalConditional5</i>	Value of regional conditional variable # 5 for the person	<b>RCOND5</b>	Y
<i>Number of Events</i>	Number of diary events covering the simulation period for the person	<b>EVENTS</b>	Y
<i>Blood Vol</i>	The volume of blood in the body (ml)	<b>BLOODVOL</b>	Y
<i>BSA</i>	Body surface area (m <sup>2</sup> )	<b>BSA</b>	Y
<i>Energy Conversion Factor</i>	Energy conversion factor for person (L-O <sub>2</sub> /kcal)	<b>ECF</b>	Y
<i>Lung Diffusivity</i>	A lung diffusivity parameter used in the COHb (CO dose) calculation (ml/min/torr)	<b>DIFFUS</b>	Y
<i>Endogenous CO production 1</i>	Endogenous CO production rate; only used for calculating CO dose (ml/min)	<b>ENDGN1</b>	Y
<i>Endogenous CO production 2</i>	Endogenous CO production rate for women between ages of 12 and 50 for half the menstrual cycle; only used for calculating CO dose (ml/min)	<b>ENDGN2</b>	Y
<i>Hemoglobin</i>	The amount of hemoglobin in the blood (g/ml)	<b>HEMOGLOB</b>	Y
<i>METmax</i>	Maximum obtainable MET level for the person. (MET)	<b>METMAX</b>	Y

<b>Variable</b>	<b>Description</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>Maximum Oxygen Uptake</i>	Maximum obtainable oxygen uptake rate for person (L-O <sub>2</sub> /min)	<i>VO2MAX</i>	Y
<i>Maximum Oxygen Debt</i>	Maximum obtainable oxygen debt for person (ml-O <sub>2</sub> /kg)	<i>MOXD</i>	Y
<i>Physical Activity Index</i>	Median of the daily PAI values (time-averaged MET on each simulated day)	<i>PAI</i>	Y
<i>Recovery Time</i>	Time required to recover the maximum oxygen debt (hours)	<i>RECTIME</i>	Y
<i>Resting Metabolic Rate</i>	Resting metabolic rate (kcal/min)	<i>RMR</i>	Y
<i>VE Intercept</i>	Regression parameter for the ventilation routine	<i>VEINTER</i>	Y
<i>VE Slope</i>	Regression parameter for the ventilation routine	<i>VERESID</i>	Y
<i>VE Residual</i>	Regression parameter for the ventilation routine	<i>VESLOPE</i>	Y
<i>%ΔFEV1 β1 – β9</i>	Model parameters for the %ΔFEV1 Ozone calculations	<i>B1 – B9 (WRITE EACH ONE SEPARATELY)</i>	Y
<i>%ΔFEV1 Personal Variance</i>	Model parameters for the %ΔFEV1 Ozone calculations	<i>FEVU</i>	Y
<i>%ΔFEV1 age slope</i>	Regression parameters for the age fit for %ΔFEV1 Ozone calculations	<i>FEVSLP</i>	Y
<i>%ΔFEV1 age intercept</i>	Regression parameters for the age fit for %ΔFEV1 Ozone calculations	<i>FEVINT</i>	Y
<i>BMI</i>	Body-mass index (kg/m <sup>2</sup> )	<i>BMI</i>	Y
<i>Average Exposure</i>	Mean exposure concentration over the simulation (ppm, ppb, or μg/m <sup>3</sup> , as specified by <i>OUTPUTUNITS</i> on <i>Control Options</i> file)	<i>AVGEXP</i>	Y
<i>Maximum Exposure</i>	Maximum exposure concentration on one timestep over the simulation (ppm, ppb, or μg/m <sup>3</sup> , as specified in <i>Control Options</i> file)	<i>MAXEXP</i>	Y
<i>Average Dose</i>	Mean dose over the simulation. Units of dose depend on pollutant, see <i>Volume II</i> .	<i>AVGDOSE</i>	Y
<i>Maximum Dose</i>	Maximum 1-hour dose on one timestep over the simulation. Units of dose depend on pollutant, see <i>Volume II</i> .	<i>MAXDOSE</i>	Y
<i>Maximum DFEV1</i>	Personal maximum of DFEV1 lung function loss (for ozone only)	<i>MAXDFEV</i>	Y
<i>Moderate EVR cutoff</i>	Minimum event-level EVR to be considered as moderate exertion	<i>MODEVR</i>	Y
<i>Days with DFEV1&gt;10%</i>	Number of simulation days with more than 10% lung function loss on at least one event (ozone runs only)	<i>NDFEV10</i>	Y
<i>Days with DFEV1&gt;15%</i>	Number of simulation days with more than 15% lung function loss on at least one event (ozone runs only)	<i>NDFEV15</i>	Y
<i>Days with DFEV1&gt;20%</i>	Number of simulation days with more than 20% lung function loss on at least one event (ozone runs only)	<i>NDFEV20</i>	Y

The exposure and dose variables listed are written for all pollutants in a multiple-pollutant run. An example portion of a *Profile Summary* file for a theoretical 2-pollutant (Pol1 and Pol2) scenario is given in Exhibit 5-4. This file was created using the *Control Options* file command:

```
PSumList      =  PAI, AVGEXP, GROUPNAME
```

Note that each record in the file could be much longer, as many more variables could be printed.

```
APEX Diary Questionnaire File
APEX Version 4.0 (dated February 19, 2007)  Run Date = 20070321  Time = 133813
Location      = Description of Location of the Study Area
Scenario      = APEX Sensitivity Simulation
Simulation    = ! APEX Sensitivity Simulation
Pollutant     = Pol1 Pol2
Air Quality   = ! Hourly Pol1 air quality data for an example metropolitan area
Air Quality   = ! Hourly Pol2 air quality data for an example metropolitan area
P  HSect  WSect  HDis  WDis  Zone  Age  Gender  Race  Empl  Height  Weight  PAI  AvgExp-Pol1  AvgExp-Pol2
1  513    513    27   27   2    22  Male   White Works  71.908  228.339  2.09  9.956E-03  1.005E-02
2  64     64     14   14   2    19  Male   Black NoWrk  67.179  138.067  1.74  9.238E-03  9.413E-03
3  359    359    42   42   2    22  Female Other Works  61.018  173.609  1.92  9.749E-03  9.415E-03
4  222    222    39   39   2    15  Male   Black NoWrk  68.519  182.139  1.74  9.100E-03  9.131E-03
5  177    177    39   39   2    20  Female Other NoWrk  65.608  160.464  1.65  8.906E-03  9.377E-03
6  287    287    49   49   2    32  Male   White Works  65.658  155.154  2.04  9.978E-03  1.059E-02
7  688    688    28   28   2    48  Female Black Works  66.264  183.261  1.97  8.873E-03  9.257E-03
8  661    661    23   23   2    50  Female White Works  60.355  106.818  1.93  8.765E-03  8.625E-03
9  280    280    55   55   2    39  Male   Black NoWrk  69.081  209.165  1.76  9.120E-03  9.331E-03
10 793    793    17   17   2    32  Female White NoWrk  65.700  172.692  1.87  1.020E-02  1.041E-02
...

```

**Exhibit 5-4. First Part of an Example Profile Summary File**

## 5.6 Microenvironmental Results File

The *Microenvironmental Results* file contains hourly values for a number of microenvironment parameters and variables for all microenvironments, for all persons in a simulation. The file is pollutant-specific, so one file will be created for each pollutant in the simulation. The variable values are written for the “home,” “work,” “other,” locations, along with the “road,” “road work,” “near home,” and “near work” locations if specified. There is a set of microenvironment concentrations associated with each location for each profile. This file may be useful in examining/testing the effects of conditional values on microenvironment concentrations.

The creation of the file for all pollutants is controlled by the *Control Options* file variable **MRESOUT**. The files are written if **MRESOUT** = YES. The default is “NO”, as these files are very large, and writing them greatly affects the speed of the simulation! The printing of the optional variables is dictated by the *Control Options* file keyword **MRESLIST** via a comma- or space-separated list of variable keywords. The **MRESLIST** will control the writing of the *Microenvironmental Results* file for all of the simulation pollutants. The variables that may be written to the file and their corresponding keywords are provided in Table 5-6.

**Table 5-6. APEX Variables Written to the Microenvironmental Results File**

<b>Variable</b>	<b>Description</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<b>Person</b>	The number of the simulated profile	-	N
<b>Hour #</b>	Hour of the simulation. Hour ranges from -23 to 24 times the number of days in the simulation. The hours -23 to 0 are included because APEX extends the calculation of the microenvironment concentrations to include the 24 hours prior to the beginning of the simulation.	-	N
<b>Micro #</b>	Microenvironment number (See Section 4.18).	-	N
<b>Location</b>	APEX calculates concentrations for each microenvironment including: home (1), work (2), other (3), roadway (4), near home (5), near work (6), road work (7), and near last (99) locations (see <i>Volume II</i> ). Any locations that are used are listed in the file.	-	N
<b>Proximity</b>	Proximity factor: microenvironment parameter, greater than or equal to 0.	<b>PRX</b>	Y
<b>Penetration</b>	Penetration factor: microenvironment parameter, ranging from 0 to 1.	<b>PEN</b>	Y
<b>CSum</b>	Sum of concentration sources (CSource) terms ( <b>INPUTUNITS</b> )	<b>CSUM</b>	Y
<b>Ambient Concentration</b>	Pollutant concentration associated with the location sector and hour as determined from the <i>Air Quality Data</i> file ( <b>INPUTUNITS</b> )	<b>AMB</b>	Y
<b>Micro Concentration</b>	Pollutant concentration in the microenvironment ( <b>INPUTUNITS</b> )	<b>CONC</b>	Y
<b>ESum</b>	Sum of emission sources (ESource) terms (µg/hr)	<b>ESUM</b>	Y
<b>Source Strength</b>	Combined source strength for emission and concentration sources (µg/m <sup>3</sup> /hr)	<b>SOURCE</b>	Y
<b>Micro Volume</b>	Volume of microenvironment (m <sup>3</sup> )	<b>VOL</b>	Y
<b>Air Exchange Rate (AER)</b>	Rate of air exchange in microenvironment (1/hr)	<b>AER</b>	Y
<b>Removal (Decay) Rate</b>	Total removal rate of pollutant from microenvironment (1/hr)	<b>RR</b>	Y
<b>WindowRes</b>	Conditional variable value indicating whether residence windows are open or closed, as determined by profile functions file	<b>WINDOWRES</b>	Y
<b>WindowCar</b>	Conditional variable value indicating whether car windows are open or closed, as determined by profile functions file	<b>WINDOWCAR</b>	Y
<b>MaxTempCat</b>	Daily maximum temperature category conditional variable - will be same for all hours in a day, as determined by profile functions file	<b>MAXTEMPCAT</b>	Y
<b>AvgTempCat</b>	Daily average temperature category conditional variable (will be same for all hours in a day), as determined by profile functions file	<b>AVGTEMPCAT</b>	Y
<b>SpeedCat</b>	Conditional variable value indicating the speed at which a vehicle is traveling, as determined by profile functions file	<b>SPEEDCAT</b>	Y
<b>DailyConditional1</b>	Value of daily conditional variable 1 for the hour	<b>DCONDI1</b>	Y
<b>DailyConditional2</b>	Value of daily conditional variable 2 for the hour	<b>DCOND2</b>	Y
<b>DailyConditional3</b>	Value of daily conditional variable 3 for the hour	<b>DCOND3</b>	Y
<b>TempCat</b>	Hourly temperature category conditional variable	<b>TEMPCAT</b>	Y



<b>Variable</b>	<b>Description</b>	<b>Control Options File Keyword</b>	<b>Optional</b>
<i>HumidCat</i>	Hourly humidity category conditional variable	<i>HUMIDCAT</i>	Y
<i>PrecipCat</i>	Hourly precipitation category conditional variable	<i>PRECIPCAT</i>	Y
<i>WindCat</i>	Hourly wind speed category conditional variable	<i>WINDCAT</i>	Y
<i>DirCat</i>	Hourly wind direction category conditional variable	<i>DIRCAT</i>	Y
<i>Day</i>	Day of the simulation	<i>DAY</i>	Y
<i>Month</i>	Month of the year	<i>MONTH</i>	Y
<i>Air District</i>	Air district person is in (i.e., the district corresponding to the home/work/other location)	<i>DISTRICT</i>	Y
<i>DayWeek</i>	Day of the week	<i>DAYWEEK</i>	Y
<i>AQConditional1</i>	Value of AQ conditional 1 for the time step	<i>AQCOND1</i>	Y
<i>AQConditional2</i>	Value of AQ conditional 2 for the time step	<i>AQCOND2</i>	Y
<i>AQConditional3</i>	Value of AQ conditional 3 for the time step	<i>AQCOND3</i>	Y
<i>AQConditional4</i>	Value of AQ conditional 4 for the time step	<i>AQCOND4</i>	Y
<i>AQConditional5</i>	Value of AQ conditional 5 for the time step	<i>AQCOND5</i>	Y

A number of the parameters in the file are undefined for a **FACTORS** microenvironment (see Section 4.24.1). These parameters will be padded with 0 in that case.

An example of the use of **MRESLIST** in the *Control Options* file is:

```
MResList = AER, PRX, PEN, AMB, CONC, MAXTEMPCAT, AVGTEMPCAT, AQCOND1
```

The resulting example of a *Microenvironmental Results* file for a theoretical pollutant (Poll) is given in Exhibit 5-5. Note that in the file the variable values may not fall directly under the corresponding label in the file header (in order to minimize file size).

```

APEX Microenvironmental Results File
APEX Version 4.0 (dated February 19, 2007) Run Date = 20070321 Time = 152320
Location = Description of Location of the Study Area
Scenario = APEX Sensitivity Simulation
Simulation = ! APEX Sensitivity Simulation
Pollutant = Poll
Air Quality = ! Hourly Poll air quality data for an example metropolitan area
Simulation Start Date = 20040101

Person Micro Loc Hour Prx Pen Amb Conc AER WindowRes MaxTempCat AvgTempCat TempCat
1 1 1 -23 1.0000 1.0000 8.000E-03 4.227E-04 0.4104 0 1 1 1
1 1 1 -22 1.0000 1.0000 9.000E-03 6.518E-04 0.5042 0 1 1 1
1 1 1 -21 1.0000 1.0000 7.000E-03 7.044E-04 0.6962 0 1 1 1
1 1 1 -20 1.0000 1.0000 2.000E-03 2.335E-04 0.5010 0 1 1 1
1 1 1 -19 1.0000 1.0000 7.000E-03 1.645E-04 0.1511 0 1 1 1
1 1 1 -18 1.0000 1.0000 2.300E-02 1.712E-03 0.5822 0 1 1 1
1 1 1 -17 1.0000 1.0000 2.100E-02 1.922E-03 0.6193 0 1 1 1
1 1 1 -16 1.0000 1.0000 1.800E-02 1.398E-03 0.4841 0 1 1 1
1 1 1 -15 1.0000 1.0000 1.800E-02 1.207E-03 0.4370 0 1 1 1
1 1 1 -14 1.0000 1.0000 1.900E-02 1.192E-03 0.4138 0 1 1 1
1 1 1 -13 1.0000 1.0000 2.100E-02 1.496E-03 0.4923 0 1 1 1
1 1 1 -12 1.0000 1.0000 2.400E-02 3.147E-03 1.0217 0 1 1 1
1 1 1 -11 1.0000 1.0000 2.800E-02 3.395E-03 0.8523 0 1 1 1
1 1 1 -10 1.0000 1.0000 3.000E-02 3.091E-03 0.6968 0 1 1 1
1 1 1 -9 1.0000 1.0000 3.200E-02 2.390E-03 0.4730 0 1 1 1
1 1 1 -8 1.0000 1.0000 3.000E-02 1.132E-03 0.1955 0 1 1 1
1 1 1 -7 1.0000 1.0000 2.800E-02 1.395E-03 0.3449 0 1 1 1
1 1 1 -6 1.0000 1.0000 2.600E-02 1.487E-03 0.3753 0 1 1 1
1 1 1 -5 1.0000 1.0000 2.200E-02 1.440E-03 0.4299 0 1 1 1
1 1 1 -4 1.0000 1.0000 2.300E-02 1.960E-03 0.6052 0 1 1 1
1 1 1 -3 1.0000 1.0000 2.600E-02 2.610E-03 0.7176 0 1 1 1
1 1 1 -2 1.0000 1.0000 2.300E-02 1.683E-03 0.4316 0 1 1 1
1 1 1 -1 1.0000 1.0000 2.300E-02 1.279E-03 0.3517 0 1 1 1
1 1 1 0 1.0000 1.0000 2.200E-02 1.226E-03 0.3640 0 1 1 1
1 1 1 1 1.0000 1.0000 8.000E-03 6.081E-04 0.4104 0 1 1 1
...

```

**Exhibit 5-5. First Part of an Example Microenvironmental Results File**

## 5.7 Microenvironmental Summary File

The *Microenvironmental Summary* file provides the amount of time spent, mean exposure concentration, and maximum exposure concentration within each microenvironment during the period of simulation, for each simulated person. The *Microenvironmental Summary* file is pollutant-specific, and thus one is created for each pollutant in the simulation. After the six header records and one blank record, there is one record labeling the columns of the subsequent records in the file. These labels and descriptions of the values in the corresponding columns are given in Table 5-7. The first part of an example *Microenvironmental Summary* file is shown in Exhibit 5-6.

**Table 5-7. Format of the APEX Microenvironmental Summary File**

Column	Variable	Type	Description
1	<i>Person</i>	Num	Profile number—sequential index number for the simulated individual
2	<i>Micro</i>	Num	Microenvironment number—Sequential index number for each microenvironment (as designated in the <i>Microenvironment Mapping</i> file)
3	<i>Name</i>	Char	Microenvironment name (as designated in the <i>Microenvironment Mapping</i> file) (maximum of 40 characters)
4	<i>Minutes</i>	Num	Total time spent in the microenvironment by this individual (minutes)

Column	Variable	Type	Description
5	<i>Meanconc</i>	Num	Average concentration during the time spent in the microenvironment by this individual (ppm, ppb, or $\mu\text{g}/\text{m}^3$ , as specified by <i>INPUTUNITS</i> in the <i>Control Options</i> file)
6	<i>Maxconc</i>	Num	Maximum concentration during the time spent in the microenvironment by this individual (ppm, ppb, or $\mu\text{g}/\text{m}^3$ , as specified by <i>INPUTUNITS</i> in the <i>Control Options</i> file)

```

APEX Microenvironmental Summary File
APEX Version 5.0 (dated July 25, 2017)  Run Date = 20191028  Time = 152729
Location      = Example Metropolitan Area
Scenario      = Example Metropolitan Area, Jun-Aug 2010, 2010 pop, RandomSeed 4178348
Simulation    = Example Metropolitan Area
Pollutant     = Ozone
Air Quality   = ! 2010 base ozone AQ data for Example Metropolitan Area
  Person Mic Minutes   MeanConc   MaxConc
  1      0      0       0.000E+00  0.000E+00
  1      1    70969     5.913E-03  3.262E-02
  1      2    17738     3.173E-02  8.904E-02
  1      3    25779     4.011E-02  9.600E-02
  1      4   10541     3.759E-02  7.800E-02
  1      5    7453     3.702E-02  9.600E-02
  2      0      0       0.000E+00  0.000E+00

```

**Exhibit 5-6. First Part of an Example Microenvironmental Summary File**

## 5.8 Output Tables File

The *Output Tables* file can provide hundreds of summary tables, depending on the table specifications in the *Control Options* file. The *Tables* file is pollutant-specific; thus one is created for each pollutant in the simulation. APEX users should specify which tables to print by using the *TABLESLIST* variable in the *Control Options* file. This parameter takes the arguments listed below.

- ***EXPIH, EXP8H, EXPTS, EXPAVG***: prints the tables pertaining to the 1-hour, 8-hour, time-step maxima, and average daily exposures (*OUTPUTUNITS*)
- ***DOSE1H, DOSE8H, DOSE1EH, DOSETS, DOSEAVG***: prints the tables pertaining to the 1-hour, 8-hour, time-step, 1-hour end-of-hour maxima, and average daily doses. Dose units depend on the pollutant.
- ***MICROTIME***: prints the time spent in each microenvironment (minutes).
- ***DOSETIME***: prints the time spent in each dose category (minutes).
- ***CHILDREN***: prints additional tables that include only children.
- ***ACTIVE, ACTCHILD***: prints tables including active individuals, or the subset of active children.
- ***ILLNESS, ILLCHILD***: prints tables that includes the subgroup of those with the user-specified illness, for all simulated individuals or children only.
- ***MOD, HEAVY***: prints the tables for individuals with exposures in the moderate or heavy *EVR* categories.
- ***EMPLOYED***: printed tables for all employed persons

### 5.8.1 Exposure Summary Tables

APEX can write out well over 100 different exposure summary tables for each pollutant. There are 11 different types of exposure tables, listed below.

1. Minutes in each exposure interval by microenvironment
2. Minutes at or above each exposure level by microenvironment
3. Person-days at or above each daily maximum 1-hour exposure level
4. Person-days at or above each daily maximum 8-hour exposure level
5. Person-days at or above each daily maximum timestep exposure level
6. Number of simulated persons with multiple exposures at or above each daily maximum 1-hour exposure level
7. Number of simulated persons with multiple exposures at or above each daily maximum 8-hour exposure level
8. Number of simulated persons with multiple exposures at or above each daily maximum timestep exposure level.
9. Number of simulated persons with multiple exceedances (in the simulation) of the threshold timestep exposure levels.
10. Person-days at or above each daily average exposure level
11. Number of persons at or above each overall average exposure level

Table types 1, 2, 10, and 11 are generated only once for the entire population. Table types 3 to 9 are generated for seven population subgroups, under three exertion levels. Tables may be omitted if the subgroup contains no simulated persons.

The seven population subgroups are listed below.

1. All Persons. The table statistics are based on the entire population.
2. Children. The table statistics are based on the population of children, as defined by the age range given by the *Control Options* file settings **CHILDMIN** and **CHILDMAX**.
3. Active Persons. The table statistics are based on the population of people having a median **PAI** over the whole simulation period that exceeds the value designated by the *Control Options* file setting **ACTIVEPAI**.
4. Active Children. The table statistics are based on the population of active children, as determined by the *Control Options* file settings **CHILDMIN**, **CHILDMAX**, and **ACTIVEPAI**.
5. Ill Persons. The table statistics are based on the population of ill people. The population is determined by the probabilities given in the *Prevalence* file. This population is only considered if the input variable **DISEASE** is set in the *Control Options* file.
6. Ill Children. The table statistics are based on the population of ill people. The population is determined by the probabilities given in the *Prevalence* file and the *Control Options* file settings **CHILDMIN** and **CHILDMAX**. This population is only considered if the input setting **DISEASE** is set in the *Control Options* file.
7. Employed Persons. The table statistics are based on the population of all employed people.

The three exertion levels are listed below.

1. All Exertion Conditions. The table statistics are based on exposures experienced by the population subgroup under any ventilation conditions.
2. Moderate Exertion. The table statistics are based on exposures experienced by the population subgroup only during periods in which their average *EVR* is in the “moderate” range. The period of time during which *EVR* is averaged is either 1 hour or 8 hours, based on the table being generated. The “moderate” *EVR* ranges are defined by the *Control Options* file settings *MODEVR1* and *HEAVYEVRI* (for 1-hour exposures) and *MODEVR8* and *HEAVYEVRI8* (for 8-hour exposures). An individual’s *EVR* is in the moderate range if it is greater than or equal to the *MODEVR#* setting and less than the *HEAVYEVRI#* setting for the exposure period.
3. Heavy Exertion. The table statistics are based on exposures experienced by the population subgroup only during periods in which their average *EVR* is in the “heavy” range. The period of time during which *EVR* is averaged is either 1 hour or 8 hours, based on the table being generated. The “heavy” *EVR* ranges are defined by the *Control Options* file settings *HEAVYEVRI* (for 1-hour exposures) and *HEAVYEVRI8* (for 8-hour exposures). An individual’s *EVR* is in the heavy range if it is greater than or equal to the *HEAVYEVRI#* setting for the exposure period.

For each table that is generated, APEX prints out a label that identifies the table uniquely. For example, a table of type #1, for all people under all exertion conditions, has the identifier *TIME, WITHIN, ALL, ALL*. Users can reference these identifier labels in custom programs that read in and process the APEX *Tables* file.

### **Exposure Table Type #1: Minutes in each exposure interval by microenvironment**

This table lists the total minutes spent by all simulated persons in each microenvironment when exposure concentration is within various ranges. The bounds of a range are specified at the top of each column and the top of the next column to the right (Exhibit 5-7). For each microenvironment, the table provides three rows of data for the three variables defined below.

- **Minutes**: The number of person-minutes summed over all the simulated persons that are spent in the specified microenvironment and that fall within the exposure concentration range bounded by the values indicated at the top of the column and the top of the next column to the right;
- **Row\_ %**: The percent of the minutes spent in the specified microenvironment that fall within the exposure concentration range; and
- **Tot\_ %**: the percent of the total minutes that are spent in the microenvironment and that fall within the exposure concentration range.

TIME, WITHIN, ALL, ALL, ALL								
Exposure: Minutes in each Exposure interval ( ppm ), by microenvironment, for N = 100 Profiles								
Level:		0.0000	0.0100	0.0200	0.0300	0.0400	0.0500	0.0600
Micro		-----						
0	Minutes	8041.	0.	0.	0.	0.	0.	0.
0	Row_%	100.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0	Tot_%	0.0607	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	Minutes	6815807.	1525776.	397884.	137206.	55886.	20271.	4674.
1	Row_%	76.0842	17.0321	4.4415	1.5316	0.6239	0.2263	0.0522
1	Tot_%	51.4478	11.5170	3.0034	1.0357	0.4218	0.1530	0.0353
2	Minutes	116515.	281146.	456343.	559532.	419900.	282998.	170243.
2	Row_%	4.8170	11.6232	18.8662	23.1323	17.3596	11.6998	7.0382
2	Tot_%	0.8795	2.1222	3.4446	4.2235	3.1695	2.1362	1.2850
3	Minutes	30586.	57073.	108273.	158511.	177594.	162320.	118771.
3	Row_%	3.2347	6.0358	11.4506	16.7636	18.7817	17.1664	12.5608
3	Tot_%	0.2309	0.4308	0.8173	1.1965	1.3405	1.2252	0.8965
4	Minutes	7837.	11971.	20586.	29189.	26537.	29647.	21117.
4	Row_%	4.7583	7.2683	12.4990	17.7224	16.1122	18.0005	12.8214
4	Tot_%	0.0592	0.0904	0.1554	0.2203	0.2003	0.2238	0.1594
5	Minutes	53028.	69107.	87928.	90228.	106943.	136538.	106956.
5	Row_%	7.0458	9.1822	11.6830	11.9886	14.2095	18.1417	14.2112
5	Tot_%	0.4003	0.5216	0.6637	0.6811	0.8072	1.0306	0.8073

**Exhibit 5-7. Example of Exposure Table Type #1 in the Output Tables File**

**Exposure Table Type #2: Minutes in each exposure interval by microenvironment**

This table is similar to Table Type #1, except that it reports the cumulative person-minutes that are spent in a microenvironment with an exposure concentration that equals or exceeds the value indicated at the top of the column.

**Exposure Table Type #3: Person-days at or above each daily maximum 1-hour exposure level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, with a daily maximum 1-hour (hourly) average exposure concentration that equals or exceeds the value indicated at the top of the column (Exhibit 5-8). The interpretations of the variables in Table Type #3 (and other “person-days” tables) are provided in Table 5-8.

PERSONDAYS, DM1H, ALL, ALL, ALL								
Exposure: Person-Days at or above each Daily Maximum 1-Hour Exposure Level ( ppm ), for N = 100								
Profiles. Area Population = 16519458								
Group: All People								
Level:	0.00000	0.01000	0.02000	0.03000	0.04000	0.05000	0.06000	0.07000
Counts (Pop):	1.520E+09	1.491E+09	1.450E+09	1.397E+09	1.280E+09	1.019E+09	6.533E+08	3.507E+08
#Meet (Pop):	16519458	16519458	16519458	16519458	16519458	16519458	16519458	16519458
%Meet (Pop):	100.000	100.000	100.000	100.000	100.000	100.000	100.000	100.000
Mean	: 92.000	90.280	87.780	84.590	77.510	61.690	39.550	21.230
Std.Dev.	: 0.000	3.364	5.809	8.468	11.053	14.730	15.981	13.357
CV	: 0.000	0.037	0.066	0.100	0.143	0.239	0.404	0.629
Minimum	: 92.000	70.000	62.000	55.000	45.000	26.000	14.000	3.000
10.0 %ile	: 92.000	87.100	79.100	72.000	59.400	39.300	19.000	8.000
25.0 %ile	: 92.000	90.000	86.250	83.000	74.000	53.000	29.000	12.250
50.0 %ile	: 92.000	91.500	90.000	87.000	80.000	62.500	38.000	17.500
75.0 %ile	: 92.000	92.000	92.000	90.000	86.000	72.750	49.000	27.750
90.0 %ile	: 92.000	92.000	92.000	91.900	89.000	81.900	65.000	40.000
95.0 %ile	: 92.000	92.000	92.000	92.000	90.000	85.000	72.000	52.900
99.0 %ile	: 92.000	92.000	92.000	92.000	91.990	90.970	85.950	67.900
Maximum	: 92.000	92.000	92.000	92.000	92.000	91.000	86.000	68.000
Mean (%)	: 100.000	98.130	95.413	91.946	84.250	67.054	42.989	23.076
Min (%)	: 100.000	76.087	67.391	59.783	48.913	28.261	15.217	3.261
Max (%)	: 100.000	100.000	100.000	100.000	100.000	98.913	93.478	73.913
Counts (Sim):	9.200E+03	9.028E+03	8.778E+03	8.459E+03	7.751E+03	6.169E+03	3.955E+03	2.123E+03
#Meet (Sim):	100	100	100	100	100	100	100	100

**Exhibit 5-8. Example of Exposure Table Type #3 in the Output Tables File**

**Table 5-8. Interpretation of the Variables in Exposure Table Type #3 and Other "Person-Days" Based Tables**

Table Entry	Interpretation
<b>Counts (Pop)</b>	Total number of person-days at or above the level specified at the top of each column for the population (of the subgroup) in the study area (while at this exertion).
<b>#Meet (Pop)</b>	Number of persons (in the subgroup) in the study area population who have at least one exposure at or above the level specified at the top of each column (while at this exertion). NOTE: For exertion level tables, the 0.0 level count will not necessarily be equal to the population of the subgroup, since some persons may have no events at the given exertion level.
<b>%Meet (Pop)</b>	Percentage of people (in the subgroup) in the population who have at least one exposure at or above the level specified at the top of each column (while at this exertion). NOTE: For exertion level tables this may not be 100% at the 0.0 level, since some persons may have no events at the exertion level.
<b>Mean</b>	Mean number of days per person (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).
<b>Std. Dev.</b>	Standard deviation across persons (in the subgroup) in the number of days during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).
<b>CV</b>	Coefficient of variation across persons (in the subgroup) in the number of days during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).
<b>Minimum</b>	The lowest total number of days across persons (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).
<b>Percentiles</b>	The Nth percentile of number of days across persons (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).
<b>Maximum</b>	The highest total number of days across persons (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion).

<b>Table Entry</b>	<b>Interpretation</b>
<b><i>Mean (%)</i></b>	Mean number of days per person (in the subgroup) during which an exposure at or above the level specified at the top of each column (while at this exertion) is experienced, as percentage of possible days.
<b><i>Min (%)</i></b>	The lowest total number of days across persons (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion), as percentage of possible days.
<b><i>Max (%)</i></b>	The highest total number of days across persons (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion), as percentage of possible days.
<b><i>Counts (Sim)</i></b>	Total number of simulated person-days (in the subgroup) during which an exposure at or above the level specified at the top of each column is experienced (while at this exertion). NOTE: At the 0.0 level in the exertion-dependent tables, Counts (Sim) might not necessarily be equal to #Meet (Sim)*NumDays, since some persons may have no events at the exertion level.
<b><i>#Meet (Sim)</i></b>	The total number of simulated persons (in the subgroup) who experience at least one exposure at or above the level specified at the top of each column (while at this exertion).

#### **Exposure Table Type #4: Person-days at or above each daily maximum 8-hour exposure level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, with a daily maximum 8-hour average exposure concentration that equals or exceeds specified levels. The table and its interpretation are the same as Table Type #3 (Exhibit 5-8) except that the exposure metric is the daily max 8-hour average exposure concentration.

#### **Exposure Table Type #5: Person-days at or above each daily maximum timestep exposure level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, with a daily maximum timestep average exposure concentration that equals or exceeds specified levels. The table and its interpretation are the same as Table Type #3 (Exhibit 5-8) except that the exposure metric is the daily max timestep average exposure concentration. This table is not written if the timestep is equal to one hour.

#### **Exposure Table Type #6: Number of simulated persons with multiple exposures at or above each daily maximum 1-hour exposure level**

This table simply provides a count of the number of simulated persons who have at least 1 (2, 3, 4, 5, 6) days in the simulation during which they have experienced an exposure above each of the daily maximum 1-hour exposure levels. An example is shown in Exhibit 5-9.



```

MULTIPLE, DM1H, ALL, ALL, ALL
Exposure: Number of Simulated Persons with Multiple Exposures at or above each Daily Maximum 1-
Hour Exposure Level ( ppm ), for N = 100 Profiles.
Group: All People

```

Level	At least 1 Exposure	At least 2 Exposures	At least 3 Exposures	At least 4 Exposures	At least 5 Exposures	At least 6 Exposures
0.000	100	100	100	100	100	100
0.010	100	100	100	100	100	100
0.020	100	100	100	100	100	100
0.030	100	100	100	100	100	100
0.040	100	100	100	100	100	100
0.050	100	100	100	100	100	100
0.060	100	100	100	100	100	100
0.070	100	100	100	97	97	95
0.080	97	95	84	75	67	61
0.090	85	66	50	37	30	26
0.100	54	33	19	12	7	5
0.110	38	15	2	1	0	0
0.120	4	1	0	0	0	0
0.130	1	0	0	0	0	0
0.140	0	0	0	0	0	0

**Exhibit 5-9. Example of Exposure Table Type #6 in the Output Tables File**

**Exposure Table Type #7: Number of simulated persons with multiple exposures at or above each daily maximum 8-hour exposure level**

This table simply provides a count of the number of simulated persons who have at least 1 (2, 3, 4, 5, 6) days in the simulation during which they have experienced an exposure above each of the daily maximum 8-hour exposure levels. The table is the same as Table Type #6 (Exhibit 5-9) except that the exposure metric is the daily max 8-hour average exposure concentration.

**Exposure Table Type #8: Number of simulated persons with multiple exposures at or above each daily maximum timestep exposure level.** This table provides a count of the number of simulated persons who have at least 1 (or 2, 3, 6, etc.) days in the simulation during which they have experienced an exposure above each of the daily maximum timestep exposure levels. The table is the same as Table Type #6 (Exhibit 5-9) except that the exposure metric is the daily max timestep average exposure concentration. This table is not written if the timestep is equal to one hour.

**Exposure Table Type #9: Number of simulated persons with multiple exposures at or above some threshold timestep exposure level.** This table provides a count of the number of simulated persons who have at least 1 (or 2, 30, or 300, for example) timesteps in the entire simulation during which they have experienced an exceedance of each timestep threshold exposure level. The different number of exceedances to include in the table are listed in the *Control Options* file using the keyword *TSMULTILEVELS*. The threshold exposures are listed using the keyword *TSEXP*. This table is not written if the timestep is equal to one hour.

**Exposure Table Type #10: Person-days at or above each daily average exposure level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, with a daily average exposure concentration that

equals or exceeds specified levels. The table and its interpretation are the same as Table Type #3 (Exhibit 5-8) except that the exposure metric is the daily average exposure concentration.

**Exposure Table Type #11: Persons at or above each overall average exposure level**

This table provides a statistical summary of cumulative numbers of both simulated persons and people in the study area whose overall average exposure concentrations equal or exceed specified levels. The overall average exposure concentration is the average of hourly exposure concentrations over the whole period of simulation. An example of this table is provided in Exhibit 5-10.

PERSONDAYS, SAVG, ALL, ALL							
Exposure: Persons at or above each Overall Average Exposure Level (ppm), for N = 100							
Profiles. Area Population = 3976069							
Level:	0.000	0.500	1.000	2.000	3.000	4.000	5.000
Counts (Pop):	0.398E+07	0.398E+07	0.392E+07	0.386E+06	0.000E+00	0.000E+00	0.000E+00
#Meet (Pop):	3976069	3976069	3916428	385679	0	0	0
%Meet (Pop):	100.000	100.000	98.500	9.700	0.000	0.000	0.000
Counts (Sim):	0.100E+04	0.100E+04	0.985E+03	0.970E+02	0.000E+00	0.000E+00	0.000E+00
# Meet (Sim):	1000	1000	985	97	0	0	0

**Exhibit 5-10. Example of Exposure Table Type #11 in the Output Tables File**

**5.8.2 Dose Summary Tables**

APEX can write out over 100 different dose summary tables for each pollutant. There are 10 different types of dose summary tables. The contents of each table type are described in detail below. Table Types #7–10 are generated only once, for the entire population. Table Types #1–6 are each generated for six population subgroups, under three exertion levels. See the previous section on exposure tables for the definition of population subgroups and exertion levels. For the pollutant CO, dose is blood dose (%COHb), and for any PM pollutant the dose is the rate of mass deposited in the respiratory system in µg/min (See *Volume II*). For all other pollutants dose is simply exposure\*ventilation.

**Dose Table Type #1— Person-days at or above each daily max end-of-hour dose level**

This table provides a statistical summary of the cumulative person-days for both simulated persons and the population in the study area, for which the daily maximum end-of-hour dose is equal to or exceeds specified levels. The format of the table is the same as Exposure Table Type #3 (Exhibit 5-8).

**Dose Table Type #2— Person-days at or above each daily max 1-hour dose level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, for which the daily maximum 1-hour average dose is equal to or exceeds specified levels. The format of the table is the same as Exposure Table Type #3 (Exhibit 5-8). The definitions of the variables in this table can be found in Table 5-8.

### **Dose Table Type #3—Person-days at or above each daily max 8-hour dose level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, for which the daily maximum 8-hour average dose is equal to or exceeds specified levels. The format of the table is the same as Exposure Table Type #3 (Exhibit 5-8). The definitions of the variables in this table can be found in Table 5-8.

### **Dose Table Type #4: Person-days at or above each daily maximum timestep dose level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, with a daily maximum timestep average dose that equals or exceeds specified levels. The table and its interpretation are the same as Exposure Table Type #3 (Exhibit 5-8), except that the exposure metric is the daily maximum timestep average exposure concentration. This table is not written if the timestep is equal to one hour. The definitions of the variables in this table can be found in Table 5-8.

### **Dose Table Type #5: Number of simulated persons with multiple timestep doses at or above some threshold timestep dose level.**

This table provides a count of the number of simulated persons who have at least 1 (or 2, 30, or 300, for example) timesteps in the entire simulation during which they have experienced an exceedance of each timestep threshold dose level. The different number of exceedances to include in the table are listed in the *Control Options* file using the keyword **TSMULTILEVELS**. The threshold exposures are listed using the keyword **TSDOSE**. This table is not written if the timestep is equal to one hour.

### **Dose Table #6— Person-days at or above each daily average dose level**

This table provides a statistical summary of the cumulative person-days, for both simulated persons and the population in the study area, for which the daily average dose is equal to or exceeds specified levels. The format of the table is the same as Exposure Table Type #3 (Exhibit 5-8). The definitions of the variables in this table can be found in Table 5-8.

### **Dose Table #7— Persons at or above each overall average dose level**

This table provides a statistical summary of cumulative numbers of both simulated persons and the people in the study area whose overall average doses are equal to or exceed a specified level. The overall average dose is the average of hourly dose levels during the entire simulation period.

### **Dose Table #8—Person-hours at or above each end-of-hour dose level**

This table provides a statistical summary of the number of person-hours, for both simulated persons and the population in the study area, for which each end-of-hour dose level is equal to or exceeds specified levels. The format of the table is the same as Exposure Table Type #3 (Exhibit 5-8), except that the time units are hours rather than days. The definitions of the variables in this table can be found in Table 5-8.

### Dose Table #9— Minutes in each dose interval

This table provides a statistical summary of cumulative person-minutes, for both simulated persons and the population in the study area, for which the dose (for example, blood %COHb level) is within a specified range. The bounds of the dose range are specified by the levels at the top of each column and the top of the next column to the right. The definitions of the variables in this table are similar to those found in Table 5-8, except that the time units are in minutes rather than days.

### Dose Table #10— Minutes at or above each dose level

This table provides a statistical summary of cumulative person-minutes spent by both simulated persons and the population in the study area, for which the dose (for example, blood %COHb level) is equal to or exceeds specified levels. The definitions of the variables in this table are similar to those found in Table 5-8, except that the time units are in minutes rather than days.

### 5.8.3 ResponseProb Summary Tables

This table type provides the number of simulated persons estimated to experience the risk endpoint, e.g., lung function decrements over selected thresholds, at various pollutant exposure levels. The exposure statistic that the exposure-response function is applied to is indicated by the second item in the first line of the table header; in Exhibit 5-11 it is the daily maximum 8-hour (*DM8H*) level.

ResponseProb, DM8H, ALL, MOD, ALL								
FEV1 Decrement: Number of persons affected in each exposure bin, for N=						1000 profiles.	Area population = 4919330	
Group: All Ages, Moderate Exertion (EVR 13.00-99.00)								
Level:		0.0000	0.0100	0.0200	...	#Sim	Frac.Sim	#Pop
FEV1>10, 2.5 pctl :		0.0000	0.0000	0.0000	...	1.510	0.1510E-02	7428.
FEV1>10, 50.0 pctl :		0.0129	0.1416	0.3730	...	3.561	0.3561E-02	0.1752E+05
FEV1>10, 97.5 pctl :		0.0323	0.3234	0.7746	...	6.548	0.6548E-02	0.3221E+05
FEV1>15, 2.5 pctl :		0.0000	0.0000	0.0000	...	0.196	0.1960E-03	964.4
FEV1>15, 50.0 pctl :		0.0024	0.0281	0.0812	...	1.008	0.1008E-02	4961.
FEV1>15, 97.5 pctl :		0.0089	0.0931	0.2304	...	1.983	0.1983E-02	9756.
FEV1>20, 2.5 pctl :		0.0000	0.0000	0.0000	...	0.0561	0.5612E-04	276.1
FEV1>20, 50.0 pctl :		0.536E-04	0.853E-03	0.0350	...	0.1682	0.1682E-03	827.3
FEV1>20, 97.5 pctl :		0.869E-03	0.0105	0.0308	...	0.4756	0.4756E-03	2340.
Bin counts :		38.00	86.00	80.00	...	314.0	0.3140	0.1545E+07

**Exhibit 5-11. Portion of ResponseProb Table**

Columns for higher ozone levels in Exhibit 5-11 have been removed for clarity, as indicated by ellipses. The label before the colon in each line is user-supplied. Each value is the product of a probability and a person count. The last row labeled “Bin counts” has all probabilities set to one to show the number of persons in each column. The three rows for each FEV1 cut-point correspond to different estimates of the associated probability. For example, of the 38 persons in the first column, 0.0129 are expected to suffer a loss of over 10% in FEV1 lung function, using the median probability. If the ozone levels went high enough, the entries would all become zero because the person counts would be zero. That is, even if the risk factor were high, if no one

experiences that particular ozone level, then no one can have lung function loss due to ozone at that level.

The “#Sim” column is the sum of the previous columns, representing the number of affected persons in the simulation. In the above example there were 314 persons in the moderate exertion category, of whom 3,561 would be expected to lose over 10% lung function using the median probabilities. The final column extrapolates these results to the total population of the study area. In this case, 1.545 million persons would be in the moderate exertion category, and 17,520 would lose over 10% lung function using the median probabilities.

## 5.9 Sites File

The *Sites* output file lists the sectors, air districts, and zones in the study area, and identifies the mapping between them. Thus, each record contains the information shown below.

- **Sector#**: sector ID
- **Latitude**: sector latitude (decimal degrees)
- **Longitude**: sector longitude (decimal degrees)
- **SectorName**: sector name
- **Air#**: air district ID
- **AirDistance**: distance from air district to sector (km)
- **AirLatitude**: air district latitude (decimal degrees)
- **AirLongitude**: air district longitude (decimal degrees)
- **AirName**: air district name
- **Met#**: meteorology zone ID
- **MetDistance**: distance from meteorology zone to sector (km)
- **MetLatitude**: meteorology zone latitude (decimal degrees)
- **MetLongitude**: meteorology zone longitude (decimal degrees)
- **MetName**: meteorology zone name

## 5.10 Events File

The *Events* file contains a summary of the activity diary, with accompanying exposure and dose, at the diary event level. The variables printed in this file include those listed below.

- **Person**: the profile number of the simulated individual
- **Seq**: the event number for the profile
- **Day**: the day number of the simulation, incremented from Day 1 of simulation
- **Year**: the year of the event (4-digit)
- **Mn**: the month of the event (1 to 12)
- **Dy**: the day of the week the event (1 to 7)
- **Hr**: the hour of the event (1 to 24)
- **Dur**: the duration of the event (integer minutes)
- **Act**: the MET distribution code for the event activity
- **Mic**: the microenvironment code for the event
- **HW**: 1 = event in home sector, 2 = event in work sector, 3 = elsewhere

- **Exposure:** exposure level (concentration in the microenvironment) during the event (ppm, ppb or  $\mu\text{g}/\text{m}^3$ , as indicated by **OUTPUTUNITS**)

Optionally, the user can ask APEX to include the variables (bulleted below) by using the keyword **EVENTSLIST**, and then listing each of the variable names to be included (shown below).

- **MET:** MET level for the event (unitless)
- **UMET:** unmodified MET level for the event (unitless)
- **VA:** alveolar ventilation during the event (ml/min)
- **VE:** ventilation during the event (ml/min)
- **EVR:** equivalent ventilation rate
- **DFEVI:** end-of-event percent decrement in the forced expiratory volume in 1 sec
- **FEVX:** end of event X term used in the  $\%\Delta\text{FEV1}$  calculation
- **Deficit:** oxygen debt, percent of nominal
- **AMB:** (or “ambient”) the ambient concentration during the event (ppm, ppb or  $\mu\text{g}/\text{m}^3$ , as indicated by **OUTPUTUNITS**)
- **HomeAmb:** the ambient concentration in the home district during the event (ppm, ppb or  $\mu\text{g}/\text{m}^3$ , as indicated by **OUTPUTUNITS**)
- **FEVE1:** the intra-individual variability term associated with the  $\%\Delta\text{FEV1}$  model
- **FEVE2:** the E2 error term associated with the  $\%\Delta\text{FEV1}$  model

If **DODOSE** = Yes in the *Control Options* file, then two variables related to dose will be printed for all cases, and an additional two will be printed for PM exposure (shown below).

- **Dose:** average dose over the event
- **FDose:** final dose for the event
- **DepDose:** deposited mass dose for PM events
- **IntakeDose:** intake dose rate for the PM event

The distinction between average dose and final dose is only relevant for pollutants with dose modeled as a continuous function of time, such as carbon monoxide (CO). The dose for CO is measured by the carboxyhemoglobin concentration in the blood (%COHb). While the CO exposure in the lungs is assumed to be constant over one event, the blood %COHb changes continuously with time, much like the air concentration in the mass balance model. The **Dose** is the average over the event, while **FDose** is the value at the end of the event (and therefore becomes the initial value for the next event). For many pollutants, no such dose model has been built into APEX yet.

An example of the **EVENTSLIST** keyword would be:

```
EventsList = UMET VA VE EVR MET DEFICIT
```

```

APEX Events File
APEX Version 5.0 (dated July 25, 2017) Run Date = 20191029 Time = 121815
Location = Los Angeles Area
Scenario = Los Angeles (CSA 348), Jun-Aug 2010 Example case, 2010 pop, RandomSeed 4178348
Simulation = COF_LA_Short_Course.txt modified from APEX Intro Document Section 3.3; Example
APEX model run for the ISES-ISEE APEX short course.
Pollutant = Ozone
Air Quality = ! 2010 base ozone air quality data for CSA 348 : Los Angeles-Long Beach-
Riverside, CA

```

Person	Seq	Day	Year	Mn	Dy	Hr	Dur	Act	Mic	HW	UMET	MET	VA	VE	Deficit	EVR	Exp_Ozone
5	1	1	2010	6	1	1	60	5	1	1	0.999	0.999	3530.	7042.	0.00000	3.764	3.878E-03
5	2	1	2010	6	1	2	60	5	1	1	0.908	0.908	3206.	6513.	0.00000	3.482	3.963E-03
5	3	1	2010	6	1	3	60	5	1	1	0.868	0.868	3068.	6283.	0.00000	3.359	3.965E-03
5	4	1	2010	6	1	4	60	5	1	1	0.951	0.951	3359.	6764.	0.00000	3.616	4.030E-03
5	5	1	2010	6	1	5	60	5	1	1	0.887	0.887	3134.	6394.	0.00000	3.418	3.965E-03
5	6	1	2010	6	1	6	60	5	1	1	0.953	0.953	3367.	6778.	0.00000	3.623	3.850E-03
5	7	1	2010	6	1	7	60	5	1	1	0.838	0.838	2961.	6104.	0.00000	3.263	3.735E-03
5	8	1	2010	6	1	8	60	19	1	1	9.074	5.103	18029.	29385.	0.95311	15.708	3.642E-03
5	9	1	2010	6	1	9	60	19	1	1	5.628	3.516	12423.	20042.	0.99555	10.713	3.491E-03
5	10	1	2010	6	1	10	20	100	5	99	3.548	3.218	11369.	18520.	0.99861	9.900	4.000E-02

**Exhibit 5-12. Portion of an Events File**

A portion of an example *Events* file is shown in Exhibit 5-12 above. This file can become very large—about 1.4 MB per person-year if all variables are written to the file. For this reason, the user is given the option of writing the events for only a fraction of the simulated persons. This is controlled by the *Control Options* file settings *EVENTSAMPLE* and *CUSTOMSAMPLE*. See Section 4.2.3 for more information on these keywords.

## 5.11 Multipollutant File

This file can only be generated in a run containing at least two pollutants. It reports the amount of time (averaged across persons) in each combination of pollutant levels, for each combination of clock hour and microenvironment. The time is reported in hours, and cannot exceed the number of days in the simulation (because there is at most 1 hour per person, per day, in that bin). The divisor is the number of profiles with a non-zero amount of time in that bin. Also, the fraction of profiles that “visited” that bin is reported.

If there are N1 cut-points (listed on the *Control Options* file) for pollutant #1 and N2 cut-points for pollutant #2, then there will be  $24 * M * (N1 + 1) * (N2 + 1)$  rows of data on this file, where M is the number of micros. For each pollutant, the number of “bins” is one more than the number of cut-points. For example, if two cut-points are specified at 10 and 20 (in the *OUTPUTUNITS* for that pollutant), then the 3 bins are “below 10”, “10 - 20”, and “above 20”. If an exposure falls exactly on a cut-point, the time is assigned to the higher bin.

Similar to other output files, the results are tallied over the entire population if *COMMUTING* = NO or if *KEEPLEAVERS* = YES. If commuting is on and *KEEPLEAVERS* = NO, then the population under consideration is the set of profiles who both live and work in the study area, which will generally be fewer than the number requested on the *Control Options* file.

## 5.12 Diary Clustering Files

There are up to three diary clustering output files, available when the **CLUSTERDIARY** option is used for longitudinal diary assembly. The file names are specified on the *Control Options* file using the keywords **CLUSTER**, **CWEIGHT**, and **TRANSITIONS**. The file names are separate from the control switches with yes/no settings; the keywords for the switches are **CLUSTEROUT**, **CWEIGHTOUT**, and **TRANSOUT**.

The “Cluster” file contains one record per diary day in the input diary database. The pool and cluster assignments are listed, along with the score along each cluster axis. The score has two forms: a time in minutes, and a ratio of this diary time to the average time across all diaries. The cluster assignment is the largest of these ratio scores. If two ratios are exactly equal, the cluster is the smaller of the axis numbers. For writing out the values, the maximum ratio is limited to 999.

The “CWeight” file has one record for each unique combination of “from” pool, “to” pool, “from” cluster, “to” cluster, and age bin. Each record on this file has a statistical weight that is used when selecting the cluster for the next (or “to”) day.

The “Transitions” file has one record per example of a transition in the diary database, with the “from” pool, “to” pool, “from” cluster, “to” cluster, pool bin, and age bin, along with the CHADID of the “to” day, the diary age and whether it is part of an “adjacent” transition.

## 5.13 Sobol Results File

This file is generated only when a Sobol sensitivity analysis run is performed, which requires setting **SOBOLRUN** = Yes on the *Control Options* file (see CHAPTER 11 of *Volume II*). In this case, many of the other output files are suppressed. The Sobol file consists of two sections. First, each random variable is listed along with Sobol group number. This allows the user to identify which variables belong to a given group. The second section has two tables for each specified exposure metric, average day or maximum day.

The exposure metrics are listed in Table 11-3 of *Volume II*, and may be one or more of the following: **AvgExp**, **Max1Exp**, **Max8Exp**, **MaxTSExp**, **Max8EC**, **AvgDose**, **Max1Dose**, **Max8Dose**, **MaxTSdose**, **Max1FDose**, **Intake**, or **Dep**. The dose metrics are available only if APEX is asked to perform dose calculations (using **DODOSE** = YES on the *Control Options* file). The last two (**Intake** and **Dep**) are available only for particulate matter. Each selected metric is evaluated on each simulation day, for each profile. The “average day” is the average for each person, over all the simulated days. The “maximum day” is the worst daily value for that metric, for each person.

Sobol analyses measures the sensitivity of the selected exposure metric(s) to changes in the random input variables. Within each table on the output file, there is one row for each Sobol group number. There is one column for main effects, and a second for total effects.

Each main or total effect should be between zero and one. This range may occasionally be exceeded when the stochastic noise is extremely large, due to small sample size. In that case, the indices are not reliable. Otherwise, the total effect should be at least as large as the main effect,



and may be substantially larger. The main effect is a measure of the fractional importance of one specified group of variables, by themselves (including interactions among sets of variables entirely within that group). The total effect includes the main effect plus interactions between the specified group and variables in other groups.

The Sobol analysis provides sampled estimates for each “term” in a variance decomposition of the selected exposure metric. If the estimates were without error (which would only happen in the limit of infinite sample size), then the sum of all the main effects for all Sobol groups, plus each unique interaction, would equal one. The sum of all main effects (across all the Sobol groups) should never exceed one, and may be less than one because interactions between groups are left out. The sum of all total effects should be at least one, because each interaction between groups will be counted multiple times (once for each group that is part of the interaction). Larger indices reflect greater importance for the variables in the corresponding Sobol group. See Saltelli et al. (2004) for a detailed description of this method. Mokhtari et al. (2006) present an application of the Sobol method to an exposure model. The paper Glen and Isaacs (2012) provides guidance on using the Sobol method with stochastic models like APEX.

## REFERENCES

Glen W.G. and K. Isaacs (2012). Estimating Sobol sensitivity indices using correlations. *Environmental Modeling and Software* 37: 157-166.

Graham S.E. and T. McCurdy (2005). Revised ventilation rate ( $V_E$ ) equations for use in inhalation-oriented exposure models. EPA/600/X-05/008.

McCurdy, T., G. Glen, L. Smith, and Y. Lakkadi (2000). The National Exposure Research Laboratory's Consolidated Human Activity Database, *Journal of Exposure Analysis and Environmental Epidemiology* 10: 566-578 (2000).

Mokhtari, A., H.C. Frey, and J. Zheng (2006). Evaluation and recommendation of sensitivity analysis methods for application to Stochastic Human Exposure and Dose Simulation models, *Journal of Exposure Analysis and Environmental Epidemiology* (2006) 1-16.

National Research Council (1991). Human exposure assessment for airborne pollutants: advances and opportunities. Washington, DC: National Academy of Sciences.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto (2004). Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley & Sons, Ltd, Chichester, England.

U.S. Environmental Protection Agency (1999). Total Risk Integrated Methodology. Website: <https://www.epa.gov/fera>

U.S. Environmental Protection Agency (2017). An Introduction to APEX. Available at: <https://www.epa.gov/fera/apex-user-guides> and in the APEX installer.

U.S. Environmental Protection Agency (2019). THE CONSOLIDATED HUMAN ACTIVITY DATABASE (CHAD) Documentation and User's Guide, EPA-452/B-19-001. Available at: <https://www.epa.gov/fera/human-exposure-modeling-databases-support-exposure-modeling>

---

United States  
Environmental Protection  
Agency

Office of Air Quality Planning and Standards  
Health and Environmental Impacts Division  
Research Triangle Park, NC

Publication No. EPA-452/R-19-005a  
October 2019

---