

Preliminary Draft

DOCUMENTATION OF SPECIATION PREPROCESSOR PROGRAMS FOR SMOKE

June 23, 2005

William P. L. Carter

College of Engineering Center for Environmental Research and Technology (CE-CERT)

University of California, Riverside, CA 92521

Phone: (909) 781-5797

Email: carter@cert.ucr.edu

Overview

This document describes the programs that were prepared to integrate the speciation database with SMOKE and potentially other emissions processing systems. As presently formulated, these consist of four programs: the Profile Preprocessor (ProfPro), the Speciation Preprocessor (SpecPro), the Emissions Summary Processor (EmitSum), and the Mechanism Processor (MechPro). The relationships between these programs and relevant programs and files in the SMOKE and the CMAQ systems are shown in Figure 1. The types of files that are involved are described in more detail in Table 1. An overview of these programs, which are described in more detail in the following sections, is given below.

The Profile Preprocessor (ProfPro) program converts profile composition data with chemical classes used in existing profile databases to compositions in terms of specific chemical compounds. This program resolves mixture categories into compounds using the mixture composition assignments in the speciation database. The compounds are identified in the output files using the master speciation database designations, which (in most cases) are based on the compounds' CAS numbers. This program must be run to process profile data for input to the other programs, discussed below, but need be run only once for each profile. Its input and output files are mechanism-independent.

The Speciation Processor (SpecPro) program is used to produce the chemical mechanism assignments for a selected set of profiles and a selected chemical mechanism. The mechanism assignments can be output in several formats, but for use with SMOKE the principal function is the GSPRO file that SMOKE uses to make model species assignments. The program can also optionally output a file giving the VOC to TOG conversion factors based on the fraction of compounds in the profile that are on the EPA exempt list. Note that the profile composition data used as input to this program must be given in terms individual compounds using the master speciation database designation, as produced by the ProfPro program discussed above. The assignments of model species to the compounds are taken from the defaults for the given with the speciation database, but the user has the option of over-riding these assignments for selected compounds or groups of compounds, if different or explicit model representation of these compounds is desired.

The Total Emissions Profile Summation (EmitSum) program is used to produce composition file giving distribution of chemical compounds emitted into the model domain, given the total emissions amounts for all the profiles in the domain. The input file giving total emissions of profiles is produced from the SMOKE report files using a program that needs to be written. This is used by the mechanism processor program when processing mechanisms, such as SAPRC-99, where parameters for certain model

species depend on the distribution of compounds emitted into the modeling domain. This program is not needed when processing for "fixed parameter" mechanisms such as CB4 or RADM or the fixed parameter version of SAPRC-99.

The Mechanism Processor (MechPro) program is used to prepare mechanism files (giving reactions and rate constants or kinetic parameters) when processing for adjustable SAPRC-99 or other mechanisms where the parameters for some model species depend on the distribution of compounds they represent, or to prepare mechanisms where new model species have been added to represent selected compounds. The input files indicate which new species are added to the mechanism and/or give the compound distribution used to derive parameters for adjustable mechanisms and give the mechanisms or mechanism parameters for the model species in the base mechanism and those that are added, if any. The output is a complete mechanism file specifying the reactions and kinetic parameters in the format used by the CMAQ mechanism compilation system, or optionally in the format used by the SAPRC programs. Although the output files contain all the information needed to integrate the mechanism into a model, a separate program, not yet written, is needed to convert the mechanism information into the specific files needed to compile a mechanism into CMAQ.

Parameter Files. The program parameters and options for all four of the programs discussed here are taken from a "parameter input file" named on the command line, or from a file containing the default or common parameters, called MODELING.PRM. In general, the same parameter file can be used for speciation processing by all three programs, and this is recommended to assure the processing is done in a consistent manner. Some parameters and options are used by all or more than one of the programs, while others are specific to a program. Parameters or input records that are not used or recognized by a particular program are ignored.

Parameter files are ASCII files with records using the format "PARAMETER=value(s)", where PARAMETER is the keyword for the parameter, which must be specified in all caps, and no spaces must separate the parameter keyword and its value. The parameter value can be a character string and/or number(s), depending on the type of parameter. If more than one value is specified in a parameter input, then the various values are separated by spaces. Parameter values given in the parameter input file named on the command line supersede those given in MODELING.PRM. Example parameter input files distributed with the program provide examples of how parameter input is specified. Table 3 through Table 6 list and describe the parameters used by the four programs discussed here.

Current Status. At the present time, ProfPro, SpecPro, EmitSum and MechPro have been written and preliminary debugging has been carried out using the G77 compiler running in a DOS window under Windows 2000. Preliminary documentation of these programs is given in the following sections. These programs need to be adapted for running on a Unix system and integrated with the rest of the SMOKE system, and the source code is being provided rest of the project team for this purpose. A brief discussion of the source code and the files it needs is given in the sections below.

EmitSum uses an input file that that contains total emissions of mass for each profile in order to produce a total emission profile that is needed by MechPro to process adjustable parameter mechanisms. This file is presently not output by SMOKE. It is assumed that either a program will be added to SMOKE that will produce such a file given the output to Smkreport, as indicated on Figure 1, or that Smkreport will be modified to produce such output.

In addition, at the present time there is no automated procedure to integrate new mechanisms into CMAQ. Mechanisms are implemented into CMAQ using various Fortran data statements giving the arrays and kinetic and mechanistic parameters involved. A program exists to produce some of the needed

Fortran data statements given a mechanism definition file such as one output by MechPro, but at present other data statements need to be edited by hand to input a new mechanism. A program needs to be written to automate this process before this overall system can be completely integrated into CMAQ.

Program Descriptions

Profile Preprocessor (ProfPro)

The ProfPro program converts profile composition data given in various chemical class categorizations to compositions in terms of specific chemical compounds using the master emissions database categorization. The input consists of lists of profile(s) to process, folders where the input and output files are located, and assignment files, as indicated below. This program must be run to process profile data for input to the other programs, discussed below, but need be run only once for each profile. Its input and output files are mechanism-independent.

The input profile compositions (in terms of mass fractions) are given in (profile).EMI files, where (profile) is the profile file name. The categorization used is indicated in the SARTYPE=(type) record in the .EMI files or a default categorization type is used. For each categorization used, the assignments to master speciation database classes are given in the file (type)SARID.DAT, where (type) is as indicated in the SARTYPE record. For example, EMI files that use EPA SAROAD categorizations should have a "SARTYPE=EPA" record, and the EPA SAROAD classes are assigned master emissions database classes in the EPASARID.DAT file.

The file SPECASN.DAT is used to determine the individual compounds assigned to each chemical category in the emissions database. The file CMPDINFO.DAT is used to obtain descriptions for the compounds that are included in the output files for documentation and QA purposes.

The profile compositions are output in (profile).EMC files with the same profile file name as the .EMI input file. The output gives mass fractions of actual compounds that are identified in the profile. Some of the profiles contain categories that cannot be assigned to individual compounds, in those cases one of the special codes "NOASN", "INERT", or "NONVOL" is used in place of the compound ID. These refer to mixtures of unknown composition that are presumed to be reactive, unreactive, or non-volatile, respectively.

The formats of the EMI and EMC files are similar, and are described in Appendix D.

If multiple profiles are processed, the results are summarized in the output file PROFPRO.LOG. If errors were encountered in processing any profile, this file will give the corresponding error message for the profile. If the profile was processed successfully, it will give, in a comma-separated-value format, a summary of the total mass processed, and total unassigned, nonvolatile, or EPA-exempt mass.

To use the program, use one of the three allowable command methods:

```
PROFPRO InputParmFile
PROFPRO PROF=Profile
or PROFPRO PROFLIST=ProfileListingFile
```

In the first method, InputParmFile is the name of a file that contains the parameters, file names, and file locations the program uses. The format of the input parameter file and the parameters that are applicable to ProfPro is described in the Program Parameters and Options section (see Table 3). Note that

this parameter file can also be used for other programs, and thus may include parameters or input that is not applicable to ProfPro. Parameter keywords that ProfPro does not recognize are ignored. The parameter inputs that are rarely changed or are desired to be defaults for most executions of the program discussed below can optionally be put in a separate file called MODELING.PRM. However, the parameter values specified in the file named on the command line take precedence over those given in MODELING.PRM.

The other two methods can be used if it is to be run with the specified values of the PROF or PROFLIST command. The other parameter values are taken from MODELING.PRM. The first of these is useful when processing a single profile with a standard set of parameters, and the second is useful when processing multiple profiles listed in the PROFLIST file.

The main parameters that are required by ProfPro are the PROF or PROFLIST parameters that indicate which profile or set of profiles are to be processed. The other parameters that ProfPro uses are those indicating where the input or output files are located, and the default profile database type (SARTYPE), as indicated on Table 3.

The categorization method used in the EMI files should be given in the SARTYPE parameter in the EMI file, but if not given there the categorization method is taken from the SARTYPE parameter in the parameter input file. The (sartype)SARAD.DAT file is read to get the assignments of these categories to the categorization used in the speciation database, and the compounds assigned to these categories are read from the SPECASN.DAT file. In addition, the CMPDINFO.DAT file is read to obtain information about the compounds that are also output in the profiles' EMC file. The program looks for all three of these files in the location specified by the EMITDB parameter.

Speciation Processor (SpecPro)

The SpecPro program processes specified listed profiles for a specified mechanism and outputs a file containing model species assigned to those profiles. The mechanism assignments can be output in several formats, but for use with SMOKE the main use is to output the GSPRO file that SMOKE needs to make model species assignments. It can also optionally output a file giving the VOC to TOG conversion factors based on the fraction of compounds in the profile that are on the EPA exempt list. The input includes the chemical mechanism used, lists of profiles to be used, previously prepared GSPRO files that are to be used to specify additional and mechanism-independent profiles, if applicable, and the necessary mechanism and compound assignment files.

Compositions for each profile that is processed are obtained from the .EMC file for that profile, which contains the mass fractions of each chemical compound in the profile using the designations in the master speciation database. (If profile data are available in terms of SAROAD or other classifications, the ProfPro program must be run first to produce the corresponding .EMC files.)

In addition to the profile list and speciation files, the program uses the (mech)ASN.DAT to get assignments of model species to the compounds and other compound-specific information it needs, and (mech)MS.DAT to get information about the model species used, where (mech) is the mechanism ID that is specified using the MECH parameter in the parameter input file or MODELING.PRM.. These mechanism assignments can be over-written using EXPLICIT, MODSPE, and MODREP records in the input parameter file, as described in Table 4. Briefly, the EXPLICIT input can be used if it is desired to have specified compounds represented explicitly, the MODSPE command can be used to add new model species to the mechanism, and the MODREP command can be used to assign compounds to the new or different model species than they are normally represented using this mechanism.

This program also allows for control of how unknown or unassigned mass is represented in the model. In this context "unknown" mass is volatile mass that could not be assigned to known compounds, and "unassigned" mass consists of mass of known compounds that are not assigned to any model species. In each case, as indicated on Table 4, one of three options can be specified using the UNKN_ASN or UNASN_ASN parameters in the parameter input file. These are (1) representing the unknown or unassigned mass using the distributions of known, assigned compounds in the profile (i.e., by removing this mass and renormalizing), (2) assuming the unknown or unassigned mass is inert, i.e., ignoring it, or (3) specifying profiles to represent this unknown or unassigned mass. The first of the three options is the default, though it cannot be used if all the mass in the profile is unknown or unassigned.

The program can be run in two optional modes, the GSPRO or CMPfile modes. In the GSPRO mode, which is the default if the PROFLIST or ALLPROF parameter is specified, the program outputs a GSPRO file containing the model species assignments needed by the emissions processing program. This is the mode that is used when processing for the SMOKE system as indicated on Figure 1. In this mode SpecPro also outputs a summary file in CSV format giving a list of the profiles processed, total masses, moles, unknowns and other summary information for each profile, and any applicable error messages. This file can be readily located into a spreadsheet program for obtaining profile summary data. By default this output file is named SpecPro.CSV, but alternative names can be specified in the parameter input file, as indicated on Table 4.

In the CMPfile mode, which is the default if the PROF= parameter is used to process for a single profile, the program produces a separate file for each profile giving the model species assigned to it for the specified mechanism. Although not used for emissions processing with SMOKE, these files are useful for quality assurance or other purposes. The location of these profile mechanism composition files can be specified on the PROF parameter, but by default they are located in mechanism-specific subdirectories in the directory named in the PROFCMP input parameter. The program can be run in this mode for multiple profiles specified in the PROFLIST parameter if the "SPECPRO=CMPMODE" parameter is also specified.

To use the program, use one of the command below. Optional input is indicated by brackets [].

```
SPECPRO InputParmFile
SPECPRO PROFLIST=ProfileListingFile
or SPECPRO PROF=Profile [Mech] [OutputFile]
```

In the first method, InputParmFile is the name of a file that contains the parameters, file names, and file locations the program uses. The format of the input parameter file and the parameters that are applicable to SpecPro is described in the Program Parameters and Options section (see Table 4). Since this parameter file can also be used for other programs, keywords that SpecPro does not recognize are ignored. As with ProfPro and the other programs described below, the parameter inputs that are rarely changed or are desired to be defaults for most executions of the program discussed below can optionally be specified MODELING.PRM, which is read to obtain parameters that are not specified in the InputParmFile named on the command line.

The second method is used if the program is to be run with a particular profile listing file (i.e., PROFLIST parameter) specified on the command line, with all the other parameters as specified in MODELING.PRM. This is useful if the program is to be run with different sets of profiles but otherwise with the same parameters (i.e., using the same mechanism and lumping approach.)

The third format can be used to run the program in the CMPfile mode to output a mechanism composition (.CMP) file giving the model species assigned for a specified profile and mechanism. If the mechanism is not given, then the mechanism specified by the MECH parameter in MODELING.PRM, if given, is used. The program fails if no mechanism is given in either place. The full path of the mechanism composition file can optionally be given on the command line. If not, it is named profile.CMP and is located in a subdirectory named by the mechanism ID (the MECH parameter) and the PROFCMP parameter as indicated on Table 4. MODELING.PRM is read to obtain the other parameters used by SpecPro.

Appendix A discusses assumptions made concerned the GSPRO data required by SMOKE and procedures used in this program to calculate the model species assignment data for the profiles in the GSPRO files

SpecPro gets the assignments of model species to the compounds from the (mech)ASN.DAT file, where "(mech)" refers to the mechanism code given in the MECH parameter or on the command line. Information about the model species for the mechanism is read from the (mech)MS.DAT file. The program looks for both of these files are in the location specified by the EMITDB parameter.

Total Emissions Profile Summation Program (EmitSum)

The EmitSum program is used to produce a compound composition profile (.EMC) file giving total emissions of compounds into a model domain, given total emissions of profiles into the domain and the compounds in each of the profiles. This is used by the MechPro program to derive the mechanistic parameters that best represent the distribution of emitted species, as specified in this file.

To use the program, use one of the following two command methods,

EmitSum InputParmFile
or EmitSum FILES=EmissionsSummaryInputFile EmissionsSummrayProfile ProfileListingFile

In the first method, InputParmFile is the name of a file that contains the names of the input and output files and the location of the .EMC files giving the compounds in each profile. The format of the input parameter file and the parameters that are applicable to EmitSum is described in Appendix A. Since this parameter file can also be used for other programs, parameter keywords that EmitSum does not recognize are ignored.

In the second method the names of the emissions summary input file, the emissions summary profile output file, and the profile listing file are specified in the command line, and the other parameters are read from MODELING.PRM. This is equivalent to specifying these (and only these) using the EMITSUM-IN, EMITSUM, and PROFLIST parameters, respectively, in the input parameter file.

As indicated on Table 5, the EMITSUM-IN parameter gives the input file listing the mass emissions of the profiles, and the EMITSUM parameter gives the location of the emissions summary profile that is output. These are not needed if the second command method is used since these parameters are specified in the command line. The profile listing file, specified using the PROFLIST file, must also be input in order to get the profile names associated with the profile numbers, because SMOKE identifies profiles only by the profile numbers, so this is how they are referenced in the EMITSUM-IN file.

The output emission summary profile is placed in the same location as the .EMC files for the profiles (as indicated by the PROFEMC option) unless an EMITSUM-LOC option is specified giving a

different location. The output file has the same format as the .EMC files for the individual profiles. Comments and reference information can be placed in that file using text specified by EMITSUM-DESC, EMITSUM-REF and EMITSUM-NOTE parameters.

The program also outputs a file, named EMITSUM.CSV by default, giving the names of the profiles that were processed, and the total mass and total assigned mass in each. This may be useful for diagnostic purposes.

The program gets the compounds in each profile from its .EMC file and sums up the total amounts for each compound. The output emissions total .EMC file gives the mass fractions of total emissions of each known compound, normalized to the total amounts of known compounds. The total mass fractions of unknowns are also output. Additional information about the compounds, such as their descriptions and VOC status, is also included in the output file. This is obtained from the CMPDINFO.DAT file in the location specified by the EMITDB parameter.

Mechanism Processing Program (MechPro)

The MechPro program is used to produce the mechanism definition files needed when processing using adjustable parameter mechanisms such as SAPRC-99, or when new VOC species need to be added to the mechanism to represent selected compounds differently or explicitly. To use the program, use the command:

MECHPRO InputParameterFile

where InputParameterFile is the file giving the parameters and files used by the program.

The input parameter file (which generally should be the same one that is used by SpecPro to assure consistent assignments) specifies the mechanism, relevant lumping assignments and new model species used (if applicable), and indicates where files are to be found giving the reactions, mechanism parameters, and other mechanism definition information. The output files produced by MechPro are named (file).LOG, (file).MSP, and either (file).MEC or (file).RXN, where (file) is the file name specified in the MECHPRO-OUT parameter. The .LOG file gives diagnostic output, the .MSP file lists the model species in the mechanism (including any that may be added) and their parameters such as molecular weights. The .MEC or .RXN file gives the reactions and rate information in the mechanism, either in the standard CMAQ format (MEC) or in the format used by the SAPRC modeling programs (RXN), as specified by the RXNFMT parameter. The default for RXNFMT is MEC, so this need not be specified when processing for CMAQ.

When MechPro is used for processing mechanisms that have model species that have adjustable parameters that depend on the compounds being represented, the needs an emissions summary profile input file giving a distribution of compounds to serve as a basis of the adjustment. This is specified by the EMITSUM parameter. Although this would normally be produced using the EmitSum program discussed above, the user also has the option to use other profiles representing different distributions of compounds being represented, should this be appropriate. The mechanistic parameters the adjustable parameter model species are derived from the distribution of compounds in EMITSUM and the mechanistic parameters for those compounds given in various mechanistic parameter files that are listed in the input parameter file as discussed below.

If the input parameter file indicates that new model species are to be added, MechPro can derive their mechanisms in one of three ways. If the new model species is indicated as representing a certain

compound explicitly (by the EXPLICIT input parameter) and mechanism parameters are available for the compound, then MechPro generates the compound's mechanism from these parameters. If the new model species represents a mixture of compounds, then the mechanism can be derived based on the distribution of compounds in the EMITSUM profile and their parameters if this is desired and the parameter information is available. If the mechanism being processed is a fixed parameter mechanism such as CB4, if the necessary parameters or mixture information is not available, or if use of a different set of reactions for the species is desired, then the reactions of the compound can be read in a file. The location and names of these files are as indicated below.

MechPro reads the model species assignment and model species characteristics for the default mechanism from the (mech)ASN.DAT and the (mech)MS.DAT files in the location indicated by the EMITDB parameter, in the same way this is done by SpecPro. The other mechanism dependent files used are all read from the location indicated in the MECHLOC parameter. At a minimum, this includes the (mech)BASE.MEC or the (mech)BASE.RXN file, which give the reactions and rate parameters for the default or base mechanism. If the reactions of a model species are to be given explicitly, these are expected to be in a file named (mech)(modspe).MEC or .RXN, where (modspe) is the name given to the model species. If no such file can be found for a model species not in the base mechanism, and the reactions cannot be derived using mechanistic parameters, then the MechPro .MEC or .RXN output file contains a warning that necessary mechanism information is missing.

Additional input is needed if mechanisms are to be derived using adjustable parameters. The files containing mechanism parameters for the individual compounds are named in MECHPRM records in the parameter input file. More than one MECHPRM record can be used if there are multiple parameter files, as is the case for the SAPRC-99 mechanism. These mechanism parameter files are also expected to be in the location named by the MECHLOC parameter. In addition, in order to derive mechanisms from parameters, MechPro also needs to read (mech)PMS.DAT in order to determine which product model species correspond to the various parameter codes used in the mechanistic parameter files, and (mech)PHOT.DAT in order to determine which sets of absorption cross section and quantum yield data correspond to which photolysis parameter codes.

The base mechanism reaction definition file does not have the reactions of adjustable parameter model species because these are produced by MechPro and appended to the base mechanism reaction definition file. The default mechanisms and reactions of these model species are given in separate files, named (mech)(modspe).MEC or .RXN, where (modspe) is the name of the species. Therefore, if the model is processed without parameter input, or if a EMITSUM profile is not specified for deriving adjustable parameters, these files are read to obtain the default or fixed parameter version of the mechanism.

Details concerning the parameters used when running MechPro are given in Table 6, below. The various types of files that MechPro uses are summarized in Table 1.

Note that MechPro need not be run for fixed parameter mechanisms if there are no new model species added. The files necessary to compile CMAQ for such mechanisms are already included in the distribution.

Distributed Files, Source Code and Programming Considerations

Table 7 lists the source files for the ProfPro, SpecPro, and EmitSum programs discussed above, and example input and output provided with these files for testing and debugging purposes. Also included are executable files that should run in a DOS windows under Windows 98/2000/XP on PC systems, that

were compiled and linked using a version of GNU G77 that was adapted to run in Windows systems. (See <http://www.geocities.com/Athens/Olympus/5564/> for the minimal set of files needed for G77 v0.5.25 for Win32. Since the geocities site is often unavailable, the G77 files actually used for this work were obtained from <http://web.kcuu.ac.jp/pub/win/g77/>.)

The files are distributed in a ZIP file produced using ZIP. Note that for the programs to run with the distributed parameter and input files, the distributed .EMI files and most of the distributed .EMC files must go into the PROFEMI and PROFEMC subdirectories, as indicated on Table 7. Therefore, when unzipping, be sure to use the option that will extract the files into the subdirectories that are stored in the ZIP file.

If the files are unzipped properly onto a PC system, the distributed executables should run using the provided input files. For example, with PKUNZIP the "-d" options should be used.

The following example calculations can be run using the distributed files.

```
PROFPRO S99test.S32
EMITSUM S99test.S32
SpecPro S99test.S32
SpecPro CB4test.S32
```

The file TEST.BAT is provided to run these examples.

These programs may need some modification before they can be run on Unix systems and interfaced with the SMOKE system. As far as we are aware, the source code should be compatible with the G77 on Unix and other Unix Fortran 77 compilers, but they have not yet been tested under Unix. The one change that would have to be made is in the definition of the character variable "FILDLM" in SpecPro.for, which is used to indicate subdirectory names. This is presently set at "/" for use on PC systems, and should be changed to "\" for use with Unix.

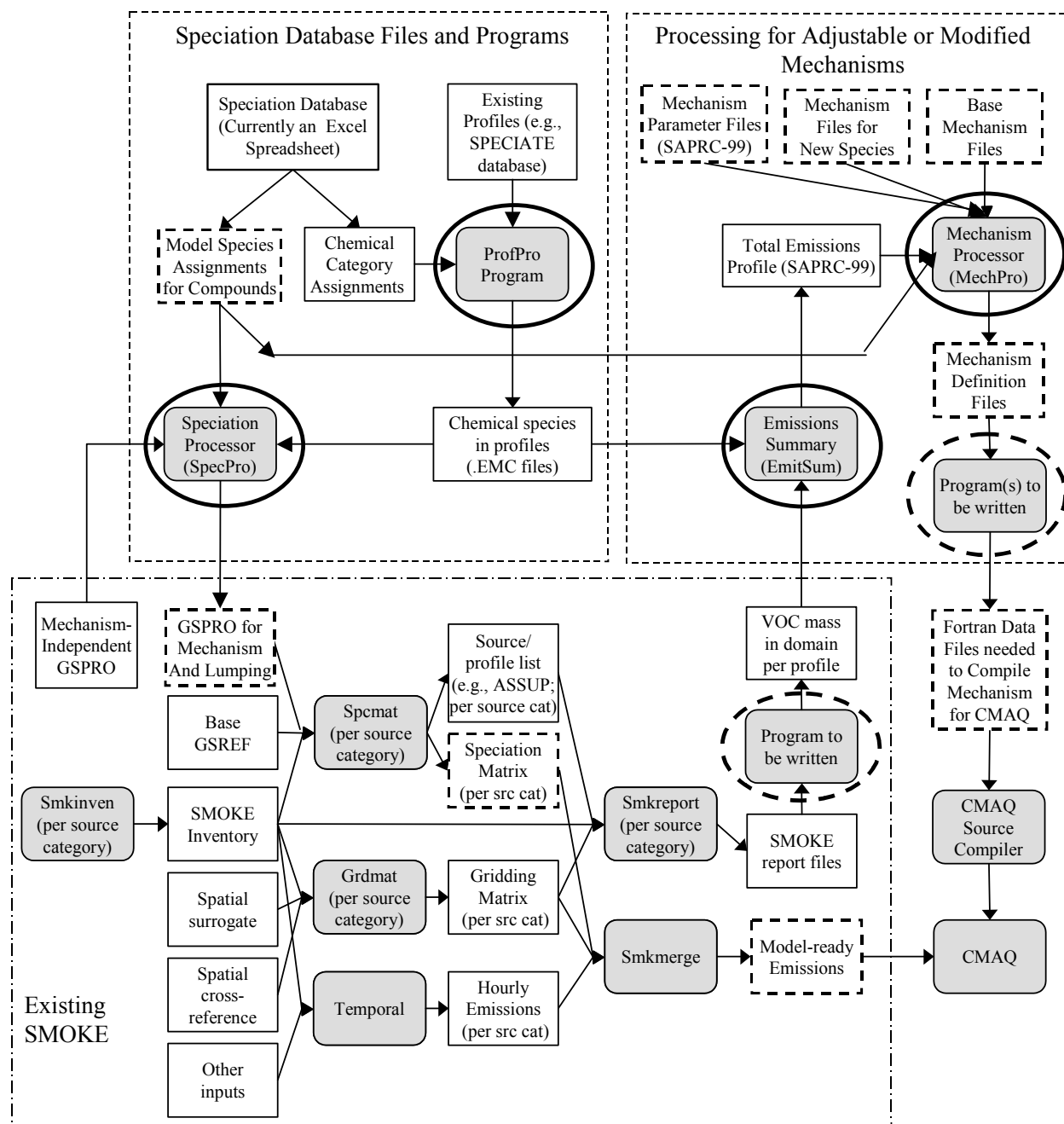


Figure 1. Diagram showing relationships between SMOKE and the speciation database programs and files. Mechanism-dependent files are indicated by dashed borders. The programs documented here are indicated with bold ovals. (The dashed ovals show programs that need to be written to complete the system.)

Table 1. Speciation assignment files produced or used by the Profile and Speciation Preprocessor programs developed for this project.

File	Contains	Produced by	Used by
Parameter Input File (named on command line)	File giving parameters controlling execution of program. See Table 3 through Table 6 for a description of the parameters used by the various programs.	User or User Interface	ProfPro, SpecPro, EmitSum, MechPro
MODELING.PRM	File giving default or common parameters controlling execution of the programs. Parameters are read from this file before reading the parameter input file named on the command line, so if parameter values given on the parameter input file on the command line supercede any conflicting values in MODELING.PRM.	User or User Interface	ProfPro, SpecPro, EmitSum, MechPro
(type)SARID.DAT [a]	Assignments of chemical categories used in EPA, California, Texas, or Speciate 3.2 database to chemical category codes used in the speciation database developed in this work. Each record contains (1) the SAROAD number or category code in the other database, and (2) the category code used in this database. [c]	EmitDB.XLS [b]	ProfPro
SPECASN.DAT	Assigns individual chemical compounds to the mixture and other chemical categories used in the speciation database. Each record contains (1) the category code for the compound or mixture, (2) the category code for the compound, and (3) the weight fraction of the compound in the category being assigned. [c,d]	EmitDB.XLS	ProfPro
COMPINFO.DAT	Gives the category code, molecular weight, carbon numbers, and description of all chemical compounds in the database. [c]	EmitDB.XLS	ProfPro, EmitSum
(mech)ASN.DAT [e]	Gives the assignments of model species for a particular mechanism to individual compounds for which assignments have been made. Also gives molecular weights and carbon numbers, to eliminate the necessity of getting this information from the COMPINFO.DAT file. Three formats are used, depending on the type of mechanism. [c,g]	EmitDB.XLS	SpecPro, MechPro
(mech)MS.DAT [e]	Gives the name, number of carbons, and molecular weight assigned to each model species in the mechanism that are assigned to VOCs in the corresponding (mech)ASN.DAT files. [c]	EmitDB.XLS	SpecPro, MechPro

Table 1 continued)

File	Contains	Produced by	Used by
(profile).EMI [h]	Profile compositions given as mass fractions for chemical categories used in the EPA, California, Texas, or Speciate 3.2 databases. The comments records include a "SARTYPE=" record that indicates the categorization system employed (EPA, CA, TX, or S32). Each record contains (1) the SAROAD number of Speciate "symbol" for a category, (2) the mass fraction in the category, and (3) the description as used in the database from which the profile was obtained. These are derived directly from the profiles as received from the various sources, with no modifications other than reformatting. [c,j]	ProfDB.XLS [f]	ProfPro
(profile).EMC [h]	Profile compositions given as mass fractions of individual chemical compounds in terms of the categorization used in the speciation database. Each record contains (1) the database code for a compound or, for unspiciated mixtures, the codes "NOASN" , "INERT" or "NONVOC" [d], (2) the weight fraction of the compound or unassigned mass, and (3) the compound description as used in the database. [c,j]	ProfPro	SpecPro, EmitSum, MechPro
PROFPRO.CSV (or file named by PROFPRO-OUT parameter)	Log summarizing the results of the execution of the Profile Preprocessor (ProfPro) program. Gives messages if any errors were encountered when profiles processes. Each record for successfully processed profiles gives the profile ID [h] and the total weight input and output (which should be the same within roundoff error) and masses of unknowns, nonvolatiles, and exempts. This is given in a comma-separate-value format so this information can be easily imported into spreadsheets.	ProfPro	User [k]
Output GSPRO file (Name specified by GSPRO parameter)	Assignments of model species to emissions profiles. See description of the SpecPro program and Appendix A. File names specified in the SpecPro "GSPRO" parameter.	SpecPro	Emissions processing programs

Table 1 continued)

File	Contains	Produced by	Used by
Input GSPRO file (name specified by GSPRO-IN parameter)	Assignments of non-VOC model species to non-TOG profiles, or default profile assignments. This is an optional input to SpecPro and if included any profiles not listed in the These are included in the output GSPRO file for profiles not otherwise processed by SpecPro so the output GSPRO file can also include profiles not dependent on VOC model species assignments.	Emissions processing database or user	SpecPro
SPECPRO.CSV (or file named by SPECPRO-OUT parameter)	Log summarizing the results of the execution of the Speciation Preprocessor (SpecPro) program. Gives messages if any errors were encountered when processing any profiles. For each successfully processed profile it gives the profile identifications, total mass input and output as model species, average molecular weight, unknown and unassigned mass, and other information as indicated in the header record in the file. This is given in a comma-separate-value format so this information can be easily imported into spreadsheets.	SpecPro	User [k]
VOC-to-TOG factor file (name specified by TOGVOC-FAC parameter)	File giving total non-EPA-exempt mass ratios for converting "VOC" input emissions masses to "TOG" masses. Consists of records containing profile number and conversion factor. No factors are given for profiles consisting only of EPA-exempt compounds.	SpecPro	Emissions processing programs
Profile listing file (named by PROFLIST parameter)	List of profiles to be processed. For ProfPro all that is needed is the profile ID that is used to identify the EMI file to process, one per record. For SpecPro, and EmitSum, the profile number should also be given on the record with the profile name (separated by a space), unless it is the same as the last 4 characters of the profile name. (The source distribution includes a file S32-PROF.LIS, which lists all the Speciate 3.2 profiles.)	Emissions processing database or User	ProfPro, SpecPro, EmitSum
Total profile emissions file (named by the EMITSUM-IN parameter)	Total mass emissions of each profile into the domain. Each record contains the profile number, the type of profile (e.g., "TOG"), and the mass emissions of the profile. Only TOG profiles and profiles included in the profile listing file are processed. (Note that it is permissible for more than one record to specify mass emissions of a given profile. In this case, the mass data in the multiple records are summed up.	SMOKE programs	EmitSum

Table 1 continued)

File	Contains	Produced by	Used by
Total emissions or representative emissions profile .EMC file (named in the EMITSUM parameter)	Distribution of compounds to be used to represent the total or typical emissions into the domain. Produced by EmitSum from the total profile emissions information. Used by MechPro to when processing for adjustable parameter mechanisms to derive the optimum parameters to represent the distribution of compounds in this profile.	User or EmitSum	MechPro
MECHPRO.MEC or .RXN [m] (or as specified using the MECHPRO-IN parameter) [n]	The mechanism definition file giving the reactions and the kinetic parameters for the mechanism as processed by MechPro. This includes the reactions of the base mechanism plus the reactions of any adjustable parameter or added model species as specified by MechPro input. Format and extension depends on RXNTYP parameter [m].	MechPro	Airshed model mechanism compiler
MECHPRO.MSP (or as specified using the MECHPRO-IN parameter) [n]	Lists all the VOC or TOG model species used in the mechanism as processed by MechPro, giving their name, carbon number, molecular weight, and other information.	MechPro	Airshed model mechanism compiler
Species table input file (named by the SPECIES-TABLE parameter)	Species table file containing deposition parameters and other information about the VOC model species in the base mechanism. This is used as input to create the species table file for the modified mechanism, depending on the input to MechPro. See documentation for the ChemMech program for the contents of this file.	Mechanism and model developers	MechPro
Species table output file (named by the SPECIES-TABLE parameter)	Species table file containing deposition parameters and other information about the VOC model species in the mechanism being processed. This is based on the information in the species table input file, described above, modified as appropriate based on the added species, if applicable.	MechPro	Airshed model mechanism compiler
MECHPRO.LOG (or as specified using the MECHPRO-IN parameter) [n]	Diagnostic output of the execution of the MechPro program, indicating which files were read and written and what new or adjustable parameter model species were processed, and giving any applicable warnings or error messages.	MechPro	User [k]

Table 1 continued)

File	Contains	Produced by	Used by
(mech)BASE.MEC or .RXN [e,m] (S99LBASE.MEC, S99LBASE.RXN, & CB4BASE.MEC are currently available)	Mechanism definition file giving the reactions and kinetic parameters for the base mechanism. This is the entire mechanism for fixed-parameter mechanisms such as CB4 or consists of the reactions of the non-adjusted-parameter model species in adjustable parameter mechanisms such as SAPRC-99. Format and extension depends on RXNTYP parameter [m].	Mechanism developer	MechPro
(mech)(modspe).MEC or .RXN [m,p]	Portion of a mechanism definition file giving the reaction(s) of a particular model species. This can either be the default reactions of a adjustable parameter model species in an adjustable parameter mechanism that is to be used when the total emissions profile is not given, or the reactions of a model species that is added to the to the mechanisms that cannot be (or are chosen not to be) derived from parameter values. The currently distributed files contain such files for each of the SAPRC-99 adjustable parameter species [q] and (as an example) for explicitly representing benzene in CB4. Format and extension depends on RXNTYP parameter [m].	Mechanism developer or user	MechPro
(parmfile).GNA (named in MECHPRM parameter input)	Mechanistic parameter files for adjustable parameter mechanisms. These given the kinetic and mechanistic parameters for the all the detailed model species for which adjustable parameter values have been derived. Currently, such files are only available for the SAPRC-99 mechanism.	Mechanism developer	MechPro
(mech)PMS.DAT (currently only S99LPMS.DAT)	This gives the product model species associated with each parameter in the mechanism parameter files. The first column gives the two-character parameter, the second gives the carbon number for the species, and the third and fourth give the model species names as used in .RXN and .MEC files [m], respectively. Currently only available for SAPRC-99.	Mechanism developer	MechPro
(mech)PHOT.DAT (currently only S99LPHOT.DAT)	This gives the absorption cross section and quantum yield sets associated with each photolysis parameter code used in mechanism parameter files for photolysis reactions. Currently only available for SAPRC-99.	Mechanism Developer	MechPro

[a] "(type)" is a code that refers to the emissions categorization or SAROAD code used. Currently supported type codes are EPA, CA, and TX, for the SAROAD codes used for compounds and mixtures in the EPA, California, and Texas speciation profiles, respectively, and S32 for the "symbol"

Table 1 continued)

categorization used for compounds or mixtures in the Speciate 3.2 database.

- [b] EmitDB.XLS is the Excel file containing the speciation database and various macros for outputting the data as ASCII files for use by other programs.
- [c] The first set of records in these files are documentation records that are treated as comments and except as indicated are ignored by programs reading them. These are terminated by a "." record.
- [d] The special codes "NOASN", "INERT", and "NONVOC" are used in the place of the compound codes for unspiciated mixtures, where INERT and NONVOC is used for mixtures judged to be inert or non-volatile, respectively, and NOASN is used for all others.
- [e] "(mech)" refers a chemical mechanism code. Currently supported mechanisms are "S99L", "S99D", "RADM", and "CB4" for the SAPRC-99 lumped and detailed mechanism and the RADM-2 and CB4 mechanism, respectively. The capability to output assignments for other mechanisms may be added later in this project.
- [f] ProfDB.XLS is the Excel file currently containing the profile data currently available for this project.
- [g] The type of mechanism is indicated by the "FORMAT=" record in the comments section of the file. Format "1" is used for mechanisms with no reactivity weighting, e.g, detailed SAPRC-99, where the records contain the compound code, molecular weight, carbon number, the model species, and the compound description. Format "2" is used for mechanism with reactivity weighting such as RADM-2 and lumped SAPRC-99, where the records contain a reactivity weighting factor after the model species, giving the number of moles of model species to use for the compound. A "D" appended to the format code for Format 1 or 2 files indicate that detailed model species assignments are also given in the file. Format "3" is used for mechanisms, such as CB4, that can use multiple model species to a compound. In this case, the first record contains the list of model species, and the subsequent records contain the compound ID, molecular weight, carbon number, and number of moles of each model species to represent a mole of the compound, followed by an assignment code (A = assigned VOC, U = unassigned VOC, N = nonvolatile) and then the compound description. In all cases, only one record is given for each compound that is assigned in this mechanism.
- [h] "(profile)" refers to refers to the designation of the profile. Profiles taken from emissions databases have designations (type)-(number), where "(type)" indicates the database from which the profile was obtained (EPA, CA, TX, or S32), and "(number)" indicates the four-digit profile number as used in the database. For example, "EPA-0004" refers to EPA profile #4, used for "External Combustion Boiler - Refinery Gas". In addition to these, there are special profiles used for defining complex mixtures used in the older speciation databases, or produced for other purposes. These designations have no restrictions other than they can be no more than 8 characters and must be legal as DOS file names.
- [j] The comments sections of .EMI and .EMC files can include the following documentation information about the profile, obtained from the database from which the profile was taken. These are indicated by a keyword followed by an "=" (without spaces), then text or a number containing the information. The keywords or types of information or data are listed on Table 2.
- [k] This file is produced for information, debugging, or QA purposes only and is not required by any programs.
- [m] File extension depends on RXNTYP parameter. The .MEC extension is used for mechanism definition files in the CMAQ format, which is the default if the RXNTYP parameter is not specified. The format of these files is given with the CMAQ documentation. The .RXN extension is used for mechanism definition files in the SAPRC format, which is used if the RXNTYP=RXN option is specified. A discussion of the format of RXN files is beyond the scope of this document.

Table 1 continued)

- [n] The MECHPRO-IN parameter can change only the first part of the file name, not the extension.
- [p] "(mech)" refers to a non-detailed mechanism such as lumped S99L or CB4, and (modspe) refers to the model species name, which should have no more than four characters. For example, the default reactions of the SAPRC-99 lumped species ARO2 in CMAQ format would be in a file named S99ARO2.MEC.
- [q] The adjustable parameter model species in SAPRC-99 are: ALK1, ALK2, ALK3, ALK4, ALK5, OLE1, OLE2, ARO1, ARO2, and TERP.

Table 2. Profile information included in the description/documentation section of the original profile composition (.EMI) and profile compound composition (.EMC) files.

Keyword	Source [a]	Type of Documentation or Information
DESC	1 [a]	Profile description from profile database.
REF	1	Reference citation(s) from profile database. May be more than one.
NOTE	1	Documentation note(s) or comment(s) from profile database
QUALITY	1	Speciate "quality" assignment, if present (for Speciate profiles).
CONTROLS	1	Speciate "controls" assignment, if present (for Speciate profiles)
DATE	1	Date assigned to profile, if given (for Speciate profiles)
SARTYPE	1	Type of speciation assignment used in .EMI file. (Only in .EMI files)
FORMAT	1,2	Format of the mass distribution data in this file, given as a FORTRAN format code. Different for .EMI and .EMC files.
SAR_ASSIGN	2	Information about categorization assignments used to when processing original native profile data (.EMI) files. This includes the dates the assignment file was produced from the EMITDB.XLS file.
CMPD_ASSIGN	2	Information about the categorization assignments used for compounds present in the mixture categorizations in the (.EMI) files. This includes the dates the assignment file was produced from the EMITDB.XLS file.
TOTWT	2	Total mass of compounds in this file. This should be close to 1 if the file is properly normalized, but small differences ($\ll 0.1\%$) may result from round off errors in files with many components.
TOTWT_INPUT	2	Total mass of compounds in input .EMI file, if it is different from the TOTWT value. These should be the same unless there are errors in the assignment files. However, small differences ($\ll 0.1\%$) may result from round off errors in files with many mixture assignments.
FRCT_UNKNOWN	2	Fraction of mass that is not assigned to any compound, excluding unassigned mass that is designated as nonvolatile. (Absent if zero.)
FRCT_UNK_NVOL	2	Fraction of mass that is not assigned to any compound but that is designated as nonvolatile. (Absent if zero)
FRCT_NONVOL	2	Fraction of mass that is designated as nonvolatile (either nonvolatile unspciated mixtures or compounds designated as nonvolatile).
FRCT_EXEMPT	2	Fraction of mass that is assigned to compounds that are on the EPA's exempt list (i.e., non-VOC's) as of the date indicated on the CMPD_ASSIGN record.

[a] Data source codes: 1 = Obtained from the profile databases and contained in both .EMI and .EMC files; 2 = Computed by ProfPro and present only in .EMC files.

Table 3. List of parameters and options that can be specified in parameter input files for the ProfPro program

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
PROF	S	Profile name (char)	File name of a profile if only a single profile is to be processed. The input profile composition file is (profile name).EMI and the output profile compound composition file is (profile name).EMC. The locations of these files are specified by the PROFEMI or PROFEMC parameters Either this or the PROFLIST parameter must be specified.
PROFLIST	S E	Profile list file name (char)	Name of a file that contains a list of profile names to be processed. Each record of this file should contain a profile name and (optionally) a profile number (the profile number is ignored by ProfPro). Blank records or records starting with a "!" are ignored. ProfPro reads the .EMI file for each profile and produces an EMC file for it.
SARTYPE		Speciation category code (char)	Default categorization (or SAROAD) type in any input .EMI files that are read in that don't have a SARTYPE record in them. If they do, the SARTYPE indicated in the file overrides the default. If no SARTYPE record is given in the parameter file, than any EMI file processed must have a SARTYPE record.
PROFEMI		File path or location (char)	Location of the profile composition .EMI files used for input to ProfPro. If not given, the .EMI files are assumed to be in the default directory for the program
PROFPRO-OUT		CSV file name (char)	ProfPro output file summarizing the results and the characteristics of the profiles that were processed. This is in CSV format so it can be loaded into spreadsheets. If not given, the summary output file is named "PROFPRO.CSV".
PROFEMC	S E M	File path or location (char)	Location of the profile compound composition .EMC files that are output by ProfPro. If not given, the .EMC files are assumed to be in the default directory for the program
EMITDB	S E M	File path or location (char)	Location of the emissions database assignment files used by the programs. These are files that reflect assignments in EmitDB.XLS. If not given they are assumed to be in the default directory for the program.

[a] Programs in addition to ProfPro that use the parameters: S=SpecPro, E= EmitSum, M=MechPro

[b] Parameter type codes: (char) = character string, (num) = floating point number, (int) integer. Options in quotes must be given exactly as shown (case sensitive) if chosen. Character parameters cannot have spaces unless indicated otherwise.

Table 4. List of parameters and options that can be specified in parameter input files for the SpecPro program.

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
PROF	P	Profile name (char) [mech, file]	File name of a profile if only a single profile is to be processed. The profile.EMC file is read to get the profile composition in terms of compounds. The locations of these files are specified by the PROFEMC parameters Either this or the PROFLIST parameter must be specified when running SpecPro. The mechanism name and the output profile composition file can optionally also be given. If the mechanism is not specified here, it must be given by a MECH input parameter. If the output file is not given, the output file is named (profile name).CMP and is located as indicated by the mechanism ID (MECH) parameter and the PROFCMP parameter as described below.
PROFLIST	P E	Profile list file name (char)	Name of a file that contains a list of profile names to be processed. Except as indicated, each record contains a profile name and (if necessary) a profile number. If no profile number is given, the number is taken from the last 4 digits of the profile name is used. (The profile number is used to identify the profile in GSPRO files.) Blank records or records starting with a "!" are ignored. The location of the EMC are specified by the PROFEMC parameter
ALLPROF		File with profile data	Name of a file containing combined profile data. Composition information for all profiles to be processed are given in this file. For each profile to be processed, the first record consists of the string ".." to indicate the start of data for a profile, the next record contains the profile ID, and the subsequent records contain the speciation database code for a compound and the weight fraction of the compound in (A16,E10.4) format. (Records prior to the first ".." record are treated as comments.) In this case, all profile compositions (as mass fractions of compounds) are given in this file, rather than from .EMC files. (The PROFEMC parameter is ignored.) This is useful when multiple profiles for which use of separate files would be cumbersome. The ProfProE program (not documented here) produces such files.
MECH	M	Mechanism code (char)	Code for the mechanism used for representing emissions in the model (4 character maximum). For example, the mechanism assignments for individual compounds used by SpecPro are expected to be in a file named (mech)ASN.DAT and the assignments for the model species are expected to be in (mech)ASN.DAT in the folder specified by the EMITDB= input parameter, where "(mech)" refers to the mechanism code.
GSPRO		GSPRO File name (char) [file format code (char)]	Name for the GSPRO output file produced by SpecPro that contains the model species assignments for the processed profiles. Optionally a parameter code can follow the name. Currently the only recognized code is "CSV" to produce the GSPRO file in a .CSV format. If no code is given, the default SMOKE format is used.

Table 4 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
GSPRO-IN		Existing GSPRO file name (char)	Optional input GSPRO file containing model species assignments for profiles that are to be merged into the output GSPRO file. Generally they contain mechanism-independent profile input, but they can contain previously-processed mechanism-dependent GSPRO files provided that they are for the same mechanism (this isn't checked). This is processed AFTER the profiles listed in the PROFLIST input file are processed. If this contains a (profile number, type) combination that was that was input previously, it is ignored. This is ignored if the program is running in CMPfile mode or if the "CSV" format is specified with the GSPRO parameter. Note: this option is not supported if the ALLPROF option is used, and is ignored in this case.
TOGVOC-FAC		VOC-to-TOG factor file name (char)	Name for an output file to contain VOC-to-TOG factors for the profiles to be processed. The default is for no such file to be output. This file contains the profile numbers and the ratios of total mass to non-EPA-exempt mass in the profiles. The assignments of which compounds are EPA exempt are read from the (mech)ASN.DAT file, which in turn were taken from EMITDB.XLS. Note that these assignments are mechanism independent.
SPECPRO=CMPMODE		none	Used to indicate that SpecPro is to be run in the CMPfile mode when processing multiple profiles. When run in this mode, the model species assignments for the profiles are output in a profile.CMP file for each profile, rather than in a single GSPRO file. The default is to run in this mode when the PROF parameter is specified, but this parameter must be specified to run in the CMPfile mode when the PROFLIST parameter is specified.
PROFEMC	P E M	File path or location (char)	Location of the profile compound composition .EMC files that are input by SpecPro. If not given, the .EMC files are assumed to be in the default directory for the program
EMITDB	P E M	File path or location (char)	Location of the emissions database assignment files used by the programs. These are files that reflect assignments in EmitDB.XLS. If not given they are assumed to be in the default directory for the program.
SPECPRO-OUT		CSV file name (char)	SpecPro output file summarizing the results and the characteristics of the profiles that were processed. This is in CSV format so it can be loaded into spreadsheets. If not given, the summary output file is named "SPECPRO.CSV".

Table 4 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
PROFCMP		File path or location	Root location for profile composition files for specific mechanisms that are output when the PROF= option or SPECPRO=CMPMODE option is used when running SpecPro. Since the contents of the profile.CMP depend on the mechanism as well as the profile, the .CMP files for each mechanism are placed in subdirectories to the PROFCMP directory that are named by the mechanism code.
UNKN-ASN		Profile name, "INERT", "NORM", or "NONE" (char)	This either gives the name of the special profile to specify a mixture of compounds to be used to represent unknown mixtures or indicates that the unknown mass is to be treated as if it were inert or as if it had the same composition as the rest of the profile. If a profile is given, the model species assigned to this profile, multiplied by the total unknown mass, are added to those for the input profiles. Note that this profile cannot itself contain unknown mass or unassigned compounds. If "INERT" is specified (all caps required) the unassigned mass is treated the same as inert unspeciaded mass. If "NORM" is specified, the unknown mass to be represented by the distribution of known, assigned compounds in the profile, i.e., their mole and mass emissions are increased by a factor of unknown to total mass. If "NONE" is specified, the unassigned mass is ignored and the total mass is not renormalized. "NONE" is the default.
UNASN-ASN		Profile name, "INERT", "NORM", or "NONE" (char)	This either gives the name of the special profile to specify a mixture of compounds to be used to represent unassigned compounds or indicates that the unassigned mass is to be treated as if it were inert or as if it had the same composition as the rest of the profile. The options are the same as given above for " UNKN-ASN". "NONE" is the default.
EXPLICIT	M	Compound ID (char), model species (char)	This optional input can be used to indicate that a certain compound is to be represented explicitly. "Compound ID" indicates the compound, given in terms of the speciation database ID code and "model species" indicates the model species name to be used by the model (maximum of 8 characters or the maximum number of characters recognized by the model, whichever is less). Note that the model species name must be different from that used for any model species in (mech)MS.DAT and given in any MODSPE= or other EXPLICIT= input to this program. The model species assignment for this compound over-rides that given in the (mech)ASN.DAT file.

Table 4 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
MODSPE	M	Model species (char), molecular wt (num), carbon no. (int)	This optional input can be used to add new model species to the mechanism, other than those defined using the EXPLICIT record, or to change the molecular weight or carbon number of a previously defined model species. "model species" is the model species name (8 characters maximum), "molecular wt" is the molecular weight in grams per mole, and "carbon number" is the number of carbons assigned to the model species, which must be an integer. (Note that this is not needed for explicit model species, since EXPLICIT= records are sufficient for specifying these, unless different carbon numbers or molecular weights are to be used.) The compounds represented by this model species must be specified in separate MODREP= records, as indicated below.
MODREP (for lumped molecule mechanisms) [c]	M [d]	Compound ID (char), model species (char), [weighting factor (num)]	Used to make model species assignments to over-write the defaults in the (mech)ASN.DAT file, as would be the case if the compound is represented by a new model species defined using a MODSPE= record. "Compound ID" is the compound given in terms of the speciation database ID code, "model species" indicates a model species name that is defined either in the (mech)ASN.DAT file or a previously input MODSPE= record, and the optional "weighting factor" is the number of moles of model species representing one mole of compound. If not given the default weighting factor is one model species per mole of compound. If a nonunit weighting factor is given, an appropriate number of inert model species is added so all molecules are represented.
MODREP (for lumped structure mechanisms) [e]		Compound ID (char), model species 1 (char), weighting factor 1 (num), model species 2, factor 2, etc.	Used to make model species assignments to over-write the defaults in the (mech)ASN.DAT file. "Compound ID" is the compound given in terms of the speciation database ID code, and pairs of "model species" and "weighting factor" input are given for each model species used to represent the compound. In this case, the weighting factor is required and maximum number of allowed model species is 12, or whatever can be defined in an 80-character record, whichever is less. An appropriate number of inert model species is added so all carbons are represented, and if any inert model species factors are given they are ignored.
CMPOPT		Option code "C" or "M"	Code to indicate type of normalization (if any) to be done if CMP files are to be output (see PROF and SPECPRO=CMPMODE parameters). "C" = normalize to give 1 mole model species carbons, "M" normalize to give 1 gram output". No normalization is done if this option not given or any other character is specified.

[a] Other programs that use the parameters: P=ProfPro, E = EmitSum or M=MechPro.

[b] Parameter type codes: (char) = character string, (num) = floating point number, (int) integer. Options

Table 4 (continued)

in quotes must be given exactly as shown (case sensitive) if chosen. Character parameters cannot have spaces unless indicated otherwise. If more than one parameter specified they must be separated by a space. Brackets ("[...]") indicate optional parameters.

- [c] Parameters for MODREP depend on the type of mechanism being processed. Parameters and description given here are applicable to "lumped molecule" mechanisms that generally use one model species per mole of compound represented, and whose (mech)ASN.DAT file has the FORMAT=1 or 2 option. The SAPRC and RADM mechanisms are lumped molecule mechanisms.
- [d] Only used by MechPro for adjustable parameter model species when parameters for this species are to be derived based on the composition in the EMITSUM file.
- [e] Parameters for MODREP depend on the type of mechanism being processed. Parameters and description given here are applicable to "lumped structure" mechanisms that generally use several model species to represent different parts of the molecule, and whose (mech)ASN.DAT file has the FORMAT=3 option. The Carbon Bond mechanisms are lumped molecule mechanisms.

Table 5. List of parameters and options that can be specified in parameter input files for the EmitSum program.

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
PROFLIST	P S	Profile list file name (char)	Name of a file that contains a list of profile names and profile numbers. EmitSum uses this to associate profile names from the profile numbers given in the emissions total output file produced by SMOKE. If the profile number is not given it is assumed to be the last 4 characters of the file name.
EMITSUM-IN		Emissions summary file (char)	Name of a file containing the total emissions of profiles into the domain. Each record in the file contains a profile number or code, the profile type, and the total mass emissions of the profile. Only profiles with type "TOG" and whose numbers are listed in the PROFLIST file.
EMITSUM	E	Emissions summary .EMC file (char)	Name of a profile composition (.EMC) file containing the total mass emissions of compounds into a domain as computed by EmitSum, or any profile that is to be used to represent total mass emissions by the MechPro program. The file is located in the location specified either by the EMITSUM-LOC option or, if that option is not specified, by the location specified by the PROFEMC option.
EMITSUM-LOC	E M	Location of emissions summary .EMC file (char)	Location of the emissions compositions summary .EMC file if it is to be different from that specified by the PROFEMC option.
EMITSUM-OUT		Summary output file (char)	Name of summary output file if not EMITSUM.CSV. This gives the names of the profiles that were processed, and the total mass and total assigned mass in each, in CSV format.
PROFEMC	P S M	File path or location (char)	Location of the profile compound composition .EMC files that are input to get profile compositions, and also the location of the output emissions summary profile if the EMITSUM-LOC parameter is not specified. If not given, the .EMC files are assumed to be in the default directory for the program
EMITDB	P S M	File path or location (char)	Location of the CMPDINFO.DAT file that EmitSum reads to obtain information about the compounds that is included in the output file.
EMITSUM-DESC	E	Description (text)	Label or description of the EMITSUM file when it is output. This is given as a DESC= record in the output EMC file.
EMITSUM-REF	E	Comment (text)	Text to be inserted as reference documentation in the EMITSUM file when it is created by EmitSum. Spaces are permitted. Maximum length is 250 characters. No more than 10 EMITSUM-REF plus EMITSUM-NOTE records can be processed. This is output as a REF= record in the output .EMC file.

Table 5 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
EMITSUM- NOTE	E	Comment (text)	Text to be inserted as note documentation in the EMITSUM file when it is created by EmitSum. See discussion of EMITSUM-REF. This is output as a NOTE= record in the output .EMC file.

[a] Programs that use the parameters: P=ProfPro, S=SpecPro, E= EmitSum, M=MechPro. Program code is underlined if the keyword is required input.

[b] Parameter type codes: (char) = character string, (num) = floating point number, (int) integer. Options in quotes must be given exactly as shown (case sensitive) if chosen. Character parameters cannot have spaces unless indicated otherwise. If more than one parameter specified they must be separated by a space. Brackets (" [...]") indicate optional parameters.

Table 6. List of parameters and options that can be specified in parameter input files for the MechPro program.

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
MECHPRO-OUT		Name for output files (char)	Name to give output files with the mechanism and model species definitions and also a log file summarizing the results of running MechPro. The reaction definition file is named (file).MEC or (file).RXN, depending on the RXNTYPE input, the model species listing is given in (file).MSP, and the log output is given in (file).LOG, where (file) is the MECHPRO-OUT parameter. If not given, then the output files are MECHPRO.MEC (or .RXN), MECHPRO.MSP, and MECHPRO.LOG.
MECH	S	Mechanism code (char)	Code for the mechanism used for representing emissions in the model (4 character maximum).
RXNTYPE		Reaction format (char)	Format code for reaction mechanisms to be computed for a variable parameter mechanism such as SAPRC-99. Can be "MEC" for CMAQ format (default), or "RXN" for SAPRC format
EMITSUM	E	Emissions summary .EMC file (char)	Name of a profile composition (.EMC) file containing the total mass emissions of compounds into a domain that is to be used to represent total mass emissions by the MechPro program. The file is located in the location specified either by the EMITSUM-LOC option or, if that option is not specified, by the location specified by the PROFEMC option. This is not needed if a fixed parameter mechanism is being processed, or if the reactions for all new model species are to be input explicitly.
MECHPRM		Mechanism parameter files (char)	Location of files containing parameters for detailed model species for a variable parameter mechanism such as SAPRC-99. These files are located in the directory named in the "MECHLOC" parameters. Many such files may be specified if parameters for different types of compounds are given in different files.

Table 6 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
EXPLICIT	S	Compound ID (char), and model species (char), parameter code (char), like model species (char)	This optional input can be used to indicate that a certain compound is to be represented explicitly. "Compound ID" indicates the compound, given in terms of the speciation database ID code and "model species" indicates the model species name to be used by the model (maximum of 8 characters or the maximum number of characters recognized by the model, whichever is less). If a parameter code is given, it is used to derive the reaction labels for the reaction of the species if it is derived from its mechanistic parameters. If the mechanism cannot be derived from parameters then it will read the reactions from a (mech)(modspe).MEC or .RXN file in the MECHLOC directory. (The type of file read will depend on the RXNTYP input, or will be .MEC if no such input is given.). If the "like model species" is given, then this species is given the same species table parameters (other than molecular weight) as the named model species if a species table is produced. (To specify a like model species without specifying a parameter code, use two commas.)
MODSPE	S	Model species (char), molecular wt (num), carbon no. (num), parameter. Code (char), like model species (char).	This optional input can be used to add new model species to the mechanism, other than those defined using the EXPLICIT record, or to change the molecular weight or carbon number of a previously defined model species. "model species" is the model species name (8 characters maximum), "molecular wt" is the molecular weight in grams per mole, and "carbon number" is the number of carbons assigned to the model species, which must be an integer. If a parameter code is given, then MechPro will attempt to derive a mechanism for the model species provided that the compounds it represents are indicated in separate "MODREP" records, mechanistic parameters are given for the compounds, and a compound distribution is given in an EMITSUM profile. Otherwise, MechPro will look for the reactions of the species in a (mech)(modspe).MEC or .RXN file in the MECHLOC directory. If a parameter code is NOT given, then the MechPro will not attempt to derive the mechanism from parameters and species distributions, and it will read the mechanism from the (mech)(modspe).MEC or .RXN file. .). If the "like model species" is given, then this species is given the same species table parameters (other than molecular weight) as the named model species if a species table is produced. (To specify a like model species without specifying a parameter code, use two commas.)

Table 6 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
MODREP	S [c]	Compound ID (char), model species (char), [weighting factor (num)] [e]	Used to make model species assignments for “lumped molecule” mechanisms [e], to over-write the defaults in the (mech)ASN.DAT file, as would be the case if the compound is represented by a new model species defined using a MODSPE= record. "Compound ID" is the compound given in terms of the speciation database ID code, "model species" indicates a model species name that is defined either in the (mech)ASN.DAT file or a previously input MODSPE= record, and the optional "weighting factor" is the number of moles of model species representing one mole of compound. This is used by MechPro only if the parameters for the model species are to be derived based on the compounds being represented.
MECHLOC		Location of mechanism files (char)	Location of most of the mechanism files read by MechPro. These include mechanism definition files for the base mechanism, default mechanisms for adjustable parameter species, and mechanisms for new species that are given explicitly. It contains the files specified in the MECHPRM, PRODPRM, and PHOTPRM inputs.
EMITSUM-LOC	E	Location of emissions summary .EMC file (char)	Location of the emissions compositions summary .EMC file if it is to be different from that specified by the PROFEMC option.
PROFEMC	P S E	File path or location (char)	Location of the profile compound composition .EMC files that are input to get profile compositions, and also the location of the output emissions summary profile if the EMITSUM-LOC parameter is not specified. If not given, the .EMC files are assumed to be in the default directory for the program
EMITDB	P S E	File path or location (char)	Location of the emissions database assignment files used by the programs. These are files that reflect assignments in EmitDB.XLS. If not given they are assumed to be in the default directory for the program. MechPro uses the (mech)MS.DAT and (mech)ASN.DAT files in this location.
NOBASE		(none)	If this option is given, then the reactions of the base mechanism are not included in the output mechanism definition file. This used mainly for debugging.

Table 6 (continued)

Parameter Keyword	Pgm s [a]	Parameter Values [b]	Description and Defaults.
PHOTPRM		Photolysis parameter file name (char)	Name of file where parameter designations and their corresponding photolysis file names are given for each type of photolysis mechanism. If not given, program looks for a file named (mech)PHOT.DAT. This file only needs to exist if the MECHPRM files include parameters for photolysis reactions. This file expected to be in the directory named in the "MECHLOC" parameter. Only one such file is input.
PRODPRM		Product parameter definition file (char)	Name of file that associates product yield parameter codes with model species names for products for a variable parameter mechanism such as SAPRC-99. If not given, default of (mech)PMS.DAT is used. The file expected to be located in the directory named in the "MECHLOC" parameter. Only one such file is input.
SPECIES_ TABLE		Input file name, output file name	Species table input and output file for use by the mechanism compilation program for CMAQ. The input file contains the species table information for all the model species in the mechanism, with the possible exception of those specified in EXPLICIT and MODSPE input parameters. This is copied to the output file, except that the molecular weights from the (mech)MS.DAT or calculated for the lumped species is used, and any new model species specified in EXPLICIT or MODSPE parameters are included. If the input molecular weight is negative and an average molecular weight can be calculated from the distribution in the EMITSUM file, then the calculated molecular weight is used. These files are in CSV format. See the documentation for the mechanism compilation program.

[a] Programs that use the parameters other than MechPro are as indicated: P=ProfPro, S=SpecPro, and E= EmitSum.

[b] Parameter type codes: (char) = character string, (num) = floating point number, (int) integer. Options in quotes must be given exactly as shown (case sensitive) if chosen. Character parameters cannot have spaces unless indicated otherwise. If more than one parameter specified they must be separated by a space. Brackets ("[...]") indicate optional parameters.

[c] Parameters for MODREP depend on the type of mechanism being processed. However, MechPro can process parameters only for lumped molecule mechanism, so only that format is applicable for this program.

Table 7. List of distributed source, input, and output files for the ProfPro, SpecPro, and EmitSum programs.

File Name	Description
	<u>Documentation</u>
pgmdoc.doc	Microsoft Word file with the preliminary documentation for these programs (this file)
pgmdoc.pdf	As above, but in Adobe PDF format.
	<u>Source and Executable Files</u>
PROFPRO.FOR	ProfPro source file. (Calls subroutines in STSUBS.FOR)
SPECPRO.FOR	SpecPro main source program. (Calls subroutines in RDASN, RDPRO, and STRSUBS.FOR, and uses specifications in SPECSPEC.FOR.)
RDASN.FOR	SpecPro and MechPro subroutine that reads species assignment data. (Also used by MechPro)
RDPRO.FOR	SpecPro subroutine that reads and processes profile data.
SPECSPEC.FOR	Specifications used by SpecPro and MechPro.
EMITSUM.FOR	EmitSum source file. (Calls subroutines in STSUBS.FOR)
MECHPRO.FOR	MechPro main source program. (Calls subroutines in RDASN, RXNOUT, and STRSUBS.FOR, and uses specifications in SPECSPEC.FOR and DMSSPC.FOR.)
RXNOUT.FOR	MechPro subroutine that outputs reaction definition information. Calls GETXAR.)
STSUBS.FOR	Source file for string parsing subroutines used by all programs.
BLDALL.BAT	Batch file to compile and link ProfPro, SpecPro, and EmitSum using G77.
SPECPRO.EXE	SpecPro executable. (Tested using a DOS box under Windows 2000.)
PROFPRO.EXE	ProfPro executable. (Tested using a DOS box under Windows 2000.)
EMITSUM.EXE	EmitSum executable. (Tested using a DOS box under Windows 2000.)
MECHPRO.EXE	MechPro executable. (Tested using a DOS box under Windows 2000.)
	<u>Example Input Files</u> (See Table 1 for additional descriptions of these files.)
TEST.BAT	A batch file to run examples of ProfPro and EmitSum using parameters from the command line (and in MODELING.PRM), and to run SpecPro and MechPro using parameters in the following two input files.
S99TEST.IN	An example of an input parameter file for running SpecPro, MechPro and EmitSum using the SAPRC-99 mechanism, and running ProfPro and EmitSum for any mechanism. This can be used for initial testing and debugging of these three programs. See Table 4 for a discussion of the parameters in this file.
CB4TEST.IN	An example of an input parameter file for running SpecPro and MechPro using the Carbon Bond mechanism with benzene represented explicitly. (Does not contain input for testing ProfPro or EmitSum.)

Table 7 (continued)

File Name	Description
MODELING.PRM	Default and common parameter values used when running the example calculations.
S32-PROF.LIS	List of all Speciate 3.2 profiles. Used by ProfPro, SpecPro to determine which profiles to process and by EmitSum to associate profile names with profile numbers.
SpecAsn.DAT	Gives assignments of compounds to mixtures in the speciation database. Produced by macros in the speciation database Excel file. Used by ProfPro
CmpdInfo.DAT	Gives descriptions and other information about individual compounds in the speciation database. Produced by macros in the speciation database Excel file. Used by ProfPro and EmitSum.
S32SARID.DAT	Gives speciation database category assignments to all the chemical categories used in the Speciate 3.2 database. Used by ProfPro.
EPASARID.DAT	Gives speciation database category assignments to all the chemical categories used in the EPA database. Used by ProfPro.
CASARID.DAT	Gives speciation database category assignments to all the chemical categories used in the California database. Used by ProfPro.
S99LMS.DAT	Lists model species in the SAPRC-99 mechanism and gives their carbon numbers, molecular weights, and parameter codes (where applicable). Used by SpecPro and MechPro when processing for this mechanism.
CB4MS.DAT	As above, but for the CB4 mechanism.
S99Lasn.DAT	Gives assignments of lumped and detailed model species to all compounds in the speciation database for which such assignments have been made. Gives additional information about the compounds that is also in CmpdInfo.DAT. Used by SpecPro and MechPro when processing for this mechanism.
CB4asn.DAT	As above, but for the CB4 mechanism. In this case, there is only lumped model species assignment, and a different format is used because in general a compound is represented by multiple model species.
S99spec.CSV S99Aspec.CSV	CMAQ species table for the version of the SAPRC-99 mechanism with no aerosol or aerosol species, respectively. Can be used as input to MechPro to produce a species table .CSV file with the specifications for the new model species added.
NOTOG.GSP	SAPRC-99 model species assignments for non-TOG emissions profiles or emissions profiles not included in the SpecPro input profile list (e.g., S32-PROF.LIS) that can optionally be added to the GSPRO file output by SpecPro.
S32-EMIT.DAT	An example of what EmitSum expects for a total emissions file produced by SMOKE giving total mass emissions for each profile used. This particular example has equal emissions for all profiles, so if used it will give a profile with the distribution of compounds weighing all compounds equally. Used by EmitSum.
S32-nnnn.EMI	<u>In PROFEMI subdirectory.</u> Profile speciation files from the Speciate 3.2 database using the categorization of that database.

Table 7 (continued)

File Name	Description
CA-1930.EMI	<u>In PROFEMI subdirectory.</u> An example profile speciation files from the California database using the categorization of that database.
REPUNK-1.EMC	<u>In PROFEMC subdirectory.</u> Example of a profile that might be used to represent unidentified compounds when making model species assignments of profiles with unknowns. This particular example was derived from the distribution of C>6, Mwt>120 compounds in Speciate all-profile average that was produced by EmitSum using S32-EMIT.DAT. See comments in the file for additional information on how it was derived. Used by SpecPro if referenced in the UNKN-ASN input parameter (see Table 4).
REPNOA-1.EMC	<u>In PROFEMC subdirectory.</u> Example of a profile that might be used to represent known compounds that have not been assigned model species in the mechanism being processed. This particular example is the molecule-weighted average of averages in each SAPRC-99 chemical class, using all compounds on the speciation database that are both by the SAPRC-99 detailed mechanism and the CB4 mechanism, See comments in the file for additional information on how it was derived. Used by SpecPro if specified in the UNASN-ASN option (see Table 4).
*.GNA	<u>In S99L subdirectory.</u> Mechanism parameter files for SAPRC-99 detailed model species.
S99LPMS.DAT S99LPHOT.DAT	<u>In S99L subdirectory.</u> Files associating parameters in the SAPRC-99 mechanism parameter files with product model species or photolysis data. See Table 6.
S99LBASE.MEC S99LBASE.RXN S99L*.MEC S99L*.RXN	<u>In S99L subdirectory.</u> Mechanism definition files for the SAPRC-99 base mechanism and for the default mechanisms for the adjustable mechanism model species.
CB4BASE.MEC, CB4BENZ.MEC	<u>In CB4 subdirectory.</u> Base mechanism definition file and mechanism for the benzene model species for the CB4 mechanism processing example.
	<u>Example Output Files</u> (See Table 1 for additional descriptions of these files.)
S32-nnnn.EMC	<u>In PROFEMC subdirectory.</u> Profile speciation files from the Speciate 3.2 database giving the distribution of individual compounds as specified in the master speciation database, as derived from the distribution of categories in the corresponding .EMI file and the assignments of compounds to these categories. Output by ProfPro and used as input by SpecPro and EmitSum.
PROFPRO.CSV	Summarizes the results of running ProfPro when processing multiple profiles. Gives the total, unknown, exempt, and nonvolatile mass in the profiles that were processed, and gives error messages for those that could not be processed. The distributed file includes this for all the Speciate 3.2 profiles listed in S32-PROF.LIS.
VOC2TOG.S32	Gives the VOC to TOG conversion factors for the Speciate 3.2 profiles that were output by SpecPro when processing these profiles.

Table 7 (continued)

File Name	Description
S32-S99L.GSP	Gives the GSPRO file when processing the Speciate 3.2 profiles for SAPRC-99 using the example input parameter file S99test.S32. Note that this file has examples for non-default lumping of benzene and some other compounds, and is not a standard GSPRO file for processing emissions for this mechanism. Includes the non-TOG profiles read from GSPRO.OLD.
S32-CB4.GSP	As above, but for the Carbon Bond mechanism, as produced by SpecPro when using CB4test.S32 as the input parameter file. Does not include non-TOG profiles because input for such profiles was not specified in that file.
S99-S32.CSV	Gives a summary of the results of processing the Speciate 3.2 profiles by SpecPro for the SAPRC-99 mechanism. Includes totals of all, volatile, non-exempt, unknown, and unassigned masses, the average molecular weight, and other information.
CB4-S32.CSV	As above, but for the Carbon Bond mechanism.
S32-AVG.EMC	Profile derived by EmitSum giving the distribution of individual compounds (specified by master speciation database categories) derived by EmitSum using S32-EMIT.DAT as the total mass emission of each profile. Since this file specifies equal mass emissions for each profile, this profile gives the profile-weighted average of compounds in all the profiles.
S99TEST.MEC	Mechanism definition file produced by the MechPro example for SAPRC-99
S99TEST.MSP	Model species listing file produced by the MechPro example for SAPRC-99
S99TEST.CSV	Species table output file produced by the MechPro example for SAPRC-99
S99TEST.LOG	Program summary and diagnostic output file produced by the MechPro example for SAPRC-99
CB4TEST.MEC	As above, but for the MechPro example for CB4.
CB4TEST.MSP	
CB4TEST.LOG	

Appendix A. Computation of Profile-to-Model Species Assignment Data by the SpecPro Program

Assumptions Made Concerning Treatment of Data in GSPRO Files by the Emissions System.

By default, the GSPRO file produced by SpecPro contains six columns of data. The first two represent the profile number and type of profile, respectively, the third represents the model species used in part to represent the profile, and the last 3 are the assignment data as described below. There is one record for each model species that is assigned nonzero emissions for each profile.

Smoke uses the data in the GSPRO file to compute moles of emissions of model species per gram of emissions of profile as shown in Equation (I). The data in Column 5 is intended to represent an average molecular weight of the compounds in the profile, while the data in Column 4 is intended to represent the moles of model species per mole of profile (mole split factor).

$$\text{Moles of emissions of model species} = \text{Inventory mass emissions of profile} \times \frac{\text{Value in Column 4 in GSPRO file}}{\text{Value in Column 5 in GSPRO file}} \quad (\text{I})$$

Of course, this is only valid if the profile is normalized properly and that all the mass in the profiles is volatile and therefore appropriately represented by some model species. However, some profiles (at least in the CARB inventories) are deliberately not normalized in order to correct for mass emissions errors, and some profiles (in all inventories) have non-volatile species that are actually not emitted as VOCs, and therefore not appropriately represented by any VOC model species, even the inert species. Therefore, if it is desired to have a molecular weight number in Column 5 that actually represents a molecular weight, it is more appropriate to think of the value in Column 4 as the mole split factor \times the fraction of mass in the profile that is emitted as VOCs.

The data in Column 6 is the mass of the emitted species that are represented by the model species.

Computation of Data in GSPRO Files by the SpecPro program.

The profile number and type are taken from the input record for the profile and copied to the GSPRO file without modification. The third column gives the model species that have nonzero emissions attributed to the profile. The assignment data in the last three columns are computed as described below.

SpecPro computes the following quantities when processing a given profile (all caps give the name of the variable used in the source code):

TOTMAS	The total number of grams of chemical categories in 1 gram of profile. Should be 1 but isn't exactly 1 in all cases. Not normalized because in some cases a different mass is deliberately used to correct for total mass measurement errors. (unitless)
--------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

NOVMAS	Mass of compounds or uncharacterized mixtures in the profile that are judged to be nonvolatile. Since these are presumed not to be emitted as they are not counted either in the mass or moles emitted. (unitless)
VOCMAS	Grams VOC emitted per gram profile. Given by TOTMAS – NOVMAS (unitless)
TOTMOL	The total number of moles of VOC compounds in one gram of profile. Note that moles of species judged to be non-volatile species are not included. Note that the present program represents unidentified mass and unassigned compounds using special profiles for this purpose. Therefore, all VOC mass is assigned to compounds that can be represented by model species, even if some is unknown or unassigned in the mechanism. (moles/gram)
MSMOL(i)	The total number of moles of model species (i) in one gram of the profile. This is computed as discussed below. (moles/gram)
MSMAS(i)	The portion of the VOC mass in the profile that is represented by model species (i). This is computed as discussed below. (unitless)

For lumped molecule mechanisms such as SAPRC or RADM, the MSMAS and MSMOL quantities for each model species are computed by summing up the following for each compound that is represented by model species (i), as follows:

$$\text{Contribution of compound to MSMOL(i)} = \text{Reactivity Factor (if applicable)} \times \frac{\text{Mass of compound in 1 gram of profile}}{\text{Molecular weight of compound}} \quad (\text{II})$$

$$\text{Contribution of compound to MSMAS(i)} = \text{Reactivity Factor (if applicable)} \times \text{Mass fraction of compound in profile} \quad (\text{III})$$

The reactivity factor is normally unity unless reactivity weighting is used. If reactivity weighting is used, the factor apportions the mass between the selected model species and the inert model species, with the contribution to the inert model species being computed using

$$\text{Contribution of compound to MSMAS for the inert species} = 1 - \text{Reactivity factor} \times \frac{\text{Mass of compound in 1 gram of profile}}{\text{Molecular weight of compound}} \quad (\text{II})$$

$$\text{Contribution of compound MSMAS for the inert species} = 1 - \text{reactivity factor} \times \text{Mass fraction of compound in profile} \quad (\text{III})$$

For lumped structure mechanisms such as CB4, the contributions of each compound to the MSMAS and MSMOL quantities for model species (i) are computed as follows:

$$\text{Contribution of compound to MSMOL(i)} = \text{Moles model species (i) per mole compound} \times \frac{\text{Mass of compound in 1 gram of profile}}{\text{Molecular weight of compound}} \quad (\text{IV})$$

$$\text{Contribution of compound to MSMAS(i)} = \frac{\text{Mass of compound in 1 gram profile}}{\text{Moles model species (i) per mole compound}} \times \frac{\text{Number of carbons in model species}}{\text{Number of carbons in compound}} \quad (\text{V})$$

Note that for the mass to be distributed properly, the total carbon number of the model species representing the compound must equal the carbon number of the compound. This rule should be adhered to in any model species assignments for such mechanisms. (Note that this is not necessarily the case for assignments in lumped molecule mechanisms.)

The quantities discussed above are used to compute the data in GSPRO for model species (i) for the profile being processed as follows:

$$\text{Number in Column 4 for model species (i)} = \frac{\text{Fraction of mass that is VOC}}{\text{Mole split factor}} = \text{VOCMAS} \times \frac{\text{MSMOL(i)}}{\text{TOTMOL}} \quad (\text{III})$$

$$\text{Number in Column 5 for model species (i)} = \frac{\text{Average molecular weight of VOCs in profile}}{\text{Mole split factor}} = \frac{\text{VOCMAS}}{\text{TOTMOL}} \quad (\text{IV})$$

$$\text{Number in Column 6 for model species (i)} = \text{Mass represented by this model species per gram profile emitted} = \text{MSMAS(i)} \quad (\text{IV})$$

Note that $\text{TOTMAS} = \sum_i \text{MSMAS(i)} + \text{NOVMAS}$, so if there is non-volatile mass the sum of mass given in Column 6 for a profile will be less than the total mass. It will also be different from 1 if the profile is not normalized, since the program does no normalization.

If a profile is entirely nonvolatile, the numbers in Columns 4 and 6 will be zero and the number in Column 5 will be one. This should result in no model species or volatile mass being assigned to the profile, as is appropriate.