PB89-181101

United States
Environmental Protection
Agency

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

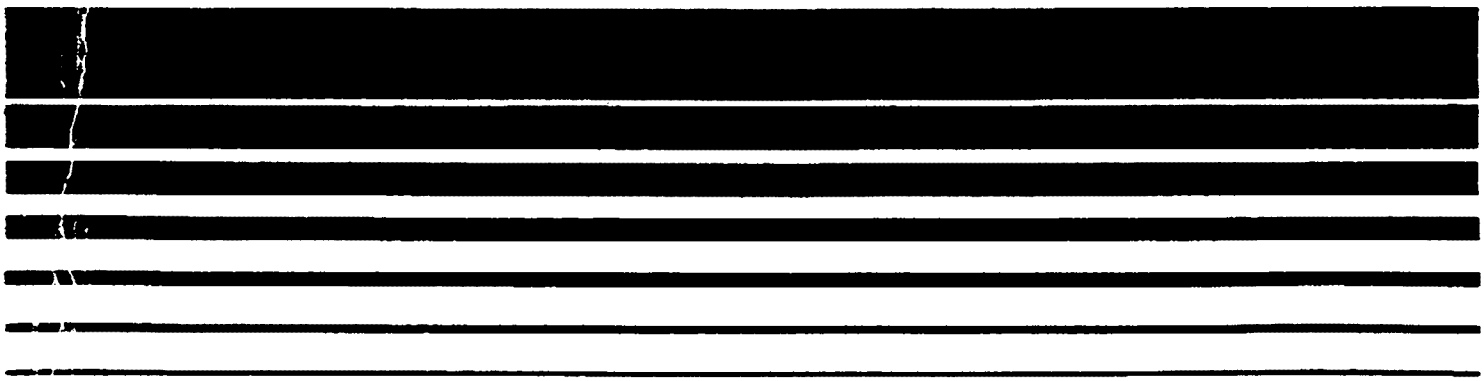
EPA-450/4-83-014R
May 1987

Air

EPA

Receptor Model Technical Series, Volume III (Revised)

CMB User's Manual (Version 6.0)



EPA-450/4-83-014R

RECEPTOR MODEL TECHNICAL SERIES, VOLUME III (REVISED)

CMB USER'S MANUAL (VERSION 6.0)

By
Kenneth Axetell
PEI Associates, Inc.
Cincinnati, Ohio 45246

and

Dr. John G. Watson
Desert Research Institute
Reno, Nevada 89506

Contract No. 68-02-3890

EPA Project Officer: Thompson G. Pace

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office Of Air And Radiation
Office Of Air Quality Planning And Standards
Research Triangle Park NC 27711
May 1987

This report has been reviewed by the Office Of Air Quality Planning And Standards, U.S. Environmental Protection Agency, and approved for publication as received from the contractor. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, neither does mention of trade names or commercial products constitute endorsement or recommendation for use.

EPA-450/4-83-014R

CONTENTS

Figures iv
Tables. v

1. Introduction. 1

What the model does. 1
How the model works. 3
Purpose of user's manual 3
Organization of user's manual. 4

2. Getting the CMB Model Up and Running. 5

Hardware requirements. 5
Starting 5
Saving results 7
Trial runs with test data. 8

3. Building CMB Input Files. 10

File types and formats 10
File generation methods, ... 21

4. Moving Through an Interactive Session 23

Calling up data files. 23
Using available commands to set up the CMB runs. 24
Selecting initial sources. 26
Selecting initial species. 27
Selecting test site and day. 28
Selecting size fraction. 29
Initiating a CMB run 29
Recovery summary information 30

5. Evaluating CMB Model Outputs. 32

Source contribution estimates display. 32
Uncertainty/similarity display 35
Species concentrations display 37
Additional diagnostics--SSCONT command 39

References. 41

Appendices

A. Theory of chemical mass balance model A-1
B. Diagnosis of error messages B-1
C. Printout of test data set for PACS1 . . . . . C-1

FIGURES

Number		Page
1	CMB model--inputs and outputs	2
2	Example source names file	12
3	Example chemical species codes file	14
4	Recommended species codes based on atomic numbers	15
5	Example fine particulate source profile file	17
6	Example ambient concentration data file	20
7	Source contribution estimates	33
8	Uncertainty/similarity display	36
9	Species concentrations	38
10	Species-source contributions display	40

TABLES

Number		Page
1	Format for Fine and Coarse Source Profile Files	16
2	Format for Ambient Concentration Data File.	19
3	Commands in CMB Program	25

SECTION 1

INTRODUCTION

1.1 WHAT THE MODEL DOES

The Chemical Mass Balance (CMB) Model uses the chemical composition of an ambient particulate sample to estimate the relative contributions of different source categories to the measured particulate concentration. The chemical composition of each source category's emissions (source profile) must also be known in order to run the model. The information required by and provided by the model are shown in Figure 1.

The CMB model identifies chemically distinct source categories, not individual emission sources. For example, one run might indicate that 6.7 $\mu g/m^3$ are from residual oil combustion, but this contribution cannot be further resolved into concentrations attributable to Power Plant 2, Industrial Boiler 3, Hospital Heating Plant 6, etc. Also, sources that have nearly identical chemical compositions (e.g., unpaved roads and construction activities) may not be distinguishable in the model.

The model can be run for size fractions of the particulate sample, such as PM_{10}, if ambient and source chemical composition data are both available for one or more size fractions. Ideally, separate runs are made for fine and coarse size fractions.

The CMB model provides source contribution estimates only for the day of sampling. Exact sources to be included in the model should be specific to that date and must be determined by the user from meteorological and emission inventory data separate from the model's required input. Because a CMB run is applicable to a single location, a single date, and possibly a single size fraction, most users perform a series of model runs in order to generate a comprehensive source-receptor analysis.

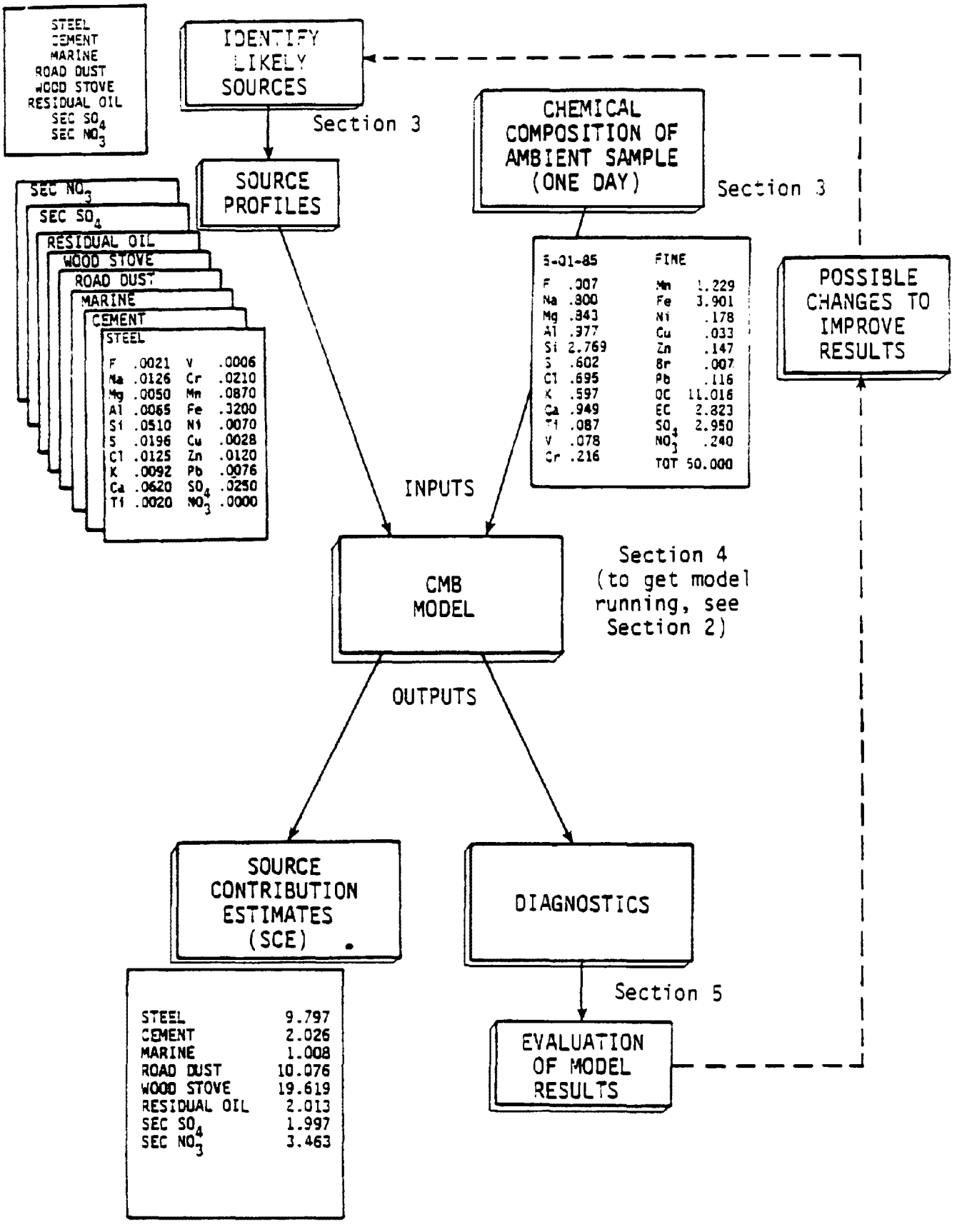


Figure 1. CMB model--inputs and outputs.

1.2 HOW THE MODEL WORKS

CMB is a statistical model in which a separate mass balance equation is developed for each chemical species. The equations are of the form:

$$C_i = F_{i1}S_1 + F_{i2}S_2 + \ldots + F_{iJ}S_J$$

where C_i = Particulate concentration of chemical species i measured in the ambient air, $\mu g/m^3$

F_{ij} = Fraction of particulate emissions from source j that is of chemical i

S_j = Ambient particulate concentration resulting from source j; also referred to as source contribution estimate (SCE); this is the unknown in the model

J = Total number of sources.

The number of equations is equal to the number of chemical species; the number of unknowns (S values) is equal to the number of sources (J). In order to solve equations in CMB, there must be more chemical species than source categories.

The simultaneous equations are solved by a procedure known as least squares estimation. Actual calculations of the least square fit are done in the computer program by matrix algebra.

1.3 PURPOSE OF USER'S MANUAL

This manual describes Version 6.0 of the Chemical Mass Balance Receptor Model. It is designed to allow persons to constructively use the CMB model with only a few hours learning time. Because of this emphasis on rapid command of modeling procedures, the manual covers primarily the mechanical aspects of running the model. Users seeking more information on the statistical or theoretical bases of chemical mass balance receptor modeling are referred to Appendix A, Theory of Chemical Mass Balance Model.

The manual is intended for wide use by State and local air pollution control agency personnel in developing State Implementation Plans (SIP's) for PM_{10}. The U.S. Environmental Protection Agency (EPA) has published two

companion documents to this manual that are also SIP-oriented and that should be consulted for this application. The first, Protocol for Applying and Validating the CMB Model (USEPA 1987A), provides guidance on applicability, assumptions, and interpretation of results. The protocol provides a practical strategy for obtaining valid results. The second document, Reconciling CMB and Dispersion Model Results (USEPA 1987B), recommends a procedure for examining and reconciling differences between dispersion and receptor models' results.

Changes made to Version 6 of CMB subsequent to this user's manual will be described briefly in a README.DOC file located on the same diskette as the updated CMB program. You may read this file by using either the TYPE or PRINT commands in DOS.

1.4 ORGANIZATION OF USER'S MANUAL

Version 6 of the CMB model is interactive. It provides prompts and explanations in the sequence of displays during a run. However, a new user will still need to refer to this manual for selecting options and formatting responses until he/she is thoroughly familiar with the model's capabilities.

The four sections of the manual following this introduction and their respective contents are:

Section	Title	Contents
2	Getting the CMB Model Up and Running	Basic instructions to get the computer program operating
3	Building CMB Input Files	Formats and instructions for ambient and source profile data entry
4	Moving Through an Interactive Session	Possible responses and consequences for all interactive queries
5	Evaluating CMB Model outputs	Descriptions and interpretations of output statistics

SECTION 2

GETTING THE CMB MODEL UP AND RUNNING

2.1 HARDWARE REQUIREMENTS

The CMB program runs on an IBM PC or IBM-compatible personal computer. A minimum memory capacity of 320 K is needed for the program.

Version 6 of CMB is compiled to take advantage of an 8087 coprocessor. The coprocessor reduces calculation times by a factor of 10 to 20. If an 8087 coprocessor is not available for the PC on which the program is to be run, the program will automatically run without the 8087. Run times may be unacceptably long (on the order of several minutes in some cases) without the 8087 coprocessor.

The program software and test data files are contained on one 5¼-inch double density floppy diskette formatted by the MS/DOS or PC/DOS operating system. The CMB program is named CMB6.EXE on this diskette. The CMB software may be run from this diskette, but its limited space for input and output files makes this inconvenient. It is recommended that the program and data files be transferred to a hard disk directory, if available, for routine modeling.

2.2 STARTING

2.2.1 Disk Operating System (DOS)

After the PC and screen units are turned on, the computer is "booted" with a DOS operating system by placing the DOS diskette into drive A. DOS Version 2.0 or later is required. The standard DOS configuration must be modified so that the CMB program can open up to 14 files simultaneously. This is accomplished by modifying or creating a file entitled CONFIG.SYS and placing

it on the DOS diskette. If the PC has a hard disk, the CONFIG.SYS file should be placed in the root directory.

An existing CONFIG.SYS file is modified by adding the line 'FILES=14' (see instructions below). If none exists, a new CONFIG.SYS is created by moving to the appropriate drive (A> for DOS on diskette, usually C> for a hard disk) and using the text editor EDLIN (supplied with the operating system) to add the following:

```
A> or C> EDLIN CONFIG.SYS (ENTER)
         I                (ENTER)
         FILES=14         (ENTER)
         (CONTROL)(BREAK)
         E                (ENTER)
```

The sequence above activates EDLIN, names the new file CONFIG.SYS, enters the Insert mode (I), increases the number of files in memory to 14, exits the Insert mode, exits EDLIN, and leaves the new file where it will be part of the DOS operating system when the computer is subsequently booted.

If the CMB program is to be used immediately, the computer must be re-started by simultaneously pressing (CONTROL)(ALT)(DEL) so that the new CONFIG.SYS file can expand the maximum number of files to 14. In subsequent starts, no special action is necessary. However, always be sure to use the modified DOS diskette when running CMB.

2.2.2 The CMB and Data File Diskettes

The CMB executable program and test data files are on a single diskette, but they may be copied onto separate diskettes to free more space for output. The placement of these diskettes in drives depends on the configuration of the system being used--hard disk-drive, dual disk-drive, or single disk-drive.

If the PC is equipped with a hard disk, the CMB program should be copied onto the hard disk (into a specified subdirectory) from the floppy disk drive with the COPY command:

```
C:
CD/
MKDIR CMB
CD CMB
COPY A:*.*
```

This command will also transfer the input data files. Be sure that the desired subdirectory has been accessed before using the copy command.

With a dual drive, CMB and data file diskettes can be placed in separate drives. If the PC has only a single disk drive, the data files to be used in a run (six files) must normally reside on the same diskette as the CMB program. However, with output files, the capacity of the diskette will soon be exceeded.

2.3 SAVING RESULTS

2.3.1 Printouts

A permanent record of each CMB run can be obtained in one of two different forms:

- o A printout of the entire interactive session, including all routine CMB program outputs
- o Printouts of just the information from each run that the user wants.

The complete printout is obtained by simultaneously pressing (CONTROL)(PRT SC) at the beginning of the run; the printer must be connected to the computer and its power switch turned on. Pressing (CONTROL)(PRT SC) a second time switches off this feature. Selected information (second method above) is printed by pressing (SHIFT)(PRT SC) when desired data from the run is displayed on the screen.

If several runs are anticipated, the volume of printout generated by the first method may be so great that analysis becomes unwieldy. Also, the time required to complete a run is much longer if all input and output are printed. It then becomes more expedient to use the second method, typing (SHIFT)(PRT SC) whenever useful information is on the screen.

2.3.2 Storage of Outputs

The results of any CMB run can be permanently stored in a disk file by using the WRITE command at the completion of the run. With multiple runs, the output files should be appropriately named so they can be distinguished from one another at a later date.

The two disk files generated by the WRITE command--CMBOUT and SUMMRY--are in text format. Therefore, these files can be input to another software program (e.g., a word processor or text editor) for editing and enhancement. They can be read using either the TYPE or PRINT commands in DOS.

2.4 TRIAL RUNS WITH TEST DATA

A test data set has been distributed with the CMB program to assist new users in gaining familiarity with it prior to entering their own data. The CMB program and test data files should be loaded into the computer as described in Section 2.2. The program is activated by addressing the drive (and directory, if any) with CMB in it, typing CMB6, and pressing the ENTER key (↵). The CAPS LOCK key must be on for the interactive session.

The program begins an interactive format at this point by asking:

```
DISK FILE FOR INITIAL INPUT?
IF NOT ENTER CARRIAGE RETURN
IF SO ENTER NAME OF DISK FILE
```

The appropriate response to bring up the test data is 'INPORT.DAT', which should be followed by pressing the ENTER key again.

Another query, 'DO YOU WISH TO RENAME CMBOUT?', follows. Press ENTER again. The title screen for CMB appears next, followed by the statement:

```
THE 'HELP' COMMAND LISTS COMMANDS.
USE COMMANDS AE, DE, AS, DS FOR CHANGES.
ENTER COMMAND
```

Enter the command 'SELECT'. This allows the user to specify the sampling site and date of ambient data for the run. Three days of ambient data are included in the test set:

Site	Year	Date
PACS1	77	0813
PACS2	78	0124
PACS3	77	0807

Any of these three dates can be entered for the initial run in response to the program queries:

```
ENTER DESIRED CMB SITE CODE:  XXXXXXXXXXX
ENTER YEAR:  YY
ENTER DATE:  MMDD
```

The user is also asked to select the size fraction for analysis, and can respond with either 'F' for fine or 'C' for coarse.

The next user command should be 'CMB', which initiates the least square calculations and presents the results. The program produces three displays of output statistics and then permits entry of another command to modify input data or obtain additional output for the completed run.

Eight source categories have been preselected as the fitting sources for the initial run. They are listed in the first output display, Source Contribution Estimates. The names (and source numbers) of 12 additional sources with complete input data in the test files can be found by using the command PINFO. The third output screen--Species Concentrations--lists up to 35 chemical species for which complete input data are available and notes the 21 fitting species used in the run with asterisks.

Detailed instructions for the interactive portion of the CMB program are in Section 4 of this user's manual. These instructions should be followed closely during practice sessions with the test data.

A printout of the initial run with the test data set for PACS1 is presented in Appendix C. If any difficulties are encountered following the instructions in Section 4, Appendix C should be consulted. Several additional runs for the same site and sampling date should be performed to determine whether predictions can be improved by adding or deleting some source(s) or species(s).

SECTION 3

BUILDING CMB INPUT FILES

3.1 FILE TYPES AND FORMATS

Six different files are required as input to CMB as listed below:

File	Usual File Name
List of Input Data File Names	IN------.DAT
List of Source Names	SO------.DAT
List of Chemical Species Codes	PO------.DAT
Source Profiles, Fine Particle	FS------.DAT
Source Profiles, Coarse Particle	CS------.DAT
Ambient Concentration Data	DA------.DAT

Each file except the Ambient Concentration Data file contains a single line format (i.e., entries on sequential lines all have the same format). The formats for the six files are presented in separate subsections below. The notations A, I, and F refer to alphanumeric, integer, and floating-point formats, respectively. A floating-point format designated F8.6, for example, indicates that the field is 8 characters wide and if no decimal point appears in the field it is assumed that six characters are to the right of the implied decimal. Frequently, this decimal point is included in the data in a floating point field; in that case, the actual decimal location overrides the implied location in the "F" format. Data entries in the I format should always be right justified (i.e., aligned to the right side of the field).

The six files can be created and stored in any order on the diskette or hard disk because they are called individually by name into the CMB program. In general, only one SO, PO, FS, and CS file is required for a CMB analysis of a site or regional set of sites. However, if different FS or CS files are desired for separate runs, this can easily be accommodated by naming the alternate files differently. The two-letter prefixes for file names shown in

the table above are recommended to avoid confusion, but are not required for the program to run. Up to six additional alphanumeric characters can be used in the file names to distinguish different data sets.

3.1.1 List of Input Data File Names (IN------.DAT)

This file contains the names of the other files used by the CMB program in the exact order that they are called into the program. The order of files in the listing is as follows:

```
FS------.DAT
CS------.DAT
DA------.DAT
SO------.DAT
PO------.DAT
```

The names must correspond to those of the user's five files, including any file name extensions.

The line format for this brief file is a 12 or fewer character alphanumeric name left oriented in columns 1-12 of lines 1-5.

The purpose of this file is to save the effort of keying in the individual file names in response to program queries each time CMB is run. If this file is not present, the program will request the five file names in a series of interactive queries.

3.1.2 List of Source Names (SO------.DAT)

This file has a 2-digit integer source code number in columns 1-2 and an 8-character alphanumeric source name in columns 5-12. If the source is to be included as an initial fitting source (see discussion in Section 4.2), an asterisk should be placed in column 15. All other columns on each line are usually blank, but columns 16-80 can be used for comments, which will appear only if the Source Names file is printed or displayed on the screen.

The maximum number of sources (lines) that can be listed in this file is 35. The maximum number of sources in the file that can be included in a single CMB run is 16. Source names need not be listed in numeric order.

The format for the Source Names file and example input for 15 sources are shown in Figure 2.

I2 SOURCE CODE				A8 SOURCE NAME								COMMENT AREA												
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	1			R	E	S	I	D	O	I	L			*										
0	2			W	O	O	D	S	T	O	V													
0	4			F	I	E	L	D	B	R	N			*										
0	7			R	O	A	D	D	U	S	T			*										
0	8			A	S	P	H	A	L	T	P													
1	0			T	R	A	F	F	I	C														
1	1			M	A	R	I	N	E					*										
1	2			E	L	E	V	A	T	O	R													
1	5			F	E	R	T	I	L	I	Z													
1	8			C	O	A	L	P	W	R	P			*										
2	1			S	M	E	L	T	E	R														
2	3			C	O	N	S	T	R	U	C													
2	4			C	E	M	E	N	T					*										
2	6			S	E	C	S	O	4					*										
2	9			S	E	C	N	O	3															

Figure 2. Example source names file.

3.1.3 List of Chemical Species Codes (PO------.DAT)

This file has an identical format to the Source Names file described above--a 2-digit species code in columns 1-2, an 8-character alphanumeric species name in columns 5-12, an asterisk in column 15 if the species is to be an initial fitting species, and all other columns blank. No species should be given the code 01; this code is reserved for particulate matter concentration measured at the sampler, and is termed TOTAL. TOTAL could be included in the species listing (with species code 01), but generally it is not used as a fitting species because this would tend to force the model to make the sum of source contribution estimates (SCE's) equal the measured concentration.

The maximum number of chemical species (lines) that can be listed in this file is 35. The maximum number of species that can be included in a single CMB run is 21. Chemical species do not have to be listed in numeric order.

In order to maintain consistency with the species coding in the EPA source profile library, atomic numbers should be used as the species codes for chemical elements. This is not required by the program, but strongly recommended. The example format in Figure 3 provides some of these numbers. Codes for all species can be found in Figure 4. The three-digit species codes in the source library must be changed to two-digit codes to conform to the input format of CMB. Change codes 201-204 to codes 91-94.

3.1.4 Fine Particle Source Profile (FS------.DAT)

The line format for this file has four required input fields plus an unstructured comment area that is frequently used to indicate what data are recorded in those four fields. The columns and formats for the fields are as shown in Table 1. All source code numbers and species codes used in this file must be listed in the previous two files, the List of Source Names and List of Chemical Species.

The number of lines in the Fine Particle Source Profile (FS) file should be equal to the number of source profiles times the number of species per profile. For example, if 20 source profiles are needed for the CMB analysis and each has an average of 22 chemical species, the FS file would have 440

I2 SPECIES CODE		A8 SPECIES NAME										COMMENT AREA																	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	9			F										•															
1	1			N	A									•															
1	2			M	G									•															
1	3			A	L									•															
1	4			S	I									•															
1	6			S										•															
1	7			C	L									•															
1	9			K										•															
2	0			C	A									•															
2	2			T	I									•															
2	3			V										•															
2	5			M	N									•															
2	6			F	E									•															
2	8			N	I									•															
3	0			Z	N									•															
3	5			B	R									•															
8	2			P	B									•															
9	1			O	C									•															
9	2			E	C									•															
9	3			S	O	4								•															
9	4			N	O	3								•															
2	4			C	R																								
2	9			C	U																								
3	4			S	E																								
9	0			T	C																								
	1	T	O	T	A	L																							

Figure 3. Example chemical species codes file.

Species code	Species name
4	BE
5	B
9	F
11	NA
12	MG
13	AL
14	SI
15	P
16	S
17	CL
19	K
20	CA
21	SC
22	TI
23	V
24	CR
25	MN
26	FE
27	CO
28	NI
29	CU
30	ZN
31	GA
32	GE
33	AS
34	SE
35	BR
37	RB
38	SR
40	ZR
47	AG
48	CD
50	SN
51	SB
55	CS
56	BA
58	CE
80	HG
82	PB
90	Total carbon
91	OC
92	EC
93	SO4
94	NO3
SUM (%)	TOTAL

Notes: OC = organic carbon, EC = elemental carbon, NA = not analyzed, and NR = not reported.

Figure 4. Recommended species codes based on atomic numbers.

TABLE 1. FORMAT FOR FINE AND COARSE SOURCE PROFILE FILES

Column	Format	Contents
1-2	I2	Source code number
5-6	I2	Species code
9-16	F8.X	Fraction of fine (or coarse) source emissions from indicated species
19-26	F8.X	Uncertainty of species emission estimate (standard error of fraction)
27-80		Can be used for comments; usually to identify size fraction, source name, and species name

lines. This file, along with the corresponding Coarse Particle Source Profile (CS) file, constitutes the bulk of the input data.

The source profile data are often obtained from published references, such as the Receptor Model Source Composition Library (USEPA 1984), and others (Hopke, 1985 and Gordon, 1987). However, profile data should be selected carefully. Source profiles from published references may be adequate for screening runs, but airshed-specific profiles are strongly recommended for final CMB runs.

If the emissions of a chemical species from a source are zero, that line should still be included in the source profile file. Also, every species must have an uncertainty value specified in columns 19-26, and that value cannot be zero. Estimated values for uncertainties can be conservatively assumed to be the same as the species fractions in columns 9-16 (unless the species fraction is zero).

All source profiles should include data for a species if it is to be included as a fitting species. The CMB analysis might yield biased results if, for instance, some source profiles measured carbon but carbon data were not reported (but actually present) for other source profiles.

Some sources emit only one or a few chemical species. Single constituent source types, such as those used to apportion secondary sulfate and nitrate, have only one species each and should be entered in the FS file as shown in Figure 5. A fine source profile for one other source, RESIDOIL, is also shown in Figure 5.

SOURCE CODE NO.				SPECIES CODE NO.				FRACTION OF EMISSIONS FROM SPECIES										STANDARD ERROR OF FRACTION								COMMENT AREA																				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
0	1				9			.	0	0	0	5	3					.	0	0	0	1	9				F	I	N	E		R	E	S	I	D	O	I	L		F					
0	1			1	1			.	0	3	5	0	0					.	0	1	7	0	0				F	I	N	E		R	E	S	I	D	O	I	L		N	A				
0	1			1	2			.	0	0	0	0	0					.	0	3	0	0	0				F	I	N	E		R	E	S	I	D	O	I	L		M	G				
0	1			1	3			.	0	0	5	3	0					.	0	0	2	4	0				F	I	N	E		R	E	S	I	D	O	I	L		A	L				
0	1			1	4			.	0	0	9	6	0					.	0	0	4	8	0				F	I	N	E		R	E	S	I	D	O	I	L		S	I				
0	1			1	6			.	1	3	3	0	0					.	0	2	4	0	0				F	I	N	E		R	E	S	I	D	O	I	L		S					
0	1			1	7			.	0	0	0	0	0					.	0	0	1	0	0				F	I	N	E		R	E	S	I	D	O	I	L		C	L				
0	1			1	9			.	0	0	2	8	0					.	0	0	1	0	0				F	I	N	E		R	E	S	I	D	O	I	L		K					
0	1			2	0			.	0	1	5	8	0					.	0	0	6	4	0				F	I	N	E		R	E	S	I	D	O	I	L		C	A				
0	1			2	2			.	0	0	1	1	0					.	0	0	0	3	8				F	I	N	E		R	E	S	I	D	O	I	L		T	I				
0	1			2	3			.	0	3	4	4	0					.	0	0	7	5	0				F	I	N	E		R	E	S	I	D	O	I	L		V					
0	1			2	5			.	0	0	0	4	6					.	0	0	0	1	3				F	I	N	E		R	E	S	I	D	O	I	L		M	N				
0	1			2	6			.	0	2	9	7	0					.	0	0	6	1	0				F	I	N	E		R	E	S	I	D	O	I	L		F	E				
0	1			2	8			.	0	5	3	6	0					.	0	1	2	1	0				F	I	N	E		R	E	S	I	D	O	I	L		N	I				
0	1			3	0			.	0	0	4	0	0					.	0	0	1	8	0				F	I	N	E		R	E	S	I	D	O	I	L		Z	N				
0	1			3	5			.	0	0	0	1	3					.	0	0	0	2	1				F	I	N	E		R	E	S	I	D	O	I	L		B	R				
0	1			8	2			.	0	0	1	1	0					.	0	0	0	6	4				F	I	N	E		R	E	S	I	D	O	I	L		P	B				
0	1			9	1			.	0	7	8	0	0					.	0	8	2	0	0				F	I	N	E		R	E	S	I	D	O	I	L		O	C				
0	1			9	2			.	0	2	4	2	0					.	0	2	5	0	0				F	I	N	E		R	E	S	I	D	O	I	L		E	C				
0	1			9	3			.	4	8	1	0	0					.	1	1	9	0	0				F	I	N	E		R	E	S	I	D	O	I	L		S	O	4			
0	1			9	4			.	0	0	6	5	0					.	0	0	4	4	0				F	I	N	E		R	E	S	I	D	O	I	L		N	O	3			
0	1			2	4			.	0	0	0	4	7					.	0	0	0	1	5				F	I	N	E		R	E	S	I	D	O	I	L		C	R				
2	6			9	3			.	7	2	7	0	0					.	0	5	0	0	0				F	I	N	E		S	E	C		S	O	4			S	O	4			
2	4			9	4			.	7	7	5	0	0					.	0	5	0	0	0				F	I	N	E		S	E	C		N	O	3			N	O	3			

Figure 5. Example.fine particulate source profile file.

3.1.5 Coarse Particle Source Profile

The format and data requirements for this file are the same as those for the Fine Particle Source Profile file (Table 1 and Figure 5).

If only PM_{10} size range of source profile data are available, they should be entered as coarse and the fine file left blank. However, a named file (containing no data) should still be created for the Fine Particle Source Profile file to prevent a file opening error when CMB is run.

Source profiles for the fine size fraction are for particulate 0-2.5 μm and for the coarse size fraction are for particulate 2.5-10 μm. If a single source apportionment for PM_{10} is desired from the CMB model, the fine and coarse source profiles can be combined into a PM_{10} source profile. The procedure for calculating the PM_{10} species fractions and uncertainties from coarse and fine profile data is described on pages II-33 to II-36 of the Receptor Model Source Composition Library publication (USEPA 1984).

3.1.6 Ambient Concentration Data (DA------.DAT)

This file contains two line formats for each day of sampling data. The header format (identified by the numbers 03 in columns 1 and 2) provides information on the receptor site and date and is limited to a single line. The second type of format (identified by the numbers 30 in columns 1 and 2) records ambient concentrations for the above site and date, with a separate line for each chemical species. The detailed formats are shown in Table 2. An example Ambient Concentration Data file for one sampling date is presented in Figure 6. Note that the uncertainty of air ambient species measurement cannot be equal to zero. It should be set equal to the minimum analytical detection limit for this species.

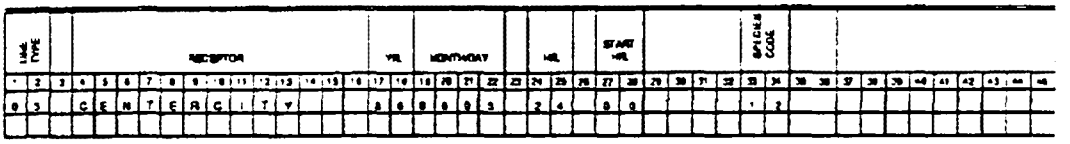
The numbers of sites/days of ambient data in this file are not limited. The CMB program uses each day's data in a separate run when the unique combination of receptor site identification and date are specified in the interactive portion of the program. Thus, several dates for both size fractions can be included behind each other in the same file.

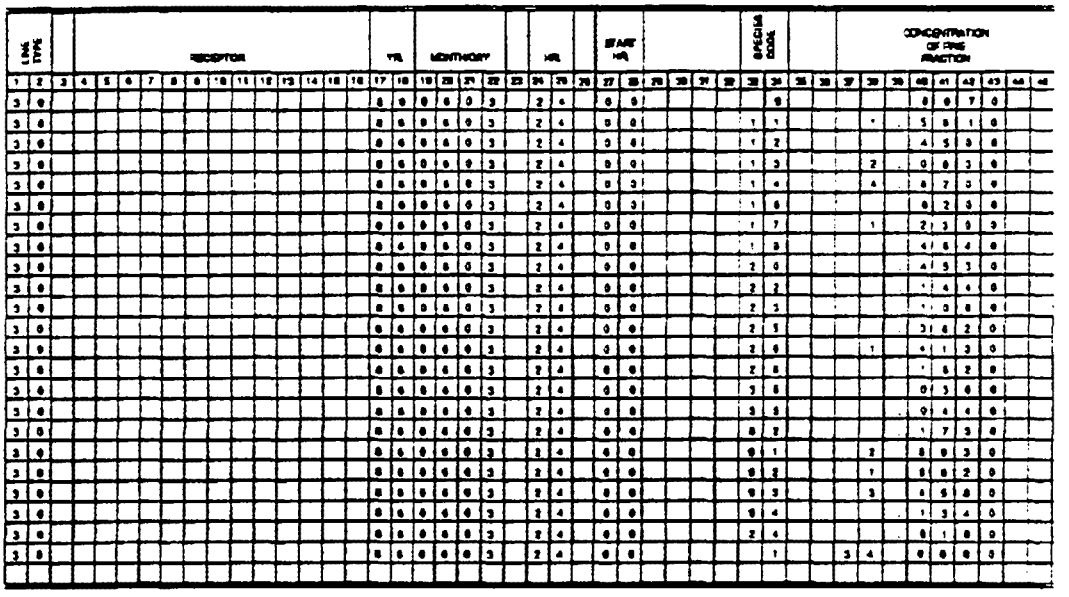
During data entry, the same species codes as in the List of Chemical Species Codes (PO) file must be used. Also, concentration uncertainties (standard errors in columns 48-56 and 70-78) must be greater than zero.

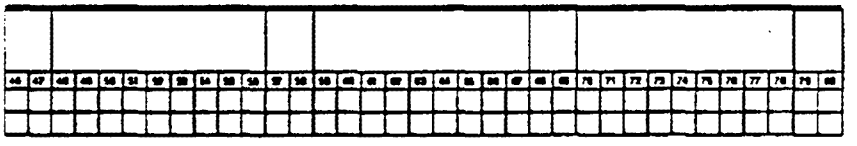
TABLE 2. FORMAT FOR AMBIENT CONCENTRATION DATA FILE

Line type	Column	Format[a]	Contents
Header	1-2	I2	'03'
	4-15	A12	Receptor identification
	17-18	A2	Year, YY
	19-22	A4	Month and day, MMDD
	24-25	I2	Duration of sample, hours
	27-28	I2	Starting hour of sample
	33-34	I2	Size fractions on next lines 12 = fine and coarse 13 = fine and total
	35-80		Must be blank
Concentrations	1-2	I2	'30'
	4-15	A12	Receptor identification
	17-18	A2	Year, YY
	19-22	A4	Month and day, MMDD
	24-25	I2	Duration of sample, hours
	27-28	I2	Starting hour of sample
	33-34	I2	Species code
	37-45	F9.X	Concentration of fine fraction
	48-56	F9.X	Standard error of fine fraction (measurement uncertainty)
	59-67	F9.X	Concentration of coarse fraction or total
	70-78	F9.X	Standard error of coarse fraction or total (measurement uncertainty)

[a] I - Integer (I2 ∿ 0 through 99)
A - Alphanumeric (A12 ∿ up to 12 characters, A-Z, 0-9, any keyboard symbol)
F - Floating point (F9.X ∿ up to 8 integers and a decimal point). Note: It is very helpful in reading the file to have the decimal points aligned vertically. A decimal point must be included in the field.







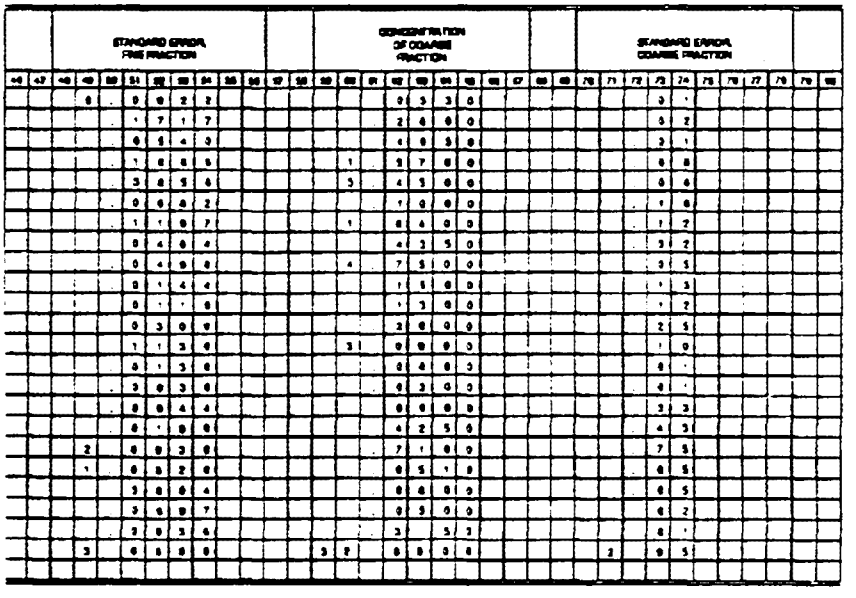


Figure 6. Example ambient concentration data file.

If only PM_{10} data are available, all fine particle concentrations should be set to zero and the "fine and total" designation (13) should be entered in columns 33-34 of the header record. CMB analyses should be performed only on the total size fraction.

3.2 FILE GENERATION METHODS

There are three common methods of creating the CMB input files: manually entering the data in the correct format using a text editor or word processing program; editing existing input files with a text editor or word processing program; or transferring files from the EPA source library.

3.2.1 Keyboard Input

A word processing program can be used to type entire input files. Figures 2 through 6 should be followed in establishing the formats for the files. Tabs should not be set to facilitate spacing between fields. Completed files should be appropriately named and then copied onto a diskette or hard disk as described in Section 2.2.2.

Word processing programs generally must be used in the text mode (with a carriage return after each line). EDLIN, the line editor supplied with MD/DOS or PC/DOS, may also be used.

3.2.2 Editing Existing Files

It is often preferable to start with existing input files such as the test data set and make changes in them with a word processing program or a text editing program such as EDLIN to obtain the desired files. This method usually reduces format errors and saves time if only a few changes are needed. Because of differences in individual editing programs, the user is advised to consult the manual for the editing program to be used for directions on opening a copy of the existing file, deleting and adding material, saving the changes, and renaming the file.

3.2.3 Transferring Files from EPA Library

More than 100 source profiles for fine and coarse particulate (components of FS and CS files, the files with most number of lines) are available on

diskette from EPA. These source profiles can be copied directly onto the diskette or hard disk being used to assemble input data files for Version 6 with a software utility program supplied with the diskette version of the EPA source composition library. Separate documentation accompanies the source composition library diskette.

Whenever the source profile files are transferred from EPA library files and merged with Source Name, Chemical Species Code, and Ambient Concentration Data files generated elsewhere, make sure that source and species codes are consistent throughout the five files. The diskette version EPA library profiles all use atomic numbers for elements, 01 for mass concentration, 91 for organic carbon, 92 for elemental carbon, 93 for SO_4, and 94 for NO_3.

SECTION 4

MOVING THROUGH AN INTERACTIVE SESSION

The CMB program and its data files are loaded as described in Section 2. The discussion in this section follows the order of prompts in the interactive session exactly. For additional clarification, the user can review CMB program queries and example responses shown in Appendix C.

4.1 CALLING UP DATA FILES

After the user types CMB6, the program responds with 'MAKE SURE YOUR CAPS LOCK IS ON!!!' if CMB is in the directory under that name. Be sure to check that the CAPS LOCK key is on at this time. The program only recognizes upper case instructions.

The first interactive statement requiring a response is:

```
DISK FILE FOR INITIAL INPUT?
IF NOT ENTER CARRIAGE RETURN
IF SO ENTER THE FILE NAME
```

Assuming that the six input files have been prepared as described in Section 3, the user should type in the name of the file that contains the other input file names (e.g., IN------.DAT) at this time. Include the three-character extender (.DAT) in the response or the program will print an error message indicating that the file cannot be found.

The program then searches for each of the five input files. If they are all located and properly named, the program automatically continues with the next prompt. If an error is found, an appropriate error message is displayed and the user must correct the files before performing the run. The session has to be initiated again (by typing in CMB6) after correcting the error.

The second prompt sets up the output file:

```
DO YOU WISH TO RENAME CMBOUT?
IF NOT ENTER A CARRIAGE RETURN.
IF SO ENTER THE FILE NAME.
```

A permanent record of the last CMB run is created under CMBOUT or the specified file name (if one is specified) by using the 'WRITE' command. The results in CMBOUT are lost when a new session is started; printouts become the only record.

After the second response, the title screen appears. This display remains for almost one minute before the third prompt is shown. The title screen shows revision number and release date of the Version 6 CMB program being used. The availability of later editions can be determined by contacting the principal authors or the USEPA Project Officer identified on the title screen. CMB results will not differ in later editions, but convenience, efficiency and error trapping may be added in response to user suggestions.

4.2 USING AVAILABLE COMMANDS TO SET UP THE CMB RUNS

Several alternatives can be exercised beginning at this point. The prompt is:

```
THE 'HELP' COMMAND LISTS COMMANDS
USE COMMANDS AE, DE, AS, DS FOR CHANGES.
ENTER COMMAND
```

Available commands and their actions are presented in Table 3.

If the user is uncertain of the complete lists of source categories and chemical species for which data have been input, or which have been designated for the initial fit, he or she may want these shown and/or printed out before proceeding further. This can be accomplished with the command 'PINFO'.

A second option at this time (either before or after using the PINFO command) would be to add or delete sources or species from the fit designated in the input files. This can be done prior to making a CMB run with the commands AE, DE, AS, and DS, as described in the following two sections.

In most cases, the appropriate response to the third prompt (ENTER COMMAND) is 'SELECT'. This command allows the user to select the ambient data set (site and date) for the run, as described in Section 4.6.

Most of the commands in Table 3 are for use after at least one run has been completed, to change some of the input data for a subsequent run or to obtain specific results of the completed run. When the program returns to this point for an additional run or an alternate command, all of the input

TABLE 3. COMMANDS IN CMB PROGRAM

General purpose of command	Command	Specific action
Runs CMB calculations	CMB	Performs CMB calculations
Changes input data	SELECT	Allows user to select ambient data set (site and date) for CMB run
	AE	Adds one or more chemical species to the fit
	DE	Deletes one or more chemical species from the fit
	AS	Adds one or more sources to the fit
	DS	Deletes one or more sources from the fit
	SIZE	Changes size fraction
Displays results	PINFO	Shows current and available input data--ambient data set, fitting species, fitting sources
	PMATRIX	Displays requested data from input files--Source Profile, Ambient Concentration, Source Names, or Chemical Species Codes
	SSCONT	Shows calculated contribution of each source to each species concentration
	PCOMP	Writes computed averages for a series of CMB runs at a site to an output file
	WRITE	Sends current CMB results to diskette (CMBOUT)
Miscellaneous	HELP	Displays information on available commands similar to this table
	EXIT	Terminates interactive session; user is returned to operating system

data (fitting sources, fitting chemical species, ambient data set, and size fraction) remain at their previous values unless changed by the user. This feature eliminates the need for repetitive entries.

There are six possible commands to initate changes in input data for additional runs--AE, DE, AS, DS, SELECT, and SIZE. The commands AE or DE require that the user specify the species to be added or deleted; similarly, AS or DS require that the user specify the source(s) to be added or deleted. The command SELECT allows the user to specify a different sampling site/date. The command SIZE changes the size fraction (from fine to coarse or coarse to fine) without any further user response.

4.3 SELECTING INITIAL SOURCES

A group of sources has already been designated as initial fitting sources in the List of Source Names input file (SO------.DAT). These should be reviewed for their applicability to the specific run to be made using the criteria discussed below.

Selection of source types is limited to the source code numbers in the Source Name file and to 16 entries. The user should evaluate information other than the CMB input data to select these sources. Recommended information sources are point and area source emission inventories of the area surrounding the sampling site, meteorological data for the sampling day, and microscopic analysis of the sampled particulate.

All source categories in the Source Names file that could reasonably be expected to contribute to ambient concentrations on the sampling day should be included in the initial run. This approach is theoretically correct because overspecification (inclusion of sources with no impact) results in unbiased SCE's and larger standard errors than a correctly specified model. In contrast, underspecification (omission of sources that actually have an impact) results in biased estimates.

Sources are added to the fit with the command AS and deleted with the command DS. The program responds to these commands with:

```
SIZE IS FINE
INPUT CODE OF ADDED (or DELETED) SOURCE
```

Type the source code of the source to be added or deleted and press ENTER. The program repeats its prompt to provide an opportunity to add or delete another source. This interaction is repeated until all source codes

associated with the changes have been entered. Enter a blank line after the last prompt to signal that the changes are complete.

In subsequent runs, sources may appropriately be deleted if prior-round model results indicate a negligible contribution. However, sources should not be added just to improve model diagnostics, unless some new external information indicates a possible impact from the source(s). One proper reason for adding a source is that there was inadequate capacity to include it in the first run as one of the initial 16 sources.

If sulfate is to be specified as a fitting species, it should definitely be included in the fit as a single species source. If it is removed as either a source or species in a subsequent run, sulfate should be completely excluded from the fit (both source and species removed).

4.4 SELECTING INITIAL SPECIES

The initial fitting species are designated in the List of Chemical Species Codes input file (P------.DAT). Changes can be made in these prior to the initial run with the AE or DE commands in a manner exactly analogous to adding or deleting sources. The fitting species are the only ones for which mass balance equations are developed during the CMB run.

In general, 21 chemical species are used in a model run. By entering the maximum number of species (21), the amount of information available for the least squares solution is maximized.

The user has responsibility for assuring consistency between the fitting species specified at this point in the program, the species available for each source in its source profile, and the species analyzed in the ambient sample. For example, fluoride (F) should only be selected as a fitting species if a fluoride concentration has been measured for the ambient sample and all of the source profiles include a fluoride component.

If a species is selected for fitting but no source profile data are available for that species in the profile data file, the model assumes a "zero" value and no error message appears. However, serious errors can occur in the CMB model results. As a consequence of potential problems associated with missing species data, data availability is often the limiting factor in determining which species will be selected for inclusion in a CMB run.

The species Cl, Br, NH_4, and NO_3 should be used with caution because of atmospheric reactions that have been shown to change their solid phase concentrations.

4.5 SELECTING TEST SITE AND DAY

After the command SELECT has been entered, the program responds with the request:

ENTER DESIRED CMB SITE CODE: XXXXXXXXXXX

The user should then type and enter the receptor identification name for the desired site exactly as it appears in columns 4-15 of the Ambient Concentration Data file. The program then requests:

ENTER YEAR: YY

and

ENTER DATE: MMDD

to which the user should respond (in two steps) with the 2-digit year, month, and date corresponding to the sampling data set desired. Again, the date specified should be exactly the same as it appears in columns 17-18 and 19-22 of the Ambient Concentration Data file.

The program then asks for desired size fraction (see Section 4.6 below) and states:

DATA SEARCH BEGUN FOR
SITE: XXXXXXXXXXX YEAR: YY DATE: MMDD

If the requested Ambient Concentration Data file is found, the program immediately responds with another ENTER COMMAND prompt. If the specified site code and date do not match those of a sampling data set in the Ambient Concentration Data file, the program lists all sampling data sets on file and requests:

```
DATA SET NOT FOUND FOR
     SITE:  XXXXXXXXXXX   YEAR:  YY   DATE:  MMDD
USE ONE OF THOSE LISTED ABOVE
ENTER COMMAND
```

The user must then enter 'SELECT' again and the site code and date for one of the listed sampling data sets.

If the contents or exact names of data sets in the Ambient Concentration Data file are unknown when the interactive session begins, any names and dates can be input in response to the program's prompts. After the program fails to find the spurious data set, it provides a complete listing from which the desired data sets can be chosen.

4.6 SELECTING SIZE FRACTION

In order to completely define the desired ambient data set, the user must specify whether the fine or coarse size fraction is to be analyzed. This is done by typing 'F' or 'C' in response to the program query:

```
INPUT DESIRED SIZE FRACTION:  (FINE OR COARSE)
```

If PM_{10} data are substituted for coarse fraction in all files, 'C' is still specified at this point. Records of the size change must be kept outside the model.

Check that data for the desired size fraction are available in the source profile and ambient data files. If data files are named but contain no data, the program will run through the least squares analysis and report mainly zeros in the summary tables.

4.7 INITIATING A CMB RUN

At any time after the ambient data set has been selected, a CMB least squares solution procedure can be obtained by typing 'CMB' in response to ENTER COMMAND. The program makes no response while the calculations are being performed (usually about one minute), but displays three screens of results after the calculations are complete with no additional commands.

To stop the displays from scrolling, the keys (CTRL) and (NUMLOCK) can be pressed simultaneously. To restart, press any key or the space bar.

Another ENTER COMMAND prompt is displayed after the three screens have been viewed. The user then has the following options:

1. Enter one of the six commands that change input data in preparation for an additional run (AE, DE, AS, DS, SELECT, or SIZE)

2. Obtain additional statistics on the run just completed

3. Display results or input data in different formats

4. Store the results in a disk report file with the WRITE command

5. Exit the CMB program.

Commands under Option 1 have been discussed in this section. Option 4 is discussed in the following subsection; Options 2 and 3 are discussed in Section 5.

4.8 RECOVERING SUMMARY INFORMATION

Printouts of data from CMB runs are normally obtained with the PRT'SC key on an intermittent or continuous basis during program execution, as described in Section 2.3.1.

The results of each run may also be placed in permanent storage for later use or display by employing the WRITE command following the CMB command. A permanent record of each run in a session is maintained separately. Results from different sessions can be distinguished by renaming the CMBOUT file at the beginning of each session. The WRITE command must be repeated after every run to save the results of that run. Storage of data is confirmed by the program response 'WRITTEN'.

To retrieve these permanent output files, the CMB program must be exited. Then, with the computer still in the same subdirectory (with hard disk), the printer must be put in condensed (17.5 characters per inch or 132 column) mode. This is accomplished by following instructions in your printer manual. A printout of an expanded version of the screen displays 'PRINT CMBOUT' or the renamed file for CMBOUT.

If both a fine fraction and coarse fraction run were performed for the same site/date, an additional report can be obtained. By typing the command 'PRINT SUMMRY', a summary of the fine, coarse, and total source contributions for that date is generated and printed.

SECTION 5

EVALUATING CMB MODEL OUTPUTS

Three different tabular summaries of model outputs and performance measures are automatically produced after the CMB command is entered:

1. Source Contribution Display
2. Uncertainty/Similarity Display
3. Species Concentration Display.

The model outputs and guidelines for their evaluation are summarized below. Interpretation is discussed elsewhere (USEPA 1987A). Mathematical formulae and significance levels are presented with a more detailed explanation in Appendix A.

5.1 SOURCE CONTRIBUTION ESTIMATES DISPLAY

An example Source Contribution Table is shown in Figure 7. Six outputs and diagnostics plus a restatement of the ambient sampling data and conditions are included in this display.

5.1.1 Reduced Chi Square

The reduced chi square, which will be called a chi square in this manual, is a measure of the goodness-of-fit for a CMB run. The value of chi square is inversely proportional to the squares of the uncertainty terms for species in the source profiles and ambient data. Therefore, a data set with relatively high uncertainties would have a lower chi square than a corresponding data set with relatively low uncertainties.

Chi square for a sample can be any positive value, with values approaching 1.0 indicating the best fit. A value greater than 4.0 suggests that the model has not explained ambient measurements well.

SOURCE CONTRIBUTION ESTIMATES - SITE: PACS3 DATE: 0807 77 VERSION: 6.0
SAMPLE DURATION 24 START HOUR 0 SIZE: COARSE
R SQUARE .97 PERCENT MASS 102.2
CHI SQUARE 1.50 DF 13

*	SOURCE TYPE	SCE(UG/M3)	STD ERR	TSTAT
1	MARIN	15.5726	2.4487	6.3595
3	UDUST	61.8332	4.9872	12.3984
4	AUTPB	35.1579	4.4520	7.8972
5	RDOIL	13.2395	2.0769	6.3746
8	KRAFT	8.4592	8.8923	.9513
11	ALPRO	4.4485	2.7329	1.6278
12	STEEL	-2.1408	1.6307	-1.3128
13	FERMN	1.3753	.8020	1.7149

MEASURED CONCENTRATION FINE/COARSE/TOTAL:
135.0+- 13.5/ 135.0+- 13.5/ 270.0+- 19.1

Figure 7. Source contribution estimates.

5.1.2 R-Square

R-square is the variance in the ambient species data as explained by the calculated species concentrations. It is determined by a linear regression of measured versus model-calculated values for each of the fitting species.

R-square values can range from 0.0 to 1.0. The closer the value is to 1.0, the better the model has explained ambient species concentrations for that data set. If R-square is less than 0.8, the model is considered to have poorly predicted species concentrations and, therefore, source contributions. R-square is mathematically related to the reduced chi square as shown in Appendix A. It will generally approach a value of 1.0 when the reduced chi square indicates a "good fit."

5.1.3 Percent Mass

Percent mass is the ratio of the sum of model-calculated species concentrations (for the particle size fraction being analyzed) to the measured mass concentration. The model should account for essentially all of the measured mass (percent mass = 100.0). If the model under- or overpredicts mass by more than 20 percent, it has not been well constructed. Percent mass values less than 80 percent may indicate a missing source type or species not included in the fit. Values over 120 percent may indicate inclusion in the fit of a source type that actually has no impact or collinearity problems, as discussed in Section 5.2.

Even though all three statistics that measure overall CMB model performance with a given data set (R-square, chi square, and percent mass) have values in the "acceptable" range, the fit may still be improved by justifiable changes in the input data. Improvement is indicated by change in these statistics toward their optimum values of 1.0 for R-square, 1.0 for chi square, and 100.0 for percent mass.

5.1.4 Degrees of Freedom (DF)

Degrees of freedom is the number of species in the fit (maximum 21) minus the number of source types (maximum 16). The model does not run if DF is less than one. Many of the model statistics may deteriorate if DF is less than five. Number of degrees of freedom must be known for statistical interpretation of the chi square value.

5.1.5 Source Contribution Estimates (SCE)

Source contribution estimates are the main product of the CMB model. SCE's are the concentrations in $\mu g/m^3$ attributed to each of the specified sources. The sum of these concentrations should approximate the measured mass concentration.

Negative SCE values are not physically meaningful, but they can occur as a result of collinear source profiles or when their absolute values are less than their corresponding standard errors. These latter negative values are of no concern. If a negative SCE is reported, source combinations (in the next display) including this source should be examined to identify the collinear sources and properly quantify their combined contribution.

5.1.6 Standard Error (STD ERR)

The standard error of estimate is the uncertainty in the associated SCE. The standard error is interpreted by use of the t-statistic described below. As an approximation, as the standard error in $\mu g/m^3$ approaches the value of the SCE, the SCE may no longer be significantly different from zero and is therefore not meaningful.

5.1.7 T-Statistic (TSTAT)

The TSTAT measures the statistical significance of an SCE by comparing the SCE to its standard error. A TSTAT value of less than about 2.0 indicates that the SCE is not significantly different than zero. The presence of low TSTAT values for several sources within a run may be due to collinearities among source profiles or to uncertainties in the input data.

5.2 UNCERTAINTY/SIMILARITY DISPLAY

The display following the source contribution estimates presents uncertainty/similarity clusters. Figure 8 shows an example of this output.

The first column of this display contains the clusters, one cluster on each row. Each cluster is identified by the code numbers associated with a source profile. The clusters are formed if two criteria are met: 1) two or more source types in an eigenvector derived from the singular value

UNCERTAINITY/SIMILARITY CLUSTERS		SUM OF CLUSTER SOURCES	
1	11	20.02+-	3.67
12	13	-.76+-	.86
12	13	-.76+-	.86

Figure 8. Uncertainty/similarity display.

decomposition exceed 0.25 (these terms are explained in Appendix A), and 2) the t-statistic for any one of these source types is less than or equal to 2.0. These uncertainty/similarity clusters are caused by excessive similarity (collinearity) among the source profiles in the cluster or by high uncertainties in the individual source profiles. The standard errors associated with the source contribution estimates of one or more sources identified in a cluster are usually very large, often too large to allow an adequate decision to be made.

If collinearity is the cause of these excessive uncertainties, then the uncertainty of the sum of the source contributions for a cluster may be smaller than the uncertainty of any single source contribution in the cluster. The sum of source contributions and the standard error of the sum are shown in the final two columns of this display. This sum may be more useful than the individual source contribution estimates if the standard error of the sum is substantially lower than the standard errors of each SCE. The sum does not allow differentiation among the contribution estimates of sources contained in the cluster.

An example of the uncertainty/similarity clusters display is shown in Figure 8. In it, sources 1 and 11 form one cluster. The estimated sum of the SCE's for these two sources is 20.02 ± 3.67. A second cluster is formed by sources 12 and 13. The repeat of a cluster should be ignored.

In an actual example, the cluster numbers will not appear if the two criteria above are not met. This would be interpreted as meaning that there are no source clusters in that model run that are sufficiently similar to cause the model estimates to have high uncertainties.

5.3 SPECIES CONCENTRATIONS DISPLAY

An example Species Concentrations display is shown in Figure 9. The data presented in the table heading are the same as in the heading of the SCE display (Figure 7). The body of the table shows measured and model-calculated concentrations for each chemical species so that the model's fit by species can be examined. Species used in the fit are marked with an asterisk in the column labeled "I." The column labeled "M" contains an "M" if data are missing from the receptor file. This "M" allows the user to distinguish missing data from nondetectable levels of a species.

There are two diagnostics included in the Species Concentrations display. The column labeled RATIO R/U contains the ratio of the signed difference between the calculated and measured concentration (the residual) divided by the uncertainty of that residual. The R/U ratio is calculated for all species, even those not included in the fit. The lower the absolute value of the ratio, the better the model has explained the species (e.g., a low residual with a high uncertainty will provide a low RATIO R/U). A high RATIO R/U (e.g., >2.0) indicates that the residual is significant. If the RATIO R/U is <-2.0, a missing source may contain a significant quantity of that species.

The column marked RATIO C/M shows the ratio of calculated to measured concentration and the standard error of that ratio for every chemical species with measured data (including some that are not fitting species). The ratios should all be near 1.00 if the model has predicted observed species concentrations well.

If the C/M ratio for a species is considerably lower than 1.00, the model is unable to find sources with emissions containing that species to fully explain its measured presence. This may indicate a missing source that is rich in the species with the low ratio. If the C/M ratio for a species is considerably greater than 1.00, the best fit for all species has forced contributions of the overpredicted species to the high level. There are several possible explanations for this occurrence (USEPA 1987A).

SPECIES CONCENTRATIONS - SITE: PACS3 DATE: 0807 77 VERSION: 6.0

SAMPLE DURATION		24	START HOUR	0	SIZE: COARSE					
R SQUARE		.97	PERCENT MASS	102.2						
CHI SQUARE		1.50	DF	13						

SPECIES	I	M	MEAS		CALC		RATIO C/M		RATIO R/U
1 TOT			135.00000+-	13.50000	137.94540+-	8.53619	1.02+-	.12 TOT	.2
9 F			.01000+-	.00100	.20279+-	.09532	20.28+- 9.75	F	2.0
11 NA		*	7.59000+-	.75900	8.34527+-	.66827	1.10+-	.14 NA	.7
12 MG		*	1.46000+-	.14600	1.68068+-	.47624	1.15+-	.35 MG	.4
13 AL		*	6.29000+-	.62900	5.94886+-	.51657	.95+-	.13 AL	-.4
14 SI		*	17.10000+-	1.71000	17.65057+-	.93566	1.03+-	.12 SI	.3
16 S			2.12000+-	.21200	2.67595+-	.38234	1.26+-	.22 S	1.3
17 CL		*	9.97000+-	.99700	7.54866+-	1.59989	.76+-	.18 CL	-1.3
19 K		*	1.37000+-	.13700	1.07583+-	.07430	.79+-	.10 K	-1.9
20 CA		*	2.14000+-	.21400	2.67331+-	.24078	1.25+-	.17 CA	1.7
22 TI		*	.48200+-	.04800	.63881+-	.11806	1.33+-	.28 TI	1.2
23 V		*	.51900+-	.05200	.47296+-	.09936	.91+-	.21 V	-.4
24 CR		*	.02700+-	.00300	.03099+-	.03546	1.15+- 1.32	CR	.1
25 MN		*	.11900+-	.01200	.12400+-	.02528	1.04+-	.24 MN	.2
26 FE		*	5.85000+-	.58500	4.19096+-	.35283	.72+-	.09 FE	-2.4
28 NI		*	.78800+-	.07900	.73140+-	.16032	.93+-	.22 NI	-.3
29 CU		*	.04900+-	.00500	.05995+-	.01355	1.22+-	.30 CU	.8
30 ZN		*	.19700+-	.02000	.22672+-	.05656	1.15+-	.31 ZN	.5
35 BR		*	1.38000+-	.13800	1.80798+-	.59785	1.31+-	.45 BR	.7
82 PB		*	7.85000+-	.78500	7.25928+-	1.05885	.92+-	.16 PB	-.4
91 OC		*	24.60000+-	2.46000	22.03127+-	3.68929	.90+-	.17 OC	-.6
92 EC		*	2.31000+-	.23100	2.94891+-	.78235	1.28+-	.36 EC	.8
93 SO4		*	8.47000+-	.84700	9.50631+-	1.70589	1.12+-	.23 SO4	.5
94 NO3		*	.44000+-	.04400	.49675+-	.19718	1.13+-	.46 NO3	.3

Figure 9. Species concentrations.

5.4 ADDITIONAL DIAGNOSTICS--SSCONT COMMAND

The command SSCONT produces a table that shows the fraction of each species' calculated ambient concentration contributed by each source in the fit. An example SSCONTS display is shown in Figure 10.

By pointing out major contributing sources according to the model's fit, some potentially incorrect source profiles or incorrect extra sources responsible for a large positive residual for a species may be identified.

```
ENTER COMMAND
SSCONT
```

INDIVIDUAL RATIO = CALC SPECIES(PER SOURCE) / MEAS SPECIES(ALL SOURCES)

SPEC\SOURCE	1	3	4	5	8	11	12	13
1 TOTAL	.115	.458	.260	.098	.063	.033	-.016	.010
9 F	.000	.495	.000	.702	.000	18.684	.000	.399
11 NA	.821	.143	.000	.061	.059	.014	-.004	.006
12 MG	.512	.652	.000	.000	.000	.082	-.095	.000
13 AL	.000	.649	.061	.011	.004	.221	-.002	.001
14 SI	.000	1.012	.017	.007	.001	.000	-.006	.001
16 S	.242	.000	.066	.831	.132	.000	-.020	.011
17 CL	.625	.000	.106	.000	.025	.005	-.004	.001
19 K	.159	.465	.018	.027	.025	.000	-.014	.105
20 CA	.102	.867	.205	.098	.014	.017	-.062	.008
22 TI	.000	1.296	.000	.030	.000	.007	-.009	.001
23 V	.000	.032	.000	.878	.000	.003	-.002	.001
24 CR	.000	1.031	.000	.230	1.504	.026	-1.665	.021
25 MN	.000	.520	.000	.051	.037	.000	-1.565	1.999
26 FE	.000	.606	.126	.067	.027	.003	-.117	.005
28 NI	.000	.003	.008	.901	.024	.012	-.019	.000
29 CU	.000	.379	.524	.203	.104	.127	-.122	.010
30 ZN	.000	.345	.625	.269	.000	.002	-.130	.040
35 BR	.023	.004	1.274	.001	.003	.004	.000	.002
82 PB	.000	.029	.896	.002	.000	.000	-.002	.000
91 OC	.000	.084	.715	.038	.054	.000	.000	.005
92 EC	.000	.415	.578	.178	.066	.031	.000	.009
93 SO4	.184	.005	.054	.752	.118	.009	-.006	.007
94 NO3	.000	.028	.727	.196	.000	.000	.000	.178

Figure 10. Species-source contributions display.

REFERENCES

Belsley, D. A., K. Edwin, and R. E. Welsch. (1980) Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. John Wiley, New York.

Benarie, M. M. (1980) Urban Air Pollution Modeling. MIT Press, Cambridge, Massachusetts.

Bevington, P. R. (1969) Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, New York.

Britt, H. I., and R. H. Luecke. (1973) The Estimation of Parameters in Non-linear, Implicit Models. Technometrics, 15:223-247.

Currie, L. A., R. W. Gerlach, C. W. Lewis, et al. (1984) Interlaboratory Comparison of Source Apportionment Procedures: Results for Simulated Data Sets. Atm. Env., 18(8):1517-1537.

DeCesar, R. T., and J. A. Cooper. (1982) Evaluation of Multivariate and Chemical Mass Balance Approaches to Aerosol Source Apportionmnet, Using Synthetic Data and an Expanded PACS Data Set. Proceedings, Receptor Models Applied to Contemporary Pollution Problems, Specialty Conference, Air Pollution Control Association, Denver, MA.

DeCesar, R. T., S. A. Edgerton, M. A. K. Khalil, and R. A. Rasumssen. (1985a) Sensitivity Analysis of Mass Balance Receptor Modeling: Methyl Chloride as an Indicator of Wood Smoke. Chemosphere, 14, 1495.

DeCesar, R. T., S. A. Edgerton, M. A. Khalil, and R. A. Rasmussen. (1985b) A Tool for Designing Receptor Model Studies to Apportion Source Impacts With Specified Precisions. Receptor Methods for Source Apportionment: Real World Issues and Applications, Air Pollution Control Association, Pittsburgh.

Friedlander, S. K. (1973) Chemical Element Balances and Identification of Air Pollution Sources. Environmental Science and Technology, 7:235.

Gordon, G. E., W. H. Zoller, G. S. Kowalczyk, and S. W. Rheingrover. (1981) Composition of Source Components Needed for Aerosol Receptor Models. Atmospheric Aerosol: Source/Air Quality Relationships (edited by Macias and Hopke), ACS Symposium Series No. 167, American Chemical Society, Washington, D.C. 1981, pp. 51-74.

Gordon, G. E. (1987), and Ann Sheffield, University of Maryland, Chemistry Building, College Park, Maryland. Personal communication to Tom Pace, U.S. Environmental Protection Agency, April 1987.

Henry, R. C. (1982) Stability Analysis of Receptor Models That Use Lowest Squares Fitting. Receptor Models Applied to Contemporary Pollution Problems, edited by S. L. Dattner and P. K. Hopke, Air Pollution Control Association, Pittsburgh, PA. 141.

Hidy, G. M., and S. K. Friedlander. (1972) The Nature of the Los Angeles Aerosol. 2nd IUAPPA Clean Air Congress, Washington, D.C.

Hopke, P. K. (1985) Receptor Modeling in Environmental Chemistry, John Wiley and Sons, New York, 1985.

Javitz, H. S., and J. G. Watson. (1986) Feasibility Study of Receptor Modeling for Apportioning Utility Contributions to Air Constituents, Deposition Quality and Light Extinction. Draft Report for Electric Power Research Institute, prepared by SRI International, Menlo Park, CA.

Kneip, T. J., M. T. Kleinman, and M. Eisenbud. (1972) Relative Contribution of Emission Sources to the Total Airborne Particulates in New York City. 3rd IUAPPA Clean Air Congress.

Kowalczyk, G. S., C. E. Choquete, and G. E. Gordon. (1978) Cehmical Element Balances and Identification of Air Pollution Sources in Washington, D.C. Atmospheric Environment, 12:1143.

Pasquill, F. (1974) Atmospheric Diffusion. 2nd Ed. Halsted Press, John Wiley, New York.

Seinfeld, J. H. (1975) Air Pollution: Physical and Chemical Fundamentals. McGraw-Hill, New York.

U.S. Environmental Protection Agency. (1983) Receptor Model Technical Series, Vol. III, Chemical Mass Balance Receptor Model User's Manual. EPA-450/4-83-014.

U.S. Environmental Protection Agency. (1984) Receptor Model Source Composition Library. EPA-450/4-85-002, November.

U.S. Environmental Protection Agency. (1987A) Protocol for Applying and Validating the CMB Model. EPA-450/4-87-010. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. May 1987.

U.S. Environmental Protection Agency. (1987B) Procedures for Reconciling Differences in Receptor and Dispersion Models. EPA-450/4-87-008. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. May 1987.

Watson, J. G. (1979) Chemical Element Balance Receptor Model Methodology for Assessing the Sources of Fine and Total Suspended Particulate Matter in Portland, Oregon. Ph. D. dissertation, Oregon Graduate Center, Beaverton, Oregon.

Watson, J. G., and N. F. Robinson. (1984) A Method to Specify Measurements for Receptor Models. Proceedings of the National Symposium on Recent Advances in Pollutant Monitoring of Ambient Air and Stationary Sources. U.S. Environmental Protection Agency, Research Triangle Park, N.C. April 1984.

Watson, J. G. (1984) Overview of Receptor Model Principles. Journal of the Air Pollution Control Association, 34:620.

Watson, J. G., A. W. Gertler, G. H. Prowell, J. C. Chow, S. V. Hering, L. W. Richards, and D. S. Blumenthal. (1985) Causes of Secondary Aerosol in the San Joaquin Valley Determined from the WOGA Aerosol Data Base. DRI Document No. 6687.402, prepared for the Western Oil and Gas Association, Reno, NV. Energy and Environmental Engineering Center, Desert Research Institute.

Watson, J. G., and J. C. Chow. (1986) Volume 4 - An Evaluation of Ambient Aerosol Chemistry in the Western United States. Draft report of Western States and Deposition Project, Phase I of SYSAPP-86-129, prepared for Western Governor's Association, Denver, Colorado.

Winchester, J. W., and G. D. Nifong. (1971) Water Pollution in Lake Michigan by Trace Elements from Aerosol Fallout. Water, Air, and Soil Pollution, 1:50.

APPENDIX A

THEORY OF CHEMICAL MASS BALANCE MODEL

A.1 NOMENCLATURE

The following symbols will be used to describe the CMB model solutions, diagnostics, and results.

C_i = The measured ambient concentration of species i

$\underset{\sim}{C}$ = $(C_1 \ldots C_I)^T$, a column vector whose $\underline{i}$th component is C_i

F_{ij} = The fraction of source j composed of species i, the source composition coefficient

$\underset{\sim}{F}$ = An I x J matrix of F_{ij}, the source composition matrix

I = Number of species used in the least squares calculation

J = Number of sources used in the least squares calculation

σ_{C_i} = Standard deviation of the C_i measurement

$\sigma_{F_{ij}}$ = Standard deviation of the F_{ij} measurement

σ_{R_i} = Standard deviation of $\left(\sum_{j=1}^{J} F_{ij} S_j - C_i\right)$

$\underset{\sim}{V}_e$ = Diagonal matrix of effective variances,

$$(V_e)_{ii} = \sigma_{C_i}^2 + \sum_{j=1}^{J} \sigma_{F_{ij}}^2 S_j^2$$

S_j = Source contribution of source j

$\underset{\sim}{S}$ = $(S_1 \ldots S_J)^T$, a column vector whose $\underline{j}$th component is S_j

χ^2 = Chi square

χ^2_{I-J} = Reduced chi square

D_j = Dispersion parameter for source j

E_j = Emission rate for source j

C_t = Total measured mass

A.2 THE PHYSICAL BASIS FOR RECEPTOR MODELS

A receptor oriented model must always be a representation of reality and not just a series of equations. The relationship between a source model and the CMB receptor model is simple and is presented here.

In general, the aerosol mass concentration at a receptor during a sampling period of length T due to a source j with constant emission rate E_j is

$$S_j = D_j E_j \tag{1}$$

where $$D_j = \int_{\emptyset}^{T} d\ (\vec{u}(t), \sigma(t), \vec{x}_j)dt \tag{2}$$

is a dispersion factor depending on wind velocity, $\vec{u}(t)$, atmospheric stability, $\sigma(t)$, and the location of source j with respect to the receptor, $\vec{x}_j$. u and σ, and possibly even $\vec{x}_j$ in the case of a mobile source, will vary with time; the instantaneous dispersion factor, d, must be integrated over the sampling period. When $\vec{u}\cdot\vec{x}_j = -1$, the source of emissions E_j is not in an upwind quadrant with respect to the receptor and $d \to \emptyset$.

Various forms for d have been proposed (Pasquill, 1974, Benarie, 1976, Seinfeld, 1975), some including provisions for chemical reactions, removal and specialized topography. None are completely adequate to describe the complicated, random nature of dispersion in the atmosphere. The advantage of receptor models is that an exact knowledge of D_j is unnecessary.

If a number of sources, J, exists and there is no interaction between their aerosols to cause mass removal, the total aerosol mass measured at the receptor, C, will be a linear sum of the contributions from the individual sources

$$C = \sum_{j=1}^{J} D_j E_j ; = \sum_{j=1}^{J} S_j \qquad (3)$$

Similarly, the concentration of elemental component i, C_i will be

$$C_i = \sum_{j=1}^{J} F_{ij} S_j = \sum_{j=1}^{J} F_{ij} D_j E_j, \; i = 1,I \qquad (4)$$

where F_{ij} is the fraction of source contribution S_j composed of element i.

In the source model each S_j is known from Equations 1 and 2, the F_{ij} are determined by chemical analyses of representative samples from source j, and C_i is calculated from Equation 4. The receptor model starts with the measurements of C_i, uses some knowledge about the chemical composition of the sources and attempts to quantify S_j, or at least to make a statement about its variability or significance as a contributor to the total mass concentration, C.

If the C_i and the F_{ij} at the receptor for all J of the source types suspected of affecting the receptor are known, and $J \leq I$, a set of simultaneous equations exists from which the source type contributions, S_j, may be calculated. Several solution methods have been proposed and tested. The method selected for Version 5 of the CMB receptor model is the effective variance weighted least squares solution (Watson et al., 1984).

The assumptions of the CMB with an effective variance solution are:

1. Compositions of source emissions are constant over the period of ambient and source sampling.

2. Chemical species do not react with each other, i.e., they add linearly.

3. All sources with a potential for significantly contributing to the receptor have been identified and have had their emissions characterized.

4. The number of sources or source categories is less than or equal to the number of components.

5. The source compositions are linearly independent of each other.

6. Measurement uncertainties are random, uncorrelated, and normally distributed.

A.3 THE CMB MODEL AND MULTIPLE LINEAR REGRESSION

The effective variance solution of the CMB equations is an application of multiple linear regression (MLR) in which a least squares solution to the equations is obtained. Several different modifications of the least squares method have been evaluated during development of versions, as explained in Section A.8 below.

The measured concentrations of individual chemical species in an ambient particulate sample constitute the array of response or dependent variables. The fractions of these chemical species contained in the emissions from each source type constitute the array of explanatory or regressor variables. These source chemical composition descriptions are called source profiles, source signatures, or source fingerprints.

The regression coefficients calculated by multiple linear regression are the estimated contributions from each identified source type, or the Source Contribution Estimates (SCE's). The main goal of the CMB modeling effort is to obtain meaningful values for the SCE's.

MLR has been widely used and researched by statisticians for the past half century. However, the emphasis has almost always been to fit the data to a given set of variables rather than to estimate the unknown values in a physical model. The CMB application is derivable from physical laws and is not just an empirical fit to acquired data.

A.4 HISTORY OF CMB SOFTWARE DEVELOPMENT

The Chemical Mass Balance receptor model was first developed quasi-independently by Hidy and Friedlander (1972), Kneip et al. (1972), and Winchester and Nifong (1971). The original applications used unique chemical species associated with each source type, the so-called "tracer" solution. Friedlander (1973) introduced the ordinary weighted least squares solution to the CMB equations, and this had the advantages of relaxing the constraint of a unique species in each source type and of providing estimates of uncertainties associated with the source contributions.

Dr. G. E. Gordon and his students at the University of Maryland (e.g., Kowalkzyk et al., 1978) subsequently applied this method to a wider range of

chemical species that they measured in source and receptor samples. The ordinary weighted least squares solution was limited in that only the uncertainties of the receptor concentrations were considered; the uncertainties of the source profiles, which are typically much higher than the uncertainties of the receptor concentrations, were neglected.

The first user-oriented software for the CMB model was programmed in 1979 by Dr. John Watson at the Oregon Graduate Center in FORTRAN IV on a PRIME 300 minicomputer. All subsequent versions have followed the same basic pattern of this original, and some of the constraints imposed by limited computer power and storage at the time of initial development have not been modified. Version 1 incorporated a variety of solutions, including the ordinary weighted least squares solution used by Drs. Friedlander and Gordon, an effective variance procedure which used the uncertainties of the source compositions as well as the uncertainties of the receptor concentrations, and a theoretically (but not necessarily practically) "exact" solution which would change the source compositions toward more statistically correct values with each iteration. Extensive tests with simulated ambient data verified the effective variance solution's source contribution estimates and their uncertainties.

The Portland Aerosol Characterization Study (PACS) and ancillary studies in the State of Oregon were ultimately used for the development of SIP's for TSP. These studies led to the first enhancements of Version 1 by Mr. Pat Hanrahan of the Oregon Department of Environmental Quality (DEQ).

Version 2 of the CMB Receptor Model was installed on EPA's UNIVAC system by Mr. John Core of Oregon DEQ during an assignment with EPA in 1980. This version was identical to Version 1 except that it operated on more versatile input data files. Receptor model workshops were conducted in Washington, D.C., Denver, CO, and Houston, TX, for State and local pollution control agencies using remote tie-ins to the UNIVAC.

Version 3 was implemented by Radian, Inc. by Messrs. Hugh Williamson and Dennis DuBose in 1982 on the EPA UNIVAC. The software was completely rewritten in FORTRAN 77, though the command structure and features of Version 1 were retained. A ridge regression option was added to address the issue of collinearity in addition to the effective variance solution. The collinearity issue (i.e., two or more source types having source compositions which are

too similar to be separated from each other by the model) was found to be a major limitation of the CMB in practical applications. Excessive collinearity in an application increases the uncertainty of CMB model results to intolerable levels.

Dr. Ronald C. Henry of the University of Southern California theoretically examined the collinearity issue in his presentation at the 1982 Receptor Model Specialty Conference in Danvers, MA (Henry, 1982). He presented the singular value decomposition of the weighted source matrix as a method of determining the degree of collinearity in a single application. Dr. Henry also demonstrated that a ridge regression solution could not return realistic solutions in cases of excessive collinearity.

Version 4 of the CMB was created at the Desert Research Institute (DRI) of the University of Nevada, under DRI and Oregon Department of Environmental Quality (DEQ) sponsorship in 1984. This version transferred the Version 3 computer code to the IBM/XT and made the modifications required for it to run under MicroSoft FORTRAN, added the original solution subroutines of Versions 1 and 2, and programmed Henry's singular value decomposition. Most of the data file structures and reporting formats of Version 3 were retained to take advantage of the documentation in the EPA user's manual prepared for that version. Various revisions to this version were made at DRI and by Mr. Pat Hanrahan of Oregon DEQ in response to user feedback about "bugs" and convenience features.

Dr. Richard T. Decesar of Oregon Graduate Center and Dr. Luke Wijnberg of PEI Associates, Inc. (PEI) began investigating theoretical methods of identifying collinearity related to CMB applications in 1984, and they identified additional diagnostics for multiple linear regression solutions which are applicable to the CMB (Belsley et al., 1980). Dr. Wijnberg specified these diagnostics, and they were implemented in the CMB program at DRI during the latter half of 1985. The utility of these diagnostics for practical applications is a topic of current research.

This Version 5 software was distributed to a number of scientists, and regulators who applied it to both synthesized and real data sets. These reviewers conversed at a workshop in May 1986, to exchange their results,

recommend changes to the software, and to construct applications, validation, and model reconciliation protocols. Their recommendations have been implemented via a series of revisions to Version 5, resulting in the current Version 6 described in this manual.

A.5 EFFECTIVE VARIANCE WEIGHTED LEAST SQUARES SOLUTION ALGORITHM

The effective variance weighted least squares solution to a set of CMB equations requires estimates of the source contributions, S_j, which are the unknowns to be calculated. The solution algorithm is, therefore, an iterative one that calculates a new set of S_j based on the S_j estimated from the previous iteration. The first iteration is performed with the S_j arbitrarily set to zero, and it is equivalent to the ordinary weighted least squares CMB solution proposed by Friedlander (1973). Derivations of the effective variance solution to overdetermined sets of linear equations are presented by Britt and Luecke (1973) and Watson et al. (1984). The algorithm was originally implemented in Version 1 of the CMB software, and is carried out by the following steps expressed in matrix notation. A superscript k is used to designate the value of a variabl e at the kth iteration

1. Set initial estimate of the source contributions equal to zero

$$S_j^{k=0} = 0, \quad j=1 \ldots . J$$

2. Calculate the diagonal components of the effective variance matrix, $\underline{V}_e$. All off-diagonal components of this matrix are equal to zero,

$$V_{eii}^{k} = \sigma_{C_i}^2 + \sum_{j=1}^{J} S_j^{k2} \sigma_{F_{ij}}^2$$

3. Calculate the k+1 value of S_j

$$\underline{S}^{k+1} = (\underline{F}^T(\underline{V}_e^k)^{-1}\underline{F})^{-1} \underline{F}^T(\underline{V}_e^k)^{-1}\underline{C}$$

4. Test the (k+1)th iteration of the S_j against the kth iteration. If any one differs by more than 1 percent, then perform the next iteration. If all differ by less than 1 percent, then terminate the algorithm.

If $\frac{S_j^{k+1} - S_j^k}{S_j^k} > 0.01$, go to Step 2

If $\frac{S_j^{k+1} - S_j^k}{S_j^k} \leq 0.01$, go to Step 5.

5. Assign the $\underline{(k+1)}$th iteration to S_j and σ_{S_j}. All other calculations are performed with these final values of S_j and σ_{S_j}

$$\sigma_{S_j} = \left[(\underset{\sim}{F}^T(\underset{\sim}{V}_e^{k+1})^{-1}\underset{\sim}{F})^{-1}_{jj}\right]^{\frac{1}{2}}, \quad j = 1 \ldots\ldots J$$

The number of iterations required to meet the test in Step 4 normally ranges from three to five. Version 6 allows as many as 20 iterations to take place before returning a result. A warning message is given if a 1 percent convergence has not been reached within 20 iterations. This lack of convergence is usually indicative of an unstable solution caused by collinearity of source profiles, and other performance measures given by the Version 6 software will normally confirm this. The source contribution estimates should not be believed when convergence is not obtained. A different combination of species and sources in the solution should be used in this case.

A.6 EFFECTS OF DEVIATIONS FROM CMB MODEL ASSUMPTIONS

Assumptions 1 through 6 stated in Section A.2 are fairly restrictive, and they will never be strictly complied with in actual practice. Fortunately, the CMB model can tolerate reasonable deviations from these assumptions, though these deviations increase the stated undertainties of the source contribution estimates.

The CMB model has been subjected to a number of tests to determine its abilities to tolerate deviations from model assumptions (Watson, 1979; Gordon et al., 1981; Currie et al., 1984; DeCesar et al., 1985a,b; Watson et al., 1985; Watson and Robinson, 1984; Henry, 1982; Dzubay et al., 1984; Javitz and Watson,

1986; Watson et al., 1986). These studies all point to the same basic conclusions regarding deviations from the above-stated assumptions.

With regard to assumption 1, source compositions are known to vary substantially among sources, and even within a single source over an extended period of time. These variations are both systematic and random and are caused by three phenomena: 1) transformation and deposition between the emissions point and the receptor, 2) differences in fuel type and operating processes between similar sources or the same source in time, and 3) uncertainties or differences between the source profile measurement methods. Evaluation studies have generally compared CMB results from several tests using randomly perturbed input data and from substitutions of different source profiles for the same source type.

The general conclusions drawn from these tests are:

> The error in the estimated source contributions due to biases in all of the elements of a source profile is in direct proportion to the magnitude of the biases.
>
> For random errors, the magnitude of the source contribution errors decreases as the number of components increases.

With regard to the nonlinear summation of species, assumption 2, no studies have been performed to evaluate deviations from this assumption. While these deviations are generally assumed to be small, conversion of gases to particles and reactions between particles are not inherently linear processes. This assumption is especially applicable to the end products of photochemical reactions and their apportionment to the sources of their precursors. Further model evaluation is necessary to determine the tolerance of the CMB mocel to deviation from this assumation.

A major challenge to the application of the CMB is the identification of the primary contributing sources for inclusion in the model, assumption 3. Watson (1979) systematically increased the number of sources contributing to his simulated data from four to eight contributors while solving the CMB equations assuming only four sources. He also examined including more sources in the least squares solutions than those which were actually contributors, with the following results:

Underestimating the number of sources had little effect on the calculated source contributions if the prominent species contributed by the missing source were excluded from the solution.

When the number of sources was underestimated, and when prominent species of the omitted sources were included in the calculation of source contributions, the contributions of sources with properties in common with the omitted sources were overestimated.

When source types actually present were excluded from the solution, ratios of calculated to measured concentrations were often outside of the 0.5 to 2.0 range, and the sum of the source contributions was much less than the total measured mass. The low calculated/measured ratios indicated which source compositions should be included.

When the number of sources was overestimated, the sources not actually present yielded contributions less than their standard errors if their source profiles were significantly distinct from those of other sources. The over-specification of sources decreased the standard errors of the source contribution estimates.

Recent research suggests that assumption 3 should be restated to specify that source contributions above detection limits should be included in the CMB. At this time, however, it is not yet possible to determine the "detection limit" of a source contribution at a receptor since this is a complicated and unknown function of the other source contributions, the source profile uncertainties, and the uncertainties of the receptor measurements. Additional model testing is needed to define this "detection limit."

With regard to assumption 4, it is very likely that the true number of individual sources contributing to receptor concentrations is much larger than the number of species that can be measured. It is therefore necessary to group sources into source types of similar compositions so that this assumption is met. For the most commonly measured species, meeting assumption 5 practically defines these groupings.

The linear independence of source compositions required by assumption 5 has become a subject of considerable interest since the publication of Henry's (1982) singular value decomposition (SVD) analysis. This analysis provides quantitative measures of collinearity and the sensitivity of CMB results to specific receptor concentrations, which can be calculated analytically in each application. Henry (1982) also proposed an optimal linear combination

of source contributions that have been determined to be collinear. These "regression diagnostics" have been summarized by Belsley et al. (1980) and have been applied to the CMB by DeCesar et al. (1985 a,b). Version 6 of the CMB has incorporated these regression diagnostics to test their ability to identify potential deviations from model assumptions. Section A.7 describes them as they have been implemented in Version 6. At this point, the exact values of these diagnostics that correspond to intolerable degrees of multicollinearity are unknown, though this is an area of active research. These diagnostics are available to advanced users of Version 6, and for research purposes only, via a set of undocumented commands which may be obtained from the author.

Tests performed on simulated data with obviously collinear source profiles typically result in positive and negative values for the collinear source types. Unless the source profiles are nearly identical, the sum of these large positive and negative values very closely approximates the sum of the true contributions.

With most commonly measured species (e.g., ions, elements, and carbon) and source types (e.g., motor vehicle, geological, residual oil sea salt, steel production, wood burning, and various industrial processes), from five to seven sources are linearly independent of each other, in most cases.

Gordon et al. (1981) found instabilities in the ordinary weighted least square solutions to the CMB equations when species presumed to be "unique" to a certain source type were removed from the solution. Using simulated data with known perturbations ranging from 0 to 20 percent, Watson (1979) found: "In the presence of likely uncertainties, sources such as urban dust and continental background dust cannot be adequately resolved by least square fitting, even though their compositions are not identical. Several nearly unique ratios must exist for good separation." It is widely recognized at this time that a substantial amount of research is needed to minimize the effects of source profile collinearity in the CMB. The Version 6 software has been structured to facilitate this research as well as to alert regulatory users to the potential of excessive collinearity in routine applications.

With respect to assumption 6, the randomness, normality, and uncorrelated nature of measurement uncertainties, there are no results available from

verification or evaluation studies. Every least squares solution to the CMB equations requires this assumption, as demonstrated by the derivation of Watson et al. (1984). In reality, very little is known about the distribution of errors for the source compositions and the ambient concentrations. If anything, the distribution probably follows a log-normal rather than a normal distribution, since none of these values can ever be negative, while a normal distribution allows for a substantial proportion of negative values. For small errors (e.g., less than 20%) the actual distribution may not be important; but for large errors it probably is important. A symmetric distribution becomes less probable as the coefficient of variation of the measurement increases. This is one of the most important assumptions of the solution method that requires testing.

A.7 CMB MODEL PERFORMANCE MEASURES

Though the CMB model evaluation studies cited in Section A.6 provide some general conclusions regarding the effects of deviations from model assumptions, these studies cannot assess the validity of a single CMB model application. Several analytical performance measures have been incorporated into Version 6 software to allow the confidence in the model solution to be evaluated. The values taken by these performance measures will suggest actions which might be taken to correct or add model input data, change fitting sources or species, or to combine source profiles.

A.7.1 Reduced Chi Square

The reduced chi square and its associated degrees of freedom indicate the "goodness-of-fit" of a solution to the CMB equations.

$$\text{Reduced chi square} = \chi^2_{I-J} = \frac{1}{I-J} \sum_{i=1}^{I} \left[\frac{C_i - \sum_{j=1}^{J} F_{ij} S_j}{V_{eii}^{\frac{1}{2}}} \right]^2$$

Degrees of freedom = DF = I - J

For a perfect least squares fit with normally distributed random deviations in the measurements of C_i and F_{ij}, the reduced chi square will equal 1 (Bevington, 1969).

Values of reduced chi square which exceed 10 indicate that one or more ambient species concentrations calculated from the source contributions differs from its measured value by more than the measurement and source contribution uncertainties. The calculated-to-measured ratios and residual to uncertainty measures indicate which species are most responsible for the large reduced chi square.

A.7.2 Percent Mass

Percent mass is the fraction of the measured mass concentration which is accounted for by the source contribution estimates,

$$\text{Percent mass} = \left(\sum_{j=1}^{J} S_j/C_t \right) \times 100$$

A percent mass less than 80 percent usually means that not all source types have been included in the CMB.

A.7.3 R-Square

R-Square is a measure of correlation between the calculated and measured species concentrations. The definition of R-Square in the CMB assumes a least squares solution with a zero intercept.

$$e_i = \frac{C_i - \sum_{j=1}^{J} F_{ij} S_j}{V_{e_{ii}}^{\frac{1}{2}}}$$

$$\text{R-Square} = 1 - \frac{\sum_{i=1}^{I} e_i^2}{\sum_{i=1}^{I} \left(\frac{C_i}{V_{e_{ii}}^{\frac{1}{2}}} \right)^2}$$

A.7.4 Measures of Residuals

The calculated to measured concentration ratio (RATIO C/M_i), indicates how closely the species concentrations calculated from the source contribution estimates compare with each other. Only the $RATIO_i$ is displayed

$$\text{RATIO } C/M_i = \sum_{j=1}^{J} F_{ij}S_j/C_i$$

$$\text{RATIO } R/U_i = \left(\sum_{j=1}^{J} F_{ij}S_j - C_i \right)/\sigma_{R_i}$$

A.7.5 Singular Value Decomposition (Uncertainty/Similarity Clusters)

The Singular Value Decomposition of the weighted $\underset{\sim}{F}$ matrix is given by

$$\underset{\sim}{V}_e^{-\frac{1}{2}} \underset{\sim}{F} = \underset{\sim}{U}\underset{\sim}{D}\underset{\sim}{V}^t \qquad (1)$$

where $\underset{\sim}{U}$ and $\underset{\sim}{V}$ are IxI and JxJ orthogonal matrices, respectively, and $\underset{\sim}{D}$ is a diagonal matrix with J nonzero and positive elements called the singular values of the decomposition. They are also given as the square roots of the eigenvalue of the symmetric matrix $\underset{\sim}{F}^T \underset{\sim}{V}_e^{-1} \underset{\sim}{F}$ with the columns of $\underset{\sim}{V}$ the associated eigenvectors.

It is these eigenvectors which give rise to the uncertainty/similarity clusters since rewriting (1) gives

$$\underset{\sim}{V}_e^{-\frac{1}{2}} \underset{\sim}{F} \underset{\sim}{V} = \underset{\sim}{U} \underset{\sim}{D}. \qquad (2)$$

Thus, if a singular value in $\underset{\sim}{D}$ is "small" one has a "small" or nearly degenerate linear combination of the columns of the weighted $\underset{\sim}{F}$ matrix with coefficients given by the elements of the column of $\underset{\sim}{V}$ corresponding to the "small" singular value. In the CMB model the sources appearing in the uncertainty/similarity clusters arise from these eigenvectors and are characterized by:

1. Eigenvector elements which have absolute value greater than or equal to 0.25

2. At least two sources satisfy (1)

3. At least two of the sources satisfying (1) have a TSTAT value less than or equal to 2.0.

A.8 SOLUTION METHODS

Version 6 of the CMB model is the first that contains no user options for using the least squares solution procedure (or fitting method). The current version employs the effective variance weighted least squares method exclusively. In previous versions, the following solution methods have been used:

- Ordinary weighted least squares - choice in Versions 1 through 4. Versions 3 and 4 provided for an implicit intercept
- Effective variance weighted least squares - choice in Versions 1 through 4. Versions 3 and 4 provided for an implicit intercept
- Britt - Luecke exact solution choice in Versions 1 and 4
- Ridge regression - choice in Versions 3 and 4.

All solution methods use the same multiple linear regression approach and matrix solution of the simultaneous equations. The differences lie in handling of the error term(s). In ordinary weighted least squares, each of the i equations is modified by dividing both sides of the equation by the square of the standard error of the species concentration ($\sigma^2_{C_i}$) to weight the more accurate concentration data more heavily. However, the errors in the source profile fractions ($\sigma_{F_{ij}}$), which are usually larger than σ_{C_i}, are ignored in the ordinary weighted least squares method.

The error term $V_{e_{ii}}$ in the effective variance weighted least squares method includes both σ_{C_i} and $\sigma_{F_{ij}}$ components, as explained in Section A.5. Effective variance is an estimate of total uncertainty, which takes into

account errors in all input data. The same error term $V_{e_{ii}}$ is also used in the Britt-Luecke solution, which is intended to provide an exact solution of the CMB equations. The Britt-Luecke procedure is not employed in Version 6 because it has not been shown to significantly improve the CMB, despite its greater complexity and computational requirements.

Both the effective variance least squares and Britt-Luecke solutions require the iterative process for estimating S_j values because the error term, $V_{e_{ii}}$, is a function of S_j. The ordinary weighted least squares method has no $\sigma_{F_{in}}$ term, and therefore does not require a convergent solution.

The ridge regression fitting method is described in the User's Manual for Version 3 of the CMB Model (EPA-450/4-83-014, July 1983).

APPENDIX B

DIAGNOSIS OF ERROR MESSAGES

Error message	Probable cause	Corrections
File system error in file CMBOUT Error Code 1034, Status 0004	CONFIG.SYS file not on DOS used to boot system or not in root directory	Increase number of files to 14 in CONFIG.SYS file
Error opening (XXXXXX.XXX) What is the name of your ____ ____	1) File name does not exist or is incorrect 2) File name extender not used 3) CONFIG.SYS has insufficient number of files	Leave interactive session using CTRL-ALT-DEL and correct file name or CONFIG.SYS
AKT*JEFFIN*AK MATRIX IN SUBROUTINE CEB2 NEEDS IMPROVEMENT	1) Too many fitting sources for number of fitting species (No. species minus No. sources must be $\geq$ 1) 2) A single constituent source (e.g., secondary sulfate or nitrate) is included, but that constituent is not a fitting species	Reduce sources or add species Add omitted species if ambient data are available otherwise delete the single species source
Some receptor concentration standard errors are less than or equal to zero. Weighted regression cannot be done in this case	A receptor data set has not been selected for modeling and each ambient concentration is set at a default value of 0.0	In response to ENTER COMMAND, type SELECT and then specify site, date, and size range
File not found in file ____ Error Code 1032, Status 000A	File name extension not included	Specify full file name as given in directory, e.g., XXXXXXX.DAT
RECORD NUMBER XX OF SOURCE NAME FILE XXXXXX.XXX DID NOT MATCH THE REQUIRED FORMAT (I2, 2X, A8, 2X, A1) AND WAS IGNORED	If input data are improperly formatted, these data will not be read into memory and may result in erroneous data	Change the record (line) in the XXXXX.XXX file to match the required format

(continued)

Error message	Probable cause	Corrections
RECORD NUMBER XX OF POLLUTANT NAME FILE XXXXXX.XXX DID NOT MATCH THE REQUIRED FORMAT (I2, 2X, A8, 2X, A1) AND WAS IGNORED	Input data are improperly formatted	Change the record (line) in the XXXXX.XXX file to match the required format
RECORD NUMBER XXX OF SOURCE COMPOSITION FILE XXXXXX.XXX DID NOT MATCH THE REQUIRED FORMAT AND WAS IGNORED	Input data are improperly formatted	Change the record (line) in the XXXXX.XXX file to match the required format
AN UNIDENTIFIED POLLUTANT CODE # XX WAS FOUND IN FILE XXXXXX.XXX SOURCE - SO4	A pollutant code found in the source composition file did not appear in the pollutant name file	The model will ignore the extra pollutant code
FITTING ELEMENT # XX HAS NON-POSITIVE UNCERTAINTY - .00000 PLEASE REPLACE WITH POSITIVE DETECTION LIMIT PROGRAM TERMINATED Stop - Program terminated	The model does not allow a nonpositive uncertainty in the ambient concentration file. Fitting element No. XX refers to the XXth fitting species in the species list	Correct the uncertainty associated with that fitting species and rerun CMB5
COMMAND NOT EXECUTED - TRY AGAIN	The command entered was not valid	Type HELP TO VIEW THE VALID COMMAND LIST
***** in any output field	The results are too large for that field	Use the CMB Application and Validation Protocol to help locate the problems with the input data that would lead to large (+ or -) results for that field
No convergence after 20 iterations - ENTER a carriage return TO VIEW RESULTS	Major collinearity, missing source, incorrect source profile or data has kept the model from meeting its internal criteria for completing its iteractive calculations	Enter <CR>, view the results and use the CMB Application and Validation Protocol to help locate the problem

(continued)

Error message	Probable cause	Corrections
SSCONT output wraps around so each species takes two lines instead of 1. The top of the table scrolls off of the screen	More than nine sources are included in the source list	Reenter the SSCONT command and press <CTRL><NUMLOCK> simultaneously to stop the scroll. Press any key to continue the scroll
Species number has zero ambient data uncertainty	The referenced ambient species has an uncertainty < or = to zero	Make uncertainty equal to detection limit

APPENDIX C

PRINTOUT OF TEST DATA SET FOR PACS1

(<<**** USER RESPONSE)

```
A>CMB6                                                                  <<****
MAKE SURE THAT YOUR CAPS LOCK IS ON ! ! !


DISK FILE FOR INITIAL INPUT?
IF NOT ENTER CARRIAGE RETURN
IF SO ENTER NAME OF DISK FILE
INPORT.DAT                                                              <<****

DO YOU WISH TO RENAME CMBOUT?
IF NOT ENTER A CARRIAGE RETURN.
IF SO ENTER THE FILE NAME.
CMBEXAM1                                                                <<****

*****************************************************************************
*                                                                           *
*              U. S. EPA CHEMICAL MASS BALANCE RECEPTOR MODEL               *
*              *** IBM-PC VERSION 6.0  (REVISED  7/20/87) ***               *
*                                                                           *
*       EPA PROJECT MGR: THOMPSON G. PACE III, PE                           *
*                        U.S. ENVIRONMENTAL PROTECTION AGENCY               *
*                        OFFICE OF AIR QUALITY PLANNING AND STANDARDS       *
*                        RESEARCH TRIANGLE PARK, NC                         *
*                        (919)-541-5585                                     *
*                                                                           *
*                                                                           *
*     PRINCIPAL AUTHOR:     DR. JOHN G. WATSON                            *
*                           DESERT RESEARCH INSTITUTE                       *
*                           UNIVERSITY OF NEVADA SYSTEM                     *
*                           (702)-972-1676                                  *
*     CONTRIBUTING AUTHORS:                                                 *
*     DR. J.C. CHOW         MR. P.L. HANRAHAN       DR. N.F. ROBINSON       *
*     MR. J.E. CORE         DR. R.C. HENRY          DR. H.J. WILLIAMSON     *
*     MR. D.A. DUBOSE       MR. T.G. PACE           DR. L. WIJNBERG         *
*                                                                           *
*****************************************************************************
```

```
THE "HELP" COMMAND LISTS COMMANDS.
USE COMMANDS AE,DE,AS,DS FOR CHANGES.
ENTER COMMAND
HELP                                                              <<****

HELP-LISTS THESE COMMANDS

---- DATA ACCESS AND SEQUENCING ----
SELECT-SELECT DATA SET FOR CMB
SIZE-CHANGE SIZE FRACTION
EXIT-CLOSE FILES AND LEAVE

---- CMB OPERATIONS ----
AE-ADD A POLLUTANT TO THE FIT
DE-DELETE A POLLUTANT FROM THE FIT
AS-ADD A SOURCE TO THE FIT
DS-DELETE A SOURCE FROM THE FIT
CMB-PERFORM CMB

---- SCREEN DISPLAY ----
PINFO - PRINT CURRENT STATUS ON SCREEN
PMATRIX - PRINT SOURCE SIGNATURE OR RECEPTOR CONCENTRATIONS
SSCONT - PRINT SOURCE CONTRIBUTIONS TO ELEMENTAL CONCS. TO SCREEN
PCOMP - PRINT COMPUTED AVERAGES OF CMB SERIES

--- DATA STORAGE ----
WRITE - WRITE CMB RESULTS TO DISK

SELECT SHOULD NORMALLY BE THE FIRST COMMAND
ENTER COMMAND
SELECT                                                            <<****

ENTER DESIRED CMB SITE CODE: XXXXXXXXXXX
PACS1                                                             <<****

ENTER YEAR: YY
77                                                                <<****

ENTER DATE: MMDD
0813                                                              <<****

INPUT DESIRED SIZE FRACTION:(FINE OR COARSE)
F                                                                 <<****

DATA SEARCH BEGUN FOR
     SITE: PACS1          YEAR: 77  DATE: 0813

     SITE: PACS1          YEAR: 77  DATE: 0813     FC
```

```
ENTER COMMAND
CMB                                                                    <<*****

SOURCE CONTRIBUTION ESTIMATES -  SITE: PACS1        DATE: 0813 77   VERSION: 6.0
SAMPLE DURATION        24      START HOUR         0        SIZE: FINE
       R SQUARE       .98    PERCENT MASS      98.7
     CHI SQUARE      1.12              DF        13

  SOURCE
  * TYPE      SCE(UG/M3)     STD ERR       TSTAT
-----------------------------------------------
  1 MARIN       12.3882      2.2452      5.5177
  3 UDUST        9.5918      1.3876      6.9127
  4 AUTPB       10.0834      1.4941      6.7486
  5 RDOIL       11.0601      1.9238      5.7490
  8 KRAFT        4.6928      5.0424       .9307
 11 ALPRO       10.6022      3.5896      2.9536
 12 STEEL        8.6728      1.3771      6.2980
 13 FERMN       11.8754      1.8321      6.4820
-----------------------------------------------

MEASURED CONCENTRATION FINE/COARSE/TOTAL:
      80.0+-      8.0/      80.0+-      8.0/     160.0+-     11.3

       UNCERTAINTY/SIMILARITY CLUSTERS  VERSION: 6.0     SUM OF CLUSTER SOURCES
-------------------------------------------------------------------------------

        1  8                                                17.081+-      4.238
        1  5  8                                             28.141+-      3.829
-------------------------------------------------------------------------------
```

SPECIES CONCENTRATIONS - SITE: PACS1 DATE: 0813 77 VERSION: 6.0

SAMPLE DURATION		24	START HOUR		0	SIZE: FINE				
R SQUARE		.98	PERCENT MASS		98.7					
CHI SQUARE		1.12	DF		13					

SPECIES	I	M	MEAS		CALC		RATIO C/M		RATIO R/U
1 TOT			80.00000+-	8.00000	78.96679+-	4.82217	.99+-	.12 TOT	-.1
9 F			.88300+-	.08800	.67644+-	.24792	.77+-	.29 F	-.8
11 NA	*		6.93000+-	.69300	6.97037+-	.56443	1.01+-	.13 NA	.0
12 MG	*		.43000+-	.04300	1.60949+-	.62626	3.74+-	1.50 MG	1.9
13 AL	*		4.66000+-	.46600	4.02416+-	.88919	.86+-	.21 AL	-.6
14 SI	*		3.02000+-	.30200	2.92212+-	.13329	.97+-	.11 SI	-.3
16 S			2.95000+-	.29500	3.02499+-	.31806	1.03+-	.15 S	.2
17 CL	*		5.95000+-	.59500	5.69358+-	1.24829	.96+-	.23 CL	-.2
19 K	*		1.64000+-	.16400	1.73088+-	.46411	1.06+-	.30 K	.2
20 CA	*		1.78000+-	.17800	1.43535+-	.11366	.81+-	.10 CA	-1.6
22 TI	*		.08300+-	.00800	.10088+-	.01630	1.22+-	.23 TI	1.0
23 V	*		.37200+-	.03700	.39756+-	.08308	1.07+-	.25 V	.3
24 CR	*		.31500+-	.03200	.20977+-	.12151	.67+-	.39 CR	-.8
25 MN	*		2.99000+-	.29900	2.82843+-	.14115	.95+-	.11 MN	-.5
26 FE	*		4.53000+-	.45300	4.24446+-	.33269	.94+-	.12 FE	-.5
28 NI	*		.76500+-	.07700	.68248+-	.13427	.89+-	.20 NI	-.5
29 CU	*		.04400+-	.00400	.05274+-	.00510	1.20+-	.16 CU	1.3
30 ZN	*		.22500+-	.02300	.26786+-	.03966	1.19+-	.21 ZN	.9
35 BR	*		.41900+-	.04200	.56133+-	.17386	1.34+-	.44 BR	.8
82 PB	*		2.53000+-	.25300	2.13748+-	.30300	.84+-	.15 PB	-1.0
91 OC	*		7.54000+-	.75400	8.50981+-	1.35631	1.13+-	.21 OC	.6
92 EC	*		1.42000+-	.14200	1.33579+-	.34011	.94+-	.26 EC	-.2
93 SO4	*		10.30000+-	1.03000	9.78932+-	1.47512	.95+-	.17 SO4	-.3
94 NO3	*		.63800+-	.06400	.88401+-	.35938	1.39+-	.58 NO3	.7

ENTER COMMAND
SSCONT <<****

$$\text{INDIVIDUAL RATIO} = \frac{\text{CALC SPECIES(PER SOURCE)}}{\text{MEAS SPECIES(ALL SOURCES)}}$$

SPEC\SOURCE	1	3	4	5	8	11	12	13
1 TOTAL	.155	.120	.126	.138	.059	.133	.108	.148
9 F	.000	.000	.000	.007	.000	.720	.000	.039
11 NA	.715	.017	.000	.056	.086	.063	.016	.053
12 MG	1.383	.290	.000	.000	.069	.690	1.311	.000
13 AL	.000	.182	.024	.013	.003	.614	.012	.016
14 SI	.000	.708	.027	.035	.002	.012	.144	.039
16 S	.139	.012	.014	.499	.186	.050	.058	.068
17 CL	.833	.000	.051	.000	.014	.024	.027	.008
19 K	.106	.060	.004	.019	.043	.014	.049	.760
20 CA	.097	.131	.071	.098	.000	.020	.302	.087
22 TI	.000	.740	.000	.147	.003	.051	.209	.066
23 V	.000	.006	.000	1.023	.000	.018	.014	.008
24 CR	.000	.014	.000	.017	.042	.000	.578	.016
25 MN	.000	.004	.000	.002	.000	.000	.252	.687
26 FE	.000	.127	.047	.073	.012	.011	.613	.055
28 NI	.000	.001	.002	.775	.008	.026	.079	.000
29 CU	.000	.065	.167	.189	.022	.106	.552	.097
30 ZN	.000	.047	.157	.197	.014	.007	.463	.306
35 BR	.059	.005	1.203	.003	.015	.009	.000	.045
82 PB	.000	.014	.797	.005	.000	.001	.026	.002
91 OC	.000	.150	.669	.103	.011	.055	.000	.142
92 EC	.000	.125	.270	.241	.007	.172	.000	.125
93 SO4	.120	.004	.013	.516	.182	.045	.021	.048
94 NO3	.000	.000	.144	.113	.000	.068	.000	1.061

ENTER COMMAND
WRITE <<****

WRITTEN
ENTER COMMAND
SIZE <<****
SIZE IS COARSE

ENTER COMMAND
CMB <<*****

SOURCE CONTRIBUTION ESTIMATES - SITE: PACS1 DATE: 0813 77 VERSION: 6.0
SAMPLE DURATION 24 START HOUR 0 SIZE: COARSE
R SQUARE .97 PERCENT MASS 100.6
CHI SQUARE 1.44 DF 13

*	SOURCE TYPE	SCE(UG/M3)	STD ERR	TSTAT
1	MARIN	10.6029	1.8239	5.8132
3	UDUST	9.5985	1.2616	7.6081
4	AUTPB	9.0906	1.3961	6.5113
5	RDOIL	9.7127	1.6107	6.0302
8	KRAFT	12.3264	7.9091	1.5585
11	ALPRO	11.0997	2.2441	4.9462
12	STEEL	8.1587	1.5239	5.3540
13	FERMN	9.8720	1.6165	6.1072

MEASURED CONCENTRATION FINE/COARSE/TOTAL:
80.0+- 8.0/ 80.0+- 8.0/ 160.0+- 11.3

UNCERTAINTY/SIMILARITY CLUSTERS VERSION: 6.0 SUM OF CLUSTER SOURCES

SPECIES CONCENTRATIONS - SITE: PACS1 DATE: 0813 77 VERSION: 6.0
SAMPLE DURATION 24 START HOUR 0 SIZE: COARSE
R SQUARE .97 PERCENT MASS 100.6
CHI SQUARE 1.44 DF 13

SPECIES	I	M	MEAS		CALC		RATIO C/M		RATIO R/U
1 TOT			80.00000	8.00000	80.46153	6.36333	1.01	.13 TOT	.0
9 F			.73400	.07300	.50073	.22362	.68	.31 F	-1.0
11 NA		*	6.33000	.63300	6.07759	.47089	.96	.12 NA	-.3
12 MG		*	1.48000	.14800	1.48676	.52389	1.00	.37 MG	.0
13 AL		*	4.84000	.48400	4.40992	.55081	.91	.15 AL	-.6
14 SI		*	3.27000	.32700	3.38783	.15779	1.04	.11 SI	.3
16 S			2.50000	.25000	2.41255	.28216	.97	.15 S	-.2
17 CL		*	4.67000	.46700	5.19693	1.07658	1.11	.26 CL	.4
19 K		*	1.12000	.11200	1.44197	.38620	1.29	.37 K	.8
20 CA		*	1.52000	.15200	1.47195	.10333	.97	.12 CA	-.3
22 TI		*	.14000	.01400	.13692	.03204	.98	.25 TI	-.1
23 V		*	.27700	.02800	.34841	.07289	1.26	.29 V	.9
24 CR		*	.00800	.00100	.24531	.11640	30.66	15.05 CR	2.0
25 MN		*	2.47000	.24700	2.43813	.12229	.99	.11 MN	-.1
26 FE		*	5.41000	.54100	4.11644	.30571	.76	.09 FE	-2.1
28 NI		*	.77900	.07800	.63016	.11802	.81	.17 NI	-1.1
29 CU		*	.05000	.00500	.06613	.00683	1.32	.19 CU	1.9
30 ZN		*	.21400	.02100	.23739	.03413	1.11	.19 ZN	.6
35 BR		*	.52000	.05200	.51379	.15795	.99	.32 BR	.0
82 PB		*	1.78000	.17800	1.93077	.27347	1.08	.19 PB	.5
91 OC		*	10.10000	1.01000	8.38185	1.29686	.83	.15 OC	-1.0
92 EC		*	1.68000	.16800	1.34287	.39741	.80	.25 EC	-.8
93 SO4		*	8.10000	.81000	8.11918	1.25659	1.00	.18 SO4	.0
94 NO3		*	1.13000	.11300	.71048	.41362	.63	.37 NO3	-1.0

```
ENTER COMMAND
SSCONT                                                                    <<****

                                 CALC SPECIES(PER SOURCE)
          INDIVIDUAL RATIO = ------------------------
                                 MEAS SPECIES(ALL SOURCES)

SPEC\SOURCE      1      3      4      5      8     11     12     13
 1 TOTAL       .133   .120   .114   .121   .154   .139   .102   .123
 9 F           .000   .001   .000   .007   .000   .635   .000   .039
11 NA          .670   .027   .000   .054   .103   .042   .016   .048
12 MG          .344   .100   .000   .000   .000   .202   .358   .000
13 AL          .000   .131   .021   .011   .007   .718   .011   .013
14 SI          .000   .822   .023   .029   .005   .003   .125   .030
16 S           .140   .000   .015   .517   .163   .000   .064   .067
17 CL          .908   .000   .058   .000   .077   .029   .032   .009
19 K           .133   .088   .006   .024   .044   .000   .067   .925
20 CA          .098   .189   .075   .101   .029   .059   .333   .084
22 TI          .000   .692   .000   .076   .000   .060   .117   .032
23 V           .000   .009   .000  1.206   .000   .016   .018   .009
24 CR          .000   .540   .000   .571  7.396   .222 21.417   .518
25 MN          .000   .004   .000   .002   .003   .000   .287   .691
26 FE          .000   .102   .035   .053   .042   .008   .483   .038
28 NI          .000   .000   .002   .668   .035   .030   .073   .000
29 CU          .000   .058   .133   .146   .148   .311   .457   .071
30 ZN          .000   .049   .149   .182   .000   .005   .457   .268
35 BR          .041   .001   .874   .002   .013   .026   .000   .030
82 PB          .000   .020  1.021   .006   .000   .000   .035   .002
91 OC          .000   .032   .450   .067   .193   .000   .000   .088
92 EC          .000   .089   .206   .179   .132   .106   .000   .088
93 SO4         .131   .001   .015   .577   .180   .023   .025   .051
94 NO3         .000   .002   .073   .056   .000   .000   .000   .498

ENTER COMMAND
WRITE                                                                     <<****

WRITTEN
WRITTEN

ENTER COMMAND
EXIT                                                                      <<****

Stop - Program terminated.
```

*** CMB SOURCE CONTRIBUTION SUMMARY ***

RESULTS FOR CMB SITE: PACS1 YEAR: 77 DATE: 0613 VERSION: 6.0
SAMPLING DURATION: 24 HRS. WITH START HOUR: 0

SOURCE	FINE				COARSE				TOTAL			
	UG/M3		I		UG/M3		I		UG/M3		I	
MARIN	12.388	+- 2.245	15.485	+- 3.205	10.603	+- 1.824	13.254	+- 2.637	22.991	+- 2.893	14.369	+- 2.074
UDUST	9.592	+- 1.388	11.990	+- 2.109	9.599	+- 1.262	11.998	+- 1.982	19.190	+- 1.875	11.994	+- 1.447
AUTP8	10.083	+- 1.494	12.604	+- 2.253	9.091	+- 1.396	11.363	+- 2.082	19.174	+- 2.045	11.984	+- 1.533
RDOIL	11.060	+- 1.924	13.825	+- 2.774	9.713	+- 1.611	12.141	+- 2.351	20.773	+- 2.509	12.983	+- 1.817
KRAFT	4.693	+- 5.042	5.866	+- 6.330	12.326	+- 7.909	15.408	+- 10.006	17.019	+- 9.380	10.637	+- 5.910
ALPRO	10.602	+- 3.590	13.253	+- 4.679	11.100	+- 2.244	13.875	+- 3.129	21.702	+- 4.233	13.564	+- 2.814
STEEL	8.673	+- 1.377	10.841	+- 2.034	8.159	+- 1.524	10.198	+- 2.161	16.832	+- 2.054	10.520	+- 1.484
FERMN	11.875	+- 1.832	14.844	+- 2.729	9.872	+- 1.616	12.340	+- 2.368	21.747	+- 2.443	13.592	+- 1.804
CALC.MASS	78.967	+- 4.822	98.708	+- 11.566	80.462	+- 6.363	100.577	+- 12.823	159.428	+- 7.984	99.643	+- 8.634
MEAS.MASS	80.000	+- 8.000			80.000	+- 8.000			160.000	+- 11.314		

RESULTS FOR CMB SITE: PACS1 YEAR: 77 DATE: 0813 VERSION: 6.0
FINE PARTICULATE FRACTION
SAMPLING DURATION: 24 HRS. WITH START HOUR: 0
R-SQUARE: .98
CHI SQUARE: 1.12
DF: 13

#	TYPE	UG/M3		%	
1	MARIN	12.388+-	2.245	15.485+-	3.205
3	UDUST	9.592+-	1.388	11.990+-	2.109
4	AUTPB	10.083+-	1.494	12.604+-	2.253
5	RDOIL	11.060+-	1.924	13.825+-	2.774
8	KRAFT	4.693+-	5.042	5.866+-	6.330
11	ALPRO	10.602+-	3.590	13.253+-	4.679
12	STEEL	8.673+-	1.377	10.841+-	2.034
13	FERMN	11.875+-	1.832	14.844+-	2.729
	TOTAL:	78.967+-	4.822	98.708+-	11.566

UNCERTAINTY/SIMILARITY CLUSTERS:	SUM OF CLUSTER SOURCES	
1 8	17.081+-	4.238
1 5 8	28.141+-	3.829

SPECIES	INCL	MISS FLG	FINE SUSPENDED PARTICULATE MEAS. UG/M3		PERCENT		CALC. UG/M3		RATIO R/U	
1 TOTAL			80.00000+-	8.00000	100.00000+-	14.14214	78.96679+-	4.82217	-.1	TOTAL
9 F			.88300+-	.08800	1.10375+-	.15583	.67644+-	.24792	-.8	F
11 NA		*	6.93000+-	.69300	8.66250+-	1.22506	6.97037+-	.56443	.0	NA
12 MG		*	.43000+-	.04300	.53750+-	.07601	1.60949+-	.62626	1.9	MG
13 AL		*	4.66000+-	.46600	5.82500+-	.82378	4.02416+-	.88919	-.6	AL
14 SI		*	3.02000+-	.30200	3.77500+-	.53387	2.92212+-	.13329	-.3	SI
16 S			2.95000+-	.29500	3.68750+-	.52149	3.02499+-	.31806	.2	S
17 CL		*	5.95000+-	.59500	7.43750+-	1.05182	5.69358+-	1.24829	-.2	CL
19 K		*	1.64000+-	.16400	2.05000+-	.28991	1.73088+-	.46411	.2	K
20 CA		*	1.78000+-	.17800	2.22500+-	.31466	1.43535+-	.11366	-1.6	CA
22 TI		*	.08300+-	.00800	.10375+-	.01441	.10088+-	.01630	1.0	TI
23 V		*	.37200+-	.03700	.46500+-	.06556	.39756+-	.08308	.3	V
24 CR		*	.31500+-	.03200	.39375+-	.05613	.20977+-	.12151	-.8	CR
25 MN		*	2.99000+-	.29900	3.73750+-	.52856	2.82843+-	.14115	-.5	MN
26 FE		*	4.53000+-	.45300	5.66250+-	.80080	4.24446+-	.33269	-.5	FE
28 NI		*	.76500+-	.07700	.95625+-	.13568	.68248+-	.13427	-.5	NI
29 CU		*	.04400+-	.00400	.05500+-	.00743	.05274+-	.00510	1.3	CU
30 ZN		*	.22500+-	.02300	.28125+-	.04022	.26786+-	.03966	.9	ZN
35 BR		*	.41900+-	.04200	.52375+-	.07416	.56133+-	.17386	.8	BR
82 PB		*	2.53000+-	.25300	3.16250+-	.44725	2.13748+-	.30300	-1.0	PB
91 OC		*	7.54000+-	.75400	9.42500+-	1.33290	8.50981+-	1.35631	.6	OC
92 EC		*	1.42000+-	.14200	1.77500+-	.25102	1.33579+-	.34011	-.2	EC
93 SO4		*	10.30000+-	1.03000	12.87500+-	1.82080	9.78932+-	1.47512	-.3	SO4
94 NO3		*	.63800+-	.06400	.79750+-	.11296	.98401+-	.35938	.7	NO3

MEASURED AMBIENT MASS (UG/M3): FINE: 80.0+- 8.0 COARSE: 80.0+- 8.0 TOTAL: 160.0+-11.3

RESULTS FOR CMB SITE: PACSI YEAR: 77 DATE: 0813 VERSION: 6.0
COARSE PARTICULATE FRACTION
SAMPLING DURATION: 24 HRS. WITH START HOUR: 0
R-SQUARE: .97
CHI SQUARE: 1.44
DF: 13

#	TYPE	UG/M3	%
1	MARIN	10.603+- 1.824	13.254+- 2.637
3	UDUST	9.599+- 1.262	11.998+- 1.982
4	AUTPB	9.091+- 1.396	11.363+- 2.082
5	RDOIL	9.713+- 1.611	12.141+- 2.351
8	KRAFT	12.326+- 7.909	15.408+-10.006
11	ALPRO	11.100+- 2.244	13.875+- 3.129
12	STEEL	8.159+- 1.524	10.198+- 2.161
13	FERMN	9.872+- 1.616	12.340+- 2.368
	TOTAL:	80.462+- 6.363	100.577+-12.823

UNCERTAINTY/SIMILARITY CLUSTERS: SUM OF CLUSTER SOURCES

SPECIES		INCL	MISS FLG	COARSE SUSPENDED PARTICULATE MEAS. UG/M3		PERCENT		CALC. UG/M3		RATIO R/U	
1	TOTAL			80.00000+-	8.00000	100.00000+-	14.14214	80.46153+-	6.36333	.0	TOTAL
9	F			.73400+-	.07300	.91750+-	.12940	.50073+-	.22362	-1.0	F
11	NA	*		6.33000+-	.63300	7.91250+-	1.11900	6.07759+-	.47089	-.3	NA
12	MG	*		1.48000+-	.14800	1.85000+-	.26163	1.48676+-	.52389	.0	MG
13	AL	*		4.84000+-	.48400	6.05000+-	.85560	4.40992+-	.55081	-.6	AL
14	SI	*		3.27000+-	.32700	4.08750+-	.57806	3.38783+-	.15779	.3	SI
16	S			2.50000+-	.25000	3.12500+-	.44194	2.41255+-	.28216	-.2	S
17	CL	*		4.67000+-	.46700	5.83750+-	.82555	5.19693+-	1.07658	.4	CL
19	K	*		1.12000+-	.11200	1.40000+-	.19799	1.44197+-	.38620	.8	K
20	CA	*		1.52000+-	.15200	1.90000+-	.26870	1.47195+-	.10333	-.3	CA
22	TI	*		.14000+-	.01400	.17500+-	.02475	.13692+-	.03204	-.1	TI
23	V	*		.27700+-	.02800	.34625+-	.04923	.34841+-	.07289	.9	V
24	CR	*		.00800+-	.00100	.01000+-	.00160	.24531+-	.11640	2.0	CR
25	MN	*		2.47000+-	.24700	3.08750+-	.43664	2.43813+-	.12229	-.1	MN
26	FE	*		5.41000+-	.54100	6.76250+-	.95636	4.11644+-	.30571	-2.1	FE
28	NI	*		.77900+-	.07800	.97375+-	.13780	.63016+-	.11802	-1.1	NI
29	CU	*		.05000+-	.00500	.06250+-	.00884	.06613+-	.00683	1.9	CU
30	ZN	*		.21400+-	.02100	.26750+-	.03748	.23739+-	.03413	.6	ZN
35	BR	*		.52000+-	.05200	.65000+-	.09192	.51379+-	.15795	.0	BR
82	PB	*		1.78000+-	.17800	2.22500+-	.31466	1.93077+-	.27347	.5	PB
91	OC	*		10.10000+-	1.01000	12.62500+-	1.78544	9.38185+-	1.29686	-1.0	OC
92	EC	*		1.68000+-	.16800	2.10000+-	.29698	1.34267+-	.39741	-.8	EC
93	SO4	*		8.10000+-	.81000	10.12500+-	1.43189	9.11918+-	1.25659	.0	SO4
94	NO3	*		1.13000+-	.11300	1.41250+-	.19976	.71048+-	.41362	-1.0	NO3

MEASURED AMBIENT MASS (UG/M3): FINE: 80.0+- 8.0 COARSE: 90.0+- 8.0 TOTAL: 160.0+-11.3

RESULTS FOR CMB SITE: PACS1 YEAR: 77 DATE: 0813 VERSION: 6.0
TOTAL PARTICULATE FRACTION
SAMPLING DURATION: 24 HRS. WITH START HOUR: 0
RESULTS DERIVED FROM FINE AND COARSE SAMPLES

#	TYPE	UG/M3		%	
1	MARIN	22.991+-	2.893	14.369+-	2.074
3	UDUST	19.190+-	1.875	11.994+-	1.447
4	AUTPB	19.174+-	2.045	11.984+-	1.533
5	RDOIL	20.773+-	2.509	12.983+-	1.817
8	KRAFT	17.019+-	9.380	10.637+-	5.910
11	ALPRO	21.702+-	4.233	13.564+-	2.814
12	STEEL	16.832+-	2.054	10.520+-	1.484
13	FERMN	21.747+-	2.443	13.592+-	1.804
	TOTAL:	159.428+-	7.984	99.643+-	8.634

SPECIES		MISS FLG	TOTAL SUSPENDED PARTICULATE MEAS. UG/M3		PERCENT		CALC. UG/M3		
1	TOTAL		160.00000+-	11.31371	100.00000+-	10.00000	159.42830+-	7.98406	TOTAL
9	F		1.61700+-	.11434	1.01063+-	.10106	1.17717+-	.33387	F
11	NA	*	13.26000+-	.93858	8.28750+-	.82917	13.04796+-	.73507	NA
12	MG	*	1.91000+-	.15412	1.19375+-	.12808	3.09625+-	.81650	MG
13	AL	*	9.50000+-	.67187	5.93750+-	.59380	8.43408+-	1.04597	AL
14	SI	*	6.29000+-	.44512	3.93125+-	.39328	6.30995+-	.20656	SI
16	S		5.45000+-	.38668	3.40625+-	.34121	5.43754+-	.42518	S
17	CL	*	10.62000+-	.75638	6.63750+-	.66616	10.89051+-	1.64841	CL
19	K	*	2.76000+-	.19860	1.72500+-	.17402	3.17285+-	.60378	K
20	CA	*	3.30000+-	.23407	2.06250+-	.20657	2.90730+-	.15361	CA
22	TI	*	.22300+-	.01612	.13938+-	.01410	.23781+-	.03594	TI
23	V	*	.64900+-	.04640	.40563+-	.04079	.74597+-	.11052	V
24	CR	*	.32300+-	.03202	.20187+-	.02458	.45508+-	.16826	CR
25	MN	*	5.46000+-	.38783	3.41250+-	.34202	5.26657+-	.18676	MN
26	FE	*	9.94000+-	.70561	6.21250+-	.62247	8.36090+-	.45181	FE
28	NI	*	1.54400+-	.10960	.96500+-	.09669	1.51264+-	.17877	NI
29	CU	*	.09400+-	.00640	.05875+-	.00577	.11888+-	.00852	CU
30	ZN	*	.43900+-	.03114	.27438+-	.02748	.50525+-	.05232	ZN
35	BR	*	.93900+-	.06684	.58688+-	.05868	1.07511+-	.23489	BR
82	PB	*	4.31000+-	.30934	2.69375+-	.27141	4.06825+-	.40816	PB
91	OC	*	17.64000+-	1.26040	11.02500+-	1.10829	16.69165+-	1.87655	OC
92	EC	*	3.10000+-	.21997	1.93750+-	.19409	2.67865+-	.52308	EC
93	SO4	*	18.40000+-	1.31034	11.50000+-	1.15410	17.90850+-	1.93778	SO4
94	NO3	*	1.76800+-	.12787	1.10500+-	.11268	1.59449+-	.54794	NO3

MEASURED AMBIENT MASS (UG/M3): FINE: 80.0+- 8.0 COARSE: 80.0+- 8.0 TOTAL: 160.0+-11.3

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
EPA/SW/DK-89/033 a EPA-450/4-83-014R		PB89-181101

4. TITLE AND SUBTITLE	5. REPORT DATE
Receptor Model Technical Series Volume III (Revised): CMB User's Manual (Version 6.0)	May 1987
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
Kenneth Axetell, John G. Watson, and Thompson G. Pace	

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
PEI Associates, Durham, NC Desert Research Institute, Reno, NV U.S. EPA, OAQPS, RTP, NC	11. CONTRACT/GRANT NO. 68-02-3890, WA No. 38 and CX-813087-01-1

12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
U.S. EPA OAQPS, MDAD, Mail Drop 14 Research Triangle Park, NC 27711	14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES
EPA Project Officer: Thompson G. Pace For Diskette, See: PB89-181093

16. ABSTRACT
The Chemical Mass Balance (CMB) Model uses the chemical composition of an ambient particulate sample to estimate the relative contributions of different source categories to the measured particulate concentration. The chemical composition of each source category's emissions (source profile) must also be known in order to run the model.

This manual describes Version 6.0 of the CMB Receptor Model. It is designed to allow persons to use the CMB Model constructively with only a few hours' learning time. Because of this emphasis on rapid command of modeling procedures, the manual covers primarily the mechanical aspects of running the model. Users seeking more information on the statistical or theoretical bases of chemical mass balance receptor modeling are referred to "Appendix A: Theory Of Chemical Mass Balance Model."

The manual is intended for wide use by State and local air pollution control agency personnel in developing State Implementation Plans (SIPs) for PM_{10}. The U.S. Environmental Protection Agency has published a companion document to this manual that should be consulted for this application. The Protocol For Applying And Validating The CMB Model, EPA-450/4-87-010, provides guidance on applicability, assumptions, and interpretation of results. The protocol provides a practical strategy for obtaining valid results.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Receptor Models Chemical Mass Balance Source Apportionment Least Squares Multiple Linear Regression Microcomputer Software User's Manual		

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
	Unlimited	82
	20. SECURITY CLASS (This page) Unlimited	22. PRICE A05 $15.95

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE