Chapter 8 - SMOKE Input Files

Contents

8.1 Overview

This chapter contains all formatting requirements for input files to SMOKE. Chapters 9 and 10 provide documentation on SMOKE's intermediate and output file formats. First, a few important notes about this Chapter.

Important notes about file formats

1. In the tables used to describe the file formats, column letters indicate that the record(s) should be list-formatted, and therefore need to be delimited (by a space, comma, or semi-colon). If there are text entries that have embedded spaces, commas, or semi-colons in a list-formatted file, these entries must be surrounded by single or double quotation marks to prevent SMOKE from treating those characters as delimiters. Column numbers indicate that the entry must be placed exactly within the columns listed.

2. For all cross-reference files, all of the source-characteristic fields are optional. For example, it is permissible to provide only the SCC for use in matching a profile to a source. For list-formatted files (i.e., files that do not have column-specific formats), if there are source characteristic fields that will not be used (such as country/state/county code) that come **before** the required fields in the file, these must be filled in with zeros or –9. If there are fields that come **after** the required fields in the file, they can be left blank. For column-specific formats, the optional entries can be left blank when they are not used.

3. **[add note to explain file naming conventions and exceptions]**

In this section, we first provide a summary table of all input files needed by SMOKE listed by source category. In another table, we next list by program all the default SMOKE input files that are provided with the "nctox" sample case, described in Chapter 4.

8.1.1 Where to put SMOKE input files

A common difficulty for new users is where to put the input files for their case. In SMOKE version 1.5, we have simplified the number of places users need to keep their raw input files. Table 8-2 lists the all of the input directories for the files used in the example "nctox" case. At the end of the table, the path names are defined relative to the EDSS_ROOT directory, which is the directory in which you install SMOKE (and is different for every installation). The table is a good reference for understanding where the files should be placed. In Chapter 3, we provide diagrams of the SMOKE directory structures, which help in visualizing the tree of directories that SMOKE uses. Additionally, Section 4.5 of Chapter 4 contains instructions for getting setup for new runs, including the files that need to be created and where they should go.

In summary, SMOKE input files must go into a handful of directories (see the end of Table 8-2). There are several directories to use because in some cases (such as speciation profiles and meteorology data), we want to share the data across multiple runs. In other cases (such as inventory data) the data is specific to a case and should be stored independently to prevent corrupting other modeling runs. To change the directories used for input (or any) files in SMOKE, the Assigns files that are described in Chapter 4 must be modified.

8.1.2 SMOKE input file list

In Table 8-1, we list all of the raw input files to SMOKE. Where these files are optional, this is

noted in the description column of the table. The files are grouped by type, such as inventory, speciation, gridding, emission factors, etc.

All of the ASCII files could conceivably be prepared manually. However, the last 3 columns indicate other options for the user to prepare the files. Column "CNV" indicates that there is a SMOKE converter (either Fortran program, SAS program, script, or Excel spreadsheet) to create the file. Column "ST" indicates that the SMOKE Tool in Models-3 can prepare the file. Column "Other" indicates if there is another method to create the file, or if the file is resident in the installation and is likely to not need modification for most users.

Table 8-1: List of all SMOKE input files

Logical File Name with (Program)	Category	Description	CNV	ST	Other
Inventory Files					
PTINV (Smkinven)	P	Point source raw inventory in IDA format or SMOKE list format for EMS-95 files ($PTDAT directory)		X	NEI resident
PTDAY (Smkinven)	P	Point source raw day-specific inventory data (optional, $PTDAT directory)			User
PTHOUR (Smkinven)	P	Point source raw hour-specific inventory data (optional, $PTDAT directory)			User or CEM data
ARINV (Smkinven)	A	Area source raw inventory in IDA format or SMOKE list format for EMS-95 files ($ARDAT directory)		X	NEI resident
MBINV (Smkinven)	M	Mobile source raw inventory in IDA format for nonlink data or SMOKE list format for EMS-95 or EMS-95 list directed nonlink and link data ($MBDAT directory)			NEI resident
VMTMIX (Smkinven)	M	Mobile source vehicle mix for EMS-95 and EMS-95 list directed files ($MBDAT directory)			
BCUSE	B	Biogenic county land use data for use in			Resident

Logical File Name with (Program)	Category	Description	CNV	ST	Other
(Rawbio)		RAWBIO ($GE_DAT directory)			
BGUSE (Rawbio)	B	Biogenic gridded land use data for use in RAWBIO ($BGDAT directory)			U.S. 36 km resident
Temporal Allocation Input Files					
[A|M|P]TPRO (Temporal)	A, M, P	Temporal profile file ($GE_DAT directory)		X	NEI resident
[A|M|P]TREF (Temporal)	A, M, P	Temporal cross-reference file ($GE_DAT directory)	X	X	NEI resident
Spatial Allocation Input Files					
G_GRIDPATH (Grdmat, Elevpoint, Smkmerge)	A, B, M, P	Grid and grid projection information ($BASDIR directory)			M3 Study Planner
[A|M]GPRO (Grdmat)	A, M	Area and/or mobile source spatial surrogates ($GE_DAT directory)		X	U.S. 36-km resident
BGPRO (Rawbio, Smkmerge)	B	Biogenic spatial surrogates for county land use ($GE_DAT directory) in RAWBIO		X	U.S. 36-km resident
BELD3_A BELD3_B BELD3_TOT (Normbeis3)	B	Biogenic gridded landuse for use in SMOKE-BEIS3		X	U.S. 36-km resident
[A|M]GREF (Grdmat)	A, M	Area source or mobile gridding cross-reference ($GE_DAT directory)		X	NEI resident

Logical File Name with (Program)	Category	Description	CNV	ST	Other
MGLNK	M	Link-based gridding file for gridding county VMT onto links by road type (optional, but **not yet supported**)			
Speciation Input files					
GSPRO (Spcmat,Tmpbio,Beis3)	A, B, M, P	Speciation profiles file.CB-IV and RADM2 with PM2.5 speciation for CMAQ model comes with SMOKE ($GE_DAT directory)			OR Science Manager
GSREF (Spcmat)	A, B, M, P	Speciation cross-reference file (a.k.a., mapping table) ($GE_DAT directory)	X	X	NEI resident
GSCNV (Spcmat)	A, M, P	Pollutant-to-pollutant conversion file (e.g., ROG to TOG) (optional) ($GE_DAT directory)			ROG to TOG resident
Control Input Files					
GCNTL (Cntlmat)	A, M, P	Control packets or projection packets file ($STATIC directory)		X	User
Special Merging Input Files					
FILELIST (Mrggrid)	A, B, M, P	Gridded merge 2-d file list ($SCRIPTS/run directory)			User
Meteorology Input Files					
GRID_CRO_2D (Tmpbio, Beis3)	B	Grid 2D cross-point file (optional) ($METDAT directory)			MCIP
GRID_CRO_3D(Tmpbio, Beis3, Laypoint)	P	Grid 3D cross-point file ($METDAT directory)			MCIP

Logical File Name with (Program)	Category	Description	CNV	ST	Other
MET_CRO_2D (Tmpbio, Beis3, Temporal, Premobl, Laypoint)	B, M, P	Meteorology 2D cross-point file ($METDAT directory)			MCIP
MET_CRO_3D (Tmpbio, Beis3, Laypoint)	B, M, P	Meteorology 3D cross-point file ($METDAT directory) (optional for biogenics)			MCIP
MET_DOT_3D (Laypoint)	P	Meteorology 3D dot-point file ($METDAT directory)			MCIP
METLIST (Tmpbio, Premobl)	M	Listing of meteorology files ($STATIC directory)			User
Source-specific Input Files					
B3XRF (Beld3to2)	B	BELD3 to BELD2 cross-reference file			Resident
BFAC (Rawbio)	B	Summer or winter biogenic emission factors ($GE_DAT directory)			Resident
B3FAC (Normbeis3)	B	BEIS3 emissions factors			Resident
MCODES (Smkinven, Grdmat, Spcmat, Temporal)	M	Mobile vehicle type and road codes ($MBDAT directory)			Resident /User
MEPROC (Emisfac, Spcmat, Temporal)	M	Mobile emission processes (e.g., exhaust, running loss, refueling) with associated activities and pollutants ($MBDAT directory)			Resident /User
MCREF	M	Mobile source county cross-reference			User

Logical File Name with (Program)	Category	Description	CNV	ST	Other
(Mbsetup)		($MBDAT directory)			
MVREF (Mbsetup)	M	Mobile source county settings file ($MBDAT directory)			User
M6LIST (Emisfac)	M	List of MOBILE6 input scenarios ($MBDAT directory)			User
MEFLIST (Temporal)	M	List of emission factor files ($SMK_EMISPATH directory)			User
PELVCONFIG (Elevpoint)	P	Elevated source and plume-in-grid selection configuration file. Takes the place of PGROUP and PSPLIT from previous versions of SMOKE to permit SMOKE to select sources for elevated and/or plume-in-grid treatment .			User
PSTK (Smkinven)	P	Replacement stack parameters for invalid stacks. This is not really needed when SMOKE is used in Models-3 because the SMOKE Tool will check the stack parameters. ($GE_DAT directory)			Resident
PTMPLIST (Elevpoint)	P	List of PTMP files for Elevpoint to use to compute maximum daily emissions.			User
QA, Utility, and Other Support Input Files					
ARTOPNT (Smkinven)	A, M	Pollutant and activity names and codes in order of output ($INVDIR/other directory)			Resident /user
INVTABLE (Smkinven, Emisfac, Spcmat, Temporal, Smkmerge)	A, M, P	Pollutant and activity names and codes in order of output ($INVDIR/other directory)			Resident /user

Logical File Name with (Program)	Category	Description	CNV	ST	Other
COSTCY (Smkinven, Temporal, Smkmerge, Smkreport)	A, B, M, P	Country, state, and county names and time zones ($GE_DAT directory)			Resident
NHAPEXCLUDE	A, M	NONHAP exclusions file			Resident /User
SCCDESC	A, M, P	SCC descriptions			Resident
ORISDESC	P	ORIS facility descriptions, needed for getting ORIS facility names for reporting purposes			Resident
VNAMMAP (Smk2emis)	A, B, M, P	Name mapping file for Smk2emis ($SCRIPTS/run directory)			User
REPCONFIG (Smkmerge)	A, M, P	Reports configuration file, examples are provided ($SCRIPTS/configure directory)			Resident example files

8.1.3 Input files for SMOKE nctox case

The following table summarize the SMOKE input files for the nctox sample case provided with SMOKE for processing ozone, PM, and toxics pollutants. This list can be adapted by users to keep track of the raw input files needed for their cases, and to help understand the files that need to be prepared for input to SMOKE. Chapter 4 includes a description of the nctox case in more detail. The documentation in Section 4.5 of Chapter 4 describes which input files need to be changed for adapting the default scripts to a particular case.

In Table 8-2 below, we list the raw SMOKE input files by program, with the programs listed in alphabetical order. These are the files that would have to be created by the user or provided by SMOKE, and most cannot be created with a SMOKE utility (unless otherwise noted). Those files that are typically prepared by the users are listed with their logical file names in bold, while the files that users do not usually change are not. Intermediate files (output by one program and input to another) are not included in the list, but can be found with the program-specific documentation in Chapter 6 and in the intermediate file documentation in Chapter 9. The files that are shared among several programs are listed with each program for which they are used (i.e., listed multiple times), and noted with an asterisk (*) adjacent to the logical file name. The file paths are listed using the path logical names, the definitions for which are available at the

end of the table.

Table 8-2: SMOKE raw input files for the nctox case

Logical File name and description	Source categories	Path	Physical file name
Beis3			
Same raw inputs as Tmpbio program, documented below	See Tmpbio	See Tmpbio	See Tmpbio
Cntlmat			
GCNTL projection and/or control instructions	A, M, P	$INVDIR/area/ $INVDIR/nonroad/ $INVDIR/mobile/ $INVDIR/point/	gcntl.control.txt gcntl.control.txt gcntl.control.txt gcntl.control.txt
GSPRO speciation profiles needed for reactivity controls only	A, M, P	$GE_DAT/	gspro.cmaq.cb4p25_wtox.txt
Elevpoint			
PELVCONFIG elevated source selection instructions	P	$INVDIR/point/	pelvconfig.top50.txt
GRIDDESC grid descriptions	P	$GE_DAT/	GRIDDESC
PTMPLIST optional list of PTMP SMOKE intermediate files for selecting elevated point sources based on emissions	P	$INVDIR/point/	ptmplist.txt (generated automatically by default run script)
Emisfac			
M6LIST	M	$INVDIR/mobile/	m6list.txt

Logical File name and description	Source categories	Path	Physical file name
List of MOBILE6 scenario files with path			
MEPROC mobile processes	M	$INVDIR/mobil/e	meproc.txt
Required MOBILE6 input files. (Note that there are other MOBILE6 input files that are needed if you use certain settings in the M6LIST file).	M	$SMK_M6PATH/	ASMDATA.D CUTPOINT.D IMTEST.D NLEVNE.D PMDZML.CSV PMDDR1.CSV PMDDR2.CSV PMGZML.CSV PMGDR1.CSV PMGDR2.CSV PMNH3BER.D PMNH3SDR.D TECH12.D
Grdmat			
AGPRO gridding surrogates	A	$GE_DAT/	abmgpro.m3.040402.us36+can+mex.txt
MGPRO gridding surrogates	M	$GE_DAT/	abmgpro.m3.040402.us36+can+mex.txt
AGREF gridding cross-reference	A	$GE_DAT/	amgref.m3.040402.us36+can+mex.txt
MGREF gridding cross-reference	M	$GE_DAT/	amgref.m3.040402.us36+can+mex.txt
GRIDDESC grid descriptions	A, M, P	$GE_DAT/	GRIDDESC
MCODES * mobile code definitions	M	$INVDIR/mobile/	mcodes.txt

Logical File name and description	Source categories	Path	Physical file name
Grwinven			
COSTCY * country,state,county names	A, M, P	$GE_DAT/	costcy.txt
Laypoint			
MET_CRO_2D * 2-d cross-point meteorology data	P	$METDAT/	METCRO2D_1996191 METCRO2D_1996192
GRID_CRO_3D 3-d cross-point grid data	P	$METDAT/	GRIDCRO3D
MET_CRO_3D * 3-d cross-point grid data	P	$METDAT/	METCRO3D_1996191 METCRO3D_1996192
MET_DOT_3D 3-d dod-point grid data (for winds)	P	$METDAT/	METDOT3D_1996191 METDOT3D_1996192
GRIDDESC grid descriptions	P	$GE_DAT/	GRIDDESC
Mbsetup			
MCREF county cross-reference	M	$INVDIR/mobile/	
MVREF county meteorology group settings	M	$INVDIR/mobile/	
Mrggrid			
FILELIST list of logical file names to merge	All	$INVDIR/other/	filelist.mrggrid.txt

Logical File name and description	Source categories	Path	Physical file name
Normbeis3			
B3FAC BEIS3 emissions factors	B	$GE_DAT/	b3fac.march2002.beisv3.09.txt
BELD3_A File A of biogenic gridded landuse	B	$SMKDAT/ inventory/beld3	LANDA_c_cmaq_nat.dat
BELD3_B File B of biogenic gridded landuse	B	$SMKDAT/ inventory/beld3	LANDB_c_cmaq_nat.dat
BELD3_TOT Landuse totals	B	$SMKDAT/ inventory/beld3	LAND_TOTALS_c_cmaq_nat.dat
Premobl			
METLIST temperature file list	M	$INVDIR/mobile/	metlist.premobl.txt (This list will include the paths of the 2-d or 3-d meteorology files for the entire modeling episode)
Rawbio			
BGUSE gridded land use	B	$INVDIR/ inventory/beld2/	beld.5.us36.txt
BCUSE county land use	B	$GE_DAT/	not used in this case
BFAC emission factors	B	$GE_DAT/	bfac.summer.txt bfac.winter.txt
GRIDDESC grid descriptions	B	$GE_DAT/	GRIDDESC
Smkinven			
ARINV (area) area inventory (or list of inventories)	A	$INVDIR/area/	arinv.stationary.lst
ARINV (nonroad)	A	$INVDIR/nonroad/	arinv.nonroad.lst

Logical File name and description	Source categories	Path	Physical file name
nonroad inventory (or list of inventories)			
MBINV on-road mobile inventory (or list of inventories)	M	$INVDIR/mobile/	mbinv.lst
VMTMIX vehicle mix	M	$INVDIR/mobile/	Not used in this case because of MBINV format selected.
MCODES * mobile code definitions	M	$INVDIR/mobile/	mcodes.txt
PTINV point source inventory (or list of inventories)	P	$INVDIR/point/	ptinv.lst
PTDAY point source list of day-specific data	P	$INVDIR/point/	Not used in this case
PTHOUR point source list of hour-specific data	P	$INVDIR/point/	pthour.lst (List of CEM files of interest)
NHAPEXCLUDE exclusions from NONHAPVOC calculation	A, M	$INVDIR/other/	nhapexclude.1999.txt
INVTABLE * inventory table	A, P M	$INVDIR/other/ $INVDIR/other/	invtable_nonroad.cb4.120202.txt invtable_onroad.cb4.120202.txt
ARTOPNT area-to-point assignments	A	$INVDIR/area/	artopnt.1999.txt
COSTCY * time zones by county	All	$GE_DAT/	costcy.txt

Logical File name and description	Source categories	Path	Physical file name
ORISDESC ORIS ID descriptions	P	$GE_DAT/	oris_info.txt
PSTK replacements for bad stack parameters	P	$GE_DAT/	pstk.m3.txt
SCCDESC * SCC descriptions	A, M, P	$GE_DAT/	scc_desc.txt
Smkreport			
REPCONFIG Report configuration	A, M, P	$INVDIR/other/	REPCONFIG.*.txt (several files used)
COSTCY * country,state,county names	A, M, P	$GE_DAT/	costcy.txt
SCCDESC SCC descriptions	A, M, P	$GE_DAT/	scc_desc.txt
Smkmerge			
GRIDDESC grid descriptions	All	$GE_DAT/	GRIDDESC
BGPRO Spatial surrogates with cell-area surrogate	B	$GE_DAT/	abmgpro.m3.040402.us36+can+mex.txt
Spcmat			
GSPRO speciation profiles	A (stat) Biogenic Nonroad On-road Point	$GE_DAT/	gspro.cmaq.cb4p25.txt gspro.cmaq.cb4p25.txt gspro.cmaq.cb4p25_wtox.txt gspro.cmaq.cb4p25_wtox.m.txt gspro.cmaq.cb4p25.txt
GSREF	A, M, P	$GE_DAT/	gsref.cmaq.cb4p25_wtox.txt

Logical File name and description	Source categories	Path	Physical file name
speciation cross-reference			
GSCNV pollutant-to-pollutant conversion	A, M, P	$GE_DAT/	gscnv.txt
MEPROC mobile processes	M	$INVDIR/mobile/	meproc.txt
MCODES mobile code definitions	M	$INVDIR/mobile/	mcodes.txt
Temporal			
ATPRO Temporal profiles	A	$GE_DAT/	amptpro.m3.062801.us+can.txt
MTPRO Temporal profiles	M	$GE_DAT/	amptpro.m3.062801.us+can.txt
PTPRO Temporal profiles	P	$GE_DAT/	amptpro.m3.062801.us+can.txt
ATREF Temporal cross-reference	A	$GE_DAT/	amptref.m3.032503.us+can.txt
MTREF Temporal cross-reference	M	$GE_DAT/	amptref.m3.032503.us+can.txt
PTREF Temporal cross-reference	P	$GE_DAT/	amptref.m3.032503.us+can.txt
COSTCY Daylight time exemptions	All	$GE_DAT/	costcy.txt
HOLIDAYS Holidays date list	A, M, P	$GE_DAT/	holidays.txt

Logical File name and description	Source categories	Path	Physical file name
MEFLIST	M	$SMK_EMISPATH/	meflist.txt (note: this file is created automatically when running using SMOKE scripts)
MEPROC mobile processes	M	$INVDIR/mobile/	meproc.txt
MCODES mobile code definitions	M	$INVDIR/mobile/	mcodes.txt
Tmpbio			
BIOSEASON Winter/summer designation by cell and day	B	$GE_DAT/	bioseason
GSPRO Speciation profiles	B	$GE_DAT/	gspro.cmaq.cb4p25.txt
METLIST List of MCIP outputs with temperature data	B	$INVDIR/biog/	metlist.tmpbio.txt (note: this file is created automatically when running using SMOKE scripts)
RADLIST List of MCIP outputs with solar radiation data	B	$INVDIR/biog/	radlist.tmpbio.txt (note: this file is created automatically when running using SMOKE scripts)
GRID_CRO_2D 2-d cross-point grid data	B	$METDAT/	Not used in this case (From MCIP)

As you can see from the list above, there are several directories for raw SMOKE inputs. These are:

$INVDIR/area: Raw area inputs

$INVDIR/biog: Raw biogenic inputs

$INVDIR/mobile: Raw on-road mobile inputs

$INVDIR/nonroad:	Raw nonroad mobile inputs (these can alternatively be included in the area directory)
$INVDIR/point:	Raw point inputs
$INVDIR/other:	Other raw inputs not associated with a specific source category.
$GE_DAT:	Shared data
$SMK_M6PATH:	Raw MOBILE6 input files
$SMK_EMISPATH:	Emission factor inputs (the single file that is considered a raw input file is actually prepared by the SMOKE scripts, when these are being used).
$METDAT:	Meteorology data (in UNIX, this can be linked to a meteorology directory that is also used by the air quality model).

All of these directories are based on the $SMKDAT directory, which is set by the Assigns file as $EDSS_ROOT/data/smoke. The EDSS_ROOT environment variable should be set as part of the SMOKE installation process, as described in Chapter 12.

For the nctox case, these directories (based on your EDSS_ROOT) are as follows:

$INVDIR:	$SMKDAT/inventory/nctox
$GE_DAT:	$SMKDAT/ge_dat
$SMK_M6PATH:	$SMKDAT/run_nctox/static/m6 (NOTE: Currently, this directory name must be constrained to 50 characters. The Assigns file checks this.)
$SMK_EMISPATH:	$SMKDAT/run_nctox/static/m6emfac
$METDAT:	$SMKDAT/met/run_a1a

8.2 Inventory Files:

- PTINV
- PTDAY
- PTHOUR
- ARINV
- MBINV
- BCUSE
- BGUSE

For all inventory files of a particular format. The user can force SMOKE to recognize a file as having a specific format by using the following three header entries:

#IDA indicates that the file is in IDA format
#EPS2 indicates that the file is in EPS2.0 format
#EMS-95 indicates that the file is in EMS-95 format

NOTE: The file formats for EPS2.0 input files have not yet been added to this documentation. The formats supported by SMOKE are consistent with the input file formats described in the User's Guide for the Urban Airshed Model, Volume IV: User's Manual for the Emissions Preprocessor System 2.0. These files are listed in this documentation as the AIRS Facility Subsystem (AFS) data input (Table B-1) and the AIRS Area and Mobile Source (AMS) data input (Table B-2). For the AFS format, the stack and segment fields are treated as integers if they are integers (otherwise, they are treated as character strings). In the cross-reference readers, however, these fields are not (because the format is shared with other formats). So, the cross-reference files need to have these fields specified with leading zeros. Also, SIC is treated as an integer instead of a character as the format specifies.

A NOTE ABOUT EPS2.0 POINT SOURCE FILES: the EPS2.0 point source input file documentation is confusing, because it lists "XLOC" as "latitude" and "YLOC" and "longitude." In truth, the EPS2.0 software reads latitude first, so this is what SMOKE does as well. Make sure that YLOC is in columns 78-87, and XLOC is in columns 89-98.

For all file formats, you may leave fields blank that are not required (labeled as either "recommended," "optional," or "not used by SMOKE"). The required fields are listed as required for the inventory file formats.

8.2.1 PTINV: Point source annual or average day emissions

The PTINV file is the input point source inventory data file. The SMOKE Tool can create this file in IDA format. This file can also have the same format as the SMOKE prototype's PNLST file, which can be used to import EMS-95 data.

IDA Format

This inventory file consists of a set of fixed-format detail records, preceded by a set of header records. The general format of the header records is:

```
command parm1 parm2 ... parm n
```

The *command* field starts in the first position. The remainder of the header is free-format, blank delimited. The specific header formats are:

#TYPE	file type	first nonblank character after the command to the last character of the record
#COUNTRY	country name	first nonblank character after the command to the last character of the record; valid country names are those in the COSTCY file
#YEAR	data year	number from 1900 to 2200
#DESC	description	a description of the data (SMOKE does not use)
#POLID or #DATA	pol1 pol2 ... poln	defines the pollutants that are contained on the detail records. "pol1" is the ID of the first pollutant, "pol2" is the ID of the second pollutant, etc.

Sample header records are shown below:

```
#TYPE     Point Source Inventory
#COUNTRY  US
#YEAR     1995
#DESC     Alabama 1995
#POLID    CO NH3 NOX PM10 PM2_5 SO2 VOC
```

Please note: Starting in version 1.4 of SMOKE, the reader of this file has been changed to workaround a bug in version 6.5 IRIX operating system. This workaround makes the reader more sensitive to missing pollutant data. This can cause problems reading the data if a record has missing emissions data for the last pollutant in the list given by the #POLID header. To prevent this problem, spaces must be used to fill in the missing data columns.

The detail records are fixed-format. The fields in each record are shown in the table below.

Position	Name	Type	Description
1-2	STID	Int	State Code (required)
3-5	CYID	Int	County Code (required)
6-20	PLANTID	Char	Plant Identification Code (required)

21-35	POINTID	Char	Point Identification Code (recommended)
36-47	STACKID	Char	Stack Identification Code (recommended)
48-53	ORISID	Char	DOE Plant ID (generally recommended, and required if matching to hour-specific CEM data)
54-59	BLRID	Char	Boiler Identification Code (recommended)
60-61	SEGMENT	Char	DOE ID (recommended)
62-101	PLANT	Char	Plant Name (recommended)
102-111	SCC	Char	Source Classification Code (required)
112-115	BEGYR	Int	Begin Year of Plant Operation (not used by SMOKE)
116-119	ENDYR	Int	End Year of Plant Operation (not used by SMOKE)
120-123	STKHGT	Real	Stack Height (ft) (required)
124-129	STKDIAM	Real	Stack Diameter (ft) (required)
130-133	STKTEMP	Real	Stack Gas Exit Temperature (deg F) (required)
134-143	STKFLOW	Real	Stack Gas Flow Rate (ft^3/sec) (required)
144-152	STKVEL	Real	Stack Gas Exit Velocity (ft/sec) (required)
153-160	BOILCAP	Real	Design Capacity (mmBtu/hr) (not used by SMOKE)
161-161	CAPUNITS	Char	Capacity Unit Code (not used by SMOKE)
162-163	WINTHRU	Real	Winter throughput (% of Annual) (not used by SMOKE)
164-165	SPRTHRU	Real	Spring throughput (% of Annual) (not used by SMOKE)
166-167	SUMTHRU	Real	Summer throughput (% of Annual) (not used by SMOKE)
168-169	FALTHRU	Real	Fall throughput (% of Annual) (not used by SMOKE)
170-171	HOURS	Int	Normal Operating Time (hours/day) (not used by

			SMOKE)
172-173	START	Int	Normal Operation Start Time (not used by SMOKE)
174-174	DAYS	Int	Normal Operating Time (days/week) (not used by SMOKE)
175-176	WEEKS	Int	Normal Operating Time (weeks/year) (not used by SMOKE)
177-187	THRUPUT	Real	Throughput Rate (SCC units/yr) (not used by SMOKE)
188-199	MAXRATE	Real	Max O3 Season Rate (units/day) (not used by SMOKE)
200-207	HEATCON	Real	Heat Content (mmBtu/SCC unit) (not used by SMOKE)
208-212	SULFCON	Real	Sulfur Content (mass percent) (not used by SMOKE)
213-217	ASHCON	Real	Ash Content (mass percent) (not used by SMOKE)
218-226	NETDC	Real	Max Nameplate Capacity (MW) (not used by SMOKE)
227-230	SIC	Int	Standard Industrial Classification Code (required)
231-239	LATC	Real	Latitude (decimal degrees) (required)
240-248	LONC	Real	Longitude (decimal degrees) (required
249-249	OFFSHORE	Char	Offshore Flag (fill in with “X” if source is offshore), (recommended)
250-262	ANN1	Real	pol1 Annual Emissions (short tons / year) (required)
263-275	OSD1	Real	pol1 Ozone Season Emissions (short tons/average ozone day) (optional)
276-282	CE1	Real	pol1 Control Efficiency (give value of 0-100) (recommended, if left blank, SMOKE default is 0)
283-285	RE1	Real	pol1 Rule Effectiveness (give value of 0-100) (recommended, if left blank, SMOKE default is 100)
286-295	EMF1	Real	pol1 Emission Factors (SCC units) (not used by SMOKE)

296-298	CPRI1	Int	pol1 Primary Control Equipment Code (not used by SMOKE)
299-301	CSEC1	Int	pol1 Secondary Control Equipment Code (not used by SMOKE)
302-314	ANN2	Real	pol2 Annual Emissions (short tons / year), (required)
315-327	OSD2	Real	pol2 Ozone Season Emissions (short tons/average ozone day), (recommended)
328-334	CE2	Real	pol2 Control Efficiency (0-100), see pol1 above
335-337	RE2	Real	pol2 Rule Effectiveness (0-100), see pol1 above
338-347	EMF2	Real	pol2 Emission Factors (SCC units), see pol1 above
348-350	CPRI2	Int	pol2 Primary Control Equipment Code, see pol1 above
351-353	CSEC2	Int	pol2 Secondary Control Equipment Code, see pol1 above
354-366	ANN3	Real	pol3 Annual Emissions (short tons/year) , see pol1 above
367-379	OSD3	Real	pol3 Ozone Season Emissions (short tons/average ozone day) , see pol1 above
380-386	CE3	Real	pol3 Control Efficiency (0-100) , see pol1 above
387-389	RE3	Real	pol3 Rule Effectiveness (0-100) , see pol1 above
390-399	EMF3	Real	pol3 Emission Factors (SCC units) , see pol1 above
400-402	CPRI3	Int	pol3 Primary Control Equipment Code, see pol1 above
403-405	CSEC3	Int	pol3 Secondary Control Equipment Code, see pol1 above

(Repeat for n pollutants)

EMS-95 Format

For EMS-95 inputs, the PTINV file must list the different types of EMS-95 files in a specific

order. This is called a SMOKE list format. The order of the EMS-95 files is the same as would be achieved from the UNIX command "ls". The order is: device file, emissions file, facility file, process file, and stack file, and these 5 files must be grouped together for the files that work together. For example, the North Carolina files must be together, not just any 5 of these files. The file can be created by using the UNIX command, which assume the EMS-95 files (and only these files) are all in the $PTDAT directory.

```
ls $PTDAT/*.pt > $PTINV
```

When using a list format, the PTINV file must also contain one or more **INVYEAR** packets. This packet acts as a switch that sets the inventory year for the files, each of which will get the year designated by the packet until another **INVYEAR** packet appears in the file. Multiple years of inventory data can be combined to the same SMOKE inventory file in this way.

Although it would be possible to create a single file of each type (i.e., a single device, emission, facility, process, and stack file with all records), this approach is not recommended because Smkinven will be must slower than if the files are separated by state. This is because SMOKE must match the records among the various files, and it is much faster to do the matching with subsets of the data than with all of the data at once.

Example: The following is an example of a correct PNLST file when used for EMS-95 inputs.

```
INVYEAR 1995
/epserv/apps/smoke/inventory/smraq95/al5/common/device.pt
/epserv/apps/smoke/inventory/smraq95/al5/common/emission.pt
/epserv/apps/smoke/inventory/smraq95/al5/common/facility.pt
/epserv/apps/smoke/inventory/smraq95/al5/common/process.pt
/epserv/apps/smoke/inventory/smraq95/al5/common/stack.pt
/epserv/apps/smoke/inventory/smraq95/ar5/common/device.pt
/epserv/apps/smoke/inventory/smraq95/ar5/common/emission.pt
/epserv/apps/smoke/inventory/smraq95/ar5/common/facility.pt
/epserv/apps/smoke/inventory/smraq95/ar5/common/process.pt
/epserv/apps/smoke/inventory/smraq95/ar5/common/stack.pt
```

Format for EMS-95- formatted "device.pt" file(s):

The columns listed here are only those that are used by SMOKE. These files *cannot* have a #COUNTRY entry. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Facility ID (alpha-numeric) (required)

21-32	Char	Stack ID (alpha-numeric) (required)
33-44	Char	Device ID (alpha-numeric) (required)
45-48	Char	Standard Industrial Classification (SIC) (required)
121-122	Int	Diurnal profile code (optional, you may use the PTREF file instead)
123-124	Int	Weekly profile code (optional, you may use the PTREF file instead)

Format for EMS-95- formatted "emission.pt" file(s):

The columns listed here are only those that are used by SMOKE. These files cannot have a #COUNTRY entry. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Facility ID (alpha-numeric) (required)
21-32	Char	Stack ID (alpha-numeric) (required)
33-44	Char	Device ID (alpha-numeric) (required)
45-56	Char	Process ID (alpha-numeric) (required)
57-61	Char	Pollutant ID (e.g., NOX) (required)
88-100	Real	Actual emissions (short tons/average annual day) (required)
114-115	Char	Temporal basis (required) AA = annual emissions AD = daily emissions with weekday normalizer DS = day-specific emissions
126-132	Real	Control equipment efficiency (percent, give value 0-100) (optional, SMOKE default is 0)

Format for EMS-95- formatted "facility.pt" file(s):

The columns listed here are only those that are used by SMOKE. These files cannot have a #COUNTRY entry. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Facility ID (alpha-numeric) (required)
21-24	Int	Standard Industrial Classification (SIC) (not used by SMOKE)
25-33	Real	UTM easting (m) (required)
34-42	Real	UTM northing (m) (required)
43-44	Int	UTM zone (required)
45-84	Char	Facility Name (required)

Format for EMS-95- formatted "process.pt" file(s):

The columns listed here are only those that are used by SMOKE. These files cannot have a #COUNTRY entry. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Facility ID (alpha-numeric) (required)
21-32	Char	Stack ID (alpha-numeric) (required)
33-44	Char	Device ID (alpha-numeric) (required)
45-56	Char	Process ID (alpha-numeric) (required)

57 - 64	Char	Source Classification Code (SCC) (required)

Format for EMS-95- formatted "stack.pt" file(s):

The columns listed here are only those that are used by SMOKE. These files cannot have a #COUNTRY entry. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Facility ID (alpha-numeric) (required)
21-32	Char	Stack ID (alpha-numeric) (required)
33-40	Real	Stack diameter (ft) (required)
41-47	Real	Stack height above ground surface(ft) (required)
48-54	Real	Stack exit gas temperature (deg K) (required)
55-61	Real	Stack exit gas velocity (ft/s) (required)
62-71	Real	Stack exit flow rate (actual ft^3/min) (optional)
72-80	Real	UTM easting of stack (m) (default is plant's value) (optional, SMOKE will use facility's value)
81-89	Real	UTM northing of stack (m) (default is plant's value) (optional, SMOKE will use facility's value)

8.2.2 PTDAY: Point source day-specific emissions

The PTDAY file is the input point source inventory file for day-specific data. It must be created by the user without the SMOKE Tool. The file is simply a list of day-specific data files (directory paths included) that are to be read in by SMOKE as day-specific data. The files listed in this file must be in the following format. In addition, these files can have a #COUNTRY entry at the first line of the file to set the country code; valid country names are those in the COSTCY file. If no #COUNTRY header is used, then the default country code that will be assumed by

SMOKE is 0 (for which the default COSTCY file has the US as the country code).

SMOKE requires that the sources in the PTDAY file also be in the PTINV file, and that the user has matched the sources between the inventory and the day-specific data. See the documentation for the PTHOUR file (below) for more information on how to use this file with IDA inventories.

Position	Name	Type	Description
1-2	STID	Int	State Code (required)
3-5	CYID	Int	County Code (required)
6-20	FCID	Char	Facility ID (a.k.a. plant ID)
21-32	CHAR1	Char	If PTINV in EMS-95 format: Stack ID (required) If PTINV in IDA format: Point ID
33-44	CHAR2	Char	If PTINV in EMS-95 format: Device ID (required) If PTINV in IDA format: Stack ID
45-56	CHAR3	Char	If PTINV in EMS-95 format: Process ID (required) If PTINV in IDA format: Segment
57-61	POLID	Char	Pollutant name (required)
62-69	DATE	Char	Date in format "MM/DD/YY". Years less than 70 are treated as century 2000. (required)
70-72	TZONNAM	Char	Time zone name. Valid entries GMT, ADT, AST, EDT, EST, CDT, CST, MDT, MST, PDT, PST. Add others to EMCNST3.EXT include file. (required)
73-90	DAYTOT	Real	Daily emissions total (short tons/day) (required)
92-101	SCC	Char	If PTINV in IDA format: source category code is needed to work properly.

8.2.3 PTHOUR: Point source hour-specific emissions

The PTHOUR file is the input point source inventory file for hour-specific data and profiles. The file is simply a list of hour-specific data files (directory paths included) that are to be read in by SMOKE as hour-specific data. The files listed in this file must be in the following format. In addition, these files can have a #COUNTRY entry at the first line of the file to set the country

code. The default country code is 0.

The hour-specific data files that are listed in the PTHOUR file can be provided in two ASCII formats: EMS-95 format and Continuous Emissions Monitoring (CEM) format. The EMS-95 format must be created by the user manually. The CEM format can be output from the SMOKE Tool.

Date range setting for extracting hour-specific data from a larger dataset:

The PTHOUR file may optionally contain a packet to control the range of dates that Smkinven will read in the data. The following entry should appear on the first line of the PTHOUR file when this feature is needed.

DATERANGE MMDD(start) MMDD(end)

Where MMDD is two-digit month and day of the month at the start and end of the period of interest. For example, to request July 10th through the 12th, the first line of the file should read:

DATERANGE 0710 0712

The year is implied by the CEM data itself, which can only be provided to SMOKE with a single year in the input file. The year can be broken up into smaller periods in separate files, and SMOKE can read in only part of the year as well from a file with only specific dates in it. This feature is useful when users want to use the default files and extract a certain period to output to the PHOUR file from Smkinven. Please note that reading in the whole year of data and extracting just a few days will run much slower than manually editing the CEM file to contains just the days of interest. Also note that the user is responsible for selecting days that fully cover the modeling episode after time zone differences between the facility's time zones and the modeling time zone (set by the OUTZONE option) are accounted for.

EMS-95 Format

For EMS-95 formatted data, SMOKE requires that the sources in the PTHOUR file also be in the PTINV file. The fields that are common to the hour-specific data and the inventory data need to match exactly. These fields are Facility ID, Stack ID, Device ID, and Process ID. These fields are treated as text fields (not numeric) in SMOKE, so leading zeros need to be consistent between the hour-specific and average-day or annual inventory files.

When using hour-specific data in EMS-95 format with the IDA inventory, the fields in the PTHOUR file should line up with the fields in the PTINV file as follows:

- PTHOUR Facility ID maps to PTINV PLANTID
- PTHOUR Stack ID maps to PTINV POINTID
- PTHOUR Device ID maps to PTINV STACKID
- PTHOUR Process ID maps to PTINV SEGMENT

- PTHOUR SCC (implied by Process ID in EMS-95 process file) maps to PTINV SCC

Position	Name	Type	Description
1-2	STID	Int	State Code (required)
3-5	CYID	Int	County Code (required)
6-20	FCID	Char	Facility ID (a.k.a. plant ID) (required)
21-32	SKID	Char	If PTINV in EMS-95 format: Stack ID (required) If PTINV in IDA format: Point ID
33-44	DVID	Char	If PTINV in EMS-95 format: Device ID (required) If PTINV in IDA format: Stack ID
45-56	PRID	Char	If PTINV in EMS-95 format: Process ID (required) If PTINV in IDA format: Segment
57-61	POLID	Char	Pollutant name (required)
62-69	DATE	Char	Date in format "MM/DD/YY". Years less than 70 are treated as century 2000 (required)
70-72	TZONNAM	Char	Time zone name associated with emissions data. Valid entries GMT, ADT, AST, EDT, EST, CDT, CST, MDT, MST, PDT, PST. NOTE: Add others to EMCNST3.EXT include file and recompile. (required)
73-79	HRVAL1	Real	Hourly emissions for hour 1 (short tons/hour) (required for all hours)
80-86	HRVAL2	Real	Hourly emissions for hour 2 (short tons/hour)
87-93	HRVAL3	Real	Hourly emissions for hour 3 (short tons/hour)
...			
234-240	HRVAL24	Real	Hourly emissions for hour 24 (short tons/hour)
241-248	DAYTOT	Real	Daily emissions total (short tons/day)
250-259	SCC	Char	If PTINV in IDA format: source category code is

			needed to work properly.

CEM hour-specific format

The CEM hour-specific format can be written by the SMOKE Tool in Models-3. This format can *only* be used with IDA-formatted inventory files at this time. With this format, SMOKE uses the CEM ORIS ID field and Boiler ID **[Add the header fields that can be added to limit the read. The syntax is DATERANGE MMDD1 MMDD2. Goes in LIST file!!]**

Position	8.2.3.1.1.1 Name	Type	Description
A	ORISID	Char	DOE Plant ID (required, should match the same field in the PTINV file in IDA format)
B	BLRID	Char	Boiler Identification Code (required, should match the same field in the PTINV file in IDA format)
C	YYMMDD	Int	Date of data in YY (2-digit year), MM (month), and DD (date) format (required)
D	HOUR	Integer	Hour value from 0 to 23
E	CO2	Real	Carbon dioxide emissions in pounds per hour If PTINV in IDA format: Point ID
F	SO2	Real	Sulfur dioxide emissions in pounds per hour
G	NOXRATE	Real	Rate of Nitrogen Dixoxide emissions in pounds per MMBtu
H	OPTIME	Real	Operating time (not used by SMOKE)
I	GLOAD	Real	Not used by SMOKE, but some real needs to be in the field
J	SLOAD	Real	Not used by SMOKE, but some real needs to be in the field
K	HTINPUT	Real	Heat input in MMBtu

8.2.4 ARINV: Area source annual or average day emissions

The ARINV file is the input area source inventory data file. The SMOKE Tool can create this file in IDA format. This file can also have the same format as the SMOKE prototype's ANLST file, which can be used to import EMS-95 data.

IDA Format

This inventory file consists of a set of fixed-format detail records, preceded by a set of header records. The general format of the header records is:

```
command parm1 parm2 ... parm n
```

The *command* field starts in the first position. The remainder of the header is free-format, blank delimited. The specific header formats are:

#TYPE	file type	first nonblank character after the command to the last character of the record
#COUNTRY	country name	first nonblank character after the command to the last character of the record; valid country names are those in the COSTCY file
#YEAR	data year	number from 1900 to 2200
#DESC	description	a description of the data (SMOKE does not use)
#POLID or #DATA	pol1 pol2 ... poln	defines the pollutants that are contained on the detail records. "pol1" is the ID of the first pollutant, "pol2" is the ID of the second pollutant, etc.

Sample header records are shown below:

```
#TYPE     Area Source Inventory
#COUNTRY  US
#YEAR     1995
#DESC     Alabama 1995
#POLID    CO NH3 NOX PM10 PM2_5 SO2 VOC
```

The detail records are fixed-format. The fields in each record are shown below:

Position	Name	Type	Description
1-2	STID	Int	State Code (required)
3-5	CYID	Int	County Code (required)
6-15	SCC	Char	Source Classification Code (required)

16-25	ANN1	Real	pol1 Annual Emissions (short tons) (required)
26-35	OSD1	Real	pol1 Ozone Season Emissions (short tons/average ozone day) (optional)
36-46	EMF1	Real	pol1 Emission Factors (SCC units) (not used by SMOKE)
47-53	CE1	Real	pol1 Control Efficiency (give value of 0-100) (optional, SMOKE default is 0)
54-56	RE1	Real	pol1 Rule Effectiveness (give value of 0-100), (optional, SMOKE default is 100)
57-62	RP1	Real	pol1 Rule Penetration (give value of 0-100) (optional,. SMOKE default is 100)
63-72	ANN2	Real	pol2 Annual Emissions (short tons)
73-82	OSD2	Real	pol2 Ozone Season Emissions (short tons/average ozone day)
83-93	EMF2	Real	pol2 Emission Factors (SCC units)
94-100	CE2	Real	pol2 Control Efficiency (give value of 0-100)
101-103	RE2	Real	pol2 Rule Effectiveness (give value of 0-100)
104-109	RP2	Real	pol2 Rule Penetration (give value of 0-100)
110-119	ANN3	Real	pol3 Annual Emissions (short tons/year)
120-129	OSD3	Real	pol3 Ozone Season Emissions (short tons/average ozone day)
130-140	EMF3	Real	pol3 Emission Factors (SCC units)
141-147	CE3	Real	pol3 Control Efficiency (give value of 0-100)
148-150	RE3	Real	pol3 Rule Effectiveness (give value of 0-100)
151-156	RP3	Real	pol3 Rule Penetration (give value of 0-100)

(Repeat for n pollutants)

EMS-95 Format

For EMS-95 inputs, the ARINV file must list the EMS-95 files. This format is called a SMOKE list format. The file can be created by using the UNIX command, which assumes the EMS-95 files (and only these files) are all in the $ARDAT directory.

```
ls $ARDAT/*.pt > $ARINV
```

When using a list format, the ARINV file must also contain one or more **INVYEAR** packets. This packet acts as a switch that sets the inventory year for the files, each of which will get the year designated by the packet until another **INVYEAR** packet appears in the file. Multiple years of inventory data can be combined to the same SMOKE inventory file in this way.

Format for EMS-95- formatted area source emissions file(s):

The columns listed here are only those that are used by SMOKE. These files cannot have a #COUNTRY header. The default country code is 0.

Columns	Type	Description
1-2	Int	FIPS state code (required)
3-5	Int	FIPS county code (required)
6-20	Char	Source Category Code (SCC) or Area Source Category (ASC) code (required)
21-25	Char	Pollutant ID (name, e.g. VOC) (required)
52-65	Real	Actual emissions of specified pollutant (short tons/average annual day) (required)
88-94	Real	Control equipment efficiency (give value of 0-100) (optional, default SMOKE value is 0)
95-96	Char	Temporal basis: (required) AA = annual emissions AD = daily emissions with weekday normalizer DS = day-specific (will not cause SMOKE to crash, but may not be fully supported to give same behavior as EMS-95 – use with caution)

8.2.5 MBINV: Mobile source activity and emissions

The MBINV file is the input mobile source inventory data file. The SMOKE Tool can create this file in either IDA activities format or IDA emissions format. The IDA activities format is a list-directed file format, but the IDA emissions format is a column-specific format. Although the LINK ID field is now supported by the IDA format, SMOKE cannot yet read the IDA link definitions file, which contains the coordinates for the links. So, in order to use link data with the current version of SMOKE, one must use the EMS-95 format.

The MBINV file can also be in a SMOKE list format, to read in EMS-95 link and non-link data, or a modified EMS-95 format that is list-directed. The following sections have details about these file formats.

Mobile IDA Header

The mobile IDA inventory formats consists of a set of fixed-format header records, preceded by a set of data records. The header records can appear anywhere in the file to change the meaning of the data. For example, the #COUNTRY value could change to allow for multiple countries in a single file, the #YEAR value could change to allow for multiple years, and the #DATA and #UNITS values could change to allow for different data fields in different parts of the inventory file.

The general format of the header records is:

```
command parm1 parm2 ... parm n
```

The *command* field starts in the first position. The remainder of the header is free-format, blank delimited. The specific header formats are:

Command	Purpose	Description and values
#TYPE	file type	first nonblank character after the command to the last character of the record; valid types are: Motor Vehicle Activity Inventory Motor Vehicle Emission Inventory If the word “activity” does not appear in the header, emissions are assumed.
#COUNTRY	country name	first nonblank character after the command to the last character of the record; valid country names are those in the COSTCY file (default “US”, “CANADA”, and “MEXICO”)
#YEAR	data year	number from 1900 to 2200

#DESC	description	a description of the data (SMOKE does not use)
#DATA	data1, data2, dataN	defines the data that are contained after the source identifiers on each records. "data1" is the name of the first data type, "data2" is the name of the second data type, etc.
#UNITS	units1, units2, unitsN	defines the units of the input data. "units1" is the name of the first units, and so forth. These must be included in quotes. This command is **valid only for mobile activity data**, because emissions data have units built into the file formats.

Sample header records are shown below
For activities:

```
#TYPE     Motor Vehicle Activity Inventory
#COUNTRY  US
#YEAR     1996
#DESC     1996 NET inventory, version 3.1
#DATA     SPEED VMT
#UNITS    "miles/hr" "10E6 miles/yr"
```

For emissions:

```
#TYPE     Motor Vehicle Emission Inventory
#COUNTRY  US
#YEAR     1996
#DESC     1996 NET inventory, version 3.1
#DATA     VOC NOX CO SO2 PM10 PM2_5 NH3
```

Mobile IDA Activity Data Format

The records are list-formated.

Position	Name	Type	Description
A	STID	Int	FIPS state code (required)
B	CYID	Int	FIPS county code (required)
C	LINK ID	Char	Link identifier (optional, put 0 for non-link)
D	SCC	Char	AMS Source Classification Code (SCC) (required)
E	DATA(1)	Real	Data for the first field in #DATA header list, given in the units of the first field in the

			#UNITS header list (required)
F	DATA(2)	Real	Data for the second field in #DATA header list, given in the units of the second field in the #UNITS header list (required if N>1)
G	DATA(3)	Real	Data for third field… (required if N > 2)
	DATA(N)	Real	Data for final field… (required if N > 3)

Mobile IDA Emissions Data Format

The records are fixed-format. The ozone-season data does not have to be truly "ozone-season." Any day-total value can be used in this field, if some other seasonal average or day-specific data need to be used in SMOKE.

Position	Name	Type	Description
1-2	STID	Int	State Code (required)
3-5	CYID	Int	County Code (required)
6-15	LINK ID	Char	Link identifier (optional, put zero for nonlink data)
16-25	SCC	Char	AMS Source Classification Code (required)
26-35	DATA(1)	Real	Data for the first field in #DATA header list (short tons/year) (required)
36-45	OSD(1)	Real	Ozone season data for the first field in the #DATA header list (short tons/average ozone day) (optional)
46-55	DATA(2)	Real	Data for the second field in #DATA header list (short tons/year)
56-65	OSD(2)	Real	Ozone season data for the second field in the #DATA header list (short tons/average ozone day)
66-75	DATA(3)	Real	Data for third field (short tons/year)
76-85	OSD(3)	Real	Ozone-season data for the third field (short

			tons/average ozone day)
	DATA(N)	Real	Data for final field (short tons/year)
	OSD(N)	Real	Ozone-season data for the final field (short tons/average ozone day)

EMS-95 Formats

For EMS-95 inputs, the MBINV file must list the nonlink and link EMS-95 files. This format is called a SMOKE list format. The file can be created by using the UNIX command, which assumes the EMS-95 files (and only these files) are all in the $MBDAT directory.

```
ls $MBDAT/*.mv > $MBINV
```

When using a list format, the MBINV file must also contain one or more **INVYEAR** packets. This packet acts as a switch that sets the inventory year for the files, each of which will get the year designated by the packet until another INVYEAR packet appears in the file. Multiple years of inventory data can be combined to the same SMOKE inventory file in this way.

The data files that are referenced by the SMOKE list formatted MBINV files may also contain some header fields. They may contain #COUNTRY, #YEAR, and/or #DATA header fields, as described for IDA-formatted area source files. The default country is 0 and the default year is set by the INVYEAR packet. The #DATA header must appear, even if the only data in the file are VMT data. **The #UNITS header must NOT appear.**

EMS-95 nonlink format (column-specific)

Line	Columns	Type	Description
1	A	Int	#NONLINK
2	A+	Int	#DATA VMT
3+	1-2	Int	FIPS state code (required)
	3-5	Int	FIPS county code (required)
	6-6	Int	Area type code (required)
	7-10	Int	Facility type code (required)

	11-18	Real	Vehicle Miles Traveled (VMT) (miles/average annual weekday) (required)

EMS-95 nonlink list-directed format (not column-specific)

Line	Columns	Type	Description
1	A	Char	#NONLINK
2	A+	Char	#DATA VMT PM10 PM2_5 … Last Data
3+	A	Int	Country, state, and county code (YSSCCC, required)
	B	Int	Road type code (required)
	C	Real	Vehicle Miles Traveled (VMT) (miles/average annual weekday) (required)
	D	Real	PM10 in (short tons/average annual weekday)
	E	Real	PM2_5 (short tons/average annual weekday)

(additional columns for additional data listed on the #DATA line)

EMS-95 link list-directed format (not-column-specific)

Line	Columns	Type	Description
1	A	Char	#LINK
2	A+	Char	#DATA VMT PM10 PM2_5 … Last Data
2+	A	Int	Country, state, and county code (YSSCCC, required)
	B	Int	Road type code (required)

	C	Char	Link ID (15 Character) (required)
	D	Real	Link start longitude (decimal degrees) or UTM Easting (meters) (required)
	E	Real	Link start latitude (decimal degrees) or UTM Northing (meters) (required)
	F	Real	Link end longitude (decimal degrees) or UTM Easting (meters) (required)
	G	Real	Link end latitude (decimal degrees) or UTM Northing (meters) (required)
	H	Int	UTM Zone or 0 for lat/lon coordinates (required)
	I	Real	Vehicle Miles Traveled (VMT) in (miles/average annual weekday) (required)
	J	Real	PM10 (short tons/average annual weekday)
	K	Real	PM2_5 (short tons/average annual weekday)

(additional columns for additional data listed on the #DATA line)

8.2.6 VMTMIX: Mobile vehicle mix data

The VMTMIX file is used for EMS-95 inventory import to disaggregate the data totals by road type. This file can be in EMS-95 column-specific format, or in a list-directed format based on the same. **NOTE: The road type or road class codes used in either VMTMIX file format must match the codes that are used in the mobile VMT inventory file. If roadclass codes are used for the free-formatted files, they must be used in both the inventory and VMTMIX files.**

[NOTE: Can this file use the #COUNTRY header?]

For both files, a header line is needed to indicate the number and order of vehicle types listed in the file. The standard EMS-95 structure is

```
#VTYPES   LDGV   LDGT1   LDGT2   HDGV   LDDV   LDDT   HDDV   MC
```

EMS-95 format (column specific)

Line	Columns	Type	Description
1	A	Char	#VTYPES header line (see example above)
2+	1-2	Int	State code (required)
	3-5	Int	County code (required)
	6-6	Int	Area type code (required)
	7-10	Int	Facility type code (required)
	11-20	Char	Link ID (optional, zero if nonlink)
	21-25	Real	Fraction of total VMT from gas vehicles (all fractions required)
	26-30	Real	Fraction of total VMT from light duty gas trucks 1
	31-35	Real	Fraction of total VMT from light duty gas trucks 2
	36-40	Real	Fraction of total VMT from heavy duty gas vehicles
	41-45	Real	Fraction of total VMT from light duty diesel vehicles
	46-50	Real	Fraction of total VMT from heavy duty diesel trucks
	51-55	Real	Fraction of total VMT from heavy duty diesel vehicles
	56-60	Real	Fraction of total VMT from motorcycles

EMS-95 format, list-directed (not column-specific)

Line	Columns	Type	Description
1	A	Char	#VTYPES header line (required)
2+	A	Int	Country, state, and county code (required)
	B	Int	Roadway type code or road class (required)
	C	Char	Link ID (Optional, zero if nonlink)
	D	Real	Fraction of total VMT from first vehicle type (required for all fractions)
	E	Real	Fraction of total VMT from second vehicle type

F	Real	Fraction of total VMT from third vehicle type
G	Real	Fraction of total VMT from fourth vehicle type
H	Real	Fraction of total VMT from fifth vehicle type
I	Real	Fraction of total VMT from sixth vehicle type
J	Real	Fraction of total VMT from seventh vehicle type
K	Real	Fraction of total VMT from eighth vehicle type

8.2.7 BCUSE: Biogenic county land use

BCUSE is the BEIS2-format ASCII biogenic county land use file input to the Rawbio program. BCUSE has the following structure:

Line	Columns	Description
1	A	Country/state/county code (Int)
2	A	Rural forest land area (Real)
	B	Rural forest number of subtypes (Int)
For each rural forest subtype: (These subtypes must match the emission factors in BFAC)		
3+	A	Name of land use subtype as per BEIS2 documentation
	B	Area within the county devoted to land use subtype
Repeat lines 2 and 3+ for each of urban forest, agriculture, and other.		

8.2.8 BGUSE: Biogenic gridded land use

BGUSE is the emissions prototype's BEIS2-format ASCII biogenic gridded land use file input to the Rawbio program. BGUSE is an ASCII file. Note that the data records in BGUSE all have list-directed formatting (e.g., quoted strings instead of free-form). given below:

Line	Columns	Description

1	A	#GRID header (see AGPRO and MGPRO)
optional	A	Header with text metadata sections including the following: Introduction; Grid Definition: File Description: A list of "VEGID" names and descriptions, formatted "(1X, A4, 3X, A)" Notes on VEGID: Source of Data Version:
2	A	ENDHEADER (this text is required even if there is no header)
Repeat for each cell		
3+	A	Cell column number (Integer)
	B	Cell row number (Integer)
	C	Cell area in hectares (Real)
Repeat for each land use types (rural forest, urban forest, agriculture, other)		
4+	A	Area of land use type in hectares (Real)
	B	Number of land use subtypes for this land use type in this cell (Integer)
Repeat for each land use subtype		
5+	2-5	4-character name of land use type (in quotes)
	7-80	Area of land use subtype in hectares (Real)

8.3 Temporal Allocation

8.3.1 [A|M|P]TPRO: Temporal profile file

The ATPRO, MTPRO, and PTPRO files are used for the Temporal program for area, mobile, and point sources, respectively. These files contain the monthly, weekly, and diurnal temporal profiles used to convert the annual emissions stored in the inventory file to hourly emissions. The SMOKE Tool can create these files, if you are using the Models-3 modeling system. The files can contain four types of records:

- Monthly temporal weighting factors (/MONTHLY/ packet)
- Weekly temporal weighting factors (/WEEKLY/ packet)
- Weekday Diurnal temporal weighting factors (/DIURNAL WEEKDAY/ packet)
- Weekend Diurnal temporal weighting factors (/DIURNAL WEEKEND/ packet)

The weekend diurnal temporal factors are optional. If the monthly factors are not used because all of the inventory emissions are supplied as "per day" emissions, at least one monthly profile is still needed to prevent and error. The file format is column-specific based on the columns listed in the following table. All three source categories use the same file format, and they can share the same file.

Line	Columns	Description
/MONTHLY/ packet		
1	1-20	Packet identifier = /MONTHLY/
2+	1-5	Monthly temporal profile code (Integer)
	7-9	Temporal weight for January (Integer)
	11-13	Temporal weight for February (Integer)
	...	...
	51-53	Temporal weight for December (Integer)
	54-58	Total of weights for entire year (Integer)
3	1-20	Packet terminator = /END/
/WEEKLY/ packet		

1	1-20	Packet identifier = /WEEKLY/
2+	1-5	Weekly temporal profile code (Integer)
	7-9	Temporal weight for Monday (Integer)
	11-13	Temporal weight for Tuesday (Integer)
	...	...
	31-33	Temporal weight for Sunday (Integer)
	34-38	Total of weights for entire week (Integer)
3	1-20	Packet terminator = /END/
/DIURNAL/ packet		
1	1-20	Packet identifier = /DIURNAL WEEKDAY/ or /DIURNAL WEEKEND/
2+	1-5	Diurnal temporal profile code (Integer)
	7-9	Temporal weight for hour beginning 0000 (Integer)
	11-13	Temporal weight for hour beginning 0100 (Integer)
	...	...
	99-101	Temporal weight for hour beginning 2300 (Integer)
	102-106	Total of weights for entire day (Integer)
3	1-20	Packet terminator = /END/

Example:

```
 /MONTHLY/
    1  83  83  83  83  83  83  83  83  83  83  83  83  996
/END/
/WEEKLY/
    1   1   1   1   1   1   0   0    5
    2   0   0   0   0   0   1   1    2
    3   1   1   1   1   1   0   0    5
/END/
/DIURNAL WEEKDAY/
    1   0   0   0   0   0   0   0   0   1   1   1   1   1   1   1   1   0   0
0   0   0   0   0   0    8
```

```
    2   0   0   0   0   0   0   0   0   1   1   1   1   1   1   1   1   0   0
0   0   0   0   0   0    8
   81 417 417 417 417 417 417 417 417 417 417 417 416 414 415 417 417 417 417
417 417 417 417 417 41510000
   82 388 351 315 299 299 321 348 402 441 456 457 461 457 461 461 460 464 460
456 456 455 453 450 42910000
/END/
```

8.3.2 [A|M|P]TREF: Temporal cross-reference file

The ATREF, MTREF, and PTREF files contain the area, mobile, and point source cross-reference records for assigning the temporal profiles to the sources. If needed, the same file can be used for area, mobile, or point source categories, but more fields are used for mobile than for area, and more fields are used for point than for mobile. The SMOKE Tool can create these files, if you are using the Models-3 modeling system. The description of the Temporal program in Chapter 7 provides a list of the matching hierarchies that SMOKE uses when applying the file to an inventory. **If a field is blank, it needs to be set to -9.**

Users can fill in whatever fields that theywant for the source characteristics after plant ID. It is important to ensure, however, that one is consistent between the PTINV and the PTREF file. For example, if one uses the Process ID in place of the Stack ID field in the PTINV file, the same approach would need to be used in building the PTREF file. Remember that all of the field widths are limited to what is in the original PTINV input format (IDA, EMS-95, or EPS). Also remember that these fields are character strings, and therefore leading zeros are not ignored (this has caused others problems in the past).

The pollutant field allows users to assigns a different temporal profile for each pollutant associated with a given source. This is an optional approach; therefore, if users wish not to use pollutant as a characteristic in assigning temporal profiles, users can put a "-9" in the pollutant field. In that case, all pollutants for a given source will be assigned that temporal profile.

Note that the country/state/county code is defined as follows:

```
000000      = applies to all counties, states, and counties
n00000      = applies to all states and counties in country n
nmm000      = applies to all counties in state mm and country n
nmmccc      = applies to county ccc in state mm and country n
```

8.3.2.1 Area-source temporal cross-reference file

Line	Columns	Description
1+	A	10-digit SCC code, or 7-digit with zeros filling the remaining digits (Character) (optional, put 0 for entry that is not SCC-specific)
	B	Monthly temporal profile code (Integer) (required)
	C	Weekly temporal profile code (Integer) (required)

	D	Weekday or weekend diurnal temporal profile code (Integer) (required)
	E	Pollutant name (Character) (optional, zero if not pollutant-dependent
	F	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent

8.3.2.2

Mobile-source temporal cross-reference file

Line	Columns	Description
1+	A	10-digit SCC code (Character) (optional, put 0 for entry that is not SCC-specific)
	B	Monthly temporal profile code (Integer) (required)
	C	Weekly temporal profile code (Integer) (required)
	D	Weekday or weekend diurnal temporal profile code (Integer) (required)
	E	Pollutant name (Character) (optional, zero if not pollutant-dependent)
	F	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent
	G	Link ID (Character) (optional, leave blank or put zero for entry that is not link-specific)

8.3.2.3

Point-source temporal cross-reference file

Note, for EMS-95 inventories, Characteristic 1 is the stack ID, Characteristic 2 is the device ID, and Characteristic 3 is the process ID.

Line	Columns	Description
1	A	/POINT DEFN/

	B	Number of point source characteristics in addition to plant ID in point source definition (for IDA, this will always be 4)
	C	Number of point source characteristic in addition to plant ID that is the SCC (if SCC is a part of the source definition), zero otherwise. IDA and EPS2 inventories: 4 EMS-95 inventories: 0
2+	A	10-digit SCC code, or 5-digit with zeros filling the remaining digits (Character) (optional, put 0 for entry that is not SCC-specific)
	B	Monthly temporal profile code (Integer) (required)
	C	Weekly temporal profile code (Integer) (required)
	D	Weekday or weekend diurnal temporal profile code (Integer) (required)
	E	Pollutant name (Character) (optional, zero if not pollutant-dependent)
	F	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent
	G	Plant ID (Character) (optional, leave blank for not plant-specific)
	H	Characteristic 1 (for IDA, this is Point ID or -9) (Character) (optional)
	I	Characteristic 2 (for IDA, this is Stack ID or -9) (Character) (optional)
	J	Characteristic 3 (for IDA, this is segment or -9) (Character) (optional)
	K	Characteristic 4 (blank for IDA because SCC already given in column A) (Character) (optional)
	L	Characteristic 5 (blank for IDA) (Character) (optional)

8.4 Spatial Allocation

Preparation of the spatial allocation inputs to SMOKE is one of the major stumbling blocks for most users at this time. All spatial allocation (a.k.a. "gridding") inputs need to be consistent, and each has a header to indicate the grid projection, location, dimensions, and resolution. SMOKE uses these headers to ensure all inputs are consistent. SMOKE also requires that any meteorology data inputs are consistent with these gridded data. This version of SMOKE cannot extract input data from larger data sets or aggregate data at a finer resolution, although these additions are planned for future SMOKE releases.

SMOKE does not require a Graphical Information System (GIS) to be able to install the software. However, the AGPRO, MGPRO, and BGPRO or BGUSE files require some processing with a GIS to create them. The options that users currently have to prepare these files are the following:

- Use the national 36-km Lambert surrogates provided with SMOKE (only appropriate for 36-km modeling). Users will have to manually copy and edit the 36-km file to create a file for a smaller grid.
- Use the Spatial Surrogate Tool included with EPA's MIMS interface.
- Use the "unified grid" surrogates available on-line at http://envpro.ncsc.org/emcenter. These data are hosted, but not supported, by MCNC. These data are at 4-km resolution for the Eastern US, and a program is available with the data to regrid and to aggregate the surrogate data for different grids. The output format from this program is not the SMOKE input format.
- Hire a contractor to prepare the gridding surrogates and/or gridded land use data, or prepare the data yourself using a GIS.

8.4.1 G_GRIDPATH: Grids and grid projection information

The G_GRIDPATH file is a file written by the Models-3 framework when you are using SMOKE through Models-3. Otherwise, users must manually prepare this file with a text editor (based on the sample files provided with the SMOKE installation).

Line	Columns	Description
1	A	Keyword for variable name (Character)
	B	Number of dimensions (Integer)
	C	Number of elements in the first dimension, *optional* (Integer)

	D	Number of elements in the second dimension, *optional* (Integer)
	E	Number of elements in the third dimension, *optional* (Integer)
2	A+	Value(s) of record (the number of values depends on entries in columns C, D, and E).

NOTE: The number of dimensions can be set to 0 for a single value, and then the value can appear on the same line as the keyword.

The keywords are:

- GDNAME_GD: Grid name (16 character max)
- GDDESC_GD: Grid description (80 character max)
- GDTYP_GD: Coordinate system type
  Valid entries include the following. Users are responsible for ensuring that other settings in the file are consistent with the grid selected using the I/O API documentation (see link below).
  - Latitude-longitude grids: “LAT-LON” or “GEOGRAPHIC” or “LATGRD3”
  - Lambert grids: “LAMBERT” or “LATGRD3”
  - UTM grids: “UTM” or “UTMGRD3”
- GDUNT_GD: Units of coordinate system
- P_ALP_GD: Alpha projection factor
- P_BET_GD: Beta projection factor
- P_GAM_GD: Gamma projection factor
- XCENT_GD: X center of projection
- YCENT_GD: Y center of projection
- XORIG_GD: Column offset (x-origin) in projection units
- YORIG_GD: Row offset (y-origin) in projection units
- XCELL_GD: X direction cell thickness in projection units
- YCELL_GD: Y direction cell thickness in projection units
- NCOLS: Number of grid columns (x-cells)

- NROWS: Number of grid rows (y-cells)
- NTHIK: Number of boundary cells (not used in SMOKE, but needs to be consistent between all of your files)
- NLAYS: Number of vertical layers (not used in SMOKE, see SMK_EMLAYS in Chapter 5, section on Shared Environment Variables).
- VG_TYP_GD: Type of vertical structure. (Not used in SMOKE). Possible values:
  - VGSGPH3: hydrostatic sigma-P
  - VGSGPN3: non- hydrostatic sigma-P
  - VGSIGZ3: sigma-Z
  - VGPRES3: pressure (mb)
  - VGZVAL3: Z (m) (above sea lvl)
  - VGHVAL3: H (m) (above ground)
- VGTPUN_GD: Units of vertical structure (not used in SMOKE)
- VGTOP_GD: Value of top of domain in units of vertical structure (not used in SMOKE)
- VGLVS_GD: Vertical layer structure values in units of vertical structure (not used in SMOKE)

The I/O API documentation provides more information on the settings for horizontal and vertical grids at http://envpro.ncsc.org/products/ioapi/H.GRIDS.html.

8.4.2 AGPRO and MGPRO: Spatial Surrogates Files

The spatial surrogates files contain the spatial allocation factors for all area sources and non-link mobile sources. One file may be used for both area and mobile sources, although the SMOKE programs use the two logical file names to access it.

Line	Columns	Description
1	A	#GRID
	B	Grid name
	C	X origin in units of the projection
	D	Y origin in units of the projection

	E	X direction cell length in units of the projection
	F	Y direction cell length in units of the projection
	G	Number of columns
	H	Number of rows
	I	Number of boundary cells
	J	Projection types: Latitude-Longitude: “LAT-LON” or “LATGRD3” Lambert Conformal: “LAMBERT” or “LAMGRD3” Universal Transverse Mercator: “UTM” or “UTMGRD3”
	K	Projection units
	L	Projection alpha value
	M	Projection beta value
	N	Projection gamma value
	O	X-dir projection center in units of the projection
	P	Y-dir projection center in units of the projection
2+	A	Spatial Surrogates code (area) or County feature/roadway type (mobile) (Integer) (required)
	B	Country/state/county code (6-digit Integer YSSCCC) (required):
	C	Grid column number (Integer) (required)
	D	Grid row number (Integer) (required)
	E	Spatial surrogate ratio (area) or fraction of county feature in cell (mobile) (Real) (required)

8.4.3 BGPRO: Biogenic spatial surrogate file

The BGPRO file must be provided to SMOKE when using county land use, in order to provide a means of converting county landuse to gridded landuse. The format for the BGPRO file is the same as for the AGPRO and MGPRO files (above). Typically, the same gridded surrogates file

would be used, and the Rawbio options would be set to assign the county area, agriculture, or forest surrogates to the various biogenic land use groups.

8.4.4 BELD3_A, BELD3_B, BELD3_TOT: BELD3 land use data files

BELD3_A, BELD3_B and BELD3_TOT files are I/OAPI netCDF files that contain gridded Biogenic Emissions Landcover Database version 3 (BELD3) data. The BELD3 consists of 230 land use types. Since the I/OAPI limit on the number of variables in one netCDF file is 120, the gridded BELD3 data is broken up into two files (BELD_A and BELD_B). The BELD_TOT file contains additional information on the BELD3 data for the same grid as the BELD3_A and BELD3_B files. For a list of the BELD3 land use types look in the B3XRF or B3FAC file. The BELD3_TOT file contains the following variables and descriptions:

- FIPCODE: Dominant FIPS code for each grid cell
- URW_FLAG: Urban-Rural-Water flag
- USGS_TYPE: Dominant USGS type for each grid cell
- USGS_TOTAL: total USGS area in each grid cell
- FOREST: total forest (FIA) area in each grid cell
- AGRICULTURE: total agricultural area in each grid cell
- TOTAL_AREA: total area in each grid cell
- LAND: total land area in each grid cell

8.4.5 AGREF and MGREF: Gridding cross-reference files

The gridding cross-reference files are used to match the emissions sources with the spatial surrogates in the surrogates file. Application of the gridding cross-reference files in SMOKE will always be done so that the most-specific matches override the least-specific ones. A description of how the gridding cross-reference is applied to the inventory is provided with the description of the Grdmat program.

Line	Columns	Description

1+	A	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent
	B	10-digit SCC (Character) (optional, set to zero if not SCC-dependent)
	C	Spatial surrogate code (Integer) (required)

8.5 Chemical speciation

8.5.1 GSPRO: Speciation profile file

The speciation profile file, GSPRO, contains the factors that are used to separate aggregated inventory pollutant emissions totals into emissions of model species for the AQM. In traditional ozone air quality modeling, this file contained split factors for only VOC. In SMOKE, all inventory pollutants need to be represented in this file, even if it is only to indicate that a pollutant is not split into any model species (e.g., CO emissions would have a "split factor" of 1.0).

Line	Columns (All required)	Description
1 +	A	Speciation profile number (Character)
	B	Pollutant ID (name) (Character)
	C	Species ID (name) (Character)
	D	Split factor (Real)
	E	Divisor (Real)
	F	Mass Fraction (Real)

The columns are used by SMOKE as follows:

- The speciation profile number is matched with the number in the GSREF file. This “number” is treated as a character string, so leading zeros must be the same in both the GSPRO and GSREF files.
- The pollutant ID must match the pollutant names in the SMOKE inventory file (AREA, MOBL, or PNTS). These names are set by the #POLID or #DATA headers in the Smkinven input files.
- The species IDs are the names of the model species needed by the air quality model for which SMOKE is being used. These depend on the requirement of that model. The default SMOKE inputs will not be correct for all air quality models.
- The split factor is the numerator in the factor computed for the mole-based speciation matrix.

- The Divisor is the denominator in the factor computed for the mole-based speciation matrix.
- The mass fraction is the value put in the mass-based speciation matrix. The ratio of the mass fraction and the split factor is the average molecular weight for the species and profile, if the divisor is 1.
- For BEIS3 modeling, the user must use profile "BV309" in order to properly speciate for CB-IV modeling. There are also BEIS3 speciation profiles for the RADM2 and SAPRC99 chemical mechanisms in the $GE_DAT directory.

8.5.2 GSREF: Speciation cross-reference file (mapping table)

The speciation cross-reference files (ASREF, MSREF, and PSREF) are used to match the emission sources with the speciation profiles in GSPRO.

Because of the many changes needed for supporting multiple inventory pollutants, flexible point source definitions, 6-digit country/state/county codes, and the use of pollutant-to-pollutant-type conversion factors, the EPS2.0 format is no longer supported by SMOKE and a new format has been defined. A header has been added that is used to define the point-source unique keys. The same file can be used for all source categories. The SMOKE programs use only the applicable entries in the file. Application of the file in SMOKE will always be done so that the most-specific matches override the least-specific ones. A description of how the speciation cross-reference is applied to the inventory is provided with the description of the Spcmat program.

8.5.2.1 Area- and mobile-source speciation cross-reference file entries

When emission processes are used (as with mobile source processing using the Emisfac program), the emission types must be listed instead of the pollutants in column C. For example, if the MEPROC file contained processes EXH (exhaust), EVP(evaporative), and RFL (refueling), then in order to use pollutant-specific entries in this file for mobile sources, records would need to be added to the file that use EXH__VOC, EVP__VOC, and RFL__VOC in column C. The separator for these process-pollutant combinations must be a double underscore.

Line	Columns	Description
1+	A	Area: 10-digit SCC code, or 7-digit with zeros filling the remaining digits (Character) **OR** Mobile: 10-character SCC code (Character) **OR** set to 0 if cross-reference entry is not SCC-specific

	B	Speciation profile code (Character) (required, must match codes in the GSPRO file)
	C	Pollutant name or emission type (Character) (required, much match pollutants for profile code in GSPRO file, zero for all pollutants)
	D	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent

This file can also be in column-specific format, with the fields in the following columns:
A: 1-10
B: 12-16
C: 18-33
D: 35-40

8.5.2.2 Point-source speciation cross-reference file entries

Note, for EMS-95 inventories, Characteristic 1 is the stack ID, Characteristic 2 is the device ID, and Characteristic 3 is the process ID.

Line	Columns	Description
1	A	/POINT DEFN/
	B	Number of point source characteristics in addition to plant ID in point source definition (for IDA, this will always be 4)
	C	Number of point source characteristic in addition to plant ID that is the SCC (if SCC is a part of the source definition), zero otherwise. IDA and EPS2 inventories: 4 EMS-95 inventories: 0
2+	A	8- or 10-digit SCC code (Character) (optional, set to 0 if cross-reference entry is not SCC-specific)
	B	Speciation profile code (Character) (required, must match codes in the GSPRO file)

	C	Pollutant name or emission type (Character) (required, much match pollutants for profile code in GSPRO file, 0 for all pollutants)
	D	Country/state/county code (6-digit Integer YSSCCC) (optional): - leave blank or put zero if not country/state-dependent; - set to YSS000 for country/state-dependent; - set to YSSCCC for county-dependent
	E	Plant ID for point sources (Character) (optional, leave blank if entry is not plant-specific)
	F	Characteristic 1 (for IDA, this is Point ID or -9) (Character) (optional)
	G	Characteristic 2 (for IDA, this is Stack ID or -9) (Character) (optional)
	H	Characteristic 3 (for IDA, this is segment or -9) (Character) (optional)
	I	Characteristic 4 (for IDA, this is 10-digit SCC, or 5-digit with zeroes) (Character) (optional)
	J	Characteristic 5 (blank for IDA) (Character) (optional)

This file can also be in column-specific format, with the fields in the following columns in rows 2 and higher:
A: 1-10
B: 12-16
C: 18-33
D: 35-40
E: 42-56
F: 58-72
G: 74-88
H: 90-104
I: 106-120
J: 122-138

8.5.3 GSCNV: Pollutant-to-pollutant conversion file (optional)

The pollutant-to-pollutant conversion factors files are required when there is a mismatch between

the pollutant in the inventory and the pollutant for which the speciation profiles have been developed. One example is when the inventory has VOC (ROG), but the speciation profiles were developed for TOG. SMOKE permits such a factor to be applied to any pollutant. The same format is used by all source categories. The available combinations and hierarchies for applying the factors are the same as the speciation cross-reference, with the exception of the plant and other more specific source characteristics for point sources. This file is not needed if there is no mismatch between the inventory and the pollutants in the speciation profiles file. The Spcmat program will input this file wheh the POLLUTANT_CONVERSION environment variable is set to Y.

If emission types are being used as with mobile source processing using the Emisfac program, the emission type must be listed instead of the pollutant. For example, if exhaust VOC were being converted to exhaust TOG, then EXH_VOC and EXH_TOG would need to appear in their respective columns on the header line (line 1 in the format below). It would not be correct to put only VOC and TOG. All emission processes must be listed, or the conversion factor that will be applied is 1. This approach permits conversion factors that are specific to each emission process (ie., different conversions for EXH and EVP).

Line	Columns	Description
1	1-16	Name of pollutant or emission type converting *from* (Character) (required)
	18-33	Name of pollutant or emission type converting *to* (Character) (required)
2+	1-6	Country/State/County code (Integer) (optional, put zero if entry is not country/state/county-specific)
	8-17	Area: 10-digit SCC (full SCC is needed) (Character) **OR** Mobile: 10-character SCC (full SCC is needed) (Character) **OR** Point: 8-digit SCC (full SCC is needed) (Character) **OR** Zero if entry is not SCC-specific
	19-23	Volatile conversion factor (Real) (required)
Repeat as lines 1 and 2+ needed for multiple pollutants and/or emission types		

8.6 Controls and Projection

8.6.1 GCNTL: Files for controls and/or past and future year projection

[NOTE: MAKE SURE TO HAVE UNITS FOR ALL CONTROL PACKET FIELDS AND INDICATE WHETHER PERCENTAGES ARE GIVEN AS 0-100 OR 0-1.]

[NOTE: The emissions values should be in tons/day.]

[NOTE: The text that explains these packets is currently provided with the Cntlmat documentation]

SMOKE uses the projection and control packets. A packet is a group of inputs used to apply specific types of controls. A single control file can be shared by area and point sources and the packets that do not apply to a given source category will be ignored. The following packets are recognized by the SMOKE control programs:

- /CTG/ contains settings for control technology guideline (CTG) controls, maximum achievable control technology (MACT) controls, and reasonably available control technology (RACT) controls. Contributes to the multiplicative control matrix.
- /CONTROL/ contains settings for control efficiency, rule effectiveness, and rule penetration. Contributes to the multiplicative control matrix. Cannot appear in the same input file with an /EMS_CONTROL/ packet.
- /EMS_CONTROL/ contains settings for control control efficiency, rule effectiveness, and rule penetration for both the base year and a future year. Also contains a point source conversion factor and a control aggregated factor. Cannot appear in the same input file with a /CONTROL/ packet. This packet is for point sources only.
- /ALLOWABLE/ contains county-specific, SIC-specific, SCC-specific controls, caps, and replacement emissions. Contributes to the multiplicative control matrix.
- /REACTIVITY/ contains settings needed for reactivity-based controls. Creates the reactivity control matrix.
- /PROJECTION/ contains settings to permit projection of emission values into the future. Creates the projection matrix for *all* pollutants and activities using the same factor per source.

These packets have the formats as described in the following tables. The projection formats have been designed to handle multiple year-to-year projections. Matching to the SCCs is performed using the 4 levels of SCCs defined by EPA. These levels are defined as follows:

Point sources:
Level 1: X0000000
Level 2: XXX00000

Level 3: XXXXXX00
Level 4: XXXXXXXX (full SCC)

Area and mobile sources:
Level 1: XX00000000
Level 2: XXXX000000
Level 3: XXXXXXX000
Level 4: XXXXXXXXXX (full SCC)

In the above hierarchies, the X's represent numbers in an SCC that could be matched during cross-referencing. The zeros represent defaulting that is implemented in the cross-reference file. For example, if a cross-reference file has a Level 1 point-source entry (e.g. 10000000), this would be a default for any SCC in the point inventory that starts with 1.

Where appropriate, the packets also support the new flexible point source definitions, as noted in the tables below. All packets except for the EMS_CONTROL packet are not column specific formats.

8.6.1.1 /CTG/ Packet

Line	Columns	Description
1	A	/CTG/
2+	A	Country/State/County code or Country/state code with zero for county or 0 (Integer)
	B	8 or 10-digit SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid SCC defaulting.
	C	Pollutant ID, or –9 if not a pollutant-specific control; *in most cases, the emissions cutoff value will be a pollutant-specific value, and that pollutant name needs to be placed in this column* (Character)
	D	Control technology control factor; required; value should be a fraction 0 to 1. (Real)
	E	CTG emissions cutoff for application of CTG, MACT, or RACT factor [tons/day] (Real)
	F	Maximum achievable control technology control factor; value should be a fraction 0 to 1, or –9 if missing; (Real)
	G	Reasonably achievable control technology control factor; value

		should be a fraction 0 to 1; or –9 if missing (Real)
3	A	/END/

The valid key combinations and their order of precedence during assignment to the inventory are:

1. County, [area: 10-digit SCC, point: 8-digit SCC], pollutant
2. County, [area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
3. County, [area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
4. County, [area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant
5. Country/state, [area: 10-digit SCC, point: 8-digit SCC], pollutant
6. Country/state, [area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
7. Country/state, [area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
8. Country/state, [area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant
9. [Area: 10-digit SCC, point: 8-digit SCC], pollutant
10. [Area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
11. [Area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
12. [Area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant
13. County, [area: 10-digit SCC, point: 8-digit SCC]
14. County, [area/mobile: left 7 of SCC, point: left 6 of SCC]
15. County, [area/mobile: left 4 of SCC, point: left 3 of SCC]
16. County, [area/mobile: left 2 of SCC, point: left 1 of SCC]
17. Country/state, [area: 10-digit SCC, point: 8-digit SCC]
18. Country/state, [area/mobile: left 7 of SCC, point: left 6 of SCC]
19. Country/state, [area/mobile: left 4 of SCC, point: left 3 of SCC]
20. Country/state, [area/mobile: left 2 of SCC, point: left 1 of SCC]
21. [Area: 10-digit SCC, point: 8-digit SCC]
22. [Area/mobile: left 7 of SCC, point: left 6 of SCC]

23. [Area/mobile: left 4 of SCC, point: left 3 of SCC]

24. [Area/mobile: left 2 of SCC, point: left 1 of SCC]

25. County, pollutant

26. County

27. Country/state, pollutant

28. Country/state

29. Pollutant (no further defaulting available)

8.6.1.2 /CONTROL/ Packet

Line	Columns	Description
1	A	/CONTROL/
2+	A	Country/State/County code or Country/state code with zero for county or 0 (Integer)
	B	8 or 10-digit SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid SCC defaulting.
	C	Pollutant ID or –9 if not a pollutant-specific control (Character)
	D	Primary control equipment code (PCEC); -9 or 0 applies to all equipment (*not yet supported*)
	E	Control efficiency [value should be a percent; e.g., enter 90 for a 90% control efficiency] (Real)
	F	Rule effectiveness [value should be a percent; e.g., enter 50 for a 50% rule effectiveness] (Real)
	G	Rule penetration rate [value should be a percent; e.g., enter 80 for a 80% rule penetration] (Real)
	H	Standard Industrial Category (SIC), *optional* (4-digit Integer)**; *not supported by SMOKE at this time***
	I	Plant ID for point sources, *optional* (Character)

	J	Characteristic 1 (for IDA inventory input format, this is Point ID), *optional* (Character)
	K	Characteristic 2 (for IDA inventory input format, this is Stack ID), *optional* (Character)
	L	Characteristic 3 (for IDA inventory input format, this is segment), *optional* (Character)
	M	Characteristic 4 (blank for IDA inventory input format), *optional* (Character)
	N	Characteristic 5 (blank for IDA inventory input format), *optional* (Character)
3	A	/END/

The valid key combinations and order for precedence in applying the records to the inventory are:

1. County, plant, Char1, Char2, Char3, Char4, Char 5, SCC, pollutant
2. County, plant, Char1, Char2, Char3, Char4, Char 5, pollutant
3. County, plant, Char1, Char2, Char3, Char4, pollutant
4. County, plant, Char1, Char2, Char3, pollutant
5. County, plant, Char1, Char2, pollutant
6. County, plant, Char1, pollutant
7. County, plant, pollutant

30. County, [area: 10-digit SCC, point: 8-digit SCC], pollutant
31. County, [area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
32. County, [area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
33. County, [area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant
34. Country/state, [area: 10-digit SCC, point: 8-digit SCC], pollutant
35. Country/state, [area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
36. Country/state, [area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
37. Country/state, [area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant

38. [Area: 10-digit SCC, point: 8-digit SCC], pollutant
39. [Area/mobile: left 7 of SCC, point: left 6 of SCC], pollutant
40. [Area/mobile: left 4 of SCC, point: left 3 of SCC], pollutant
41. [Area/mobile: left 2 of SCC, point: left 1 of SCC], pollutant
42. County, [area: 10-digit SCC, point: 8-digit SCC]
43. County, [area/mobile: left 7 of SCC, point: left 6 of SCC]
44. County, [area/mobile: left 4 of SCC, point: left 3 of SCC]
45. County, [area/mobile: left 2 of SCC, point: left 1 of SCC]
46. Country/state, [area: 10-digit SCC, point: 8-digit SCC]
47. Country/state, [area/mobile: left 7 of SCC, point: left 6 of SCC]
48. Country/state, [area/mobile: left 4 of SCC, point: left 3 of SCC]
49. Country/state, [area/mobile: left 2 of SCC, point: left 1 of SCC]
50. [Area: 10-digit SCC, point: 8-digit SCC]
51. [Area/mobile: left 7 of SCC, point: left 6 of SCC]
52. [Area/mobile: left 4 of SCC, point: left 3 of SCC]
53. [Area/mobile: left 2 of SCC, point: left 1 of SCC]
54. County, pollutant
55. County
56. Country/state, pollutant
57. Country/state
58. Pollutant (no further defaulting available)

8.6.1.3 /EMS_CONTROL/ Packet (point sources only)

Line	Columns	Description
1	A	/EMS_CONTROL/
2+	1-2	State code (Integer) or zero

	3-5	County code or zero (Integer)
	6-9	Standard Industrial Category (SIC) (Integer) or blank for missing; ***not supported by SMOKE at this time***
	10-17	8-character SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid SCC defaulting.
	18-22	Pollutant ID, leave blank if not a pollutant-specific control (Character)
	23-37	Plant ID (EMS-95 Facility, leave blank if not needed) (Character)
	38-49	Characteristic 2 (EMS-95 Device, blank if not needed) (Character)
	50-61	Characteristic 1 (EMS-95 Stack, blank if not needed) (Character)
	62-73	Characteristic 3 (EMS-95 Process, blank if not needed) (Character)
	74-81	Future year SCC code (*not used by SMOKE*) (Character)
	82-85	Base year control efficiency percent [1-100] (Real)
	86-89	Base year rule effectiveness percent [0-100] (Real)
	90-93	Base year rule penetration rate percent [0-100] (Real)
	94-97	Future year control efficiency fraction [0.0-1.0] (Real)
	98-101	Future year rule effectiveness fraction [0.0-1.0] (Real)
	102-105	Future year rule penetration rate fraction [0.0-1.0] (Real)
	106-110	Point source conversion factor fraction [0.0-1.0+ (values > 1.0 are okay)] (Real)
	111-126	Control strategy code (not used by SMOKE) (Character)
	127-130	Point source ratio (not used by SMOKE) [0.0-1.0] (Real)
	151-154	Control adjustment factor [0.0-1.0+ (values > 1.0 are okay)] (Real) **[Note: this is not a error in the format, columns 131-150 are ignored by the reader to be consistent with legacy files].**

3		/END/

The control adjustment factor can be precomputed from other values in each record. The value in this column overrides all other values on the record, and SMOKE uses this value directly as the factor to adjust the emissions.

Note that the expected order of the plant characteristics is reversed in this file from other EMS-95 files. The device is specified before the stack. Also note that the number of characters available for several of the input variables are not at the SMOKE maximum, so this file format must be used cautiously when using for inventories other than EMS-95. Finally, note that the base year factors are expected to be given as percentages (1-100), while the future year factors are expected to be given as fraction (0-1)

The valid key-combinations and order of precedence is the same as for the /CONTROL/ packet, described above.

8.6.1.4 /ALLOWABLE/ Packet

Note, for EMS-95 inventories, Characteristic 1 is the stack ID, Characteristic 2 is the device ID, and Characteristic 3 is the process ID. The matching hierarchy is the same as for the CONTROL packet, above.

Line	Columns	Description
1	A	/ALLOWABLE/
2+	A	Country/State/County code or Country/state code with zero for county or 0 (Integer)
	B	8 or 10-digit SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid SCC defaulting.
	C	Pollutant ID, or –9 if not a pollutant-specific control; *in most cases, the cap or replacement value will be a pollutant-specific value, and that pollutant name needs to be placed in this column* (Character)
	D	Control factor, *no longer used by SMOKE, enter –9 as placeholder* (Real)
	E	Allowable emissions cap value [tons/day]; required if no “replace” emissions are given (Real)

	F	Allowable emissions replacement value [tons/day]; required if no "cap" emissions are given (Real)
	G	4-digit SIC for point sources (Integer)or –9 if not needed**; *not supported by SMOKE at this time***
	H	Plant ID for point sources, *optional* (Character)
	I	Characteristic 1 (for IDA inventory input format, this is Point ID), *optional* (Character)
	J	Characteristic 2 (for IDA inventory input format, this is Stack ID), *optional* (Character)
	K	Characteristic 3 (for IDA inventory input format, this is segment), *optional* (Character)
	L	Characteristic 4 (blank for IDA inventory input format), *optional* (Character)
	M	Characteristic 5 (blank for IDA inventory input format), *optional* (Character)
3	A	/END/

The valid key-combinations and order of precedence is the same as for the /CONTROL/ packet, described above.

8.6.1.5 /REACTIVITY/ Packet

Note, for EMS-95 inventories, Characteristic 1 is the stack ID, Characteristic 2 is the device ID, and Characteristic 3 is the process ID.

Line	Columns	Description
1	A	/REACTIVITY <4-*digit base year*> <4-*digit future year*>/
2+	A	Country/State/County code or Country/state code with zero for county or 0 (Integer)
	B	8 or 10-digit SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid

		SCC defaulting.
	C	Pollutant ID, required (Character)
	D	Replacement emissions for base year [tons/day] (Real)
	E	Projection factor (or 0. for none; enter fractional value; e.g., 1.2 to increase emissions by 20%) (Real)
	F	Future year 10-digit SCC code (or 0 to set equal to base case) (Character)
	G	Speciation profile number (or 0 to set equal to base case) (Real)
	H	Market penetration of new speciation [fraction/year; e.g., enter 0.2 for 20% market penetration per year] (0. to 1.0) (Real)
	I	4-digit SIC for point sources (Integer)or –9 if not needed**; *not supported by SMOKE at this time***
	J	Plant ID for point sources, *optional* (Character)
	K	Characteristic 1 (for IDA, this is Point ID), *optional* (Character)
	L	Characteristic 2 (for IDA inventory input format, this is Stack ID) , *optional* (Character)
	M	Characteristic 3 (for IDA inventory input format, this is segment) , *optional* (Character)
	N	Characteristic 4 (blank for IDA inventory input format) , *optional* (Character)
	O	Characteristic 5 (blank for IDA inventory input format) , *optional* (Character)
3	A	/END/

The valid key-combinations and order of precedence is the same as for the /CONTROL/ packet, described above.

8.6.1.6 /PROJECTION/ Packet

Note that matching using the SIC is not yet supported. If the SCC is supplied, the SIC will not be used. Also note that there is not currently a pollutant/activity field in this file format, which means that all pollutants and/or activities for a source are changed with the same factor.

[NOTE: THIS NEW FORMAT STILL NEEDS TO BE IMPLEMENTED IN THE CODE!!!!! The pollutant ID has not bee added to the Cntlmat program at this time.]

Line	Columns	Description
1	A	/PROJECTION <4-*digit from year*> <4-*digit to year*>/
2+	A	Country/State/County code or Country/state code with zero for county or 0 (Integer)
	B	8 or 10-digit SCC, or pieces of SCC with remaining digits filled in with zeroes or 0 (Character). See hierarchy list above for valid SCC defaulting.
	C	Projection factor [enter number on fractional basis; e.g., enter 1.2 to increase emissions by 20%] (Real)
	D	Pollutant ID, or –9 if not a pollutant-specific projection
	E	4-digit SIC, or 2-digit with remaining digits filled in with zeroes or blank for missing, *optional* (Integer)**; *not supported by SMOKE at this time***
3	A	/END/

The valid key combinations and order of precedence in applying to the inventory are:

1. County, [area: 10-digit SCC, point: 8-digit SCC]
2. County, [area/mobile: left 7 of SCC, point: left 6 of SCC]
3. County, [area/mobile: left 4 of SCC, point: left 3 of SCC]
4. County, [area/mobile: left 2 of SCC, point: left 1 of SCC]
5. Country/state, [area: 10-digit SCC, point: 8-digit SCC]
6. Country/state, [area/mobile: left 7 of SCC, point: left 6 of SCC]
7. Country/state, [area/mobile: left 4 of SCC, point: left 3 of SCC]
8. Country/state, [area/mobile: left 2 of SCC, point: left 1 of SCC]

9. [Area: 10-digit SCC, point: 8-digit SCC]
10. [Area/mobile: left 7 of SCC, point: left 6 of SCC]
11. [Area/mobile: left 4 of SCC, point: left 3 of SCC]
12. [Area/mobile: left 2 of SCC, point: left 1 of SCC]
13. County
14. Country/state
15. Ultimate default

8.7 Merging

8.7.1 FILELIST: Gridded merge 2-d file list

The FILELIST file provides a list of logical file names to the Mrggrid program. This list should contain the logical file names for all 2-d files to be merged by a given run. It is a very simple file, which has no limit to the number of files listed.

Line	Columns	Description
1+	A	Logical file name

8.8 Meteorology input files

The meteorological input files needed by SMOKE need to come from a MM5 post-processor. Either Models-3 MCIP, EDSS MCIP, or MCPL can be used to supply SMOKE meteorology data.

The five meteorology files that SMOKE needs for biogenic calculations, mobile calculations, and plume rise are GRID_CRO_2D, GRID_CRO_3D, MET_CRO_2D, MET_CRO_3D, and MET_DOT_3D. SMOKE computes virtual temperature from information in these files.

8.8.1 METLIST: List of meteorology files

The METLIST file is used to provide the file names of available temperature files to the Tmpbio and Premobl programs. The Premobl program must runs once for the entire modeling episode, and therefore needs all of the files at one time. The Tmpbio program can also take individual files, but this file can be used to string multiple 1-day meteorology files together. **In the default SMOKE scripts this file is generated automatically.**

Line	Columns	Description
1+	A	Full name and path of I/O API meteorology file with temperature data

8.9 Source-specific input files

8.9.1 BIOGENICS

8.9.1.1 B3XRF: BELD3 to BELD2 cross-reference file

B3XRF is an ASCII file that contains cross-reference information in order to convert Biogenic Emissions Landcover Database version 3 (BELD3) data into the 127 BELD version 2 (BELD2) land use types. This cross-reference file is used in the program Beld3to2. The B3XRF file contains 230 (number of BELD3 land use types) lines of data in the following format:

Line	Columns	Description
1+	1-4	BELD2 land use type ID (Char)
	6-8	BELD2 category (Integer)
	10-25	BELD3 land use type ID (Char)

8.9.1.2 BFAC: BEIS2 emission factors file

BFAC is the same file as the BEIS2 ASCII emissions factor file. There are two different emission factor files available for use. There is a winter and a summer emissions factor file. BFAC (winter and summer files) have the following structure:

Line	Columns	Description
1+	2-6	Vegetation ID (Char)
	7-15	ISOP emission factor (Real)
	16-24	MONO emission factor (Real)
	25-33	OVOC emission factor (Real)
	34-39	NO emission factor (Real)
	40-41	Leaf area index (Integer)

8.9.1.3 B3FAC: BEIS3 emissions factors file

This file contains a leaf area index, dry leaf biomass, winter biomass factor, indicator of specific leaf weight, and normalized emission fluxes for 4 different species/compounds (ISOP, OVOC,

MONO, and NO). This ASCII file has the following structure:

Line	Columns	Description
1+	9-24	Abbreviated name for land use type (Character*16)
	A	Leaf area index (m^2/m^2) (Integer)
	B	Dry leaf biomass (g/m^2) (Integer)
	C	Winter biomass factor (Real)
	D	Indicator of specific leaf weight (Integer)
	E	Normalized emissions flux for ISOP (Real)
	F	Normalized emissions flux for MONO (Real)
	G	Normalized emissions flux for OVOC (Real)
	H	Normalized emissions flux for NO (Real)

8.9.2 MOBILE

8.9.2.1 MCODES: Mobile source codes file

The MCODES file contains the codes for mobile source vehicle types and road classes. It is used to convert from the codes in the inventory to the codes used internally by SMOKE. The vehicle types must be in the order listed below for the current version of the code so that they properly match with emission factors from MOBILE5b.

The SCC section maps the vehicle type and road class combinations to a mobile source SCC.

Line	Columns	Description
1	A	"/VEHICLE TYPES/" header
2+	A	Name of vehicle type from MBINV file (Char)
	B	Number of vehicle type used in SMOKE files (Integer)
	C	Y – use vehicle type in modeling N – do not use vehicle type

3	A	"/ROAD CLASSES/" header
4+	A	Road class code from last 3 digits of SCC in MBINV file (Integer)
	B	Road class code used in SMOKE files (Integer)
	C	Y – use road class code in modeling N – do not use road class code
5	A	"/SCC/" header
6+	A	Road class (Integer)
	B	Vehicle type name (Char)
	C	Source Category Code - SCC (10-character)

Example:

```
/VEHICLE TYPES/
LDGV   0100   Y
LDGT1  0102   Y
LDGT2  0104   Y
HDGV   0107   Y
LDDV   3000   Y
LDDT   3006   Y
HDDV   3007   Y
MC     0108   Y
LDGT   0106   N
/ROAD CLASSES/
110 01  Y
111 01  Y
130 02  Y
150 06  Y
170 07  Y
190 08  Y
210 09  Y
230 11  Y
250 12  Y
270 14  Y
290 16  Y
310 17  Y
330 19  Y
/SCC/
110 LDGV   2201001110
130 LDGV   2201001130
150 LDGV   2201001150
170 LDGV   2201001170
190 LDGV   2201001190
```

```
210 LDGV   2201001210
230 LDGV   2201001230
250 LDGV   2201001250
270 LDGV   2201001270
290 LDGV   2201001290
310 LDGV   2201001310
330 LDGV   2201001330
110 LDGT1  2201021110
130 LDGT1  2201020130
150 LDGT1  2201020150
170 LDGT1  2201020170
190 LDGT1  2201020190
210 LDGT1  2201020210
230 LDGT1  2201020230
250 LDGT1  2201020250
270 LDGT1  2201020270
290 LDGT1  2201020290
310 LDGT1  2201020310
330 LDGT1  2201020330
...
```

8.9.2.2 MEPROC: List of mobile emission processes and associated activities and pollutants

The MPROC file is a list of the emission processes to model for mobile sources.

Line	Columns	Description
1+	A	Activity name
	B	Mobile emissions process (See "Emission processes and emission types in Chapter 3)
	C+	Associated pollutant name(s) for activity and process combination

Example:

```
VMT EXH CO NOX VOC
VMT EVP VOC
VMT RST VOC
VMT RNL VOC
VMT DNL VOC
VMT RFL VOC
```

8.9.2.3 MCREF: Mobile county cross-reference file

The MCREF file is a list-formatted file used to assign reference counties to inventory counties.

Line	Columns	Description

1+	A	Two digit state FIPS code for inventory county (Integer)
	B	Three digit county FIPS code for inventory county (Integer)
	C	Two digit state FIPS code for reference county (Integer)
	D	Three digit county FIPS code for reference county (Integer)

Example:

```
01 001  01 001
01 003  01 001
01 005  01 001
01 007  01 001
01 009  01 001
01 011  01 001
01 013  01 001
01 015  01 001
01 017  01 001
01 019  01 001
```

8.9.2.4 MVREF: Mobile source county settings file

The MVREF file is a list-formatted file that assigns settings for various options to each reference county defined in the MCREF file.

Line	Columns	Description
1+	A	Two digit state FIPS code for reference county (Integer)
	B	Three digit county FIPS code of reference county (Integer)
	C	Setting for spatial averaging of temperatures 1 – temperatures are county specific 2 – temperatures are averaged over all counties sharing a reference county
	D	Setting for temporal averaging of temperatures 1 – no averaging; create hourly temperatures for each day 2 – average temperatures over weeks 3 – average temperatures over months 4 – average temperatures over entire episode
	E	Setting for local roads 1 – model local roads using MOBILE6 local road type (uses default average speed)

		2 – model local roads using MOBILE6 arterial road type with speeds from mobile inventory

8.9.2.5 Example:

```
01 001 2 2 1
04 001 2 2 1
04 013 2 2 1
04 019 2 2 1
05 001 2 2 1
```

8.9.2.6 M6LIST: MOBILE6 input scenario list file

The M6LIST file lists the available MOBILE6 input scenario files. The scenario file name must include the six digit FIPS code of the county.

Line	Columns	Description
1+	A	Full name and path of MOBILE6 input scenario

Example:

```
/smoke/data/inventory/mobile/scenarios/summer037001.in
/smoke/data/inventory/mobile/scenarios/summer037057.in
```

MOBILE6 input scenario format:

A MOBILE6 input scenario must be provided for each reference county in the MCREF file. These files follow the same format as MOBILE6 input files; most commands are allowed except for the following:

```
MOBILE6 INPUT FILE
MOBILE6 BATCH FILE
RUN DATA
END OF RUN
POLLUTANTS
PARTICULATES
AIR TOXICS
EXPRESS HC AS *
NO REFUELING
REPORT FILE
NO DESC OUTPUT
EXPAND *
IDLE PM EMISSIONS
DATABASE *
WITH FIELDNAMES
SPREADSHEET
EMISSIONS TABLE
DAILY OUTPUT
AGGREGATED OUTPUT
```

The file must include the either the MIN/MAX TEMPERATURES or HOURLY TEMPERATURE command and must have a SCENARIO RECORD command. SMOKE will ignore SPEED VMT, AVERAGE SPEED, or PARTICLE SIZE.

Example of MOBILE6 scenario:

```
SCENARIO RECORD    : 01001 96 07
CALENDAR YEAR      : 1996
EVALUATION MONTH   : 7
DIESEL SULFUR      : 500
MIN/MAX TEMPERATURE: 70.0 90.0
ALTITUDE           : 1
PARTICULATE EF     : PMGZML.CSV PMGDR1.CSV PMGDR2.CSV PMDZML.CSV PMDDR1.CSV
PMDDR2.CSV
FUEL RVP           : 7.8
GAS AROMATIC%      : 25.0
GAS OLEFIN%        : 15.0
GAS BENZENE%       : 1.5
E200               : 50.0
E300               : 85.0
OXYGENATE          : MTBE  15.1  0.50
                   : ETBE  17.6  0.05
                   : ETOH  10.0  0.45
                   : TAME   6.0  0.00
```

8.9.2.7 MEFLIST: Mobile emission factors list file

The MEFLIST file is used to provide the file names of available emission factors files to the Temporal program. This file cannot be created very easily until the Emisfac program has been run. Therefore, **in the default SMOKE scripts this file is generated automatically.**

Line	Columns	Description
1+	A	Full name and path of emission factors file

8.9.3 POINT

8.9.3.1 PELVCONFIG: Elevated source selection file (optional)

The Elevpoint program optionally uses the PELVCONFIG file. It permits users to provide SMOKE criteria for grouping stacks and for selecting elevated and plume-in-grid (PinG) point sources. Stack groups are groups of stacks from the same facility that are treated by SMOKE and the air quality model as a single stack when computing plume rise and/or processing the stack as a PinG source. They have the effect of reducing the number of elevated and PinG stacks which makes for faster processing, particularly in the case of PinG sources in the air quality models. Users can specify stack groups regardless of whether elevated or PinG sources are also being specified.

All emissions in a stack group are summed for the purposes of comparing against the elevated source criteria and PinG source criteria. In short, the grouped sources are treated as a single source for all SMOKE features regarding elevated source selection and special outputs. The sources still can have different temporal profiles and chemical speciation profiles. For more information on how the emissions are computed for source selection, see the Elevpoint documentation in Chapter 7.

The PELVCONFIG file has four sections: (1) source category selection, (2) source grouping, (3) plume-in-grid source selection, and (4) elevated source selection. Sections 2, 3, and 4 are optional, but Elevpoint will not run if all three sections are missing, since the file would be instructing it to do nothing. The sections need not appear in any particular order, and each section is permitted only once. Comments may also be used in the file. The following subsections explain these major input features.

8.9.3.1.1 USING COMMENTS

Comments can be put in the file using a pound sign (#) at the first column of a line or a double pound sign (##) anywhere else on a line. The double pound can appear after packet labels or instructions in the file.

8.9.3.1.2 SOURCE CATEGORY SELECTION

This section is used to tell Elevpoint which source category to process, but since Elevpoint is only designed for point sources, this is very simple. This section needs to appear as follows:

```
SMK_SOURCE    P
```

8.9.3.1.3 SOURCE GROUPING

The **/SPECIFY ELEV GROUPS/** packet is used for defining source groups based on source criteria. The packet contains the following entries:

- /SPECIFY ELEV GROUPS/
  Must appear to start the packet. Only one packet of this type is permitted
- <Source grouping criteria, explained next>
- /END/
  Must appear to end the packet.

The *source grouping criteria* can be as many lines as the user wants, and contain at least three parts: (1) variable, (2) type of criterion, and (3) value for criterion in the following structure:

<variable> <type of criterion> <value for criterion>

For example, the following source grouping criterion groups all inventory records (sources) at the same facility that have stack heights within plus or minus 2 meters:

```
HT  +/-  2.0
```

The fact that sources are grouped by facility is implicit to the Elevpoint program, since we assumed that grouping at a coarser level of detail was not desirable. The packet supports Boolean OR and AND conditions. Each separate line of the packet is an OR condition. The AND conditions must appear on a single line. Any number of records are permitted in this packet and any number of AND conditions or AND/OR combinations.

The following example will help illustrate. To indicate that sources with stack heights within plus or minus 2 meters that also have diameters within plus or minus 1 meter should be grouped, as well as sources with exit velocities below 10 m/s, the packet would be:

```
/SPECIFY ELEV GROUPS/
HT  +/-  2.0  AND  DIAMETER  +/- 1.0
VE  <  10.
/END/
```

The elevated groups packet supports the following variables:

- **HT**: stack height in meters
- **DM**: stack diameter in meters
- **TK**: stack gas exit temperature in Kelvin
- **VE**: stack gas exit velocity in meters/s
- **FL**: stack gas exit flow rate in cubic meters/s
- ***<pollutant name>***: whatever pollutant names are in the inventory (e.g., NOX). We do not know why a user might want this type of grouping, but SMOKE does not prevent it.

The elevated groups packet supports the following types of grouping criteria:

- **+/- or -/+**: variables must be within the <*value of criterion*> for the variable given
- **>**: variables must be greater than the <*value of criterion*> for the variable given
- **>= or =>**: variables must be greater than or equal to the <*value of criterion*> for the variable given
- **<**: variables must be less than the <*value of criterion*> for the variable given
- **<= or =<**: variables must be less than or equal to the <*value of criterion*> for the variable given
- **= or ==**: variables must be equal to the <*value of criterion*> for the variable given

8.9.3.1.4 PLUME-IN-GRID SOURCE SELECTION

The **/SPECIFY PING/** packet is used for selecting plume-in-grid sources based on source criteria. The packet contains the following entries:

- /SPECIFY PING/
  Must appear to start the packet. Only one packet of this type is permitted.
- <Source selection criteria, explained next>
- /END/
  Must appear to end the packet.

The *source selection criteria* can be as many lines as the user wants, and contain at least three parts: (1) variable, (2) type of criterion, and (3) value for criterion in the following structure:

<variable> <type of criterion> <value for criterion>

This structure is the same as the /SPECIFIC ELEV GROUPS/ packet, and the description of the source selection criteria is the same as that of the source grouping criteria. The same variables and types of criteria are supported, with two additions, as follows.

The plume-in-grid packet supports the following variables **in addition to those described with the elevated groups packet, above**:

- **RISE**: The plume rise as computed by the Briggs algorithm documented with the Elevpoint program. The <*value for* criterion> must be provided in meters. To select all sources with plume rise greater than or equal to 75 meters, one would input:

  ```
  RISE >= 75.
  ```

- **SOURCE:** Source number from SMOKE's PSRC file output from Smkinven. This variable permits selection of specific SMOKE sources as PinG sources.
- **PLANT:** Plant ID code from SMOKE's PSRC file output from Smkinven. **Please note**: the "PLANT" variable must be used with the "IS" criteria and cannot be used with any other criteria such as "<" or "=". It is recommended that this variable be used in addition to the FIPS variable to allow selection of all records at specific plants. It can also be combine with stack parameter variables.
- **FIPS:** Country/State/County region code can be used with the PLANT variable or other variables to specific PinG sources by region. It is not recommended that this variable be used alone for specifying PinG sources, since it will result in insignificant sources being prepared for the air quality model as PinG sources. It does not make sense to process such sources as PinG sources in air quality models.

The packet also supports the following type of criterion in addition to those described with the source grouping packet:

- **TOP**: This type permits users to select sources with the largest value of the variable given. The <*value for criterion*> is the number of sources that will be selected based on the value given. This type is only permitted when using a pollutant name as the <*variable*>.
- **IS**: This type must be used to assign a value to the PLANT criteria.

The following example illustrates using some of the different variables and criteria. The elevated source selection example further below demonstrates using the rest of the variables and criteria.

```
/SPECIFY PING/
RISE >= 150.
HT >= 100.
NOX TOP 100 AND SO2 TOP 100
NOX > 500.
/END/
```

8.9.3.1.5 ELEVATED SOURCE SELECTION

The **/SPECIFY ELEV/** packet is used for selecting elevated sources based on source criteria. The packet contains the following entries:

- /SPECIFY ELEV/
  Must appear to start the packet. Only one packet of this type is permitted.
- <Source selection criteria, explained above>
- /END/
  Must appear to end the packet.

All of the source selection criteria are the same as the PinG selection criteria at this time. To use the "cutoff" method for UAM-style processing, only the RISE variable should be used with no other selection criteria. This approach is the traditional approach for elevated source selection for UAM, UAM-V, REMSAD, and CAM_X; users can decide whether they want to use this traditional approach or a more complex set of criteria.

The following example illustrates using all of the different variables and criteria.

```
/SPECIFY ELEV/
RISE >= 40. AND RISE < 150.
NOX TOP 500
NH3 TOP 50
FIPS >= 37001 AND FIPS <= 37031
SOURCE > 100 AND SOURCE < 200
/END/
```

8.9.3.2 PSTK: Point source stack replacement parameters file

The PSTK file is an ASCII file that contains the replacement stack parameters for use by Smkinven when stack parameters are missing from the inventory file(s). The format of the file permits stack parameters to be replaced based on the point sources' country/state/county code and SCC.

Line	Columns	Description
1+	A	Country/State/County code (Integer)
	B	Standard Classification Code - SCC (10-Character)
	C	Stack height from ground [m] (Real)
	D	Stack diameter [m] (Real)
	E	Stack exhaust temperature [K] (Real)
	F	Stack exhaust velocity [m/s] (Real)

8.9.3.3 PTMPLIST: List of hourly point source emissions files

The PTMPLIST file is an ASCII file containing the full name and path of available hourly point source emissions files. This file is created automatically from the default SMOKE run scripts.

Line	Columns	Description
1+	A	Full name and path of hourly point source emissions file

8.10 QA and other support input files

8.10.1 ARTOPNT: Area-to-point conversions

This file provides the point source locations to assign to stationary area and nonroad mobile sources. The file is separated into sections. Each section starts with a header that contains the SCCs that will receive the point source locations and allocation factors contained in that section. Within the section, the content and format are similar to the EMS-HAP airport allocation factor (e.g., ap_af1.sas7bdat) files.

Each section of the file starts with the following packet:

/LOCATIONS/ <SCC 1> <SCC 2> … <SCC 9>

for SCCs 1 through 9 that apply to the section. The SCCs must be full 10-digit SCCs separated by spaces. A maximum of 9 SCCs may appear in one /LOCATIONS/ packet. If more than 9 SCCs are needed for a specific sections, than multiple /LOCATIONS/ packets may be included in subsequent rows prior to the section that they apply. In other words, two or more lines in sequence may include the/LOCATIONS/ packet, and all of the SCCs from both packets will be applied to the subsequent section of the file.

Within each section delineated by the /LOCATIONS/ packets, the correct format is provided in the table below.

Column	Name	Type	Description
1-6	COID	Int	Country, state, and county code (CSSYYY, where C = country, SS=state, and YYY=county) (required)
8-9	STABBR	Char 2	Postal abbreviation for state (not used by SMOKE)
11-32	COUNTY	Char 22	County name (not used by SMOKE)
34-37	LOCID	Char 4	Airport ID (unused for now, required for ISCST3)
39-61	NAME	Char 23	Airport name (if name contains spaces, commas, or semi-colons, then the entire name must be enclosed in double-quotes)
63-84	CITY	Char 22	City name (not used by SMOKE)
86-93	ARPT_USE	Char 8	Airport purpose (not used by SMOKE)
98-109	LON	Real	Longitude for the emissions (required)
114-124	LAT	Real	Latitude for the emissions (required)
126-139	ALLOC	Real	Allocation factor for this lat/lon, county, and SCC (required)

The section ends with the start of another packet (indicated by a new /LOCATIONS/ header line or with the end of the file.

8.10.2 INVTABLE: Inventory table

This file replaces the SIPOLS and ACTVNAMS files previously used by SMOKE. Smkinven uses the inventory table to select the inventory pollutants and pollutants that will be output from Smkinven. The valid pollutants are again checked at the end of SMOKE processing by Smkmerge, to ensure the program knows about all of the pollutants and also to set the order of the output variables (output species are sorted by the order of their pollutants in the inventory table and for multiple species per pollutant, by alphabetical order) Users can drop inventory pollutants in Smkinven without modifying the inventory files, simply by modifying this file. Smkinven also uses the file when reading EPS 2.0-formatted criteria inventories, to assign pollutant names based on the SAROAD codes. In addition, the units used in the SMOKE intermediate files will be listed in this file. The SMOKE programs Spcmat, Temporal, and Emisfac use this file to determine which pollutants should be included in the NONHAPVOC or NONHAPTOG calculation.

It is possible for a user to choose two pollutants that result in double counting the emissions from the inventory. For example, the user may want to model both benzo(a)pyrene and 7-PAH which would result in the emissions associated with CAS number 205992 (for benzo(a)pyrene) to be duplicated. In this situation, the user is responsible for the choices made and the resulting changes in inventory mass

Note that the Factor field is also potentially used here to adjust the mass of the pollutant (e.g., for metals in the inventory that are reported as specific compounds such as lead sulfate) before combining it.

The inventory table is similar to the HAP table except that it uses 11-character pollutant names as the unique identifiers for each row. The format of this file is as follows:

Position	Column name	Type
1 – 11	Inventory data name	11-character string
13 – 22	Chemical Abstract Service (CAS) number (optional, except for toxic compounds)	10-character string
24 – 28	Storage and Retrieval of Aerometric Data (SAROAD) code (optional)	Integer
30	Reactivity group (not used by SMOKE)	Integer
32	Keep (Y or N)	1-character string
34 – 38	Factor	Float

Position	Column name	Type
40	VOC or TOG component (V/T/N)	1-character string
42	Model species (Y/N)	1-character string
44	Explicit in mechanism (Y/N)	1-character string
46	Activity	1-character string
48 – 50	NTI	Integer
52 – 67	Units to be used in SMOKE intermediate files (*changing not recommended*)	16-character string
69 – 108	Inventory data description – a description of the name provided in columns 1-11.	40-character string
110 – 149	CAS pollutant description – a description of the CAS or pollutant code contained in the inventory	40-character string

We have elected to use a fixed file format for this file so that we will not have to include "-9" values as missing values for columns that are unused in some cases. For example, for the line that includes VMT, we will not need placeholders for the many columns that do not apply to VMT.

The purposes for these columns are as follows:

- Inventory data name: Unique name for pollutant or activity that will be stored and processed by SMOKE. Because of I/O API restrictions, these names can only include alpha-numeric characters and underscores, and the first character must be a letter. In addition, because SMOKE uses double-underscore as a separator when appending to variable names, no double-underscores are permitted in these names.
- CAS number: This code represents the pollutant that is provided in the inventory file input into SMOKE. It is named the "pollutant code" in the NIF 2.0 format, however, we call it CAS to distinguish between a SMOKE pollutant and because this number is an actual Chemical Abstracts Service (CAS) number for cases when the pollutant contained in the inventory has a valid CAS.
- SAROAD code: Same as the SAROAD code in the HAP table. This is used by SMOKE only for reporting purposes, since the inventory data name will be used by SMOKE for assignments of profiles processing**.** It is also used when importing EPS2 formatted data, which identifies emissions using a 5-digit code.
- Reactivity group: The reactivity group for the ASPEN model. It is not used in SMOKE.
- Keep: Indicates whether SMOKE should keep the CAS number associated with this inventory data name. SMOKE will keep any inventory data with at least one CAS number with a Y.

- Factor: The adjustment factor to use in splitting the emissions associated with a CAS into different inventory data names, or reducing the mass of a metal compound associated with a CAS to account for the metal mass only

- VOC or TOG component: A "V" should be entered if the CAS number is part of VOC emissions or TOG emissions. A "T" should be entered if the CAS number is part of TOG emissions only. An "N" should be entered if the CAS number is neither a part of VOC or TOG. Any other value besides "V" or "T" will be treated as an "N".

- Activity: A "Y" should be entered if the inventory data name is a type of activity instead of emissions. Any other values will be treated as an "N".

- Model species: A "Y" should be entered if the pollutant is also a VOC model species by the air quality model being targeted for the run. Any other values will be treated as an "N".

- Explicit in mechanism: A "Y" should be entered if the pollutant is a VOC species (model-species column must be "Y" as well), and that model-species is explicitly treated in the chemical mechanism. The pollutant is only an explicit species in the chemical mechanism when the GSPRO file definitions of the lumped VOC species have been adjusted to account for the pollutants/species that are explicit in the mechanism. The current explicit in mechanism species being used in the toxics CB4 mechanism for CMAQ are: formaldehyde, acetaldheyde, acrolein and 1,3 butadiene.

- NTI: Code identifying HAP on the Clean Air Act HAP list (list of 188 HAPs). Will be used for projecting HAP emissions to future years.

- Units: The units to store the data in the SMOKE intermediate files. This should be "tons/year" for emissions values, miles/year for VMT, and miles/hr for speed. Although this is provided as an input field for possible future changes, these values should *not* be changed by the user and are unrelated to the units of the emissions in the input files.

- Inventory data descriptions: This is the description of the inventory data name listed in the first column

- CAS pollutant description: Describes the pollutant that is coming from the inventory that is input into SMOKE.

Smkinven uses this file to confirm that the pollutant names in any input inventory file are valid. Smkinven only needs this file when emissions data such as CO are read, but does not need the file only activity data are being read.

Smkmerge uses this file to double-check the input pollutants being merged with the gridding matrix and optionally the speciation matrix. The order of the pollutants in this file will dictate the order of the species output for each of pollutants. Species from pollutant data will always preceed species from activity data in the Smkmerge output files. Smkmerge will sort multiple species for the same pollutant will be sorted in alphabetical order.

8.10.3 COSTCY: Country, state, and county names and data

A header line can be used in this file to indicate that population data are contained in the file as the last column on each line of the /COUNTY/ packet. This header line and the additional population column are optional. The header needs to be on the first line of the file, if it is present in the file. The syntax of the optional header is:

#POPULATION <year>

where <year> is a 4-digit year.

The country, state, and county codes in each section of this file must be sorted in increasing order. If they are not, the reader for this file will give an error.

Line	Columns	Description
1	A	Start of country section using packet /COUNTRY/
2+	1	Country code
	3-22	Country name
3	A	Start of state section using packet /STATE/
4+	1	Country code (Integer)
	2-3	State code (Integer)
	4-5	State abbreviation (Character, not used by SMOKE)
	7-26	State name (Character)
	27-28	EPA Region (Integer, optional, not used by SMOKE)
	32-34	Standard (non-daylight) state time zone - see below for valid time zones (Character, optional)
5	A	Start of county section using packet /COUNTY/
6+	2-3	State abbreviation, should match state section (Character, not used by SMOKE)
	5-24	County name (Character)
	26	Country code, should match country section (Integer)

	27-28	State code, should match state section (Integer)
	29-31	County code (Integer)
	32-34	AEROS State Code (Integer, not used by SMOKE)
	35-38	AEROS County Code (Integer, op not used by SMOKE tional)
	40-42	Standard (non-daylight) county time zone - see below for valid time zones (Character)
	43	Blank indicates daylight savings time **is** used in the county. Any character in this field will turn off the use of daylight savings time in the county.
	44-52	County center longitude [degrees] (Real, not used by SMOKE)
	53-61	County center latitude [degrees] (Real, not used by SMOKE)
	63-74	County area [square miles] (Real, not used by SMOKE)
	76-84	Westernmost longitude [degrees] (Real, not used by SMOKE)
	86-94	Easternmost longitude [degrees] (Real, not used by SMOKE)
	95-103	Southernmost latitude [degrees] (Real, not used by SMOKE)
	104-112	Northernmost latitude [degrees] (Real, not used by SMOKE)
	114-128	Population for the county [Optional, used only by Smkreport for using the NORMALIZE POPULATION instruction (see REPCONGIG file documentation]

Valid time zones are the following.

GMT	Time zone	Description
+0:00	GMT	Greenwich Mean Time (Also known as Zulu time)
-4:00	AST	Atlantic standard time
-5:00	EST	Eastern standard time

-6:00	CST	Central standard time
-7:00	MST	Mountain standard time
-8:00	PST	Pacific standard time
-9:00	YST	Yukon standard time
-10:00	HST	Hawaiian standard time
-10:00	CAT	Central Alaska time
-11:00	NT	Nome time

8.10.4 NHAPEXCLUDE: NonHAP VOC calculation exclusions file

The purpose of this file is to allow the user a way to exclude some sources from the NONHAPVOC or NONHAPTOG calculation for both on-road and nonroad precomputed emissions. In this case, Smkinven will not use the toxics pollutants to compute NONHAPVOC or NONHAPTOG, but rather will keep the VOC or TOG emissions and naming the same. The toxic emissions will be retained or dropped by Smkinven, depending on the status of the pollutants in the Inventory Table. If the pollutant is has Y for the INVTABLE columns that indicate the pollutant is a model species and is an explicit species, then the toxic emissions will not be dropped. If the pollutant is only a model species but not an explicit species, it will be renamed with _NOI appended to the name. If it is neither a model species nor an explicit species, the toxics mass will be dropped to prevent double counting of the inventory mass.

The file contains the country/state, county, and SCC codes to indicate countries/states, counties, SCC codes that are excluded from the NONHAP calculation. Using this file, SMOKE will exclude any source listed in the file from the NONHAPVOC and NONHAPTOG calculation.

Column	Column name	Type	Description
A	Country, state, and county code	Integer	Six digit FIPS code for country, state, and county as CSSYYY where C=country, SS=state, and YYY=county
B	SCC	10-character string	Ten character source classification category (SCC) code

To specify that a key in this file should not be used in assigning the cross-reference criteria, zeros may for used. For example, a SCC-only cross-reference entry would have a zero for the

country/state/county code. For a left-2 SCC entry, the final eight characters of the SCC in the cross-reference would be zero. The following examples illustrate entries for each of the 14 cross-reference levels listed above. Please see the SCC documentation in Chapter 2 for information on the segment numbers for area-source SCCs.

1. 001001 2501080050 (select a specific county and SCC)
2. 001001 2501080000 (select a specific county and all SCCs that match levels 1, 2, and 3)
3. 001001 2501000000 (select a specific county and all SCCs that match levels 1 and 2)
4. 001001 2500000000 (select a specific county and all SCCs that match level 1)
5. 001000 2501080050 (select a whole state and SCC)
6. 001000 2501080000 (select a whole state and all SCCs that match levels 1, 2, and 3)
7. 001000 2501000000 (select a specific county and all SCCs that match levels 1 and 2)
8. 001000 2500000000 (select a specific county and all SCCs that match level 1)
9. 001001 0000000000 (select all SCCs in a specific county)
10. 001000 0000000000 (select all SCCs in a specific state)
11. 000000 2501080050 (select a specific SCC)
12. 000000 2501080000 (select all SCCs that match levels 1, 2, and 3)
13. 000000 2501000000 (select all SCCs that match levels 1 and 2)
14. 000000 2500000000 (select all SCCs that match level 1)

8.10.5 SCCDESC: Source category code descriptions

This file provides a textual description of the source category codes (SCC) for:

- Smkinven when reading hourly emissions data (for reporting in the REPINVEN output file)
- Smkreport when user requests the SCC description be included in the report.

This file **should not include any descriptions that have the pipe symbol ("|"), single-quotes, or double-quotes.** It is okay to include dashes, slashes, semi-colons, colons, and commas in descriptions.

It is a 2-column, list-formatted file as follows:

Lines	Columns	Description
1+	A	Source category code (SCC), 10-character string
	B	SCC description, 200-character string. The description should be

8.10.6 ORISDESC: ORIS description

This file provides descriptive information for the ORIS IDs, which are used when reporting the status of reading and matching CEM data in Smkinven. The information will appear in the REPINVEN output file from Smkinven. The file is a list-formatted, 11-column file with the following column information. Note that only the first 5 columns of data need to appear in the file, and the rest can be blank since SMOKE does not use any information past column 5.

Lines	Columns	Description
1+	A	DOE Plant ID (ORIS ID), 6-character string, must match the same field in the IDA-formatted PTINV file.
	B	Country/State/County code (FIPS code), 6-digit integer
	C	Latitude (optional, *not used by SMOKE*), use –9 for missing
	D	Longitude (*not used by SMOKE*), use –9 for missing
	E	Plant name, 80-character string
	F	Operating company (optional, *not used by SMOKE*)
	G	EPA Region (optional, *not used by SMOKE*)
	H	State name (optional, *not used by SMOKE*)
	I	County nam e(optional, *not used by SMOKE*)
	J	North American Electric Reliability Council Region (e.g., ASCC, ECAR, ERCOT, MAAC, MAIN, MAPP, NPCC, SERC, SPP, WSCC) (optional, *not used by SMOKE*)
	K	Acid rain program phase (optional, *not used by SMOKE*)

		• P1 = Phase I Generator regulation began in 1995-1999 • P1.5 = Phase II Generator in which source opted to begin early regulation prior to the year 2000. • P2 = Phase II Generator - generator regulation began in year 2000. • If blank then plant is likely in a non-generating status (such as retired, deferred, exempt, or future.)

8.10.7 VNAMMAP: Variable name mapping file

This file is used to rename variables output from Smk2emis, the I/O API format to 2-d-UAM "EMISSIONS" file format converter. The VNAMMAP file lists the SMOKE variable names and the variable names to be output from Smk2emis. It is needed for SMOKE support of UAM-AERO, which requires variable names that cannot be supported by the netCDF format. Please see the Smk2emis documentation in Chapter 8 for more information on how to get Smk2emis to use this optional file.

The format of the VNAMMAP file is the following:

Lines	Columns	Description
1+	A	SMOKE I/O API variable name
	B	Smk2emis output name to be written to the UAM "EMISSIONS" file binary format

8.10.8 REPCONFIG: Reports configuration file

This file is used to control the Smkreport program. Please see the Smkreport documentation for more information. Each run of Smkreport is constrained to a single source category (area, mobile, or point), at most one speciation scenario, at most one grid, and at most one hourly emissions file. All of these component datasets are optional, however, and their use is controlled by the settings in this file. This file also controls the number and types of reports created by Smkreport. Lastly, REPCONFIG cannot specify control matrices to be read by Smkreport.

This file has three sections and a large number of options for each section. The file works on the principles of "packets" and "instructions." Packets are groups of instructions or a single action, and instructions are settings that tell the program what to do. The file also can contain comments and an instruction for creating several files.

The maximum number of characters that can be read on any line in this file is 300.

8.10.8.1 USING COMMENTS

Comments can be put in the file using a pound sign (#) at the first column of a line or a double pound sign (##) anywhere else on a line. The double pound can appear after packet labels or instructions in the file.

8.10.8.2 ONE-TIME SECTION

The first section of the file is the One-time section. This section is used for setting the environment variables that control Smkreport for all reports in a given run. The variables are set by listing the environment variable as an instruction, followed by its value. No "setenv" command (used in the C shell) is needed, because SMOKE is interpreting these values directly. The instructions to control these values are the following:

SMK_SOURCE:
Set as "A", "M", or "P", for area-source reporting, mobile-source reporting, and point-source reporting, respectively.

An example of the way this setting would appear in the REPCONFIG file is the following:

```
SMK_SOURCE    A
```

8.10.8.3 DEFINE-GROUP SECTION

The second section of the file is the Define-group Section. This optional section is used for defining groups of source characteristics for use later in the create-reports section. This section uses packets for defining these groups. There are currently two types of groupings, and therefore two packet types for this section: regions and subgrids.

The **/DEFINE GROUP REGION/** packet is used for defining groups by the country, state, and county codes in the inventory. The packet contains the following entries.

- /DEFINE GROUP REGION/ <Label>
  Must appear to start the packet. A label (see below) is required for naming the group.
- INCLUDE and/or EXCLUDE instruction(s) (see below) - optional.
- Country, state, and county codes in format NSSCCC, where N is the country number, SS is the state code, and CCC is the county code. Zeros can be used in place of CCC to select all counties in a state, or in place of SSCCC to select all counties and states in a country.
- /END/
  Must appear to end the packet.

The **/DEFINE SUBGRID/** packet is used for defining subdomains based on cell counts from the full grid being applied using the gridding matrices. The packet contains the following entries.

- /DEFINE SUBGRID/ <Label>

Must appear to start the packet. A label (see below) is required for naming the subgrid.

- INCLUDE and/or EXCLUDE instruction(s) (see below) - optional.
- (X cell, Y cell):
  Cell numbers should appear for all cells in the subgrid. The number of the cells depends on the grid defined by the gridding matrix being used by Smkreport. A range can be given with the following syntax:

  (X cell, Y cell) TO (X cell, Ycell)

  If only one cell is to be specified, the range must still appear, with both entries indicating the same cell.

  Smkreport is not very sensitive to the spacing of this instruction, and additional spaces or no spaces may be placed between the parentheses, the cell numbers, and the commas. Spaces around the "TO" are recommended.

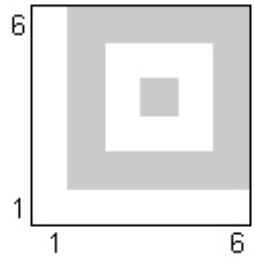
  The maximum width of these lines is 100 characters. Multiple entries of this type may appear in the packet, allowing for the grouping of spatially unconnected cells/regions or non-rectangular regions.
- /END/
  Must appear to end the packet.

The **INCLUDE** and **EXCLUDE** instructions are used for defining the meaning of the entries in the packet. If the INCLUDE instruction appears, then Smkreport interprets the subsequent entries in the packet as identifying records for *including* in the output report(s). On the other hand, if the EXCLUDE instruction appears, then Smkreport interprets the subsequent entries in the packet as identifying records to *exclude* from the output report(s). The INCLUDE and EXCLUDE instructions may appear in both the /DEFINE GROUP REGION/ and /DEFINE SUBGRID/ packets. These instructions act as toggle switches, and Smkreport will give a warning if a county previously included is later excluded. All subsequent packet entries after an instruction will be interpreted according to the instruction, until another instruction is given, or until the packet ends. The INCLUDE instruction is the default, so if neither are given, SMOKE will assume that the records in the packet are to be only those included in the report.

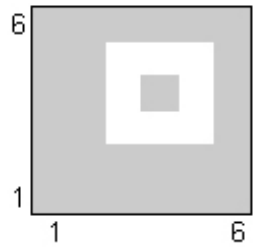
Whether an INCLUDE or EXCLUDE instruction is provided first in a group definition can change the meaning of a group definition. If INCLUDE is given first, Smkreport will initialize the group to exclude all grid cells and proceed interpreting the instructions. If EXCLUDE is given first, Smkreport will initialize the group to include all grid cells. Here is an example for a small 6 cell by 6 cell grid, which shows how the order of the INCLUDE and EXCLUDE statements will change the meaning of a group definition. If the instructions are

```
/DEFINE GROUP REGION/ Example group
    INCLUDE (2,2) TO (6,6)
    EXCLUDE (3,3) TO (5,5)
    INCLUDE (4,4) TO (4,4)
/END/
```

then the following cells will be selected as shown by the gray cells in the picture below:



If the instructions were rearranged to have the EXCLUDE statement first, then the initialized status for all cells will be INCLUDE. The additional INCLUDE statements are therefore unneeded and the definition of the group has changed to the following:



Group labels, denoted above by **<Label>**, are used to name a group for identification when creating a report in the create-report section. The text appearing in the placeholder <Label> can be any alpha-numeric combination of characters up to 200 characters, including intervening blanks. The text string (less trailing blanks) after /DEFINE …/ and before a comment (indicated with ##), is treated as the group label. References to the group label in the Create-report Section do not have to be consistent in the case of the text; SMOKE will ignore the case of the text, and treat lower- and upper-case characters as if they were the same.

In the future, other types of groups will be permitted:

- /DEFINE GROUP SCC/: Permits user-defined SCC groups
- /DEFINE GROUP PLANT/ Permits user-defined groups by country/state/county code and plant ID
- /DEFINE GROUP STACK/ Permits user-define groups by country/state/county code, plant ID, and stack ID
- /DEFINE GROUP POINT/ / Permits user-define groups by country/state/county code, plant ID, stack ID, and point ID
- /DEFINE GROUP SOURCE/: Permits user-defined source groups by SMOKE source ID
- /DEFINE GROUP HOURS/: Permits user-defined hour groupings

8.10.8.4 CREATE-REPORT SECTION

The third section of this file is called the Create-report Section, and it is responsible for providing instructions to Smkreport for generating the reports. The /CREATE REPORT/ packet is used in this section, and this packet uses several types of instructions. The instructions that can be used in the /CREATE REPORT/ packet are the following, listed in alphabetical order:

- BY
- GRIDDING
- LAYFRAC
- NORMALIZE
- NUMBER
- SELECT
- SPECIATION
- TEMPORAL
- TITLE
- UNITS

By using these instructions (and some have a number of different options), a large number of different types of outputs can be achieved by Smkreport. **The order that the instructions appear in the packet is not important**, with one exception. If multiple titles are used, they will be printed to the report in the order that they appear in the packet. The remainder of this

subsection provides descriptions of each of these instructions.

BY: The BY instruction controls what columns in addition to data columns will be printed in the output report. Multiple BY commands may appear in the same /CREATE REPORT/ packet, but not all combinations make sense. The BY instructions tell the program the basis on which to sum the emissions or activity values. The syntax of the BY instruction is:

BY <option>, with the following valid options:

- CELL: Records are printed by cell. This option forces the GRIDDING instruction.
- COUNTRY [NAME]: Records are printed by country total. This option invokes use of the COSTCY file. NAME can be appended to the option to additionally output the country name instead of only the country number.
- COUNTY [NAME]: Records are printed by county total. This option invokes use of the COSTCY file. NAME can be appended to the option to additionally output the county name instead of only the county number.
- DIUCODE: Records are printed by diurnal profile code. **This option will only work correctly when monthly temporal profiles have not been applied by pollutant.**
- ELEVSTAT: Records include the status of sources as elevated (E), plume-in-grid (P), and low-level (L). Applies for point sources only.
- HOUR: Records are printed by hour total for each day. This option forces the TEMPORAL instruction.
- LAYER: For point sources, emissions are printed by vertical layer. This option implies the LAYFRAC option that requires input of the PLAY SMOKE intermediate file. It also implies the BY HOUR instruction, and emissions will be written by hour as well as by layer. In the report, records are written first for 1 layer, then the next, and so on. The JART software can be used to sort the records in a different order. Only records with non-zero emissions in a layer are written. Since the number of records this report can generate it very large, we recommend that a source group be used to subselect the inventory records.
- PLANT [NAME]: For point sources, emissions are printed by plant total. This option cannot be used in conjunction with the BY SOURCE option, and if the REPCONFIG file contains both instructions, SMOKE will ignore the BY PLANT instruction.
- MONCODE: Records are printed by monthly profile code. **This option will only work correctly when monthly temporal profiles have not been applied by pollutant.**
- ROADCLASS: Records are printed by road class total. This option can only be used for mobile sources.

- SCC10 [NAME]: Records are printed by full 10-digit SCC totals.

  When the "NAME" option on the BY SCC10 instruction is included, the column delimiter default is automatically changed from a semi-colon to a pipe symbol ("|"). This is because the SCC descriptions include semi-colons, commas, and colons - such that some spreadsheet programs (e.g., Excel) cannot read in the report cleanly. Additionally, the SCC name is surrounded by double-quotes, which can help with the delimiter issue. If the /DELIMITER/ instruction is used, the user-specified delimiter is included in all subsequent reports, regardless of whether the BY SCC10 instruction with the NAME option is set in the report configuration (REPCONFIG) file.

- SOURCE [STACKPARM] [NAME]: Records are printed by source totals. Source characteristics are printed as well:

  Point: country/state/county code, plant ID, stack ID, point ID, segment ID, and SCC
  Area: country/state/county code and SCC
  Mobile: country/state/county code, road type code, vehicle type code, link ID, and SCC

  For area and mobile source, the SCC descriptions will automatically be included in the report. This is the same as adding the “BY SCC10 NAME” instruction to the report instructions. There is no way to turn off this feature.

  For point sources, STACKPARM can be appended to the option to also output the stack parameters (height, diameter, exit velocity, exit temperature). Also for point sources, NAME can be appended to the option to get the facility names in the output file.

- SPCCODE <POLLUTANT>: Records are printed by speciation profile codes assigned to sources by the speciation cross-reference file. **Only one pollutant can appear with this command throughout the entire REPCONFIG file.** The pollutant should be one that is in the inventory file and that was processed as part of the chemical speciation file run. This BY option does not imply the SPECIATION instruction, and it can be used without included speciated emissions in the report.

- SRGCODE [FALLBACK]: Records are printed by surrogate codes assigned to sources by the gridding cross-reference file. If no option is used, both primary and fallback surrogate codes are included in the output report. If the FALLBACK option is used, only the fallback surrogate codes are included in the output report, and a zero is used for all records that did not use a fallback surrogate (i.e., their primary surrogate did not cause emissions to go to zero). This BY option does not imply the GRDMAT instruction, and it can be used without applying the gridding matrix in the report. A value of –9 is used for fallback surrogate IDs in the resulting report.

- STATE [NAME]: Records are printed by state totals. This option invokes the use of the COSTCY file. NAME can be appended to the option to additionally output the state name instead of only the state number.

- WEKCODE: Records are printed by weekly profile code. **This option will only work correctly when monthly temporal profiles have not been applied by pollutant.**

The most detailed BY instruction will determine the resolution of the emissions and activity totals. For example, the COUNTRY, STATE, and COUNTY can be used in the same report to affect the column headers (e.g. to get the state name in the county report), but the data output will not be different because the emissions totals will be at the finest resolution specified. In other words, if the instructs are to report BY COUNTY and BY STATE, the emission totals in that report will be by county, and the BY STATE instruction will merely add the state name to the columns printed. Another example is that the BY SCC10 instruction will override the resolution of BY ROADCLASS.

Future versions of the BY command will support the following options:

- SCC[#] [NAME]: Records are printed by collections of SCCs based on the number selected from 1 to 10. For example, SCC3 would total emissions by and SCC tier with the structure ###*******, where ### is the first three digits of the SCCs, and the asterisks indicate that the rest of the SCC number is ignored for purposes of grouping. The NAME trailer could be used to include the name of the SCC as one of the records in the output report.
- STACK [NAME]: For point sources only, records are printed as facility and stack totals. NAME can be appended to the option to get the facility names in the output file.
- POINT [NAME]: For point sources only, records are printed as facility, stack, and point totals. NAME can be appended to the option to get the facility names in the output file.
- VTYPE: For mobile sources only, records are printed as vehicle-type totals.
- SIC: For point sources only, records are printed as Standard Industrial Code (SIC) totals.
- GROUP <option>: Records are printed based on the user-defined groups. Where the following options would be supported:
    - REGION: Rows in report are user-defined region groups.
    - SCC: Rows in report are user-defined SCC groups.
    - SOURCE: Rows in report are user-defined source groups, by SMOKE record number.
    - HOURS: Rows in report are user-defined hour groups.

CONTROL: The CONTROL instruction will apply the multiplicative control matrix to all pollutants and species requested in the report. This instruction does not limit the options available for resolution of the report, so reports can include any combination of BY instructions available.

GRIDDING: The GRIDDING instruction controls whether or not the gridding matrix is going to be used by the program. When the instruction is present, the gridding matrix is used in generating the report. The gridding matrix file name is determined by the environment variables set at runtime of the Smkreport program (as described in Chapter 8). When the GRIDDING instruction is given, the gridding matrix will be applied to the emissions before the totals are made. One use of this option would be the following: a report could be generated without GRIDDING to see emission totals before gridding, and another report could be generated with GRIDDING to see the totals after gridding. The totals could be compared from the two reports to see if any emissions were lost because of the gridding step.

The GRIDDING instruction is assumed when any of the following other instructions appear in the packet:

SELECT SUBGRID,
BY CELL,
NORMALIZE CELLAREA.

LAYFRAC: The LAYFRAC instruction controls whether or not the point source layer fractions file (PLAY) from the Laypoint program will be used in determining elevated and low-level sources. If the PLAY file is not used for determining elevated sources, the PELV file from the Elevpoint program will be required.

NORMALIZE: The NORMALIZE instruction controls whether or not the emission values will be normalized using some other data value. Two normalize options are currently supported; the CELLAREA option and the POPULATION option. When the emissions are normalized by cell area, they are divided by the area of each cell (or just the cells selected if a subgrid is used).

Emissions normalization can provide additional insight to the inventory instead of simply summarizing emissions values. Normalizing by grid-cell area is useful for comparing emission reports from grids at different resolutions, including nested grids, to ensure that a consistent emissions density is maintained across the multiple grid resolutions.

The syntax of the NORMALIZE instruction is:

NORMALIZE <option>, with the following valid options:

- CELLAREA: Normalize emissions based on grid cell area.

  For Lambert and UTM grids, the cell ares will be computed using the formula:

  $$8.10.8.4.1.1.1 \quad DX * DY$$

  where:

  DX is the cell x-direction size in meters
  DY is the cell y-direction size in meters

For lat-lon grids, the area of the lat-lon cell will be estimated using the formula:

$$(\cos Y)*(\pi * DX * R / 180)*(\pi * DY * R / 180)$$

where:

DX is the cell x-direction size in degrees longitude
DY is the cell y-direction size in degrees latitude
Y is the cell center in degrees latitude
R is the radius of the earth in meters

This instruction forces the GRIDDING instruction to be used, because the gridding information is needed in order to normalize the emissions by grid area. If the GRIDDING instruction is left out of the report instructions, Smkreport will add it to the report and continue.

- POPULATION: Normalize emission based on county population. Emissions in the report will be divided by county population, regardless of whether the BY COUNTY instruction has been used in the report instructions.

  This option requires that the COSTCY input file to Smkreport include an additional field and a #POPULATION header, as is described in the COSTCY documentation.

Future versions of the NORMALIZE instruction will support the following options:

- NORMALIZE COUNTYAREA: Normalize emissions based on county area.

NUMBER: The NUMBER instruction controls the format used to display *data values* in the report. The argument for the number instruction applies to all data values that are written to the report, and the structure of the argument is simply the formats supported by FORTRAN. The data values in the reports are rounded to the nearest value allowed based on the format selected. A value of five is rounded up to the next decimal place, when such rounded is needed.

The syntax of the NUMBER instruction is simply:

NUMBER <format>

where <format> is the FORTRAN format specification. The common formats one might use are the following:

Float format:
F<x>.<y>: where <x> is the number of digits for the number, including the decimal place, and <y> is the number of digits shown after the decimal place.

Exponential format:
E<x>.<y>: where <x> is the complete width of the value, including five characters for "E+00" and the decimal, and <y> is the number of significant figures.

If the NUMBER instruction does not appear in the report instructions, the E8.3 format is used.

PROJECTION: The PROJECTION instruction will apply the projection matrix to all pollutants and species requested in the report. This instruction does not limit the options available for resolution of the report, so reports can include any combination of BY instructions available.

SELECT: The SELECT instruction controls which records in the inventory are used in generating the report. It allows the user to extract particular records from the inventory, for inclusion in the report. This instruction also permits references to the groups defined in the Define-Groups section. If no SELECT instruction is used, then the other report instructions will be applied to all records in the inventory. Only one SELECT statement of each type is permitted in per report; for example, only one SELECT REGION statement and one SELECT SUBGRID statement could be used for a single report. Of course, different SELECT REGION or SELECT SUBGRID instructions could be used for different reports in the same output file.

The syntax of the command is:

SELECT <options> <label>,

where <label> is an optional field, depending on the <option> selected. The following options are available for the SELECT instruction:

- DATA <label>: Selects data values to include in the report. When no SELECT DATA instruction is used, all pollutants, activities, or species found in the input file(s) are reported. When species are being reported, the corresponding pollutants in the inventory or hourly emissions file are reported as well as the speciated emission values. The <label> must be a space-delimited list of pollutant, activity, or species names. Do not include the pollutant names in the list when species are listed, these will be included by default.

  To specify a species value when the species name is the same as the pollutant name, a "S" and a hyphen can be inserted before the species name. One can also use this method for specifying the sum of post-speciated emissions. For instance, to specify carbon monoxide after speciation, the SELECT instruction should be:

  SELECT DATA S-CO

  To specify the sum of VOC emissions after speciation, the SELECT instruction should be:

  SELECT DATA S-VOC

  Note that Smkreport does not determine whether the summation makes sense, and does not correct for species molecular weights. The user has the responsibility of ensuring that data requests make sense.

- ELEVATED [PING]: Selects all elevated point sources (recognized for point sources only). If no elevated sources are found (after applying all other selection criteria), the report will contain a warning to this effect. If the PING option appears, only plume-in-grid sources will be selected. When the ELEVATED instruction is used by any report, the PELV file (from Elevpoint) or PLAY file (from Laypoint) will be required to determine which sources or elevated or not. The latter will be used when the LAYFRAC instruction is also included in the report instructions. If the PING option is used, the PELV file will be required in all cases, although the LAYFRAC instruction can still be used to set the elevated source status based on the layer fractions.

- NOELEVATED: Selects all low-level point sources (recognized for point sources only). If both the ELEVATED and NOELEVATED commands appear in the same report, a warning will be given, and the latter setting will be used to generate the report. If no low-level sources are found (after applying all other selection criteria), the report will contain a warning to this effect. The same rules apply for use of the PELV or PLAY files as described for the ELEVATED option.

- REGION <label>: Selects a region using the country/state/county code or a region group. For this option, <label> must be either a country/state/county code in NSSCCC format, or a region group label from the Define-group section of the file.

- SUBGRID <label>: Selects a subgrid using the grid cell numbers or a subgrid group

label. For this option, <label> must by either a range of grid cells, or a subgrid label from the Define-group section of the file.

When the <label> defines the cell range, the same syntax as specified as above for the /DEFINE SUBGRID/ packet should be used. However, the INCLUDE setting is assumed, and Smkreport will not understand the INCLUDE and EXCLUDE commands as part of the <label>.

For the REGION and SUBGRID options, the value of <label> must either match a group or subgrid label or match characteristics of the inventory or grid. Also, the <label> for the REGION and SUBGRID options must be 200 characters or less.

Future versions of Smkreport will support the following additional options:

- SCC <label>: Select a single SCC or SCC group for reporting.
- PLANT <label>: Select a single point source plant, or a plant group for reporting.
- STACK <label>: Select a single point source stack, or a stack group for reporting.
- POINT <label>: Select a single point source point, or a point group for reporting.
- HOURS <label>: Select a single hour or hour group for reporting.
- SOURCE <label>: Select a single source or source group for reporting.

SPECIATION: The SPECIATION instruction controls whether or not the speciation matrix is going to be used by the program. This option controls the columns of data that will be printed in the output file. When the SPECIATION instruction is in the report instructions, the columns in the output file will include speciated emissions values as well as inventory emissions from the SMOKE inventory or hourly emissions file (depending on the use of the TEMPORAL instruction or not). When the SPECIATION instruction is not used, the columns in the output file will be inventory emissions and/or activities.

The syntax of the SPECIATION instruction is:

SPECIATION <option>

where the following options are supported

- MOLE: Uses the mole-based speciation matrix.
- MASS: Uses the mass-based speciation matrix.

TEMPORAL: The TEMPORAL instruction controls whether or not emission values are used from the hourly emission files ATMP, MTMP, or PTMP. When this instruction is used, the

default behavior of Smkreport is to sum up the emissions by day and report emissions as daily totals. To override this behavior, the BY HOUR instruction must be used.

When dates and times are given in the report, the hours are reported in the time zone from the ATMP, MTMP, or PTMP file. The time zone is listed as the *standard* time zone for the time zone stored in the header of the hourly emissions file; therefore, hours are not ever reported as *daylight* time values.

The TEMPORAL instruction is assumed when either of the following options appear in the report instructions:

```
SELECT HOURS,
BY HOUR.
```

In the future, it may be possible to enhance this instruction to permit the reported hour values to labeled by a different time zone than is in the hourly emissions file.

TITLE: The TITLE instruction allows one or more titles to be added to a report. The titles are printed into the report in the order in which they appear in the packet. Quotes are **not** needed in titles with blank spaces. The title will be ended at the last character on the line, or at a double pound sign (##), whichever appears first on the line. Double pound signs cannot be printed in the title of reports. It is recommended that titles be 80 characters wide. The maximum title is 300 characters wide. The syntax of the TITLE instruction is the following:

```
TITLE <title text>
```

UNITS: The UNITS instruction allows control of the units of the output data. The current version of Smkreport does not permit units control by data value, but rather the same units must be used for all data values, or the default units will be used. Future versions of Smkreport will permit control of the units by data variable. All unit conversions that *can* be done, *will* be done. For example, if the inventory is stored in tons/hour, and the output is requested in fathoms/day, the hour-to-day conversion will be made, while the impossible tons-to-fathoms conversion will not be made. The syntax of the UNITS instruction is the following:

```
UNITS ALL <units>
```

Valid values for the <units> option numerator are the following:

```
ton(s)
g, gm(s), gram(s)
kg(s)
mole(s), gm mole(s)
```

Valid values for the <units> option denominator are the following:

year(s), yr(s)
day
hour(s), hr(s)
min(s)
day(s)
s, sec(s)

No quotes are needed around the unit labels. Unit labels must be 30 characters or less. Some examples of the units instruction are the following:

```
UNITS ALL gm moles/s
UNITS ALL tons/hr
UNITS ALL kg/year
```

The /DELIMITER/ packet

The **/DELIMITER/** packet can be used to change the column delimiter used in the output reports from Smkreport. The default delimiter is a semi-colon (";"). Users can change the delimiter to any single-character value to suit their needs. Please note that using a space as a delimiter may result in slightly misformatted reports. Some reasons to change the delimiter from the default are:

- You are reporting "BY SOURCE" and including the source name, and it has semicolons included in one or more of the names. This would prevent those records in the output report from being read into another program using semicolon as the delimeter.
- Software you want to use to read the Smkreport output report does not support semi-colons for delimeters.
- You prefer a different appearance in the reports, and you can select another delimiter that will not conflict with other fields in your report.

The syntax of this instruction is the following:

```
/DELIMITER/ ,
```

This example would result in a comma (",") being used as the delimiter for columns of any results from /CREATE REPORT/ instructions that followed this instruction in the REPCONFIG file.

The /REPORT TIME/ packet

The **/REPORT TIME/** packet can be used to change the hour value for which daily total hourly emissions are output. The syntax of this instruction is the following:

```
/REPORT TIME/ <HHMMSS>
```

The setting uses HHMMSS format, for which HH is the hour and MMSS is 0000. The time specified is in the time zone of the input hourly data file. This setting is ignored when hourly data are not input to Smkreport. The default behavior of Smkreport is to report at the last hour of the day, hour 23 (hours are counted from 0 to 23).

8.10.8.5 The /O3SEASON/ packet

The **/O3SEASON/** packet can be used to toggle on or off the use of the ozone-season emissions from the SMOKE inventory file. The syntax of the command is the following:

```
/O3SEASON/ <option>
```

Where the <option> field can be set to Y or N. The default behavior of the program (if no such packets appears) is to *not* use the ozone-season emission in the reports. This packet *only* affects the program when the inventory emission are being reported. If the TEMPORAL instruction is given, the output report will contain whatever emission values were used to create the hourly emissions file.

8.10.8.6 The /NEWFILE/ packet

The **/NEWFILE/** packet can be used before or between any /CREATE REPORT/ packets to cause Smkreport to write out the subsequent reports to a different report file. The syntax of the command is the following:

```
/NEWFILE/ <Filename>
```

Where the placeholder <Filename> is replaced by either:

1. A physical file name, including the path.
2. The keyword LOGICAL, followed by a logical file name.

If no information is given for <Filename> the output files will be named by their automatic logical file names, REPORT[#], where [#] is replaced by the number of the report file in order of its associated /NEWFILE/ packet.

Future versions of Smkreport will be able to parse strings with environment variables if they are used in the file names, but the current version cannot.