# DEVELOPMENT OF CHEMICAL SPECIATION DATABASE AND SOFTWARE FOR PROCESSING VOC EMISSIONS

William P. L. Carter

College of Engineering Center for Environmental Research and Technology
University of California, Riverside, CA 92521

June 9, 2004

- Problems with current VOC speciation databases and software
- Project Objectives and Approach
- Speciation Database Development
- Speciation Processing Software for SMOKE
- Current Status and Additional Work Needed

# BACKGROUND

- VOC emissions consist of many hundreds of chemical compounds with widely varying atmospheric reactivities.
- These are represented in models by a much smaller number of **model species**, depending on the **chemical mechanism** used.
- The most appropriate mechanism and set of model species depends on the model application
- Most models use highly condensed mechanisms that are not appropriate for VOC reactivity or toxics assessment applications
- Emissions processing is a major obstacle to implementing new or more detailed mechanisms into airshed models.
  - Emissions and mechanism processing software have limitations that make changing mechanisms difficult
  - Difficult to make assignments for the hundreds of chemical using the current categorization systems

# PROBLEMS WITH CURRENT SPECIATION PROFILE DATABASES

- Different profile databases use different chemical categorization systems. For example
  - EPA and California SAROAD codes are incompatible
  - EPA's SPECIATE database uses different categories
- Many chemical categories are redundant, poorly defined, or refer to mixtures of uncertain composition. This means:
  - Assigning model species requires assumptions about compositions of the uncertain mixture categories
  - Inconsistent or inappropriate mechanism assignments occur
- Model species assignment files exist for only a few condensed mechanisms (CB4, RADM-2, lumped SAPRC) that are not suitable for many toxics or VOC reactivity applications
- Data bases are not set up for easy modifications of mechanisms

# PROBLEMS WITH CURRENT SPECIATION SOFTWARE

- Re-processing emissions for new or modified mechanisms is very computer and labor intensive
- Emissions processing systems do not include documented software to easily update mechanism assignment files when:
  - speciation profiles change or are updated
  - mechanisms or model species assignments are changed
  - selected compounds are to be represented explicitly
- The capabilities of the detailed SAPRC mechanisms for VOC reactivity assessment is not supported. These include:
  - Adjusting parameters in mechanism to best represent detailed compositions of VOC emissions
  - Change how VOCs are lumped and compute the mechanisms for the new lumped or explicit species

# PROJECT OBJECTIVES AND APPROACH

## OBJECTIVES

- **Specific Objective**: Implement capabilities of the SAPRC-99 detailed mechanism into the SMOKE emissions system
- **Broader Objective**: Improve speciation database and software for easier and consistent implementation of all mechanisms

## APPROACH

- Develop an improved emissions speciation database with consistent assignments for SAPRC-99, CB4 and RADM-2
- Develop software to implement speciation database and assignments in SMOKE
- Test programs to demonstrate application to VOC reactivity assessment

# SPECIATION DATABASE DEVELOPMENT

- Chemical categories used to specify chemical compositions of profiles in various databases combined into a single database
  - U.S. EPA, California, and Texas modeling databases
  - EPA's Speciate 3.2 database
  - Categories used in the SAPRC-99 detailed mechanism
- Duplicate categories removed.
- Categories classified as follows:

| Code | Description | For assignment to: | Categories |
|---|---|---|---|
| 1 | Single compound | Model species | 1387 |
| 2 | Isomeric mixture | Compounds | 367 |
| 3 | Complex mixture | Compounds | 187 |
| 4 | Poorly defined | Unassigned mass | 32 |
| | | Total: | 1973 |

# ASSIGNMENTS MADE TO SINGLE COMPOUNDS

| Assignment Type | Fraction Ass'd | |
|---|---|---|
| | Cmpds | Mass[1] |
| Atomic formula and molecular weight | All | All |
| EPA's VOC exempt status | All | All |
| CAS number | 0.92 | 0.997 |
| SAPRC-99 Detailed Model Species[2] (666)[3] | 0.75 | 0.9995 |
| Emissions Groups[2] (used to make lumped SAPRC-99 and RADM-2 assignments) (64)[3] | 0.91 | ~1 |
| Carbon Bond 4 Species[2] (11)[3] | 0.94 | ~1 |
| Compounds judged to be non-volatile | 0.04 | 0.003 |

1. Fraction of mass in all Speciate 3.2 profiles, each weighted equally
2. Assignment fractions exclude compounds judged to be non-volatile
3. Number of model species or groups

# TYPES OF MODEL SPECIES

- SAPRC-99 Detailed model species: individual compounds or groups of compounds assumed to have similar reactivities, for which reactivity values have can be calculated.
- Emissions groups: Groups of compounds expected to be lumped together in “lumped molecule” mechanisms.
  - Originally developed for emissions processing for RADM-2, and extended to lumped SAPRC-99 and RACM

| Examples of Emissions Groups (of 64) | Lumped SAPRC-99 | RADM-2 |
|---|---|---|
| Alkanes ($k_{OH}$= 1-2 x $10^4$ $ppm^{-1}$ $min^{-1}$) | ALK5 | 0.945 HC8 |
| Ethanol | ALK3 | 1.198 HC3 |
| Aromatic aldehydes | BALD | NR |

- Carbon Bond species: Used to represent parts of molecules that are assumed to react independently in Carbon Bond 4

# ASSIGNMENTS MADE TO MIXTURES

- As many mixtures as possible were assigned to distributions of individual compounds or designated “nonvolatile”.

Examples:

| | |
|---|---|
| Simple mixture:<br>$C_{12}$ alkyl benzenes | 1/3 heptyl benzene<br>1/3 1-butyl-3-ethyl benzene<br>1/3 1-methyl-2,4-dipropyl benzene |
| Complex Mixtures:<br>Distillates, naphtha, mineral spirits | Composite petroleum distillate profile derived by Censullo, et al, 1996 |
| $C_{19}$ compounds (diesel exh.) | Nonvolatile |
| “Fragrances” | Not assigned |
| Poorly defined category:<br>CB4 paraffin bond (PAR) | Not assigned |

- All “Simple Mixtures” and 60% of the “Complex Mixtures” were assigned. (92% of the mass in sum of all Speciate 3.2 profiles)

# TREATMENT OF NONVOLATILE COMPOUNDS

- Although VOC (or TOG) profiles should not contain non-volatile compounds, profiles contain compounds or mixtures such as:
  - eicosane ($C_{20}H_{42}$)
  - n-hexadecanoic acid ($C_{16}H_{32}O_2$)
  - benzo(a)pyrene ($C_{20}H_{12}$)
  - 2-propenoic acid, butyl ester, polymer with ethenyl acetate
- These are judged to be non-volatile and therefore not appropriately represented by gas-phase model species
- If these judgments are correct, these need to be either
  - moved to the PM profiles and the VOC mass reduced
  - Processed as PM species
- In the meanwhile, the alternatives are to:
  - represent the nonvolatiles as if they were inert volatiles, or
  - exclude nonvolatiles from emitted volatile mass

# PROFILE DATABASE

- Profiles from various sources combined into a single database
  - EPA's Speciate 3.2 database
  - EPA's Models-3 emissions database
    - Provided by Gipson, 11/01
    - Most but not all judged to be same as Speciate profiles
  - Texas database
    - Provided by Yarwood (Environ) in July, 2002.
    - Does not include newer profiles developed for TCAQ
    - All but 5 are same as Speciate or EPA profiles
  - California ARB database
    - Downloaded from CARB web site. Dated 3/19/03
    - Most are *different* from Speciate or EPA profiles
- Profiles judged to be same as others flagged as such
- All chemical categories in these databases assigned master speciation database categories for mechanism processing

# SPECIAL PROFILES ALSO INCLUDED IN THE PROFILE DATABASE

## USEFUL FOR ASSIGNING MODEL SPECIES TO EMISSIONS:

- "Representative Unknown #1"
  - Example of a type of profile that might be used to represent unknown mass when assigning model species to emissions
  - Consists of the mixture of all nC>6, Mwt>120 compounds in the Speciate 3.2 all-profile average
- "Representative Unassigned #1"
  - Example of a type of profile that might be used to represent compounds not assigned to model species
  - Molecule-weighed average of all compounds represented by SAPRC-99 mechanism in each SAPRC-99 chemical class.

# SPECIAL PROFILES ALSO INCLUDED IN THE PROFILE DATABASE

## USEFUL FOR DERIVING MECHANISM PARAMETERS

- "ARB 1994 Base ROG"
  - Derived from morning urban ambient air measurements
  - Used to represent base case emissions from all sources in scenarios used to derive "Carter" reactivity scales
  - Used to derive mechanistic parameters for the "fixed parameter" lumped SAPRC-99 mechanism in CMAQ
- "EPA Models-3 1998 Base ROG"
  - Derived from total anthropogenic emissions data in a 1998 EPA Models-3 emissions database
  - Included with SAPRC-99 mechanism files as an option for deriving mechanistic parameters based on emissions data

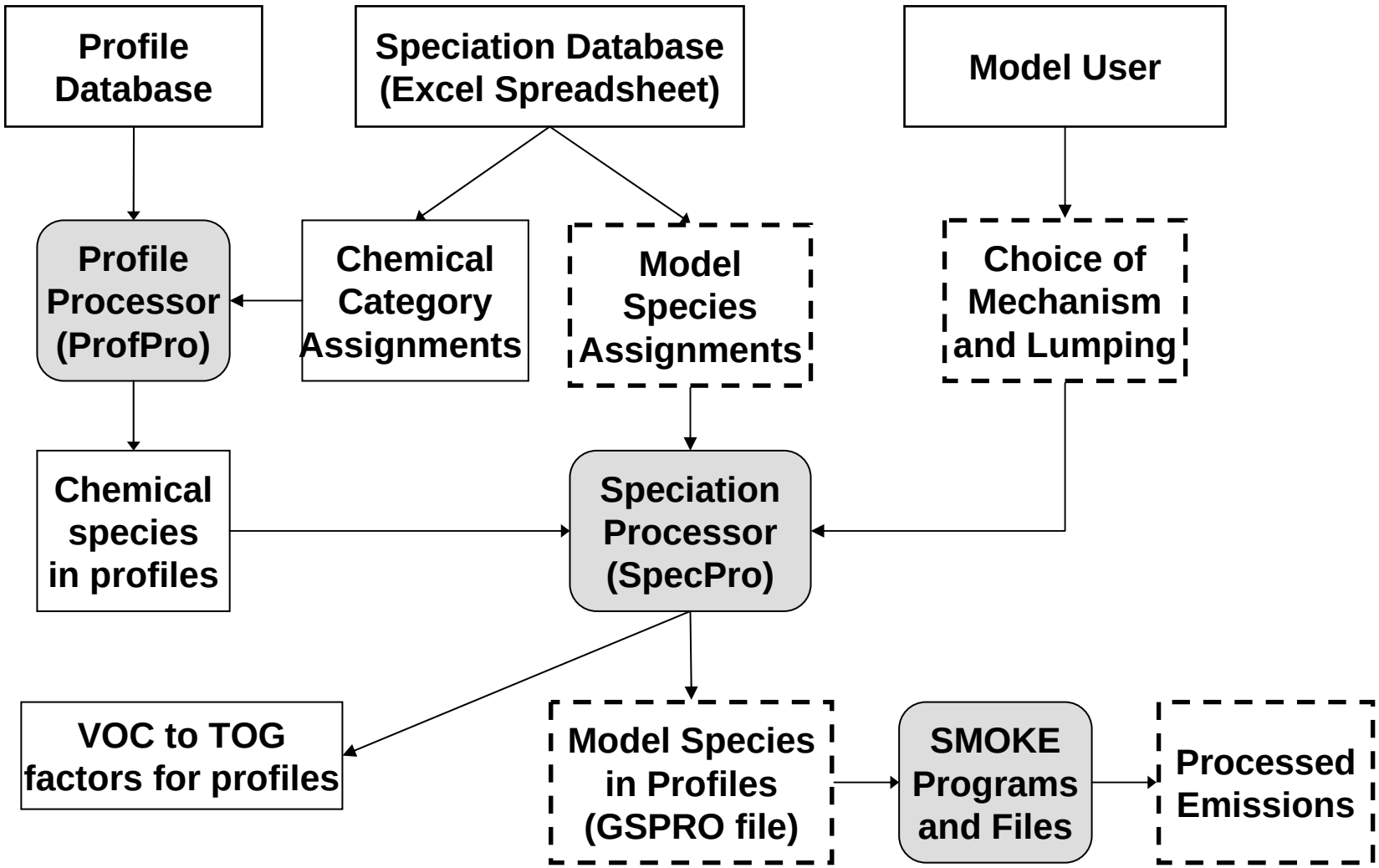
# CURRENT IMPLEMENTATION OF DATABASES

- **SPECIATION DATABASE**
  - Implemented in Excel spreadsheet file with assignments of:
    - Existing profile speciation categories to master categories
    - Mixture categories to compounds
    - Compounds to model species
    - Characteristics of model species for various mechanisms
  - Macros written for following functions:
    - Compile mixture and mechanism assignments
    - Output assignments in ASCII files for use by speciation software for SMOKE
- **PROFILE DATABASE**
  - Implemented in an Excel spreadsheet with
    - Available documentation information for profiles
    - Profile compositions (using native categories)
  - Macro to output ASCII files with profile compositions and documentation for use by speciation software

# SPECIATION PROCESSING PROGRAMS

- Profile Processor (**ProfPro**)
  - Converts profile composition data using native chemical categories to compositions of individual compounds
  - Needs only be run once for each profile unless assignments or profile compositions changes
- Speciation Processor (**SpecPro**)
  - Produces model species assignments for profiles (SMOKE GSPRO file) given profile compositions and mechanism lumping assignments
  - Allows users to specify alternative lumping or explicit representation of selected compounds
  - Gives users options for how to represent unknown or unassigned mass
  - Optionally outputs VOC to TOG factors for the profiles (mechanism independent)

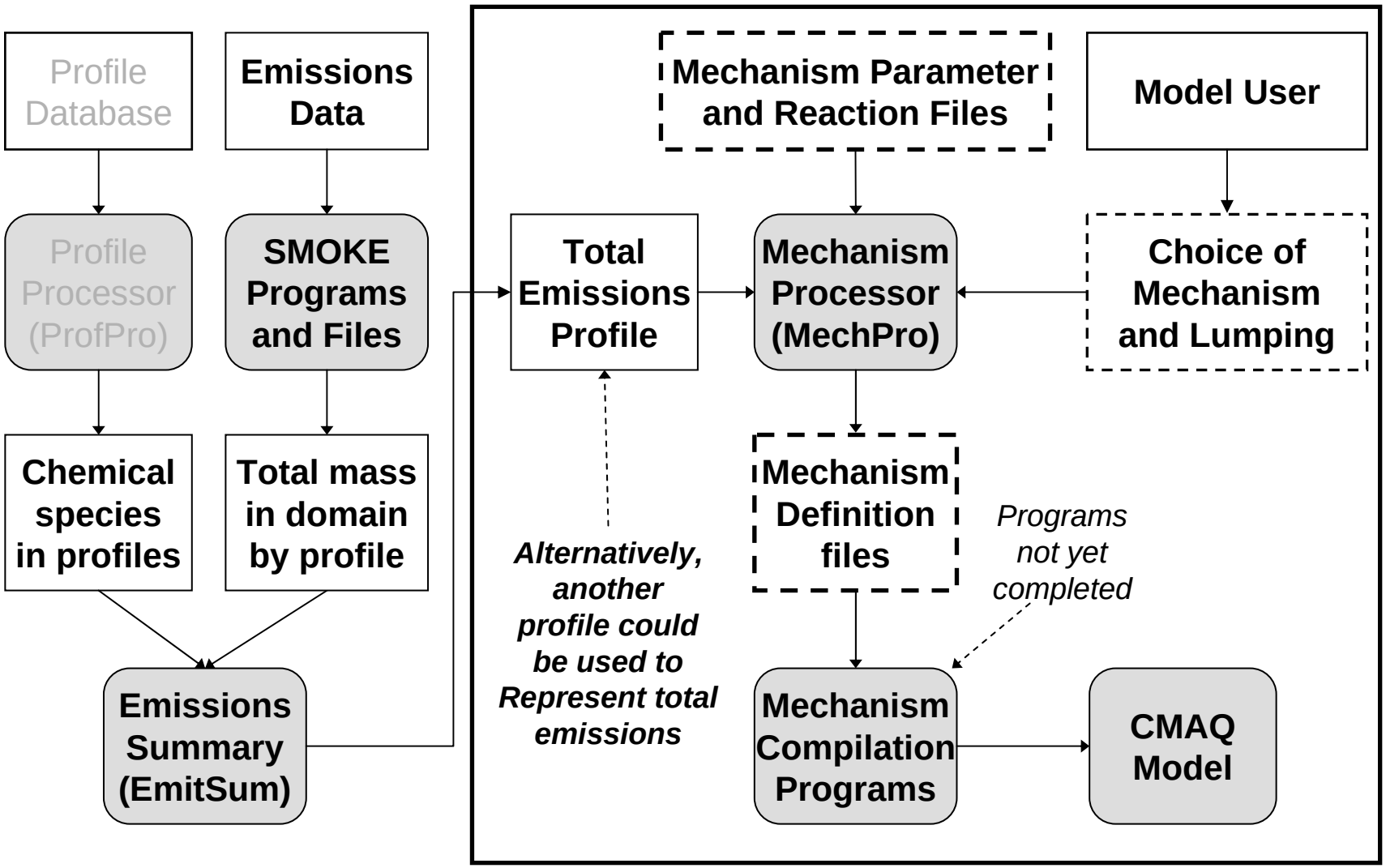
# SPECIATION PROCESSING PROGRAMS



# MECHANISM PROCESSING PROGRAMS

- Mechanism Processor (MechPro)
  - Prepare mechanism files for CMAQ for adjusted parameter mechanisms given lumping and total emissions profile
  - Can also prepare mechanism files with new model species given reactions or mechanism parameters and lumping
  - Needed to implement full capabilities of SAPRC-99
- Emissions Summary program (EmitSum)
  - Prepares total emissions profile given total mass emissions of all profiles into modeling domain and profile compositions
  - Output used as input to MechPro to derive SAPRC-99 parameters that best fit emissions data for this domain
- CMAQ mechanism compiler (under preparation)
  - Compiles MechPro output into mechanism-dependent source files to be linked into CMAQ software

# MECHANISM PROCESSING PROGRAMS



# CURRENT STATUS

- The speciation and profile database development that can be carried out for current project is essentially complete
- Software development still underway
  - ProfPro, SpecPro, EmitSum, and MechPro run on PC systems but need to be integrated into SMOKE
  - Programs to compile MechPro output into CMAQ still in preparation
  - Programs still need to be tested and debugged
- CMAQ model simulation demonstrating applicability to reactivity assessment still need to be conducted
- Database spreadsheets, software and example input and output files are available at **http://www.cert.ucr.edu/~carter/emitdb**
- Databases have been incorporated into a PostgreSQL database programmed for online access

# ADDITIONAL WORK NEEDED

- Success requires that the database be **used.** It also needs to be **updated** in to include new profiles and mechanisms
- Assignments in database need to be reviewed and improved
  - Volatility assignments – need objective basis
  - Mixture assignments – many are arbitrary guesses
  - CB4 assignments not all on consistent basis
  - Other mechanism assignments – may be mistakes
- Need to think about to represent unknown mass and unassigned compounds
- Assignments should be placed in “true” database and made available on web. (On-line PostgreSQL database a start)
- Need central clearing house and standards for profiles
- **Need to extend speciation database needs to PM profiles**

# ACKNOWLEDGEMENTS


- Funding sources
  - University of Houston (project for Dr. Anthony Haymet): initial speciation database development
  - Subcontract to the UNC Carolina Environmental Programs for a project for the American Chemistry Council: completion of database, software development, and implementation
- Collaborators on UNC/ACC project
  - Mark Houhoux (now at EPA) -- work plan development
  - Jeff Vukovich (UNC) -- SMOKE implementation and testing
  - Chao-Jung Chien (UCR) -- CMAQ mechanism compiling
- Assistance and helpful discussions in data compilation
  - Paul Allen, CARB
  - Ron Ryan, EPA
  - Gerry Gipson, EPA
  - Greg Yarwood, Environ
- Programming for on-line PostgreSQL database
  - David Boss (personal initiative)