Final Report:

Integration of the SAPRC Chemical Mechanism in the SMOKE Emissions Processor for the CMAQ/Models-3 Airshed Model

Agreement number: 1846

Prepared for:

Tina Bahadori
   American Chemistry Council
   Long Range Research Initiative Team
   1300 Wilson Boulevard
   Arlington, VA 22209

Prepared by:

Zachariah Adelman and Jeffrey Vukovich
   Carolina Environmental Program
   University of North Carolina at Chapel Hill
   Bank of America Plaza, CB# 6116
   Chapel Hill, NC 27599-6116

Dr. William P.L. Carter
   College of Engineering Center for Environmental Research and Technology
   University of California at Riverside
   Riverside, CA 92521

Contents

Document Tables

Document Figures

Introduction

Air quality simulation models are essential tools for assessing effects of emission controls on the formation of secondary pollutants such as ozone (O_3). Volatile organic compounds (VOCs) can differ significantly in their effects on O_3 formation, and control strategies aimed at reducing the O_3 formation potentials ("reactivities") of the emitted VOCs are being considered in addition to mass-based controls. The Statewide Air Pollution Research Center (SAPRC) chemical mechanism (Carter, 2000), which is well suited for VOC reactivity assessment, has been implemented in EPA's Community Multiscale Air Quality (CMAQ) (Mebust, 2003) model, but additional work is needed to allow SAPRC to be used in CMAQ for assessment of reactivity controls.

Two major challenges must be addressed before CMAQ and SAPRC can be used effectively for reactivity assessment. First, the design of current emissions speciation databases is insufficient and needs to be corrected, because the databases do not clearly distinguish among actual chemical species, complex or poorly characterized mixtures, and model species used by specific mechanisms. Second, the emission inventory data processing to support SAPRC needs to be created. Both of these challenges are addressed by this work.

We integrated the improved speciation database and the SAPRC chemical mechanism with the Sparse Matrix Operator Kernel Emissions (SMOKE) (Houyoux and Vukovich, 2000) modeling system for CMAQ. Termed the Reactivity Modeling System, we will test this SMOKE/SAPRC system to ensure support of reactivity controls and the ability to create correct emissions used in the CMAQ model. Additionally, we will run CMAQ and analyze the results in the context of simulating reactivity substitutions.

Meeting the study's objectives will result in the capability to process emission inventories to use in air quality modeling for evaluating the impact of reactivity controls. Specifically, the Reactivity Modeling System will enable changing the chemical makeup of the emissions using SMOKE and modeling the impact of this change on air quality. The specific goals and objectives of the project are as follows.

- Create the speciation database organization and assignments that appropriately represent relationships between mixtures, actual chemical compounds, and the model species in SAPRC-99 and other mechanisms.

- Develop the software and essential data to update both emissions processing methods and SMOKE for SAPRC-99 and other mechanisms using the reorganized speciation database and existing emissions datasets.

- Test SMOKE with the redesigned speciation database for SAPRC-99 and other mechanisms, and for reactivity controls using data available from one previous modeling application.

- Demonstrate application of reactivity controls to emissions using SAPRC-99 and the CMAQ modeling system.

Materials and Methods

To initiate our testing of the reactivity modeling capabilities of Models-3, CEP and CE-CERT created a suite of software for generating new VOC speciation profiles and chemical

mechanisms for SMOKE and CMAQ. Both models have specific input file formats for defining the parameters of chemical mechanisms and the reactivity of different organic pollutants. Until the completion of this work, there were no publicly available tools for generating the input files required for reactivity modeling. Using Microsoft Excel macros and stand-alone Fortran programs developed by CE-CERT for generating new speciation profiles, the CMAQ chemical mechanism compiler, and existing SMOKE programs we compiled and tested a software package, referred to herein as the Reactivity Modeling System, for performing reactivity modeling with Models-3.

The two types of VOC reactivity modeling that we tested in this work included (1) mechanism enhancements/substitutions and (2) VOC profile substitutions. In the first method we developed tools for adding new organic species or compounds, both explicit and lumped, to Models-3. The addition of new species to the models involves providing both the chemical characteristics of the pollutant to SMOKE and CMAQ and information about the gas-phase reactions that transform the pollutant. In the second method, we developed tools for creating new VOC speciation profiles for input to SMOKE and fractionally substituting these profiles into an emissions simulation to generate new emissions for input to CMAQ.

CEP designed a series of simulations for testing the capabilities of the Reactivity Modeling System to generate the necessary input files for reactivity modeling with SMOKE and CMAQ. The preliminary test of the system involved creating default CB-IV (EPA, 1990) and SAPRC-99 speciation profiles for input to SMOKE. For this test case we used the reactivity system to create SMOKE speciation profiles (GSPRO) without reactivity controls to create CMAQ-ready emissions. We included three simulations in this case to test the system for different photochemical mechanisms: (1) fixed-parameter SAPRC-99, (2) adjustable-parameter SAPRC-99, and (3) CB-IV.

In the second series of test simulations, CEP used the reactivity system to test reactivity assessments using the fixed-parameter SAPRC-99 mechanism in CMAQ. In these tests we performed different levels of speciation profile substitutions in the stationary area, nonroad mobile, on-road mobile and stationary point emissions sectors. We then ran these emissions through CMAQ and compared back to the base simulation to quality assure the results and assess the impacts of the reactivity substitutions on O_3 predictions. For one of the reactivity substitutions we replaced VOC mixtures in the base simulation with a single lumped VOC to provide a simplified illustration of the capability of the Reactivity Modeling System to perform reactivity substitutions. To check for internal consistency in the system, we also compared the results of two different techniques for performing reactivity substitutions with 100% market penetration rates. Before we discuss the design of these simulations in more detail, we will first present the components of the Reactivity Modeling System and how it can be used for performing VOC reactivity assessments with SMOKE and CMAQ.

The main parts of the Reactivity Modeling System that we developed are the Speciation Database and the assignment files it produces, the Chemical Profile and Speciation Preprocessors, ProfPro and SpecPro, respectively, various SMOKE programs, the total emissions profile summation program, EmitSum, the Mechanism Preprocessor, MechPro, the Models-3 Mechanism Compiler, ChemMech, and the CMAQ model.

CE-CERT Speciation Programs

Dr. Carter at CE-CERT developed a Speciation Database, Excel macros, and stand-alone FORTRAN programs for generating speciation profiles and chemical mechanisms from the information in the database. The current database, which is available at http://www.cert.ucr.edu/~carter/emitdb, contains chemical category assignments for speciation databases used by the U.S. EPA, California, and Texas. Output from the speciation databases are input to the programs ProfPro and SpecPro which create new chemical category assignments and then convert these assignments to the speciation profile formats required by SMOKE. When it is necessary to compile a new chemical mechanism, such as when the CMAQ simulation depends on the emissions being processed (as is the case for the adjustable-parameter version of SAPRC-99) or if alternative lumping approaches are employed, the program MechPro creates new organic chemical mechanisms for input to CMAQ. EmitSum uses an input file 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 produced by the SMOKE program Smkreport as indicated in Figure 1. This section presents detailed information about the speciation programs developed at CE-CERT for reactivity modeling.

Overview

This section 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 A-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 Processer (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 given in terms individual compounds using the master speciation database designation, as produced by the ProfPro program discussed above. The
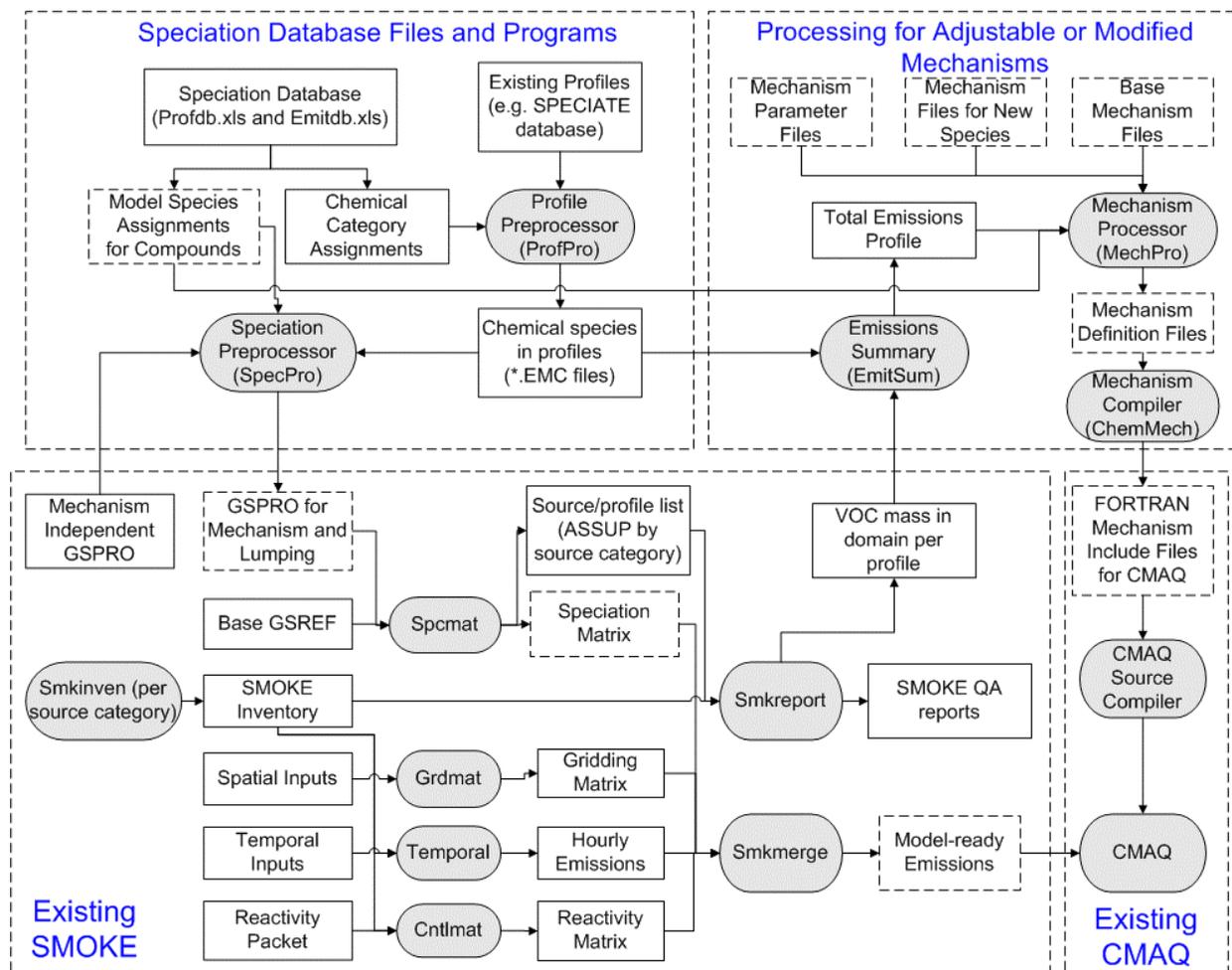
Figure 1. Processing to model a new chemical mechanism

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 from MechPro contain all the information needed to integrate the mechanism into a model, a separate program is needed to convert the mechanism information into the specific files needed to compile a mechanism into CMAQ. This function is performed by the Mechanism Compiler (ChemMech).

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 A-3 through Table A-6 list and describe the parameters used by the four programs discussed here.

Current Status. UNC-CEP ported all of the CE-CERT programs from DOS Fortran77 to comply with SMOKE coding guidelines, which included converting to Fortran90 and compiling in Linux OS, with the exception of the Mechanism Compiler. We will package the programs with a Fortran Makefile configured for compiling with the Portland Group Fortran90 compiler. A brief discussion of the source code and the files it needs is given in the sections below.

Speciation Database

The Speciation Database developed by Dr. Carter is distributed as two Excel spreadsheets: *profdb.xls* and *emitdb.xls*. In its current form, the Speciation Database includes information from the EPA, SPECIATE 3.2 (EPA, 2002), California, and Texas speciation databases about manmade VOC chemical substances that can be emitted to the atmosphere, how they are classified in each of the four databases, and how they are represented in the SAPRC-99, RADM2, and CB-IV photochemical mechanisms. The databases and associated documentation are available at http://www.cert.ucr.edu/~carter/emitdb. In terms of the relationship of the

Speciation Database to the programs and data files developed for reactivity modeling, its main features are as follows:

- **Chemical Category Assignments**. The Speciation Database gives the assignments of actual chemical compounds to the chemical categories used in the emissions profiles in the supported emissions database. Current VOC emissions speciation databases use categories that can refer to actual compounds and to mixtures of varying degrees of quality of characterization. In the latter case, the database includes assumptions about which specific compounds are contained in the mixtures. Assigning these categories to specific compounds is necessary to determine how they are represented in the model, because model species assignments are only made to actual chemical compounds.

- **Model Species Assignments**. The Speciation Database gives the assignments of model species in various mechanisms to the actual chemical compounds included in the database. The currently supported mechanisms are SAPRC-99 (both lumped and detailed[1]), RADM2, and CB-IV. Note that for the purpose of this database, the lumped and detailed SAPRC-99 mechanisms are treated as two separate mechanisms, with distinct sets of VOC model species. Although SMOKE only uses the lumped assignments when processing emissions for SAPRC-99 (except for species selected to be represented explicitly), the detailed mechanism assignments are used to determine components of the mechanism that is used by CMAQ when the adjustable parameter version of SAPRC-99 is used, or when species are chosen to be represented explicitly.

Macros in the Speciation Database spreadsheets process the information in a selected database for a selected chemical mechanism and output ASCII chemical category assignment files for each supported database (EPA, SPECIATE 3.2, California, Texas) and model species assignment files for each supported mechanism (SAPRC-99, CB-IV). The files created by the two Speciation Database spreadsheets are listed in Table 1.

---

[1] Lumped SAPRC-99 assignments designate the lumped model species that are normally used to represent the compounds in the model, unless an alternative lumping approach is employed. An example is ARO2, which represents all aromatics in the high rate constant range, such as xylenes. Detailed SAPRC-99 assignments designate the detailed model species used to represent individual compounds or groups of isomers considered to have very similar reactivities. Examples of the latter are M-XYLENE and C12-BEN2, which refer to m-xylene and 12-carbon disubstituted benzenes. Both of these examples are ARO2 in normal model applications, but these can be represented explicitly if so specified in the lumping options, and in the adjustable parameter version of SAPRC-99 the rate constant and product yield parameters used for ARO2 depend on how much M-XYLENE, C12-BEN2 and other reactive aromatic species are in the emissions they represent.

Table 1. Speciation Database output files

Spreadsheet	Output File	Description
emitdb.xls	{type}SARID.DAT	Assignments of chemical categories used in EPA, California, Texas, or Speciate 3.2 database to chemical category codes used in the Speciation Database. Each record contains the SAROAD number or category code in the other database and the category code used in the Speciation Database. {type} = CA, EPA, S32, TX, or DMS
	SPECASN.DAT	Assigns individual chemical compounds to the mixture and other chemical categories used in the Speciation Database. Each record contains the category code for the compound or mixture, the category code for the compound, and the weight fraction of the compound in the category being assigned
	CMPDINFO.DAT	Gives the category code, molecular weight, carbon number, and description of all chemical compounds in the Speciation Database.
	{mech}MS.DAT	Gives the name, carbon number, and molecular weight assigned to each model species in the mechanism that are assigned to VOC's in the corresponding {mech}ASN.DAT files. {mech} = CB4, RADM, or S99
	{mech}ASN.DAT	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. Three formats are used, depending on the type of mechanism {mech} = CB4, RADM, or S99
profdb.xls	{profile}.EMI	Profile compositions given as mass fractions for chemical categories used in the EPA, California, Texas, or Speciate 3.2 databases. Each record contains a header describing the file, SAROAD number of Speciate symbol for a category, mass fraction in the category, and the description as used in the database from which the profile was obtained.

Each profile composition file (*.EMI) output by *profdb.xls* includes a header that describes the profile, provides a reference for the source of the profile data, lists information about the profile contained in the database from which it was taken and the format of the file, and identifies the format of the speciation codes. One profile composition file is output for every profile in the selected database. The SPECIATE 3.2 database, for example, contains 567 unique profiles thus requiring 567 *.EMI files that describe each profile in the database. Only one of each of the other files output by the Speciation Database is required for reactivity modeling, depending on the mechanism and database being modeled. All of the files that are output by the Excel macros are formatted for input to the Fortran programs ProfPro, SpecPro, and MechPro.

Profile Preprocessor (ProfPro)

The Profile Preprocessor, ProfPro, is used to convert emissions profiles using various chemical characterizations, such as those used in the current speciation databases, into profiles giving the actual chemical species for a particular mechanism assignment. 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.

Input to ProfPro includes the following:

- **Detailed Profiles**. Described above as profile composition files (*.EMI), these files contain chemical speciation profiles and the associated mass fraction of their constituent chemical substances, using the chemical categorizations employed in existing databases, for example, from the SPECIATE 3.2 database. Note that these categorizations can refer to mixtures as well as single chemical compounds. More than one detailed profile file could be used as input into this program. Since different profile databases use different or mutually inconsistent chemical component codes, the specification of each detailed profile input file will include codes or other means to indicate the chemical categorization system employed. The supported categorizations are initially SPECIATE, EPA SAROAD, California SAROAD, Texas SAROAD[2], CAS numbers (for single compounds only) or Speciation Database[3] categories. The currently supported EPA, California and Texas databases use 5-digit numeric SAROAD or SAROAD-like codes.

- **Chemical Category Assignment Files**. A combination of the {type}SARID.DAT and SPECASN.DAT give the set of individual chemical compounds in the speciation database that are assigned to the chemical categories used in the detailed profile input files specified in terms of the categorization codes used in the Speciation Database. For example, the SAROAD number 45102, which refers to "isomers of xylene" in EPA and other profile databases, are assigned to 25% each of the three xylene isomers and

---

[2] Although the SAROAD codes used in the EPA, California, and Texas speciation profiles are the same for many chemical classes, there are cases where the same SAROAD number refers to different chemicals in the different databases, and some codes are used in some databases but not others. Therefore, these are treated as separate categorizations in the speciation database and for the purpose of this program.

[3] Note that currently the Speciation Database designations are based on CAS numbers in cases where these are known and well defined.

ethylbenzene, so the Speciation Database codes for these four compound will be associated with this SAROAD number, with mass weighting fractions of 25% each.

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 A.

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.

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 categorizes 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 are in the location specified by the EMITDB parameter.

The output from ProfPro is a file giving the profiles in terms of mass fractions of actual chemicals using the designations in the Speciation Database for which mechanism assignments are (or can be) made. The format of these *.EMC files is similar to the *.EMI file with the addition of fields indicating the volatility of the chemicals. Like the *.EMI files, one *.EMC file is output for every profile in the selected database. The primary function of ProfPro is as a preprocessor for the other CE-CERT reactivity processors.

Speciation Processor (SpecPro)

The Speciation Processor, SpecPro, is used to prepare the SMOKE speciation profile (GSPRO) files that contain assignments of model species to emissions profiles. GSPRO files are needed by SMOKE for allocating inventory pollutants, reported as total VOC, to model species based on one of the parameterized chemical mechanisms. SpecPro 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.

Input to SpecPro includes the following:

- **Choice of Chemical Mechanism**. The user must specify which of the available chemical mechanisms will be used to create the speciation profiles. The choices are the lumped SAPRC-99, detailed SAPRC-99, adjustable SAPRC-99 (which will cause the program to be run for both lumped and detailed species), RADM2, and CB-IV. Other mechanisms can be added if assignments for them are added to the speciation database.

- **Explicit Species Treatments**. The user can select individual chemical compounds that are not to be lumped in the normal manner, but instead are to be treated either explicitly or by another model species chosen by the user. This may be useful for reactivity assessments, toxics assessments, or assessing incremental impacts of individual compounds.

- **Model Species Assignments to Individual Compounds**. This information will be contained in the {mech}ASN.DAT file produced by the Speciation Database. The species assignment file gives the assignments of model species for a particular mechanism to individual compounds for which assignments have been made.

- **Detailed Chemical Profile Assignments**. The *.EMC files produced by ProfPro give the individual chemical compounds contained in each of the speciation profiles that SMOKE assigns to the emission inventory. The compound specifications must be given in terms of the compound codes used in the Speciation Database since these are the codes used in the model species assignment files. These data are produced from existing profiles by ProfPro, but compositions of new profiles can also be input by the user for reactivity control calculations. SpecPro can take profile input from more than one file (e.g., files containing the default profiles and one or more files containing user-defined profiles) to permit user-created input for new or modified profiles. Profiles with the same code will over-write definitions for those in files previously input, so the files with the default profiles should be input before user-defined profile files.

- **Base GSPRO file**. This is a SMOKE GSPRO file that includes profiles for inventory pollutants and species that are not otherwise processed by SpecPro, such as non-VOC or non-TOG profiles (e.g. SO_2, $PM_{2.5}$, and Hg). SpecPro will pass all profiles in the base GSPRO file that not included in the detailed profiles file to the output GSPRO file, including VOC and TOG profiles. This will allow users to change only some of the VOC or TOG profiles or add new profiles – if the profiles are not included in the detailed profiles file, any existing profiles will not be altered. However, we recommend that this file not contain VOC or TOG profiles in case some profiles with old assignments that are inconsistent with the current lumping are passed through.

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 A-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 in Table A-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 in Figure 2. 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 indicted in Table A-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.

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.

The output from SpecPro will be the GSPRO file containing the model species assignments to the emission profiles in the same format as currently used by SMOKE.

Total Emissions Profile Summation Program (EmitSum)

The EmitSum program is used to produce a 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,

(1) EmitSum InputParmFile

(2) 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 A-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 Processor (MechPro)

The Mechanism Preprocessor, MechPro, needs to be run if the mechanisms for model species used in the CMAQ simulation depend on emissions being processed (as is the case for the adjustable-parameter version of SAPRC-99) or if new species are added to the mechanism. In these cases, it is necessary to compile a new chemical mechanism to be linked into the CMAQ program, and MechPro must be run to produce the files needed by the CMAQ chemical mechanism compiler, ChemMech. The mechanism files produced by MechPro specify the rate constants and other mechanistic parameters for (1) model species whose parameters depend on the emissions they represent, for (2) new or explicit model species as specified by the user in the lumping options input and for (3) the rest of the model species in the base mechanism, whose parameters are copied to the output files. This program need not be run if CB-IV, RADM2 or the fixed-parameter version of SAPRC-99 is used and no new lumping options are specified, since the pre-compiled versions of CMAQ with the default chemical mechanisms can be used in these cases. We ran MechPro to test the capabilities of the Reactivity Modeling System to simulate an adjustable-parameter version of SAPRC-99.

Input to MechPro includes the following:

- **Choice of Chemical Mechanism.** The user must specify which of the available chemical mechanisms to use. The choices are the fixed-parameter SAPRC-99, adjustable-parameter SAPRC-99, and CB-IV. (MechPro input is not currently available for RADM2.)

- **Explicit Species Treatments.** Supported options consist of lists of individual compounds and the model species representing them. The model species can be either model species in the default mechanism, detailed model species if SAPRC-99 (either fixed or adjustable parameter) is employed, or new model species whose mechanisms will need to be defined as separate input into this program (see below). This input must be the same as the input to SpecPro that is used when preparing the SMOKE GSPRO file, and for normal applications the same file will be used as input to both programs.

- **Detailed Mechanism Parameter Files (SAPRC-99 only)**. These files contain the mechanism parameters (rate constants and product yields for all reactions) for each of the SAPRC-99 detailed model species that are not explicitly included in the base mechanism. This information is needed to derive mechanism parameters for adjustable parameter model species or to derive explicit mechanisms for selected detailed model species. Different SAPRC-99 mechanistic parameter files are used for different types of species

depending on how they are lumped, as specified by the lumping control file, discussed below. These files exist as part of the standard SAPRC-99 distribution.

- **Detailed Mechanism Files for New Model Species (any mechanism)**. If the lumping options specify use of new model species for mechanisms other than SAPRC-99, or if a species needs to be added to SAPRC-99 that is not already a detailed model species, its reactions and mechanistic parameters need to be added explicitly. If the new model species is photoreactive, photolysis rate information must also be given. The format of these files is based on that used for the existing Models-3 mechanism compiler.

- **Base Mechanism Files**. These files contain the specifications for the portion of the mechanism that is not affected by VOC lumping or adding new VOC species. These files are in the same format used by the Models-3 mechanism compiler. The existing mechanism definition files distributed with CMAQ are suitable for the fixed parameter SAPRC-99 and CB-IV. For the adjustable-parameter SAPRC-99 these files will contain the reactions of all species except those whose parameters are adjusted. The contents of these files will be copied into the output files with the reactions and parameters for the new or adjusted species added.

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 (file).MEC[4], 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 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

---

[4] Optionally, the mechanism files can be input and output in the RXN format which is used in the CE-CERT/SAPRC modeling programs, in which case they are named (file).RXN. However, since the RXN format is not currently recognized by the CMAQ mechanism compiler, that is not discussed here. See Tables A-1 and A-6.

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 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, 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 output file contains a warning that necessary mechanism information is missing.

Additional input is needed if mechanisms are to be derived using adjustable parameters or new species are added to the mechanism. 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, 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 A-6, below. The various types of files that MechPro uses are summarized in Table A-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.

The outputs from MechPro are mechanism definition files that can be input to the CMAQ Chemical Mechanism Compiler. The mechanism definition is an ASCII file consisting of a reaction listing for a particular chemical mechanism that lists the species employed, the reactions and rate constants for all reactions, and information needed to calculate rates of photolysis reactions.

CMAQ Chemical Mechanism Compiler (ChemMech)

For modeling with the adjustable-parameter version of the SAPRC-99 mechanism, which uses alternative VOC lumping schemes and the addition of explicit reaction pathways, the Reactivity Modeling System requires a tool that automates the implementation of new photochemical mechanisms into CMAQ. Photochemical reaction mechanisms are invoked in CMAQ through Fortran include files that specify (1) the parameters of the mechanism, such as the number of reactions, the types of reactions, reactants, and products, (2) mechanistic and kinetic details, such as reactant and product stoichiometries and reaction rate constants, and (3) species include files that define the physical and chemical process that affect the different model species. As these include files are in a machine-readable format, a conversion program is required to translate a human-readable mechanism listing into the format required by CMAQ. The CMAQ Chemical Mechanism Compiler (ChemMech) reads an ASCII mechanism definition file and translates the file to the CMAQ include files RXDT.EXT and RXCM.EXT.

Commonly named mech.def, the mechanism listing file can either originate as a default file distributed with CMAQ that can be edited by hand or it can be created by the Reactivity Modeling System program MechPro. The mechanism listing file is an ASCII file of the gas-phase reactions and include reactant and product stoichiometry and reaction rate information. Another input to ChemMech is the species table file, a comma separated ASCII file that can be easily opened and edited by spread sheet programs or text editors. The file contains a table of species data that control certain aspects of how the model species are treated in a model simulation and provide linkages between species in different phases. They also allow you to control the outputs that are generated in a simulation and provide links to data, such as emissions, initial concentrations, and boundary concentrations, that are generated by other CMAQ input processors.

Since different gas-phase chemical mechanisms can be used in the CMAQ system and employ different species names, it is necessary to supply the CMAQ system with the linkages among the gas-phase species to, for example, emissions, deposition velocities, and Henry's Law constants if aqueous-phase scavenging is to be modeled. The species table also allows you to select which species concentrations are written to the output files and whether or not they are to be advected or diffused. All of this information is input and utilized by ChemMech to generate the CMAQ mechanism include files.

The distributed files include species table files for the no-aerosol and aerosol versions of the standard SAPRC-99 mechanism, called S99spec.csv and S99Aspec.csv, respectively. These can be used as inputs to MechPro to produce a species table file with the added model species, if the parameters are to be the same as those for some species in the standard mechanism (see Table A-6), or as inputs to ChemMech if the lumped parameter version of SAPRC-99 is run without added model species. In other cases, the species table would have to be hand-edited to include the information about any new model species before the ChemMech program can be run. The contents of this file are discussed below.

The species table consists of columns and rows that contain detailed species information, where each row corresponds to a species and each column to some characteristic of that species. The first row contains the column headings, and the first column contains species names. Notice that the species table described here only contains gas phase species, which are included in the gas-phase mechanism only. For other types of chemical species, such as aerosol, non-reactive, and special study (tracer) species, they need to have separate species tables. Table 2 provides a

summary of the current species table column headings. These column headings are fixed; their meanings and usage are described below.

Table 2. Summary of column headings in the species table file

Column Heading	Description
SPC	Species name
MOLWT	Molecular weight of species
EMIS_SUR	Surrogate names for species specified in emissions processor
EMIS_FAC	Emission scaling factor
ICBC_SUR	Surrogate names for species specified in initial and boundary conditions processor
ICBC_FAC	Initial and Boundary conditions scaling factor
DEPV_SUR	Surrogate names for species with specified deposition velocities
DEPV_FAC	Deposition velocity scaling factor
SCAV_SUR	Surrogate names for species scavenged into aqueous phase
SCAV_FAC	Henry's law scaling factor
G2AE_SUR	Surrogate names for species linked to the aerosol species
G2AQ_SUR	Surrogate names for species used in the aqueous chemistry module
ADV	Optional advection process
DIFF	Optional diffusion process
DDEP	Dry deposition outputs to be generated for specified species
WDEP	Wet depositions outputs to be generated for specified species
CONC	Concentration outputs to be generated for specified species

SPC and MOLWT Columns

The SPC and MOLWT columns are the species name and their corresponding molecular weights that are included in the gas-phase chemical mechanism. These names should be the same as those that appear in the chemical mechanism.

EMIS_SUR and EMIS_FAC Columns

The EMIS_SUR column is a required column. The CMAQ emissions processor generates an emissions file in which emissions for each species are identified by a unique name that may be up to 16 characters in length. The emission names that are generated by the emissions processor may not necessarily correspond to the name that is used in a chemical mechanism. The emission species name that is contained on the input file should be entered in the row that corresponds to the chemical species in the mechanism. If no entry for a chemical mechanism species appears in this column, its emission will be set to zero.

The EMIS_FAC column is a column of scale factors that can be used to scale emission rates that are on the input file. This column would normally only be used for special study purposes (e.g. emission sensitivity analysis) and can usually be omitted. An example of the usage of this column is considered here. Suppose we want to examine the impact of reducing 25% of SO2 emissions without regenerating the emission input files. Then a scale factor of 0.75 could be entered in this column in the row corresponding to the chemical mechanism species name for SO2. If no scaling is to be performed, a scaling factor of 1.0 must be included for every emission species that is listed in the EMIS_SUR column.

ICBC_SUR and ICBC_FAC Columns

The ICBC_SUR column is used to map the gas-phase mechanism species to the species that are contained on the input initial condition (IC) and boundary condition (BC) input files. If a surrogate name is entered in this column, then the IC and BC files will be searched for a species with the surrogate name. If the surrogate name is not found on the file or a surrogate name is not entered for a species, the ICs and BCs for gas –phase mechanism species will be set to a nominal zero (i.e. a very small concentration).

The ICBC_FAC column provides scaling functionality for ICs and BCs similar to the EMIS_FAC column described above. The input for each entry in this column is not required if IC/BC scaling is not employed.

DEPV_SUR and DEPV_FAC Columns

The DEPV_SUR column maps gas-phase mechanism species to deposition velocities that are computed for various chemical compounds by the CMAQ MCIP processor. The list of compounds and their corresponding surrogates are described below. If no surrogate name is entered for a gas-phase mechanism species, the deposition velocity will be set to zero.

The DEPV_FAC column is used for scaling the input deposition velocities. A scale factor must be input for each surrogate name that appears in the DEPV_SUR column.

Table 3. CMAQ species that use deposition velocities

Surrogate names	Unit	Description
VD_SO2	M/S (meter per second)	DV for species sulfur dioxide
VD_SULF	M/S	DV for species sulfate
VD_NO2	M/S	DV for species nitric dioxide
VD_NO	M/S	DV for species nitric oxide
VD_O3	M/S	DV for species ozone
VD_HNO3	M/S	DV for species nitric acid
VD_H2O2	M/S	DV for species hydrogen peroxide
VD_ALD	M/S	DV for species aldehydes
VD_HCHO	M/S	DV for species formaldehyde
VD_OP	M/S	DV for species organic Peroxide
VD_ORA	M/S	DV for species organic acids
VD_NH3	M/S	DV for species ammonia
VD_PAN	M/S	DV for species peroxy acetyl nitrate
VD_HONO	M/S	DV for species nitrous acid
VD_CO	M/S	DV for species carbon oxide

G2AE_SUR Column

The G2AE_SUR column is used to link the gas-phase mechanism species to the aerosol species. If aerosols are not included in an application, this column must still be present but should have no entries.

G2AQ_SUR Column

The G2AQ_SUR column is used to link the gas-phase mechanism species to the names used in the aqueous chemistry module. It must be included if the cloud module is used in the model.

SCAV_SUR and SCAV_FAC Columns

The SCAV_SUR is used to specify whether a mechanism species is scavenged from the gas-phase into the aqueous phase by assigning a Henry's Law constant to the gas-phase mechanism species. Henry's Law constants for a number of different chemical compounds have been predefined in the CCTM and assigned names shown below. If a cloud processing module is not included in the model, this column can be omitted. A non-entry for any particular species corresponds to no scavenging.

The SCAV_FAC column is used to scale selected input Henry's Law constants. A scale factor must be input (even it's one) for every surrogate name that appears in the SCAV_SUR column. If no scaling is employed, this column can be omitted.

Table 4. Henry's Law constants for CMAQ species

Species	Henry's Law (K_H, dimensionless)
SO2	0.04
NO2	3.5
O3	2.0
H2O2	4.0×10^{-7}
HNO3	2.7×10^{-7}
NH3	6.6×10^{-4}
PAN	0.01
NO	23
HCHO	4.0×10^{-6}
ALD	2.7×10^{-3}
OP	1.7×10^{-4}
PAA	7.6×10^{-5}
ORA	1.0×10^{-8}
HONO	4.1×10^{-7}
CO	40

ADV and DIFF Columns

These required columns control the treatment of species in the advection and diffusion science processes. A YES entry indicates that the species is to be advected and diffused, and a NO entry means the species is not to be advected and diffused. Normally, all species would be advected and diffused and these columns would default to YES entries for most species.

DDEP, WDEP, and CONC Columns

These required columns control which species are included in the output concentration file and the output dry and wet deposition files. A YES entry causes the outputs to be generated for that species, and a NO entry results in no outputs being generated. Note that the entries do not affect the model simulation in any way, but only which species concentrations are written to the output files. Note also that dry and wet deposition outputs should be requested only for those species that have dry deposition velocity and scavenging surrogates assigned.

SMOKE Enhancements

An original concept of the SMOKE programs Cntlmat and Smkreport was the ability to model and QA emissions reactivity substitutions. Until this project, the functions to implement speciation profile or source category code (SCC) substitutions and inventory replacements by SCC were not available in SMOKE. The SMOKE enhancements that we completed to support reactivity modeling include:

- **Cntlmat.** Read in and apply reactivity packets and new profile specifications for reactivity controls. The SMOKE program Cntlmat is controlled using ASCII configuration files that we refer to as packets. We instrumented SMOKE to read reactivity packets that include information about the SCC or specific facility processes that will receive a reactivity control, the new SCC emissions, the future-year projection rate, a future-year for implementation of all controls, and a market penetration rate. The reactivity packet can include codes for new speciation profiles that are not in the existing database, but instead will include the composition of the new profile in terms of compounds in the Speciation Database. Cntlmat also outputs QA reports summarizing the sources that received reactivity controls, the base and replacement emissions, the base and replacement speciation profile code, and the market penetration rate. Cntlmat had to be updated to create consistency between the way VOC emissions are reported in the inventories and how they are processed internally in SMOKE.

  We modified Cntlmat to account for the reactive organic gas (ROG) to total organic gas (TOG) conversion that happens during the processing of emissions inventories. Volatile Organic Compounds in emissions inventories are commonly reported as either VOC or ROG. SMOKE converts the ROG inventory to TOG before applying temporal, spatial, and chemical profiles, as these profiles are based on the assumption that the VOC data is represented as TOG. With reactivity controls that involve substituting speciation profiles, SMOKE needs to handle both the inventory pollutants and the converted emissions that have the speciation profiles applied to them. To be consistent with these different VOC representations, we modified Cntlmat to include settings to select which inventory pollutant to use and the VOC assumption that is included in the speciation profiles. The Cntlmat settings "REACTIVITY_POL" and "REPLACEMENT_POLID" control which inventory pollutant and speciation assumption Cntlmat will use, respectively. For example, as emissions inventories are most commonly reported as ROG, this will be the setting for the Cntlmat variable "REACTIVITY_POL". However, as speciation profiles are formulated in terms of TOG, this will be the setting for the Cntlmat variable "REPLACEMENT_POLID". These settings are dependent on the SMOKE input data, and will have to be confirmed by the user before proceeding with the reactivity modeling with SMOKE.

- **Smkmerge.** The program Smkmerge uses the output of Cntlmat to apply the reactivity controls to the inventory pollutants and prepare air quality model-ready emissions. After modifying Cntlmat to integrate the reactivity control information into SMOKE, we made corresponding changes to Smkmerge to invoke these controls in the modeled emissions. Cntlmat and Smkmerge are now fully instrumented to apply reactivity substitutions for the fixed- and adjustable-parameter versions of the SAPRC-99 mechanism and the CB-IV mechanism. We did not test the reactivity controls with RADM-2.

- **Smkreport.** We used Smkreport, the SMOKE QA module, to generate summaries of VOC mass by chemical speciation profile. We modified Smkreport to read in the control matrix output from Cntlmat to apply the reactivity controls to the QA reports. The reports of VOC mass by speciation profile reflect the reactivity adjustments by (1) optionally adjusting the base year mass for all sources affected by the reactivity control, (2) optionally using different speciation profiles for those sources that have changed speciation profiles in the reactivity matrix, and (3) optionally applying growth factors to a future year using a market penetration rate assigned by the reactivity packet. We list these changes as optional because SMOKE can change any or all of these factors of the base year inventory using the reactivity packets.

Figure 1, above, is a detailed flow diagram illustrating how all of the programs that make up the Reactivity Model System integrate to model a new chemical mechanism. In the figure, the programs are rectangles with rounded corners and shaded in gray, files are white rectangles, and user-selected options are ovals. Files that depend on the mechanism and lumping options are shown with dashed borders, and those that are mechanism-independent have solid borders.

Although SMOKE must be run every time any input or options are changed, this is not necessarily true for the programs outside of SMOKE on the figure. ProfPro need only be run once for existing profiles or when these profiles are changed, and need not be run at all if profiles are given in terms of actual chemical compounds (not mixtures) using compound designations employed in the master speciation database. SpecPro need only be run if the profiles are changed or new profiles are added, or if the chemical mechanism or lumping options are changed. MechPro need only be run if the adjustable-parameter version of the SAPRC-99 mechanism is employed or if a lumped mechanism[5] is employed with compounds represented explicitly or with non-standard model species. In these cases, it must be run every time the lumping or the emissions mass or speciation profile assignments are changed.

The details of the linkages of the SMOKE programs are included for clarification for SMOKE users and implementation of our design, but those who are not concerned with these linkages on the diagram can simply ignore them. Each of the SMOKE steps shown must be run for each of the source categories processed through the SMOKE system (e.g., area, nonroad mobile, on-road mobile, and point sources) and each of the intermediate files shown will have multiple instances – one for each source category. In the design in Figure 1, the links between SMOKE and the other parts of the system are (1) the speciation profile file (GSPRO) passed from SpecPro to SMOKE's chemical speciation program Spcmat, (2) the VOC mass in the domain, summed by speciation profile, which is needed for each source category; and (3) the model-ready emissions from SMOKE, which are passed to the CMAQ program.

Figure 2 shows another flow diagram for the files and steps needed to support applying a reactivity control to the inventory and using the controlled emissions in the CMAQ. The diagram is the same as Figure 1 except that the reactivity input and the SMOKE programs and files used to process it are added. The reactivity input is shown on bottom left; it consists of the reactivity packet and new profile specifications (if any). If the new profile has a compound not included in

---

[5] The adjustable parameter version of SAPRC-99 has mechanistic parameter for certain lumped model species that depend on the emissions-weighted distribution individual compounds they represent, while the fixed parameter version uses default parameters that do not depend on emission, as is the case with other current mechanisms such as CB4 or RADM-2.

the Speciation Database, the database would need to be updated to include the compound and its mechanism assignments.

ProfPro and its input files are not shown in Figure 2 because it is not necessary to reprocess existing profiles for reactivity calculations. This also means that users should provide new speciation profiles used for reactivity controls in terms of actual compounds using the compound designations in the Speciation Database. This is our preference for how the system should be used. If the user finds it easier to specify new profiles in terms of the SAROAD numbers used in existing profiles, that user must first add the profile to the existing profile database and run ProfPro to obtain a chemical profile assignment file that includes the new profile.

Inside the SMOKE system, the Cntlmat module must be run to process the reactivity packet and create a reactivity matrix for each source category that is receiving reactivity controls. The VOC totals by speciation profile must include the changes implemented by the reactivity packet: a new speciation profile, a new base year mass, and the market penetration of such changes for the selected future year. To aid in the QA of the reactivity modeling, we changed the SMOKE program Smkreport to include reactivity information in a report of the speciation profiles with the VOC mass. The sections that follow detail the programs that we created to support reactivity modeling with SMOKE and CMAQ and include instructions for how to use these programs.
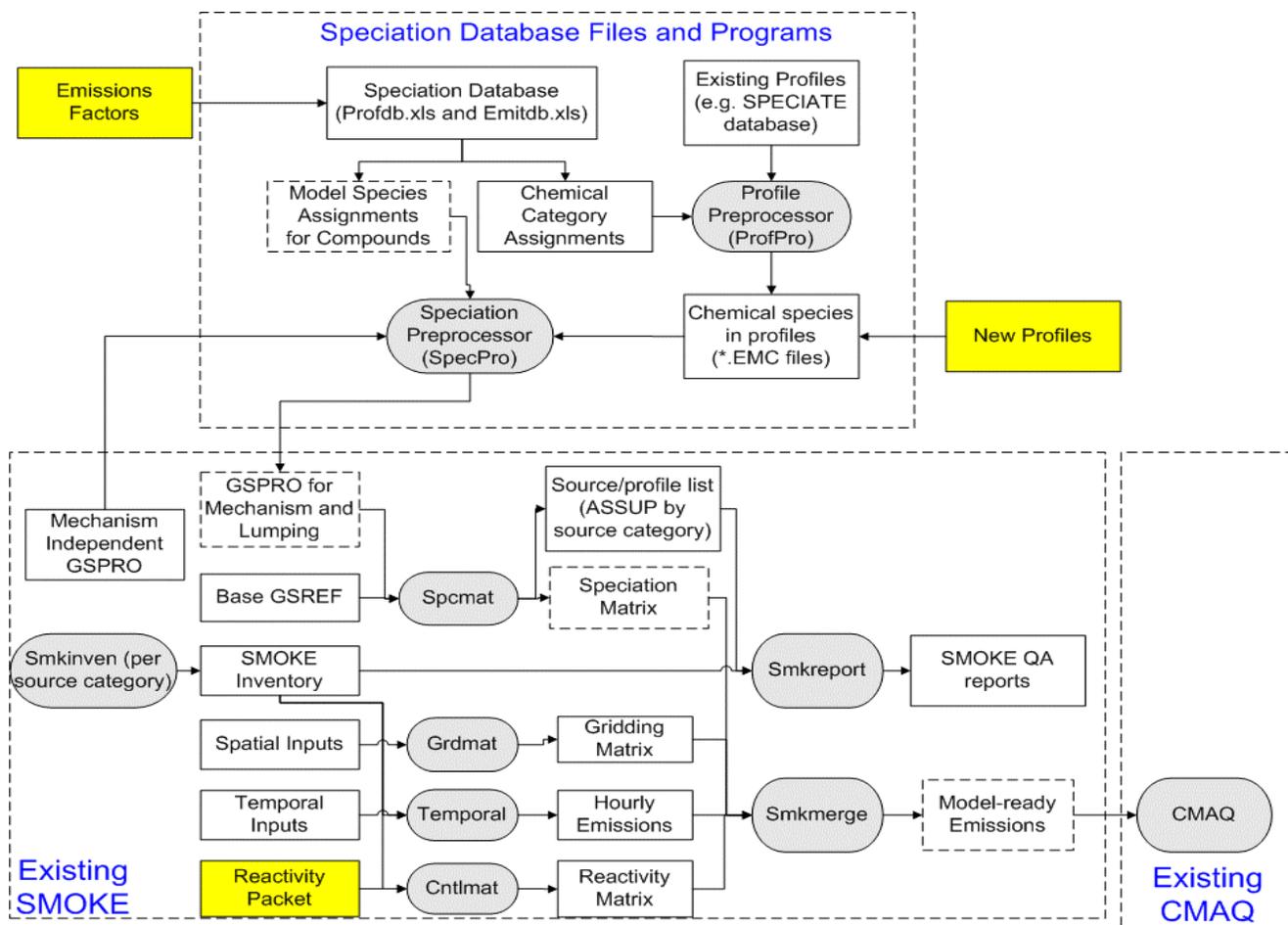

Figure 2. Processing to evaluate a reactivity control

Reactivity Modeling with SMOKE and CMAQ

With the details of the Reactivity Modeling System component programs presented above, we will now discuss how we used these programs to perform and test several difference reactivity scenarios with the system. The first test that we performed was the ability of the system to create base GSPRO files that are compatible with SMOKE. We compared these files with the default files distributed in SMOKE and generated CMAQ-ready emissions to ensure their compatibility with SMOKE and CMAQ. The second test that we performed was to create a new mechanism using the system and run a CMAQ test simulation using the new mechanism. We designed this second test to simulate reactivity modeling using an adjustable-parameter version of the SAPRC-99 mechanism. The final test that we performed included a series of reactivity substitution simulations in which we replaced default VOC speciation profiles for different SCC's in the area, on-road mobile, non-road mobile, and stationary point emissions sectors with new profiles. These substitutions tested both the functionality in SMOKE to apply the new profiles at different market penetration rates and report on these substitutions for QA purposes and the responsiveness of CMAQ to changes in VOC emissions speciation.

The test simulation that we used for the reactivity modeling is a June 19-July 1, 1996 36-km O_3 episode in central North Carolina. Figure 3 is a graphic of the domain with grid specifications. We configured SMOKE to limit the vertical plume rise to the bottom 12 model layers.

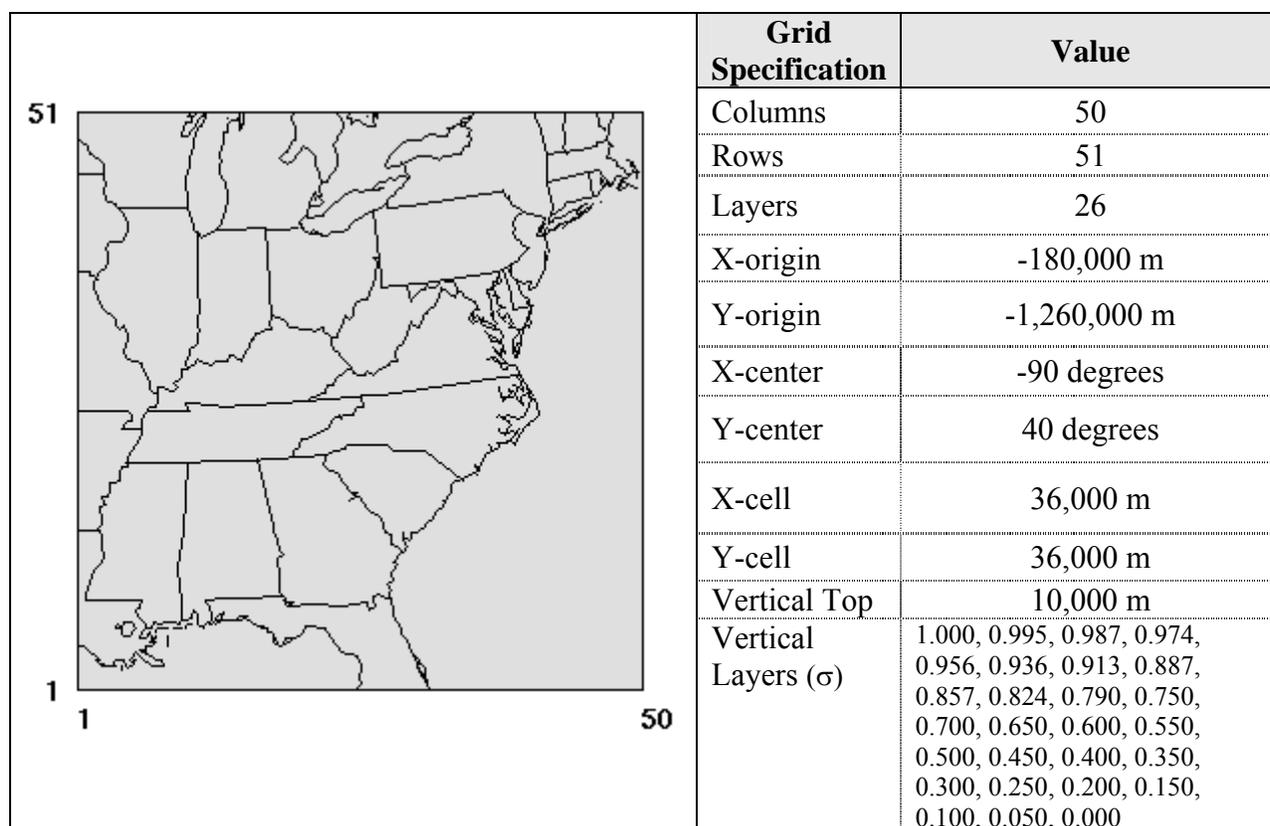| Grid Specification	Value
Columns	50
Rows	51
Layers	26
X-origin	-180,000 m
Y-origin	-1,260,000 m
X-center	-90 degrees
Y-center	40 degrees
X-cell	36,000 m
Y-cell	36,000 m
Vertical Top	10,000 m
Vertical Layers (σ)	1.000, 0.995, 0.987, 0.974, 0.956, 0.936, 0.913, 0.887, 0.857, 0.824, 0.790, 0.750, 0.700, 0.650, 0.600, 0.550, 0.500, 0.450, 0.400, 0.350, 0.300, 0.250, 0.200, 0.150, 0.100, 0.050, 0.000

Figure 3. NC 36-km modeling domain

Speciation profile generation and testing

Programs used: Speciation Database, ProfPro, SpecPro, SMOKE, CMAQ

Macros in the Speciation Database Excel files *emitdb.xls* and *profdb.xls* output the files needed by ProfPro and SpecPro to create SMOKE GSPRO files. The user configures the Speciation Database by selecting the source of speciation information from which to create VOC speciation profiles. Macros in the spreadsheet *emitdb.xls* produce files for all of the speciation databases and all mechanisms supported by the Speciation Database. The "Parameters and Commands" tab of the *emitdb.xls* contains all of the controls for running the spreadsheet macros. The "Update All" macro will create files for all mechanisms and speciation information in the database. The macro places all of the outputs in the same directory where *emitdb.xls* is located.

The files produced by *profdb.xls* are speciation database specific but mechanism independent. The "Parameters" tab of the spreadsheet contains the controls for running the macros. Configure the macro by adding the output directory paths for the profile assignment files and database listing files. Use the "Output Profile Data" and "Output Profile Listings" buttons to run the macros and produce output for all mechanisms and speciation databases contained in the spreadsheet. For the GSPRO test simulations we chose the SPECIATE 3.2 database profile data as the source of the speciation profiles; this database was the source of the default GSPRO files that are distributed with SMOKE. We normalized the source of the speciation data between the test profiles and the default SMOKE profiles in order to compare the results of the two simulations that used similar speciation profile input data. Once the macros in *emitdb.xls* and *profdb.xls* are completed, the input files for the CE-CERT speciation programs are online and ready to use.

The CE-CERT speciation programs ProfPro and SpecPro are configured and executed through scripts. In the default distribution of these programs they use DOS batch files as user interfaces. UNC-CEP ported these programs to UNIX and created C-shell scripts as user interfaces for the UNIX environment. Both sets of scripts use environment variables to pass configuration options to the speciation programs. ProfPro can be run one of three ways (These examples demonstrate how to invoke ProfPro from a command line, where PROFPRO refers to the program binary executable):

(1) PROFPRO InputParmFile. In this case, InputParmFile is the name of a file that contains the parameters, file names, and file locations that the program uses. The parameter files used by ProfPro and the other speciation database programs are ASCII files with records containing keyword=value (no spaces around the "="), with the parameters that are applicable to ProfPro listed in Table A-3 in Appendix A.

(2) PROFPRO PROF=profile. This case is useful for processing a single profile with a standard set of parameters. The input to this case is the code of a specific profile within the selected speciation database.

(3) PROFPRO PROFLIST=ProfileListingFile. This case is useful when processing multiple profiles listed in the ProfileListingFile with a standard set of parameters.

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 indicate in Table 2 above.

To run ProfPro with the profile composition files that we generated for the base GSPRO test case, we created a ProfileListingFile that points to the *.EMI files that we created with *profdb.xls*. Running ProfPro using a similar technique as case (3) above, produced a series of profile composition files given as mass fractions of individual compounds in terms of the categorization used in the speciation database. These new profile composition files are ready to use with SpecPro and MechPro for creating GSPRO and mechanism listing files, respectively.

Similar to ProfPro, SpecPro is configured and executed through scripts. SpecPro can be run in one of two optional modes, the GSPRO or CMPfile modes. The default GSPRO mode outputs a speciation profiles file for use with SMOKE. In the CMPfile mode, SpecPro produces a separate file for each profile giving the model species assigned to it for the specified mechanism. Although not used for emissions processing, the CMP files are useful for quality assurance. SpecPro can be run one of three ways (These examples demonstrate how to invoke SpecPro from a command line, where SPECPRO refers to the program binary executable):

(1) SPECPRO InputParmFile. In this case, InputParmFile is the name of a file that contains the parameters, file names, and file locations that the program uses. The input parameter file has the same format as indicated above for ProfPro, with the parameters and keywords that are applicable to SpecPro listed in Table A-4 in Appendix A.

(2) SPECPRO PROF=Profile[Mech][Output File]. This case is used to run the program in the CMPfile mode to output a mechanism composition file giving the model species assigned for a specified profile and mechanism.

(3) SPECPRO PROFLIST=ProfileListingFile. This case is used if the program is to be run with a particular profile listing file specified on the command line, with all other parameters specified in the external file, MODELING.PRM. This mode is useful for running the program with different sets of profiles but the same parameters.

We modified a default input parameters file (InputParmFile) to run SpecPro in mode (1). The parameters file specified the mechanism that we wanted to create GSPRO files for (either S99L for SAPRC99 or CB4 for the Carbon Bond IV mechanism), the name of the profile listing file output from ProfPro, the name of the GSPRO output file that we wanted to create, and the name of the SpecPro summary output file. Using similar syntax as case (1) above, we ran SpecPro to create test SMOKE speciation profiles for the SAPRCC-99 and CB-IV mechanisms.

The final step in testing the speciation profile generation function of the Reactivity Modeling System was to input the new profiles into SMOKE and create emissions for CMAQ. Before modeling with the new data we compared the new profiles with the default profiles distributed in SMOKE for both SAPRC-99 and CB-IV. The Speciation Databases outputs over 150 more speciation profiles from SPECIATE 3.2 than are contained in the default profiles distributed in SMOKE. A side-by-side comparison of the files shows that while the new files have many more profiles than the defaults, there are also several profiles that are contained in the default SMOKE files that do not come out of the Speciation Database. Presumably these missing profiles are custom data that do not exist in the SPECIATE 3.2 database but were added in the default files to represent explicit sources. The mismatch between the default profiles and the Speciation Database profiles is problematic because the Speciation Database does not generate associations between the profiles and inventory sources.

SMOKE uses two input files to perform VOC speciation (1) VOC profiles, like the GSPRO files created by the Speciation Database and (2) cross-reference files, which associate profile codes with inventory SCCs. The GSPRO file identifies each unique profile with an alpha-numeric profile code. These codes are used by the speciation cross-reference file to assign the profiles to specific inventory sources. The speciation cross-reference file is closely tied to the GSPRO file because it defines how the profiles are applied to emissions inventories. Since the Speciation Database only outputs GSPRO files and does not create accompanying cross-reference files, to test the new GSPRO files we had to default to using the cross-reference files distributed with SMOKE. Using a cross-reference file created independently of the GSPRO file is problematic because of mismatches or incompatibilities between the two files. In our test simulations, there were several profiles in the new GSPRO files that were not in the cross-reference file and vice versa, cross-reference assignments to profiles that didn't exist in the new GSPRO files. SMOKE treats the mismatches between the GSPRO and speciation cross-reference files by assigning a default VOC profile to the affected SCCs.

Strictly speaking, having the cross reference VOC or TOG files reference profiles that are not processed using the speciation programs is not acceptable, and the missing profiles need to be added to the profile input database before running SMOKE and CMAQ using this system. However, because of resource limitations we were unable to correct this problem before running the test simulations for this project.

Photochemical mechanism generation and testing

Programs used: MechPro, ChemMech, CMAQ

MechPro reads the outputs from the Speciation Database spreadsheets to generate mechanism listing files for use with the CMAQ mechanism compiler. MechPro can be run either from a script or at a command line using the following syntax (This example demonstrates how to invoke MechPro from a command line, where MECHPRO refers to the program binary executable):

> (1) MECHPRO InputParmFile, where InputParmFile is the file giving the parameters and files used by MechPro

The format of the parameter file used for MechPro is the same as discussed above for ProfPro and SpecPro, and the parameters applicable to MechPro are listed on Table A-6. In this case,the parameters file specifies the mechanism, relevant lumping assignments and new model species used, and the names and location of input and output files. For this test simulation we did not add any new species or reaction, opting to reproduce the base SAPRC-99 mechanism distributed with CMAQ. All of the input files required by MechPro are created by *emitdb.xls* and ProfPro. We configured MechPro to create a version of the SAPRC-99 mechanism for input to the CMAQ preprocessor, ChemMech.

After minor modifications to the CMAQ chemical mechanism compiler, ChemMech, we input a mechanism listing file produced by MechPro to create the mechanism include files required by CMAQ. ChemMech is a Fortran program that is compiled and run using scripts. A Makefile distributed with the source code compiles the executable and a C-shell script sets the ChemMech execution environment by defining the locations and names of the input and output files before invoking the executable. We ran the mechanism file generated with MechPro through ChemMech without any difficulty. We then used the Fortran include files produced by

ChemMech to recompile a version of CMAQ using the mechanism produced by the Reactivity Modeling System.

The first step in creating a new version of CMAQ with an updated chemical mechanism is to install the new mechanism include files in the correct location in the CMAQ directories. All of the mechanism definitions for CMAQ are located in the $M3HOME/models/include/release/$Mech directories, where $M3HOME is the root CMAQ installation directory and $Mech is a unique identifier for each available mechanism. As we based this test simulation off the default SAPRC-99 mechanism that includes aqueous chemistry, we copied the CMAQ include directory for this base mechanism to a new directory called saprc99_adj_aq. We replaced the include files (*.EXT), and mech.def files in this new directory with the files that we created with MechPro and ChemMech. We used the new mechanism include directory with the updated include files to compile CMAQ for testing the updated chemical mechanism. CMAQ compiled and executed without any difficulty and produced results for a test simulation that used a chemical mechanism generated by the Reactivity Modeling System.

Reactivity substitution modeling

Programs used: Speciation Database, ProfPro, SpecPro, SMOKE, CMAQ

We attempted to acquire real-world reactivity substitution data from the industry community. After numerous unsuccessful attempts, we designed our own substitution cases/multiple simulations to test the Reactivity Modeling System. We configured all of the reactivity simulations to use the fixed-parameter version of the SAPRC-99 mechanism and based them off the simulation that used the GSPRO file that we created using ProfPro and SpecPro. The first reactivity simulation that we completed was a hypothetical high-reactivity sensitivity in which we substituted the VOC profiles for the top 5 VOC sources by mass in the stationary area, nonroad, and stationary point emissions sectors with a pure xylene mixture. We tested the substitution at 100% and 50% market penetrations. The second reactivity simulation that we completed was a replacement of oil-based surface coatings with water-based solvents and paints and affected the stationary area emissions sector only. We tested this substitution at 100% and 25% market penetrations. The last two reactivity substitutions that we tested applied to the on-road mobile sector and included a replacement of all gasoline with a 10% ethanol blend and another replacement of all gasoline with propane or LPG. The details of these reactivity simulations follow.

Xylene reactivity substitution

By substituting a single VOC compound for a heterogeneous mixture of VOC's we can isolate the effects of the profile substitution because all of the VOC emissions for the affected sources will be speciated into one organic compound as opposed to multiple compounds. By using a speciation profile that represents pure xylene we tested if SMOKE performs reactivity substitutions correctly by replacing the speciation profiles for the top 5 VOC sources by mass in the stationary area, nonroad mobile, and stationary point emissions sectors and analyzed the resulting emissions from the affected sources. In a substitution with 100% market penetration we expect to see the VOC emissions for the affected sources go from a mixture of species in the base simulation to only xylene in the substitution scenario.

As another test of SMOKE's reactivity substitution capabilities, we also performed a brute-force profile substitution where we changed the SMOKE speciation cross-reference file to apply the xylene profile to the applicable sources. This brute-force approach is analogous to a 100% market penetration reactivity substitution and should produce comparable results.

The last test performed for the xylene substitution was to back off the market penetration rate to 50% and confirm that the output emissions and CMAQ simulations appropriately reflected this change. For the emissions we expect to see a mixture of compounds similar to the base profiles but with an added quantity of xylene in the mixture. As we referred to this sensitivity as a hypothetical high reactivity simulation, we expect to see O_3 increasing in the CMAQ simulations as we move from the base to the 50% and then 100% xylene substitution scenarios because xylene, and the corresponding SAPRC-99 species ARO2, are highly reactive relative to a mixture of VOC species.

Table 5 lists the SCC's, SCC descriptions, annual ROG inventory emissions, and the original SPECIATE 3.2 profile codes for the sources that received the xylene reactivity substitutions. We replaced the speciation profiles for all of the SCC's in Table 5 with profile 223 which have 100% of the VOC mass converted to the SAPRC-99 species ARO2.

Table 5. ROG sources used in the xylene reactivity modeling

SCC	Description	ROG Emissions (tons/yr)	Default Speciation Profile
2260001010	Mobile Sources; Off-highway Vehicle Gasoline, 2-Stroke; Recreational Equipment; Motorcycles: Off-road	1.024E+05	1101
2265004026	Mobile Sources; Off-highway Vehicle Gasoline, 4-Stroke;Lawn and Garden Equipment; Trimmers/Edgers/Brush Cutters (Commercial)	1.429E+03	1101
2265004071	Mobile Sources; Off-highway Vehicle Gasoline, 4-Stroke; Lawn and Garden Equipment; Turf Equipment	1.360E+05	1101
2282005010	Mobile Sources; Pleasure Craft; Gasoline 2-Stroke; Outboard	3.621E+05	1101
2282005015	Mobile Sources; Pleasure Craft; Gasoline 2-Stroke; Personal Water Craft	1.690E+05	1101
30100504	Industrial Processes; Chemical Manufacturing; Carbon Black Production; Oil Furnace Process: Main Process Vent	1.777E+05	1002
30300302	Industrial Processes; Primary Metal Production; By-product Coke Manufacturing; Oven Charging	5.525E+04	11
40201901	Petroleum and Solvent Evaporation; Surface Coating Operations; Wood Furniture Surface Coating; Coating Operation	6.224E+04	1003
30199999	Industrial Processes; Chemical Manufacturing; Other Not Classified; Specify in Comments Field	1.077E+05	9004
30300306	Industrial Processes; Primary Metal Production; By-product Coke Manufacturing; Oven Under firing	5.725E+04	11
2401005000	Solvent Utilization; Surface Coating; Auto Refinishing: SIC 7532;Total: All Solvent Types	1.971E+05	1194

2465000000	Solvent Utilization; Miscellaneous Non-industrial: Consumer; All Products/Processes; Total: All Solvent Types	2.032E+05	197
2401001000	Solvent Utilization; Surface Coating; Architectural Coatings; Total: All Solvent Types	4.609E+05	1003
2461800000	Solvent Utilization; Miscellaneous Non-industrial: Commercial; Pesticide Application: All Processes; Total: All Solvent Types	2.816E+05	76
2501060100	Storage and Transport; Petroleum and Petroleum Product Storage; Gasoline Service Stations; Stage 2: Total	2.195E+05	1190

Oil-based solvent reactivity substitution

The xylene substitution reactivity simulation was unrealistic by design. To test the Reactivity Systems ability to model more realistic substitutions we designed another sensitivity to replace one profile in the base GSPRO files with another profile. We did not create any new profiles using the Speciation Database, we simply used one existing profile in the SPECIATE 3.2 database to replace another existing profile. In this simulation we targeted four surface coating sources in the stationary area inventory that use oil-based solvents and paints. We replaced the speciation profile of the four SCC's listed in Table 6 with a profile that represents emissions from water-based solvents, with the hypothesis that the water based solvents are less reactive than the oil-based solvents. The emissions from the four SCC's that we used for this sensitivity compose over 10% of the total VOC emissions from area sources. We replaced the speciation profiles for these sources with profile 3013 for both 100% and 25% market penetration rates.

Note that this substitution is not totally realistic because the total mass of coatings VOC emissions was not changed, even though in practice use of most water based coatings involve emissions of different VOC mass than use of most oil based coatings. Reactivity packets do allow for changes in mass as well as profiles, but for simplicity this was not done in this example.

Table 6. ROG sources used in the oil-base solvent reactivity modeling

SCC	Description	ROG Emissions (tons/yr)	Default Speciation Profile
2401001000	Solvent Utilization; Surface Coating; Architectural Coatings; Total: All Solvent Types	4.154E+05	1003
2401008000	Solvent Utilization; Surface Coating; Traffic Markings; Total: All Solvent Types	6.662E+04	1003
2401002000	Solvent Utilization; Surface Coating; Wood Furniture: SIC 25; Total: All Solvent Types	1.408E+05	1003
2401099000	Solvent Utilization; Surface Coating; All Surface Coating Categories; Total: All Solvent Types	2.525E+05	1003

10% Ethanol/gasoline reactivity substitution

All of the previous simulations applied to the stationary area, non-road mobile, or stationary point emissions sectors. We designed this simulation to target on-road mobile source gasoline emissions modeled with the integrated MOBILE6 (http://www.epa.gov/otaq/m6.htm)/SMOKE emissions processor. Similar to the oil-based solvent substitution, for this sensitivity we replaced one set of existing SPECIATE 3.2 profiles with another to evaluate the effects of a 10%

ethanol/gasoline blend on the reactivity of on-road mobile emissions. We replaced the default SCC's for all of the MOBILE6 processes, which by default represent 1990 industry average levels, with a mixture 1990 industry average gasoline and 10% ethanol. We evaluated the effects of this reactivity substitution at 100% and 33% market penetration. As this sensitivity is our first application to on-road mobile sources in MOBILE6, we also performed a brute-force substitution at the speciation cross-reference file level for comparison with the 100% market penetration simulation. Table 7 summarizes the speciation profiles applied to the different MOBILE6 process for all on-road mobile emissions calculated with MOBILE6.

Table 7. 10% ethanol/gasoline substitution profiles

MOBILE6 Process	Description	Old Profile	New Profile
EXR	Running Exhaust	1313	1314
EXS	Starting Exhaust	1313	1314
EVR	Running Evaporative	1305	1301
RST	Resting Evaporative	1306	1302
DNL	Diurnal Evaporative	1306	1302
CRC	Crankcase Evaporative	1305	1301
RFL	Refueling Evaporative	1305	1301
HOT	Hot Loss Evaporative	1307	1303

LPG substitution

The last reactivity sensitivity that we completed tested the ability of the Reactivity Modeling System to create new speciation profiles using emissions factors for a VOC emissions source not represented in the Speciation Database. For this simulation we used data in the primary literature to develop a VOC speciation profile for liquefied petroleum gas (LPG) emissions. To generate the LPG profile, we entered emissions factors in mg VOC/km from Chang et al. (2001) into the spreadsheet *profdb.xls*, used the macros in *profdb.xls* to create a profile composition file, and ran this file through ProfPro and SpecPro to generate the new SAPRC-99 speciation profile for LPG emissions. We then used the SMOKE program Cntlmat to substitute the LPG profile for all of the MOBILE6 on-road mobile emissions processes at 100% and 50% market penetration. Table 8 lists the new LPG speciation profile that we created for this sensitivity.

Table 8. LPG VOC speciation profile (Chang et al., 2001)

Profile ID	Inventory Pollutant	SAPRC-99 Pollutant	Mass Fraction
LPG	TOG	ALK1	1.6%
LPG	TOG	ALK2	43.4%
LPG	TOG	ALK3	30.9%
LPG	TOG	ALK4	0.2%
LPG	TOG	ETHE	13.6%
LPG	TOG	OLE1	6.4%
LPG	TOG	OLE2	3.1%
LPG	TOG	ARO1	0.3%

LPG	TOG	ARO2	0.4%
LPG	TOG	NROG	0.1%

Results and Discussion

The emissions and air quality modeling results from all of the reactivity simulations completed for this project are presented in this section. In addition to the results of these simulations, problems that we encountered and their corrections are also presented here.

Speciation profile test simulation

In the absence of directly compatible speciation cross-reference files and GSPRO files created by SpecPro, we generated emissions with the available data to test if the new GSPRO files worked with SMOKE. SMOKE processed the GSPRO files output from SpecPro without any major operational problems. The emissions generated from these new profiles also ran through CMAQ without any difficulties. Figures 4 and 5 compare O_3 predicted by the base CMAQ simulations that used the default SMOKE profiles against simulations the used the new profiles for the SAPRC-99 and CB-IV mechanisms, respectively. The differences in these simulations are primarily due to the GSPRO and speciation cross-reference file incompatibilities that result in the application of default speciation profiles in the simulations that used the new profiles. Some of the differences are also due to the profile updates, but these are assumed to be small in comparison to the differences caused by the widespread application of default profiles.

Simulations that use speciation profiles generated by the Reactivity Modeling System will not be able to take advantage of the new information in the Speciation Database unless all the profiles referenced in the cross-reference files that associate these data with inventory sources are included in the profile database and can be processed using the speciation programs. In addition, if a speciation cross-reference file is to be used that originated outside of the speciation database on which the profiles are based, the consistency between the profile numbers used in the cross reference file with those associated with the profiles in the SpecPro input would need to be verified. SpecPro allows for associating different profile numbers with the various profiles as optional input (see discussion of the PROFLIST parameter in Table A-1).

Since we were unable to use the speciation programs to process all the profiles used in the speciation profile test simulation, the results are strictly speaking a test of the speciation processing system. As illustrated in Figures 4 and 5, the effects from the speciation file incompatibilities on O_3 predictions are larger in the SAPRC-99 simulation than in the CB-IV simulation. Presumably the effects of the speciation profile assignments are magnified in the SAPRC-99 simulation because, for this episode, SAPRC-99 predicted 16% more maximum O_3 than CB-IV. Consistent with both higher magnitudes and a greater spatial extent in the predicted O_3, the SAPRC-99 simulation shows corresponding larger effects of the errors in the speciation profile assignments. Figure 6 compares the maximum 1 hour O_3 predictions for the SAPRC-99 and CB-IV simulation that used the new speciation profiles, respectively.
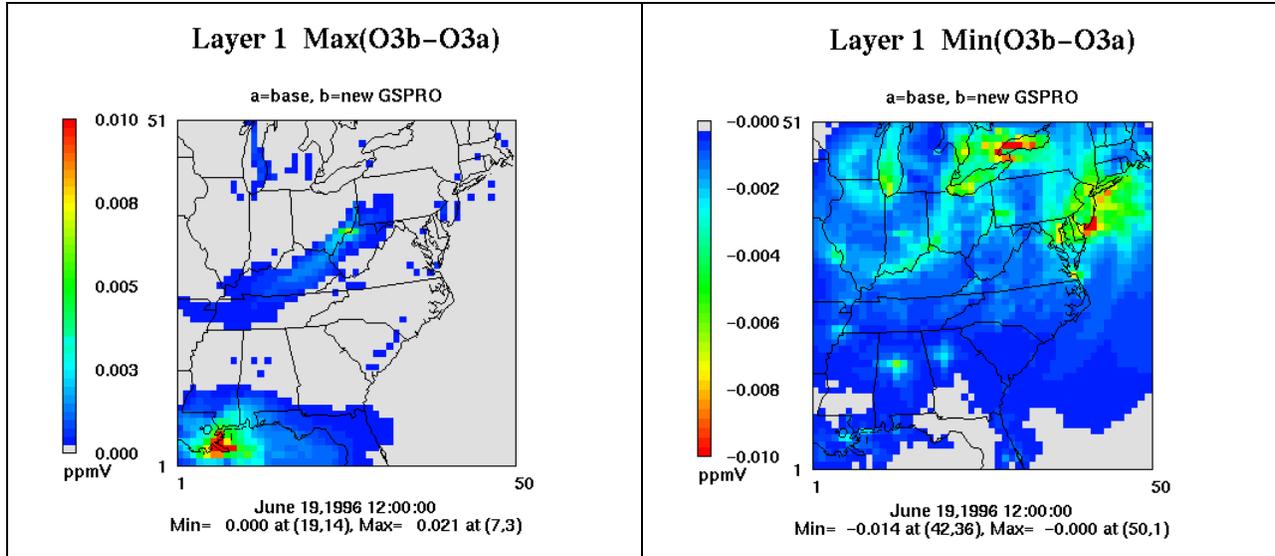
Figure 4. CMAQ maximum (L) and minimum (R) O3 differences for SAPRC-99 speciation profile test case simulation
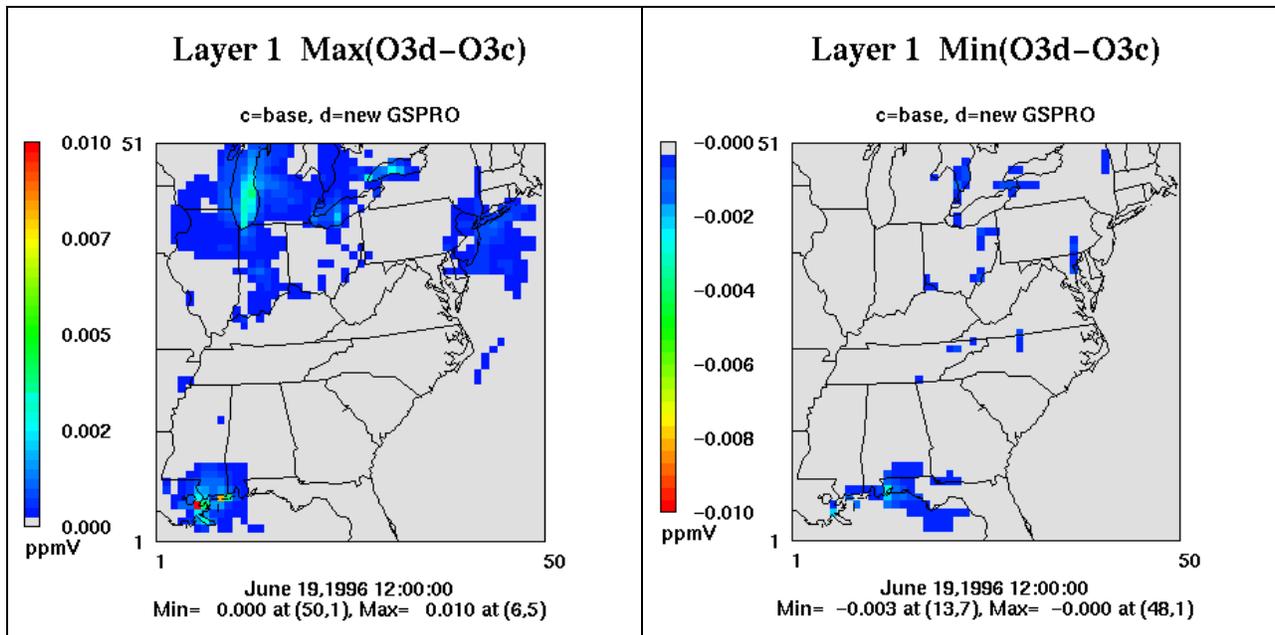

Figure 5. CMAQ maximum (L) and minimum (R) O3 differences for CB-IV speciation profile test case simulation
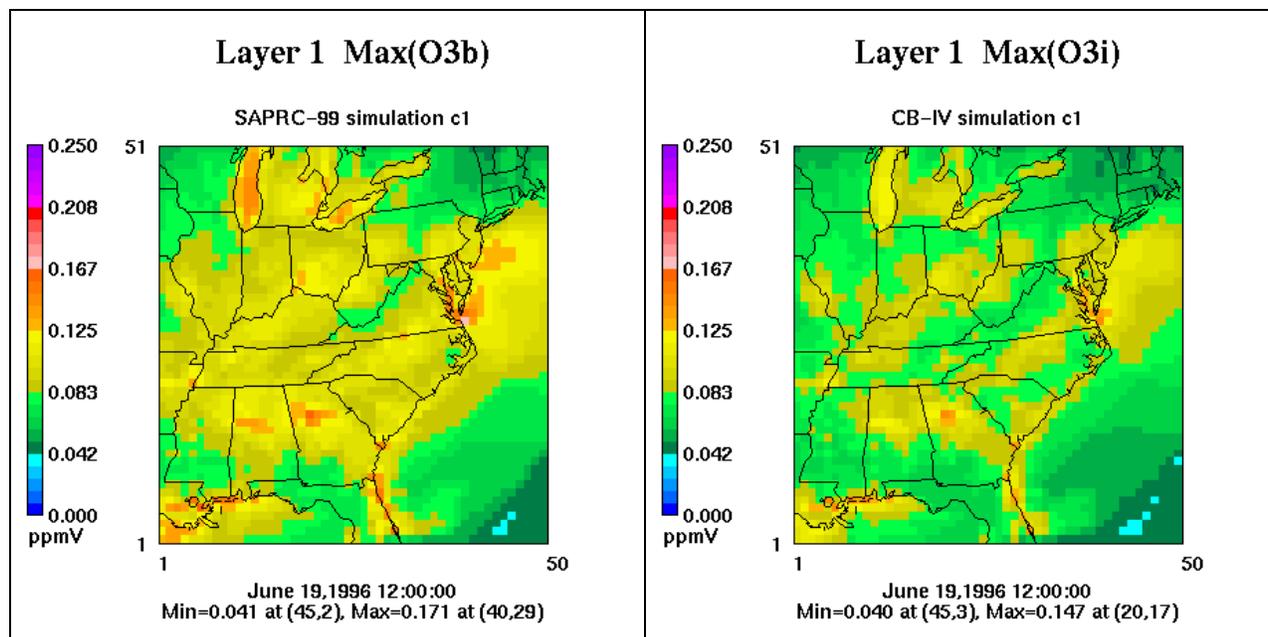
Figure 6. CMAQ maximum O₃ predictions for the speciation profile test simulation; SAPRC-99 (L) and CB-IV (R)

Photochemical Mechanism Test Simulation

After creating and configuring new SAPRC-99 photochemical mechanism include files for CMAQ we simulated the test episode and compared the results against the base SAPRC-99 simulation. Because this test simulation was carried out before the final version of the ChemMech program developed for this project was available, to use the MechPro mechanism listing file in CMAQ we had to update the names of some mechanism species and all of the photolysis rate names to be consistent with the default CMAQ data. For example, ethylene in the default CMAQ mechanism is named ETHENE; in the mechanism listing file produced by MechPro it is called ETHE. Similarly, the default CMAQ mechanism listing file uses the identifier NO2_SAPRC99 for the name of the NO₂ photolysis rate data; the mechanism listing file produced by MechPro uses the identifier NO2 for the same rate. Before processing the mechanism listing file to generate the CMAQ mechanism include files, we modified the MechPro file to use consistent species names and photolysis rate names as the default mechanism file to ensure consistency with the default CMAQ data. This should not be necessary with the current version of ChemMech distributed with these programs.

We also ensured consistency in the input files between the two simulations, including using the same SMOKE emissions inputs, which we based off the default SMOKE speciation cross-reference and profiles to eliminate the speciation file compatibility problems discussed above. The differences between the CMAQ simulations that are compared in Figure 7 are due to differences in the photochemical mechanisms.

Figure 7 illustrates that there were differences between the SAPRC-99 simulations that used the base mechanism and the mechanism generated with the Reactivity Modeling System. A comparison of the two mechanism listing files reveals both kinetic and mechanistic differences between the two mechanism versions. The version of SAPRC-99 in the version that we created

with MechPro contains different product yields and reaction rates for the ALK2, ALK3, ALK4, ALK5, OLE1, and OLE2, oxidation reactions. These mechanism differences led to the differences in O_3 predictions illustrated in Figure 7. Although there were some differences in the mechanism that we created with MechPro that lead to differences in the CMAQ predictions, through this simulation we confirmed that the Reactivity Modeling System can generate new photochemical mechanism for CMAQ.
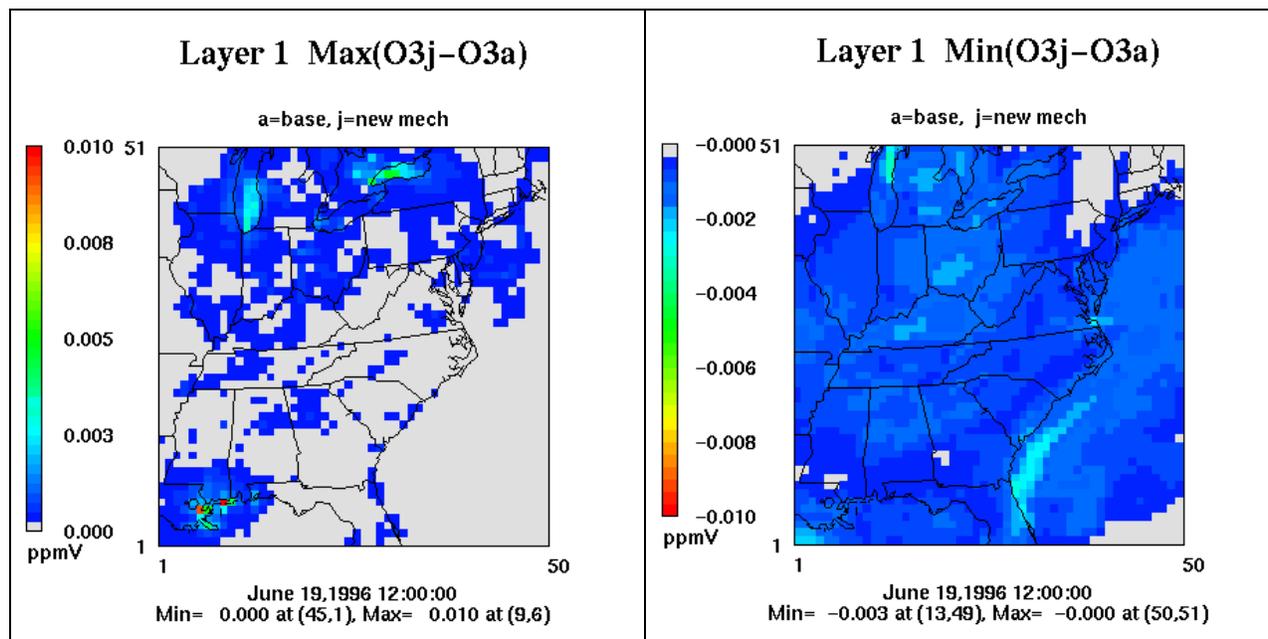

Figure 7. CMAQ maximum (L) and minimum (R) O3 differences for SAPRC-99 photochemical mechanism test case simulation

Reactivity Substitutions

All of the reactivity simulations that we completed with the Reactivity Modeling System, SMOKE, and CMAQ tested different features of the system to evaluate its utility in modeling real-world reactivity substitutions. For the xylene/high-reactivity substitution we replaced the speciation profiles in the top 5 stationary area, non-road mobile, and stationary point VOC emissions sources with a xylene profile that converted all of the VOC emissions from these sources to the SAPRC-99 species ARO2. For the simulation that used a 100% market penetration rate, as expected, we observed all of the VOC's for the selected sources speciated into ARO2 rather than the complex mixture of SAPRC-99 species created by the default profiles for the selected sources. A comparison of the emissions from the 100% market penetration rate substitution with the emissions from a brute-force substitution, where we replaced the speciation profiles for the selected SCC's in the speciation cross-reference file, originally revealed a systematic difference between the two simulations.

As direct analogues, a 100% market penetration rate simulation and a speciation cross-reference file profile replacement should produce exactly the same emissions predictions. A comparison of the results between these two simulations for the xylene substitution showed a systematic difference in the results due to the way that SMOKE converts inventory VOC's from ROG to TOG before speciating the emissions. Our original implementation of the reactivity substitution

with Cntlmat did not account for this conversion and lead to systematic underestimations in the emissions. After modifying the SMOKE programs to account for the ROG to TOG conversion when creating reactivity intermediate files, we corrected this discrepancy. Figure 8 shows maximum and minimum O_3 differences between the base and xylene substitution simulations.

The ROG to TOG factors currently being used in SMOKE were acquired from USEPA, but may be outdated. Therefore, we strongly encourage the air quality scientists and chemists to review these ROG to TOG factors when applying reactivity substitutions. Note that the SpecPro program will compute ROG to TOG factors (see discussion of the TOGVOC-FAC parameter on Tables A-1 and A-4), but procedures and programs to use them as inputs to SMOKE have not yet been developed. This should be done in future updates.

Figure 8 shows that the xylene substitution produced a maximum O_3 increase of 59 ppbV and a minimum decrease of 4 ppbV in calculating differences between the sensitivity and base simulations (note the differences in the scales of the two plots). As expected the xylene substitution simulation produced more O_3 than the base simulation because we substituted a higher reactivity VOC for a lower reactivity mixture of VOC's. An interesting pattern in these results is the spatial gradients between the O_3 increases and decreases, with the increases clustered primarily in the Northeast and Midwest and the decreases clustered in the Southeast. These gradients are likely due to differences in NO_x versus VOC limited chemical regimes for O_3 formation between these different regions of the domain. Generally, the Southeast is in a NO_x limited regime so the O_3 decreases are likely due to NO_x scavenging by the xylene oxidation products. With the Northeast and Midwest being VOC limited, the higher reactive VOC enters these NOx rich areas of the domain and reacts to readily form O_3.
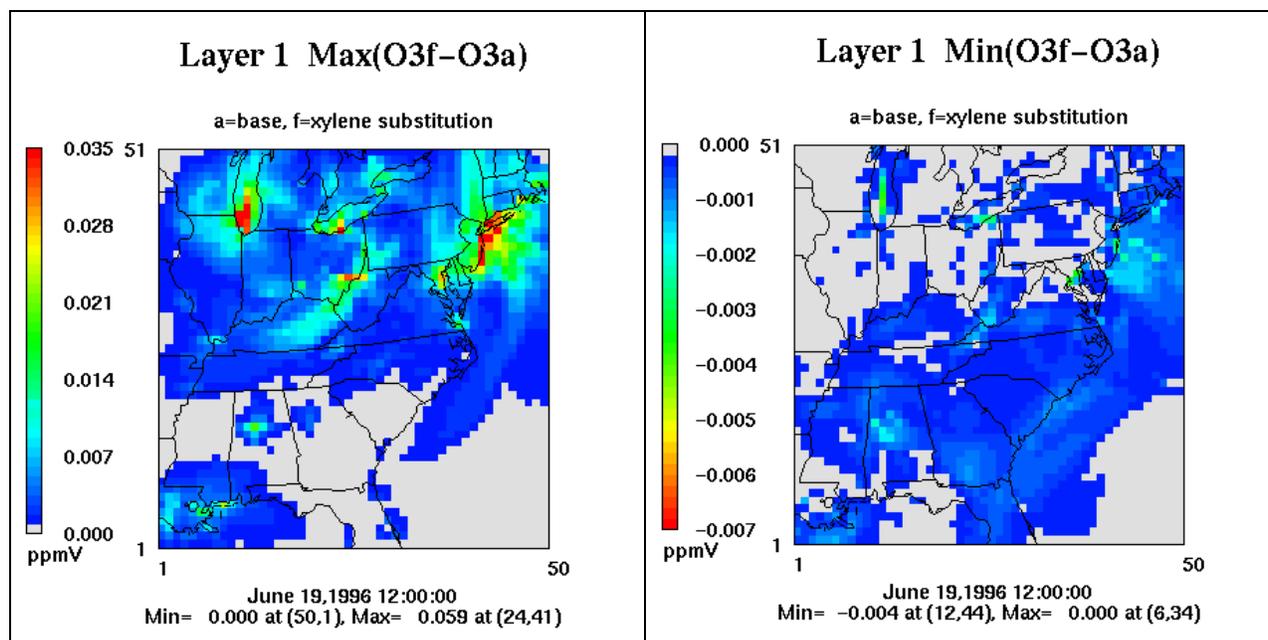

Figure 8. CMAQ maximum (L) and minimum (R) O3 differences for xylene substitution reactivity simulation (100% market penetration)

For the xylene substitution simulation with a 50% market penetration we observed similar results but with lower magnitude changes in O_3. The speciated VOC emissions for the selected SCC's

from the 50% market penetration simulation contained a similar VOC mixture as the base simulation but with higher concentrations of ARO2, the SAPRC-99 equivalent of xylene.

For the second reactivity substitution that we performed we targeted VOC emissions from oil-based surface coating operations in the stationary area source inventory. We replaced the speciation profiles for the top 4 inventory VOC sources in this emissions sector with profile representing VOC emissions for water-based solvents, with the assumption that emissions from the water-based solvents are less reactive than the emissions from the oil-based solvents. We used existing profiles from the SPECIATE 3.2 database to test this substitution at 100% and 25% market penetration rates. Figure 9 shows maximum and minimum O_3 differences between the base and water-based solvent substitution simulations.

Figure 9 shows that substituting a water-based VOC emissions profile for an oil-based profile in the top 4 oil-based surface coating operations SCC's in the area source inventory results in a maximum increase of 0.3 ppbV O_3 and a maximum decrease of 3 ppbV O_3 in calculating differences between the sensitivity and base simulations (note the differences in the scales of the two plots). These results support our hypothesis that the water-based solvent substitution would produce less O_3 than the base simulation that used oil-based solvents, with higher O_3 predicted in the base simulation, even if the substitution involved no net change in mass emissions.
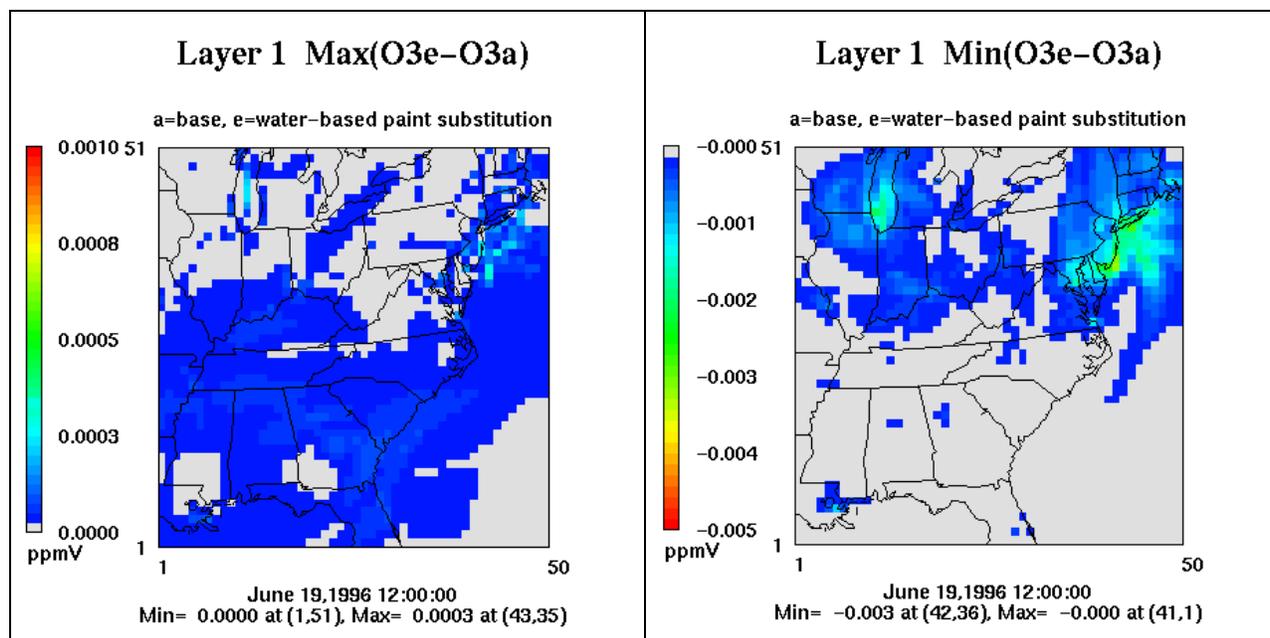
Figure 9 CMAQ maximum (L) and minimum (R) O3 differences for water-based surface coating substitution reactivity simulation (100% market penetration)

Table 9 compares the distribution of the SAPRC-99 VOC species for oil-based and water-based solvent emissions profiles used in this reactivity simulation. The major differences in these profiles are that the largest components of the oil-based profile are the aromatic species where the largest components of the water-based profile are the alkane species. The simplest explanation of the differences in the O_3 predictions from these two simulations is that the larger

fraction of higher-reactivity aromatic emissions in the base simulation leads to higher O_3. Further analysis into the simulation that may provide insight into the results may include using process analysis in CMAQ to calculate integrated reaction rates for the important VOC oxidation reactions. The integrated reaction rates would reveal the chemical conditions that are present in the model which lead to these results. The simulation that used a 25% market penetration rate yielded similar results but on a smaller scale. Simulations like this which targets a relatively small component of the entire mix of emitted VOC's (10% of the area sources), boarders on being within the range of the sensitivity of CMAQ. Changes of less than 1 ppbV O_3 are arguably within the noise of the CMAQ predictions and yield results that should be interpreted with caution.

Table 9. Oil-based solvent (1003) and water-based solvent (1013) speciation profile comparison

SAPRC-99 VOC Species	Profile 1003 Mass %	Profile 1013 Mass %
ACET	1.3	0
ALK1	0	5.5
ALK2	2.0	0.6
ALK3	0	2.2
ALK4	12.5	27.6
ALK5	31.7	61.2
ARO1	38.4	0.3
ARO2	8.5	0
MEK	0.3	0.5
OLE1	0.1	0.1
NR	0	0.6
PROD2	2.0	0.4

In the third reactivity substitution simulation we targeted on-road mobile source fuel by replacing the emissions from conventional gasoline with a 10% ethylene-gasoline mix. Similar to the water-based solvent substitution, we did not add any new profile information to the database. The profiles for the 10% ethylene mixture already exist in the SPECIATE 3.2 database and we performed this substitution by simply replacing the gasoline profiles for the different MOBILE6 processes (Table 4) with analogous profiles for the 10% ethylene mix. Figure 10 shows maximum and minimum O_3 differences between the base and 10% ethylene substitution simulations.

Figure 10 illustrates extremely small difference in O_3 as a result of the 10% ethylene mix sensitivity. With maximum and minimum differences between the sensitivity and base simulations of 0.18 ppbV O_3 and 0.32 ppV O_3, respectively, the effects of this sensitivity are well

within the noise of CMAQ predictions. While the effects of this substitution are small, these results do show that at a market penetration rate of 100% the emissions from conventional gasoline produce higher maximum O_3 relative to the 10% ethylene mixture. The simulation that used a 33% market penetration rate yielded similar results but on a smaller scale.
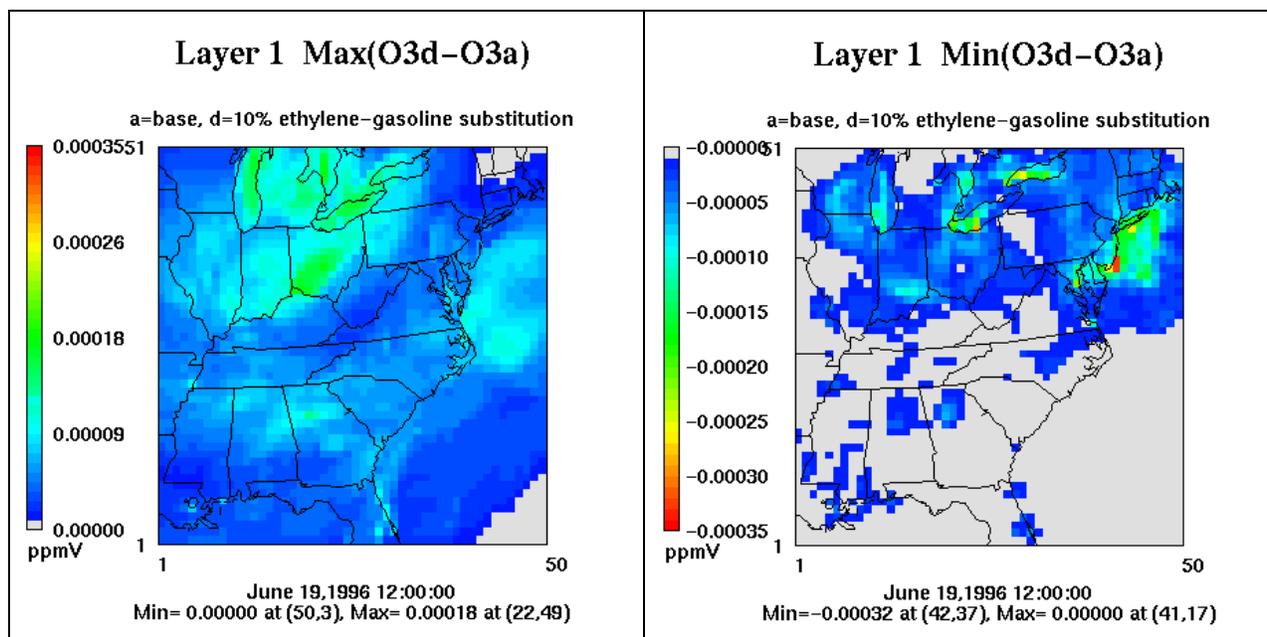

Figure 10 CMAQ maximum (L) and minimum (R) O3 differences for 10% ethanol-gasoline substitution reactivity simulation (100% market penetration)

The last reactivity substitution sensitivity that we performed also targeted on-road mobile source emissions but involved using the Speciation Database and the programs ProfPro and SpecPro to generate a new speciation profile. We used information from the literature to derive VOC profiles for liquefied petroleum gas (LPG) emissions. After deriving SAPRC-99 profiles from tabulated VOC emissions factors from Chang et al. (2001), we added a cross-reference for the new LPG profile to the base speciation cross-reference file and created new MOBILE6 emissions for CMAQ. Figure 11 shows maximum and minimum O_3 differences between the base and LPG substitution simulations.

The LPG substitution produced less O_3 than the base gasoline profiles as illustrated in the larger minimum O_3 predictions in Figure 11 (e.g. higher O_3 in the base simulation). We anticipated that the LPG emissions would produce less O_3 as it has lower-reactivity emissions relative to conventional gasoline. As Chang et al. only listed emissions factors for running exhaust we used this single profile for all MOBILE6 emissions factors. Since MOBILE6 represents several automobile emissions processes other than running exhaust, by substituting a profile developed for running exhaust for all of the processes we are misrepresenting the effects of this substitution. A more accurate substitution would be to use emissions factors for each individual mobile emissions process (e.g. evaporative, refueling, etc) to develop process-specific profiles. Despite only using a single profile for all MOBILE6 processes, we did accomplish the objective of this substitution, which was to demonstrate the ability of the Reactivity Modeling system to convert

emissions factors from the literature to SMOKE emissions profiles and model a reactivity substitution with the new data.
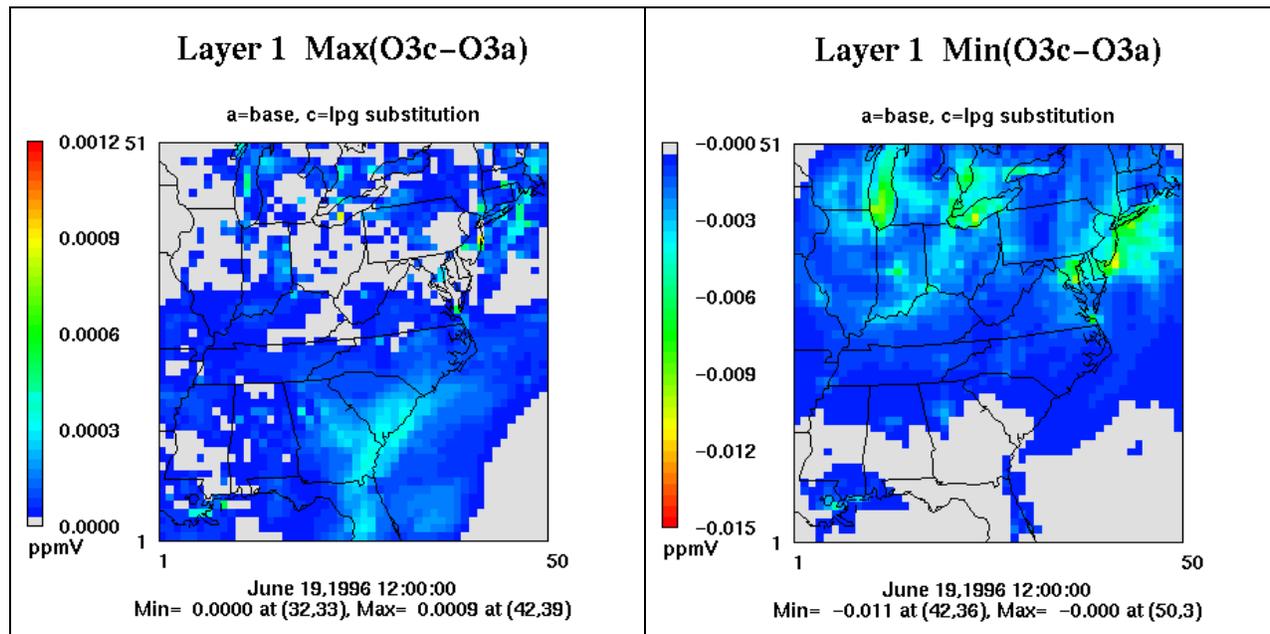

Figure 11. CMAQ maximum (L) and minimum (R) O3 differences for LPG substitution reactivity simulation (100% market penetration)

Conclusions

Through this work, we demonstrated that the Reactivity Modeling System is a functional tool for developing new speciation profiles for SMOKE to use for reactivity substitutions. SMOKE accurately applies the new profiles at different market penetration rates and reports the results of the reactivity simulations for QA purposes. Once the emissions are prepared with SMOKE, CMAQ can model the emissions perturbations that result from the reactivity substitutions. The sensitivity of CMAQ to the reactivity changes is related to the percentage of total VOC emissions that are affected by the substitution. Emissions sources that compose a larger percentage of the total VOC emissions will have a greater impact when targeted for reactivity substitutions.

Realistic reactivity simulations involve mass reductions in addition to profile substitutions. The reactivity simulations that we tested only reflect the profile substitution component and do not account for the corresponding mass adjustments. To more accurately represent the effects of the reactivity substitutions, these simulations should be recreated with the correct mass adjustments.

Unresolved Issues

- All the profiles used in existing modeling databases and cross reference files need to be added to the profile database associated with this reactivity modeling system. Alternatively, it may be useful for the reactivity modeling system to have a base speciation cross-reference file that associates all inventory sources with a default profile that can be used if the profile used by the source is not on the profile database.

In any case, compositional information for all profiles referenced in the source cross reference files must be available for the speciation database and programs to be properly used with SMOKE for creating emissions for CMAQ.

- The capability of the speciation programs to compute ROG to TOG factors needs to be integrated into SMOKE so outdated factors and inconsistencies encountered when testing these programs will no longer a source of error.

- Real-world test cases for evaluating the system are also needed. The tests documented here were hypothetical and thus produced unrealistic results. The system is now ready to apply with actual reactivity information.

- The system needs to be evaluated with all the profiles used in the cross reference file included in the profile database.

- The system needs to be evaluated with new species. We did not add new species or reactions to the photochemical mechanisms.

- The system needs to be evaluated for adjustable parameter versions of the SAPRC-99 mechanism. We did not integrate the inventory summaries of VOC mass by speciation profile into the system to create an adjustable parameter version of the SAPRC-99 mechanism.

References

Carter, W. P. L, 2004. Documentation of Speciation Preprocessor Program for SMOKE, ftp://ftp.cert.ucr.edu/pub/carter/emitdb/pgmdoc.pdf .

Carter, W.P.L. 2000. Programs and Files Implementing the SAPRC-99 Mechanism and its Associates Emissions Processing Procedures for Models-3 and Other Regional Models. January 31, 2000. http://pah.cert.ucr.edu/~carter/SAPRC99.htm

Chang, Chih-Chung, Jiunn-Guang Lo, Jia-Lin Wang, 2001. Assessment of reducing ozone forming potential for vehicles using liquefied petroleum gas as an alternative fuel. Atmospheric Environment 35, 6201-6211.

EPA. 1990. User's Guide for the Urban Airshed Model-Volume I; User's Manual for UAM(CB-IV). U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-450/4-90-007a.

EPA (2002). Speciate 3.2 Release Notes. http://www.epa.gov/ttn/chief/software/speciate/spec32_rel_notes.txt

Houyoux, M.R., J.M. Vukovich, C.J. Coats Jr., N.W. Wheeler and P.S. Kasibhatla. "Emission Inventory Development and Processing for the Seasonal Model for Regional Air Quality (SMRAQ) project." *Journal of Geophysical Research, Atmospheres* 105:D7:9079-9090 (April 16, 2000).

Mebust, M., B.K. Eder, F.S. Binkowski, and S.J. Roselle. "Models-3 Community Multiscale Air Quality Model (CMAQ) Aerosol Component: 2. Model Evaluation" *J. Geophys. Res.* (D6) 4184, doi:10.1029/2001JD001410 (2003).

Appendix A – UCR Reactivity Program Configuration Options and File Formats

Table A-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 A-3 through A-6 for descriptions 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
CMPDINFO.DAT	Gives the category code, molecular weight, carbon numbers, and description of all chemical compounds in the database. [c]	EmitDB.XLS	ProfPro, EmitSum

File	Contains	Produced by	Used by
(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 CMPDINFO.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
(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 unspeciated 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

File	Contains	Produced by	Used by
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	SMOKE
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 not yet written

File	Contains	Produced by	Used by
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.) Note that the profile number must be consistent with the profile numbers used in the cross reference file that associates SCC emissions sources to profiles, and the listing file should contain all profiles references in this cross reference file.	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
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	ChemMech

File	Contains	Produced by	Used by
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	Not used in current version (useful for QA)
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	ChemMech
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]
(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

File	Contains	Produced by	Used by
(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" 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 contains 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 unspeciated 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. Note that the present distribution does not contain the files needed by MechPro to process for RADM-2.

[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 (e.g., 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 3.

[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.

[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 A-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.

Keyword	Source [a]	Type of Documentation or Information
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 (<<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 (<< 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 unspeciated 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 A-3. List of parameters and options that can be specified in parameter input files for the ProfPro program

Parameter Keyword	Pgms [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). Blakn 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

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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 A-4. List of parameters and options that can be specified in parameter input files for the SpecPro program.

Parameter Keyword	Pgms [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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.
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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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".
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 unspeciated 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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.
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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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 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 A-5. List of parameters and options that can be specified in parameter input files for the EmitSum program.

Parameter Keyword	Pgms [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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.
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 A-6. List of parameters and options that can be specified in parameter input files for the MechPro program.

Parameter Keyword	Pgms [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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.
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 us 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.)

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.)
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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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.
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.

Parameter Keyword	Pgms [a]	Parameter Values [b]	Description and Defaults.
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 ChemMech 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.

Appendix B – 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).

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

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 x 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

60

$$
\begin{array}{cccc}
\text{Contribution of} \\
\text{compound to MSMAS} & = & 1 - \text{Reactivity} & x & \dfrac{\text{Mass of compound in}}{\text{Molecular weight of}} & \text{(II)} \\
\text{for the inert species} & & \text{factor} & & \text{compound}
\end{array}
$$

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

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

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

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

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

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

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

Note that TOTMAS = Σ_i MSMAS(i) + 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.