

Speciation in the National Emissions Inventory and Modeling Platforms

Karl Seltzer

EPA/OAR/OAQPS/AQAD/EIAG



Goals of the Training Session

Introduction to Speciation

- What pollutants require speciation?
- Why is speciation necessary?
- What does the speciation process “look” like?

Overview of Speciation Methods & Modeling Tools Used at EPA

- How are HAPs in the NEI incorporated into modeling platforms?
- SPECIATE, S2S-Tool, HAP-augmentation and PM_{2.5} speciation in EIS.

Introduction to Speciation

The National Emissions Inventory

The National Emissions Inventory

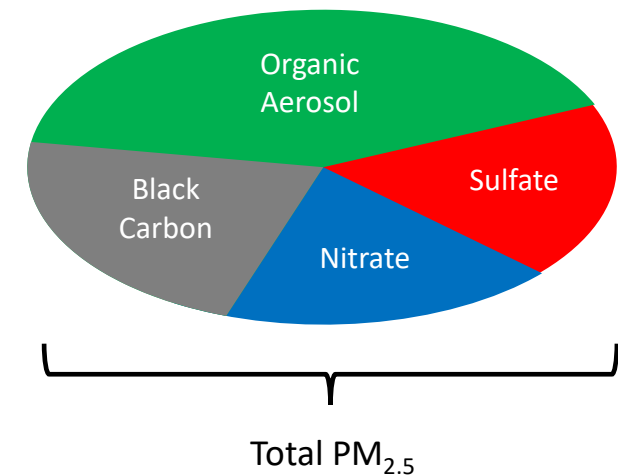
- The NEI is a national compilation of air emission estimates of criteria air pollutants (CAPs) and their precursors, and hazardous air pollutants (HAP).

CAPs in the NEI	CAP Precursors in the NEI	HAPs in the NEI
PM_{2.5} , PM₁₀ , Carbon Monoxide, Lead, Sulfur Dioxide	Ammonia (precursor to PM), Volatile Organic Compounds (precursor to ozone and PM), Nitrogen Oxides (precursor to ozone and PM)	Benzene, Manganese, + 187 others

- Particulate matter** and **volatile organic compounds** are not homogenous → composed of chemically distinct compounds.
 - Accurately speciating PM_{2.5} and VOC is essential for source attribution research, evaluating photochemical models, and atmospheric chemistry/transport processes.
- Most HAP are speciated PM_{2.5} or VOC. For example, lead and benzene, respectively.

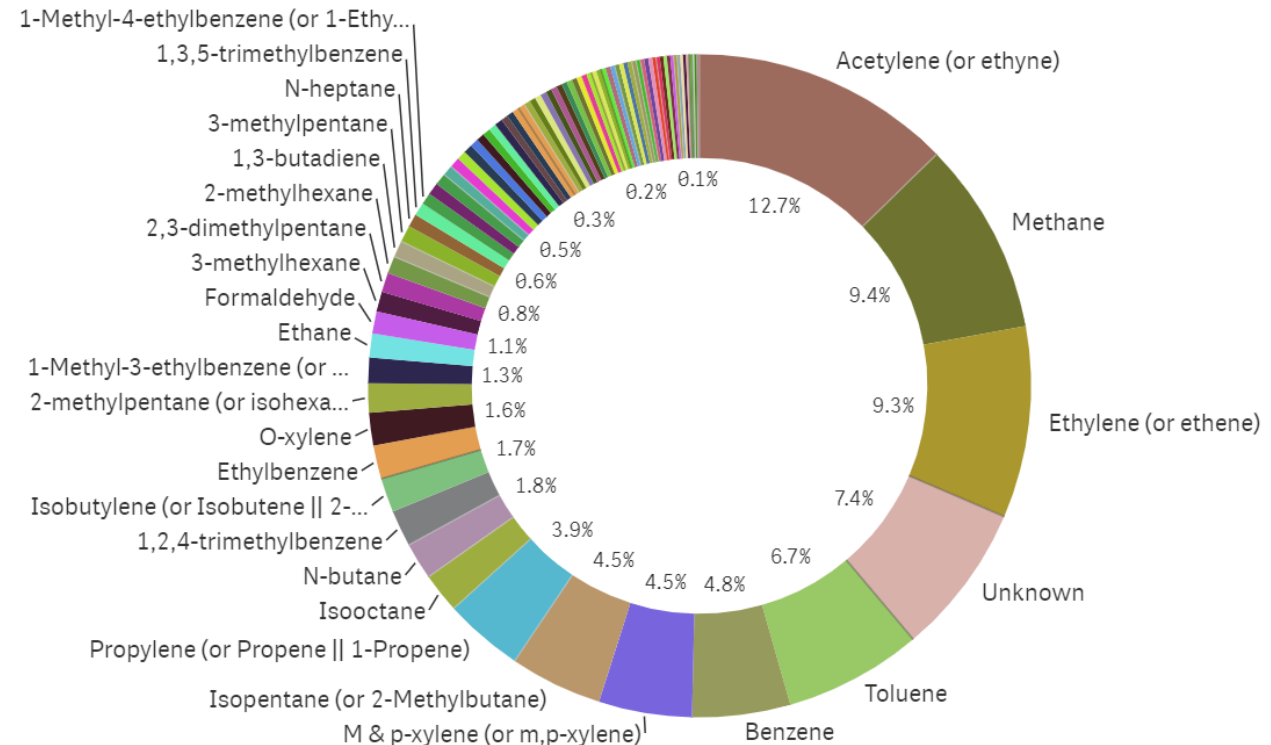
PM_{2.5} Speciation

- The NEI includes primary PM_{2.5} and select speciated PM_{2.5} components (e.g., sulfate, black carbon).
- Photochemical models transport and deposit speciated PM_{2.5}. In addition, these models estimate secondarily formed PM_{2.5} from precursors (e.g., SO₂, VOCs).
- Speciated PM_{2.5} within photochemical models is more chemically explicit than the components contained within the NEI.



VOC Speciation

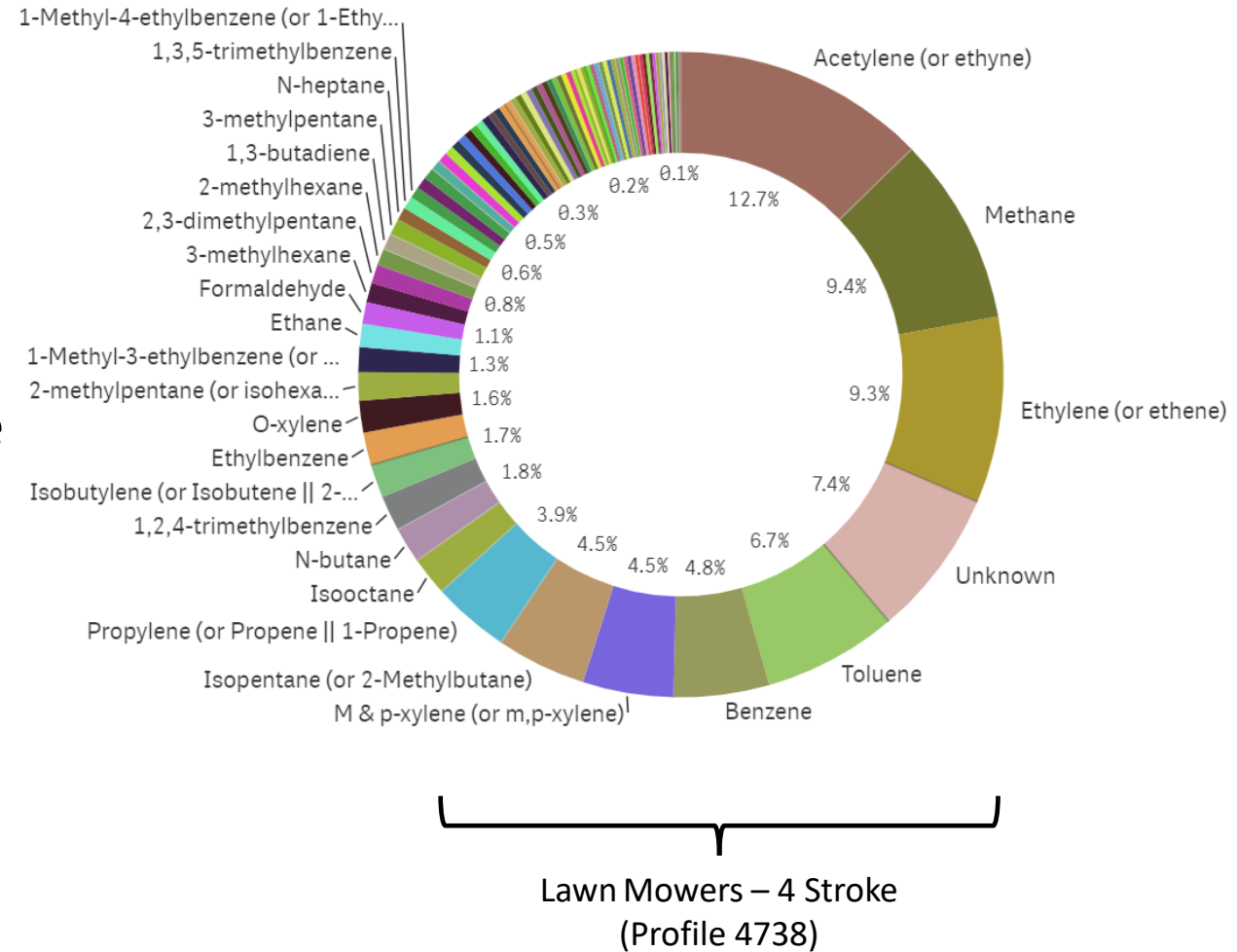
- The NEI includes total VOC emissions. Some HAP are also VOCs (e.g., benzene) and are inventoried separately.
- Photochemical models require VOC emissions to be speciated. Not all compounds equally participate in atmospheric chemistry and compounds vary in toxicity.
- Compared to $PM_{2.5}$ speciation, VOC speciation for air quality modeling features greater chemical resolution.



Lawn Mowers – 4 Stroke (Profile 4738)
 This is an explicit profile provided by SPECIATE

TOG vs. VOC

- Some VOC react slowly and have limited effects on local or regional ozone. Those VOC, determined to have low photochemical reactivity, may be excluded from the VOC definition.
- Total Organic Gases (TOG) includes these [exempt species](#) (e.g., methane, ethane).
- Photochemical models also include these gases. Inventoried VOC must be converted to TOG, and this process is accomplished using the TOG/VOC ratio of speciation profiles.



What does speciation “look” like?

- Source XYZ emits 100 tons of VOC per year.
- The total organic gas (TOG) profile of these emissions is 50% ethanol and 50% acetone.
- The photochemical modeler will be using the CB6R5_AE7 mechanism in CMAQ to model the effects of these emissions.

$$100 \frac{\text{tons VOC}}{\text{year}} \times 2 \frac{\text{tons TOG}}{\text{tons VOC}} \times 50\% \text{ ETOH} = 100 \frac{\text{tons ETOH}}{\text{year}}$$

$$100 \frac{\text{tons VOC}}{\text{year}} \times 2 \frac{\text{tons TOG}}{\text{tons VOC}} \times 50\% \text{ ACET} = 100 \frac{\text{tons ACET}}{\text{year}}$$

- In CB6R5_AE7: ethanol → ETOH; acetone → ACET

Detailed Overview of Speciation Methods Used by OAAQPS

Inventory to Modeling Platform

- A modeling platform consists of all the emissions inventories and ancillary data files (e.g., profile-to-SCC cross reference files) used for emissions modeling, as well as the meteorological, initial condition, and boundary condition files needed to run the air quality model.



Example:

100 TPY VOC



50 TPY ETOH, 50 TPY ALD2

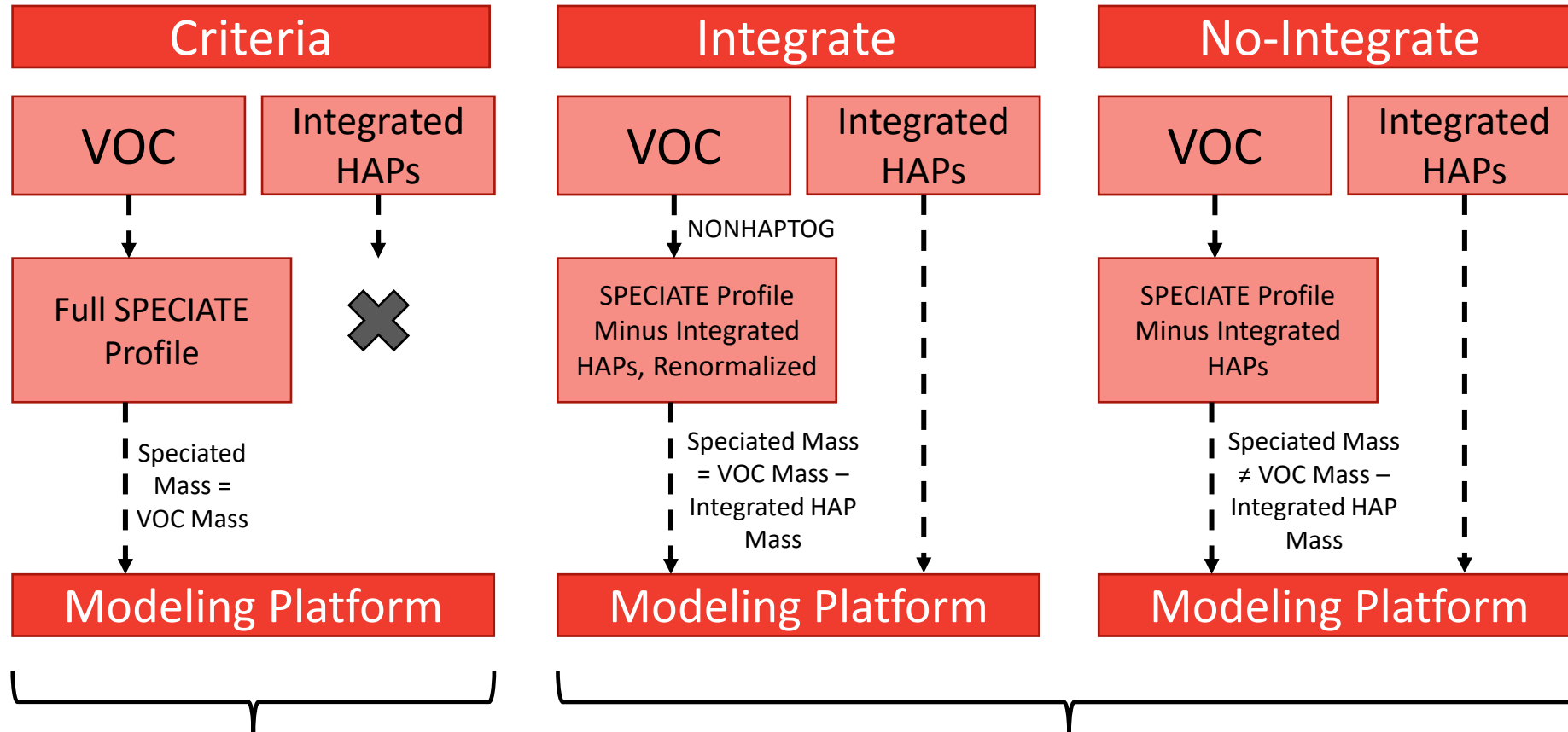
HAP Treatment in Modeling Platforms

- Recall that HAP are in the NEI and often components of VOC or PM_{2.5}. These emissions might better reflect what is emitted than what is in a speciation profile. How are these speciated components incorporated into a modeling platform?
 - **Integrate, HAP-use:** NONHAPTOG is calculated by subtracting the HAP mass from the inventory from the VOC mass. A NONHAPTOG speciation profile, renormalized to 100%, is then used to speciate the NONHAPTOG mass.
 - **No-Integrate, HAP-use:** A NONHAPTOG speciation profile, which is not renormalized to 100%, is used to speciate the VOC mass from the inventory. The reason for not renormalizing the profile will be discussed later.
- In addition, HAPs can be incorporated into a modeling platform using speciation profiles. This method is called **criteria** speciation.
- The method utilized when developing a modeling platform is sector specific.

Which HAPs are Incorporated?

- In addition, the HAPs incorporated (“integrated”) from the inventory into the modeling platform are specific to each sector and chemical mechanism.
- In recent years, CB6R3_AE7 has been the primary chemical mechanism used at the EPA for regulatory applications.
 - Naphthalene (NAPH), benzene (BENZ), acetaldehyde (ALD2), formaldehyde (FORM), and methanol (MEOH) are explicit HAPs in this mechanism. These compounds are collectively referred to as **NBAFM**.
 - Chemical mechanisms require “lumped” species to represent certain compounds due to computational restraints and knowledge gaps.
- Since NBAFM are explicitly modeled in CB6R3_AE7, these species have become the default collection of “integrated” species at the EPA.
- MOVES, the EPA’s mobile emissions model, features additional species that are explicitly modeled (e.g., ethanol). These species are explicit in the chemical mechanism.

Visualizing HAP Treatment



HAPs from the inventory are not incorporated into the modeling platform

HAPs from the inventory are incorporated into the modeling platform

Types of Modeling Platforms

- Currently, the EPA generates two types of modeling platforms, (1) HAP-CAP and (2) CAP-only. The speciation methods for these two platform types often overlap and vary by sector:

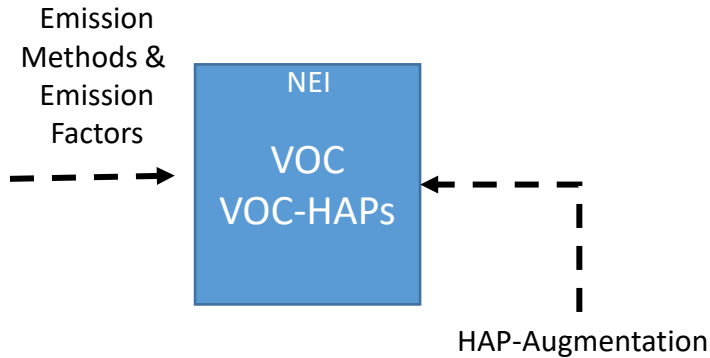
	Integrate, HAP-use	No-Integrate, HAP-use	Criteria
HAP-CAP	Onroad, Nonroad, Nonpoint	Point, Nonpoint (when HAP > VOC or no emissions of integrated species)	n/a
CAP-Only	Onroad, Nonroad, Nonpoint	n/a	Point, Nonpoint (when HAP > VOC or no emissions of integrated species)

BREAK

We will resume in 10 minutes

Detailed Overview of Speciation Modeling Tools Used at EPA

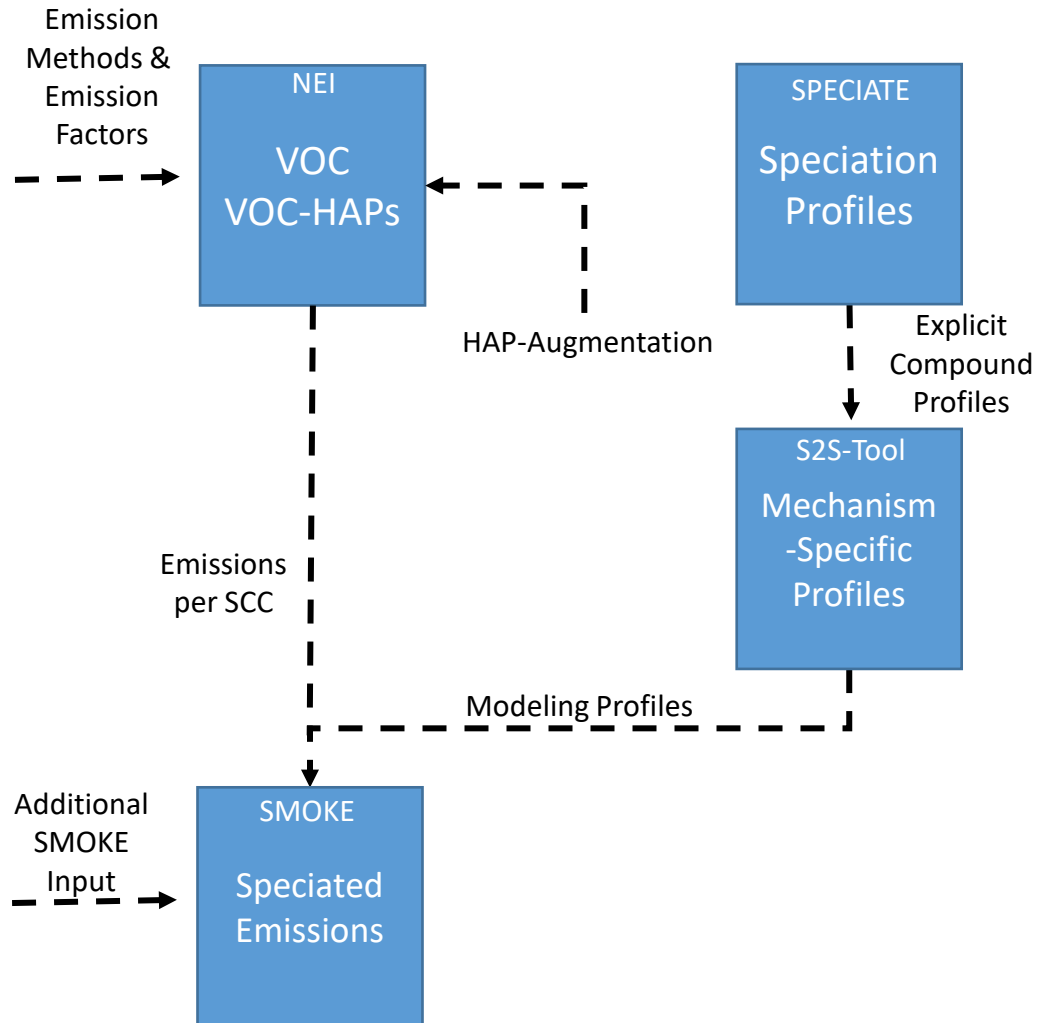
VOC Speciation Tools at the EPA: Overview



The National Emissions Inventory

- EIAG incrementally updates methods to better characterize emissions in the NEI.
- Both the VOC and VOC-HAPs are housed in the NEI and used in modeling platforms.
- NEI VOC-HAP emissions are largely generated using “HAP-augmentation” in EIS (nonpoint, point), provided by MOVES (mobile), or submitted/added from the Toxics Release Inventory (point).

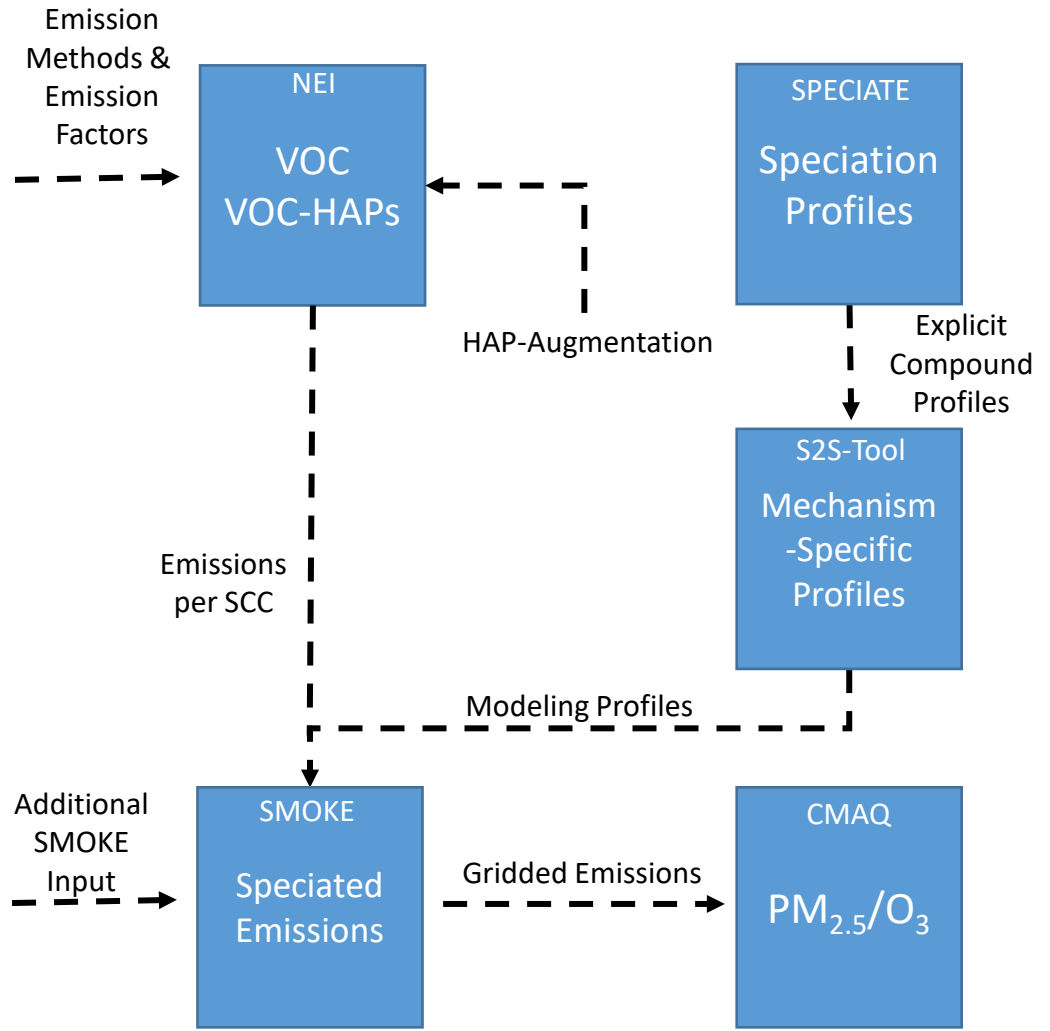
VOC Speciation Tools at the EPA: Overview



Modeling Platform VOC Speciation

- SPECIATE houses explicit VOC speciation profiles for various emissions sources.
- The S2S-Tool translates these explicit VOC profiles into chemical mechanism-specific profiles.
 - S2S: SPECIATE-to-SMOKE
 - Formerly done by the Speciation Tool.
 - e.g., translates octane to 8 PAR and ethanol to 1 ETOH for CB6R5; model species needed for CMAQ and CAMx.
- SMOKE combines data from the NEI, S2S-Tool, and other sources to generate gridded, speciated photochemical modeling files.

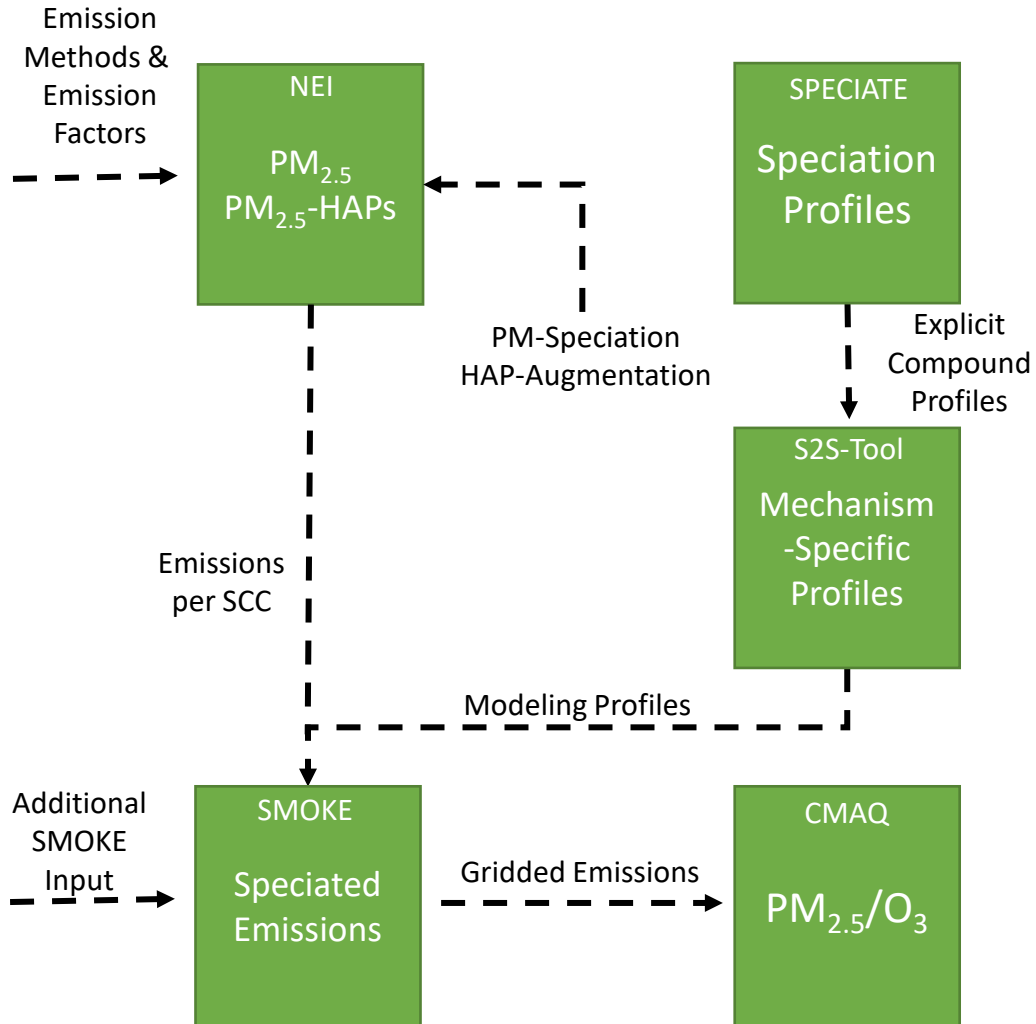
VOC Speciation Tools at the EPA: Overview



Photochemical Modeling

- Gridded emission files generated by SMOKE are used in photochemical models, such as CMAQ or CAMx, to simulate air quality.
 - Different chemical-mechanisms simulate the chemistry of the atmosphere.

PM_{2.5} Speciation Tools at the EPA: Overview

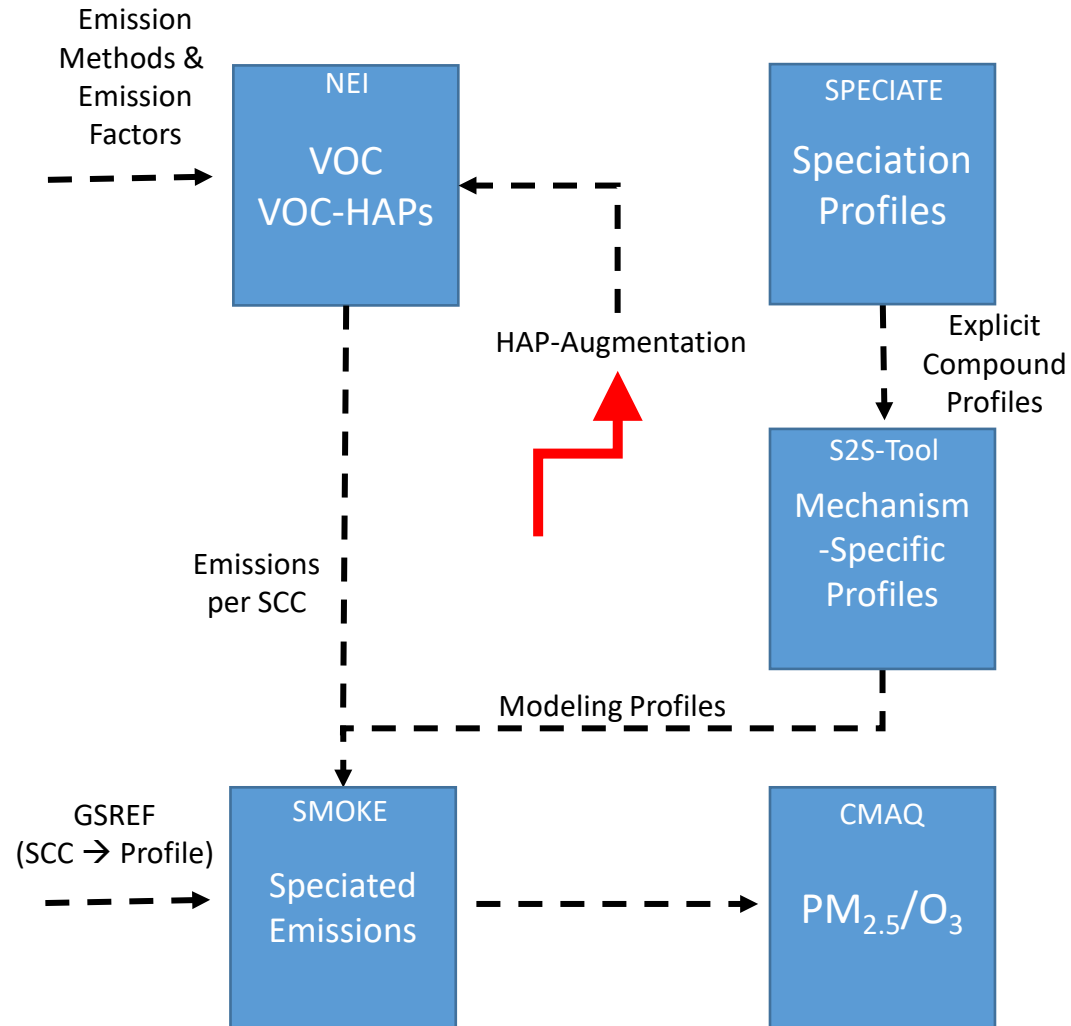


- Methods are conceptually the same as VOC speciation.
- PM-Speciation is performed in EIS to generate a broad estimate of speciated PM_{2.5} emissions (EC, OC, SO₄, NO₃, PMOTHR).

30,000 Foot Overview of Tools

- The translation of inventory pollutants (VOC and PM_{2.5}) to model-ready files requires several datasets and tools:
 - Continual updates and improvements to source composition data (SPECIATE, HAP-Augmentation).
 - Translation of source composition data into data that is compatible with the desired chemical mechanism of the modeler (S2S-Tool).
 - Processing of an emissions inventory, speciation data, and other ancillary files into model-ready gridded emissions (SMOKE).

HAP-Augmentation in our workflow



HAP-Augmentation

CAPs (e.g., VOC) are used to estimate HAP emissions within EIS. EPA uses a similar process is performed for PM-Speciation for the NEI.

**Example HAP-
Augmentation profile:**

Augmentation Type	Profile Name	Input Pollutant Code	Input Pollutant Description	Output Pollutant Code	Output Pollutant Description	Multiplication Factor	SCC Assignment	SCC Description Level 1	SCC Description Level 2	SCC Description Level 3
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	108883	Toluene	0.00157	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	111422	Diethanolamine	0.00113	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	122996	Phenyl Cellosolve	0.00103	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	171	Glycol Ethers	0.000659	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	110543	Hexane	0.00054	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products
HAP	2460100000_HAP-VOC	VOC	Volatile Organic Compounds	85449	Phthalic Anhydride	0.00027	2460100000	Solvent Utilization	Miscellaneous Non-industrial: Consumer and Commercial	All Personal Care Products













HAP-Augmentation

- If State ABC submits 100 TPY of VOC for SCC 2460100000 for County XYZ to the NEI, these emissions will be used to estimate HAP emissions for that SCC.
- SCC 2460100000 is assigned to HAP-Augmentation profile “2460100000_HAP-VOC.”
 - EIS metadata (at right) indicates this profile came from SPECIATE profile 95823.
- HAP-Augmentation would estimate 0.157 TPY of toluene for SCC 2460100000 for County XYZ.


AUGMENTATION PROFILE NAME

Profile Name: 2460100000_HAP-VOC
 Profile Description:
 Profile Source: SPECIATE v5.2; profile 95823
 Profile Comment:
 Augmentation Type: HAP
 Input Pollutant Code: VOC
 Input Pollutant Description: Volatile Organic Compounds
 Input Pollutant Comment:

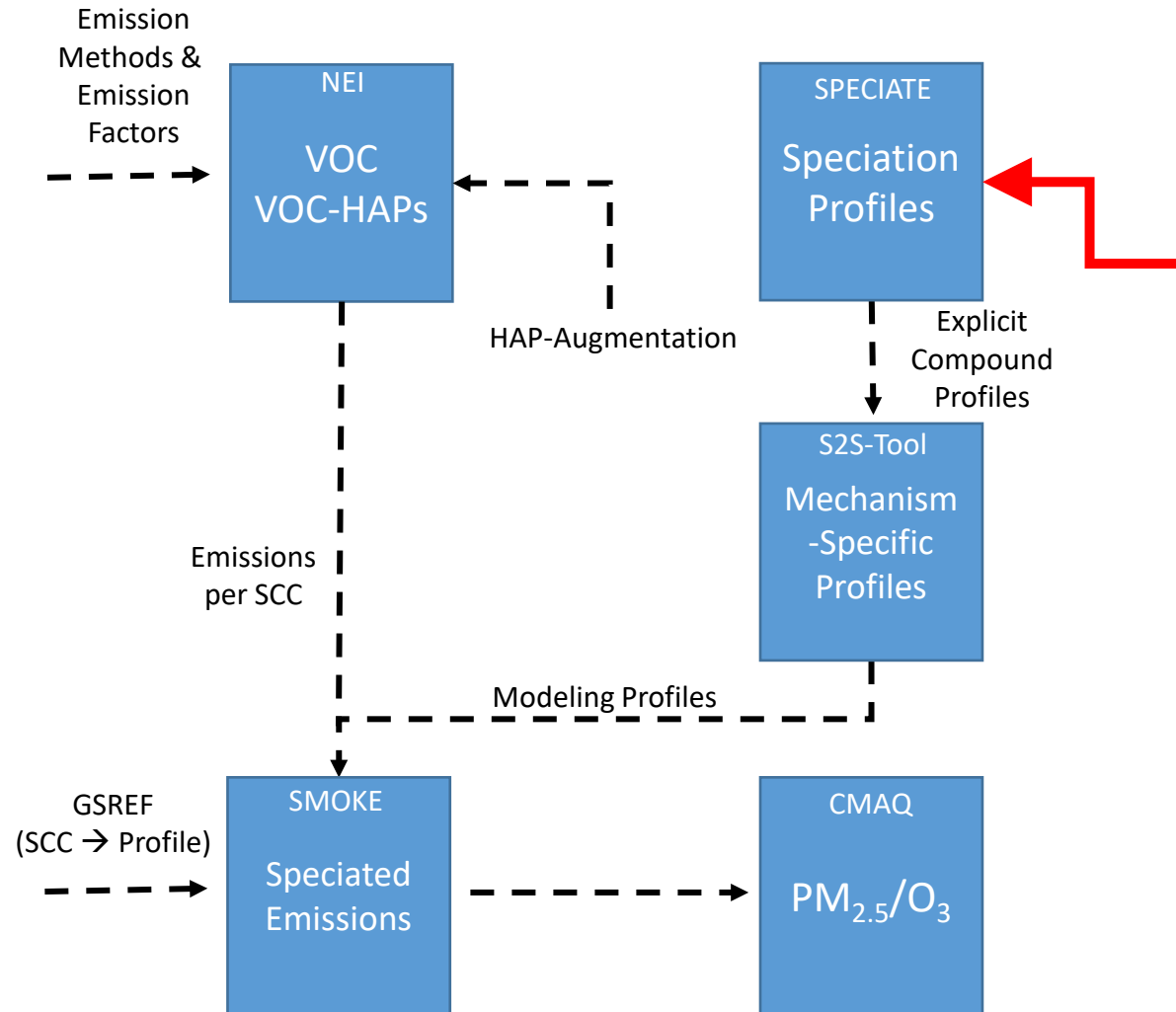
AUGMENTATION PROFILE FACTORS

Output Pollutant Code	Output Pollutant Description	Multiplication Factor		
110543	Hexane	0.00054		
122996	Phenyl Cellosolve	0.00103		
85449	Phthalic Anhydride	0.00027		
171	Glycol Ethers	0.000659		
108883	Toluene	0.00157		
111422	Diethanolamine	0.00113		

AUGMENTATION PROFILE ASSIGNMENTS

SCC Assignment	EIS Facility Assignment	EIS Process Assignment	State Abbreviation
2460100000			

SPECIATE in our workflow



SPECIATE

- SPECIATE is an Access database that provides source-specific profiles of organic gases, PM_{2.5}, and mercury.
- The database can be [downloaded](#) or explored via a [web-based](#) tool.
- SPECIATE v5.3 soon to be published.
 - 50 new profiles were added to the database; total database is now 6,895 profiles.



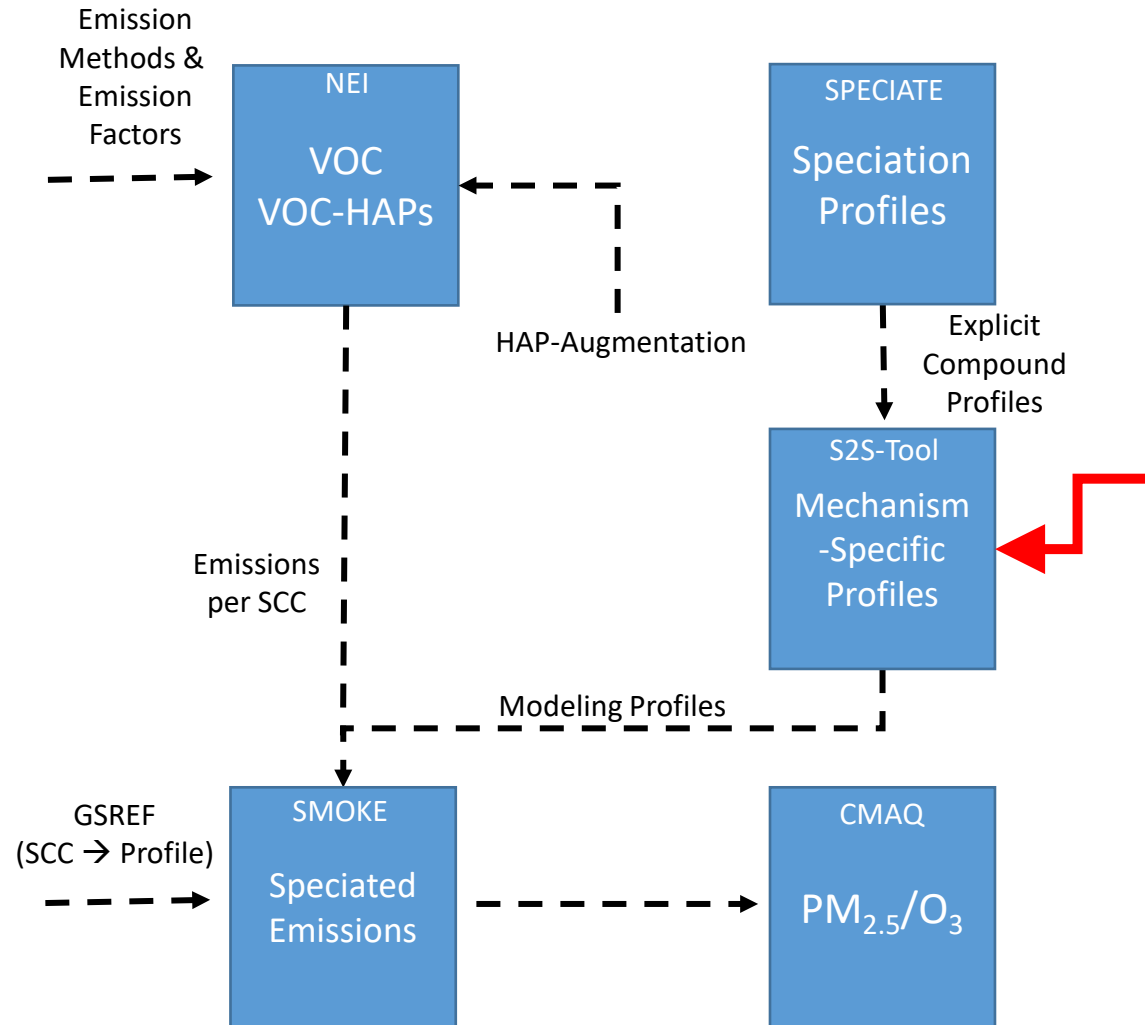
SPECIATE

- Appending profiles to SPECIATE requires EPA developers to compile a “SPECIATE workbook.”
 - Includes raw data, processed data, and other metadata that fits the format of existing SPECIATE tables.
 - This workbook and the profile(s) go through a QSCORE, which numerically grades the quality of the profile and data used to create it.
- The SPECIATE Access database includes queries that users can use to export data for downstream use, including the S2S-Tool inputs.

Results from example query:

PROFI	PROFILE_NAME	MAST	TEST_Y	TOG_to	SPECIES_NAME	WEIGHT_PER
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Acetylene (or ethyne)	12.69
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Methane	9.42
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Ethylene (or ethene)	9.34
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Unknown	7.4
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Toluene	6.73
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Benzene	4.78
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	M & p-xylene (or m,p-x	4.55
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Isopentane (or 2-Methy	4.51
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Propylene (or Propene	3.87
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Isooctane	1.94
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	N-butane	1.8
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	1,2,4-trimethylbenzene	1.79
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Isobutylene (or Isobute	1.66
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Ethylbenzene	1.64
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	O-xylene	1.58
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	2-methylpentane (or is	1.42
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	1-Methyl-3-ethylbenze	1.25
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Ethane	1.16
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	Formaldehyde	1.08
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	3-methylhexane	0.97
4738	Lawn Mowers - 4 strc	TOG	1997	1.118318	2,3-dimethylpentane	0.95

S2S-Tool in our workflow



S2S-Tool

- U.S. EPA's S2S-Tool generates speciation input files (GSPRO, GSCNV) for SMOKE.
 - Big picture: The S2S-Tool translates SPECIATE data for use in SMOKE.
- In January 2023, S2S-Tool v1.0 was publicly released. Recently, v2.0 of the Tool was published.
 - Available via the [USEPA's GitHub](#).
- Prior to the S2S-Tool's development, we used the "Speciation Tool." Advantages of the S2S-Tool include:
 - Developed/maintained in-house.
 - Python-based and includes "mechanism mapping" scripts (more on this later).
 - Generates semi-volatile primary organic aerosol profiles (more on this later).
 - Generates files to allow speciation outside of MOVES.

S2S-Tool Interface

```
#####
### S2S-Tool: U.S. EPA's SPECIATE-to-SMOKE Tool generates GSPRO and GSCNV files that are used by
### SMOKE to generate gridded emissions for photochemical modeling. All profiles are in the U.S.
### EPA's SPECIATE database, and the type of GSPRO/GSCNV files generated by the S2S-Tool are
### dependent on the speciation methods employed for the desired modeling platform. SPECIATE is
### a database of organic gas, particle, and mercury speciation profiles. These profiles provide
### greater specificity than what is needed for a chemical mechanism within a photochemical model.
### The S2S-Tool bridges this gap and translates SPECIATE data into a format that is chemical
### mechanism specific.
#####

#####
### User Input
### Photochemical mechanism; options include CB6R3_AE7, CB6R5_AE7, CB7_AE7, CRACMMv1.0, SAPRC07TC_AE7,
MECH_BASIS = 'CB6R3_AE7'
### Output type; options include VOC, PM
OUTPUT      = 'VOC'
### Select run type; options include CRITERIA, INTEGRATE, NOINTEGRATE
RUN_TYPE    = 'CRITERIA'
### Select air quality model; options include CMAQ, CAMX
AQM         = 'CMAQ'
### Assign acceptable deviation from 100% allowable. Applies only to gas profiles.
TOLERANCE   = 0.05 # 0.05 = 5%
### If applicable, input pollutant for toxics INTEGRATE scenario
TOX_IN      = 'NONHAPTOG' # NONHAPTOG ; TOM ; RESID_PM
### CAMx FCRS file name & path:
FCRS_FILE   = './input/camx_fcrs.profile.csv'
### MW file name & path:
MW_FILE     = './input/mechanism_mw.csv'
### mechanism_forImport file name & path:
M4I_FILE    = './input/mechanism_forImport_SPECIATEv5_3.csv'
### pm_mech file name & path:
PMM_FILE    = './input/mech_pm_ae5_ae6_ae8_crl.csv'
### tbl_tox file name & path:
TOX_FILE    = './input/tbl_tox_NBAFM.csv' # 'tbl_tox_NBAFM.csv' ; 'tbl_tox_TOM.csv' ; 'tbl_tox_RESID_PM
#####
```

- S2S-Tool meets the needs for:
 - Various modeling platform speciation setups, integrated HAP treatments, treatments of primary organic aerosol, mechanisms (9 gas, 2 PM), and models (2 AQMs).
 - These variations yield 64 runtime configurations.
- To run, a user simply executes the S2S.main.py script in the directory, with the “User Input” selections filled out to meet the needs of the SMOKE modeler.
- A [User’s Guide](#) is available detailing all input, modules, and methods of the Tool.

S2S-Tool Output

GSCNV:

```
#S2S_AQM          CMAQ
#S2S_CAMX_FCFS   Not Applicable
#S2S_MW           ./input/mechanism_mw.csv
#S2S_MECH_BASIS  CB6R3_AE7
#S2S_RUN_TYPE     CRITERIA
#S2S_RUN_DATE     2023-08-09
#S2S_TBL_TOX     Not Applicable
#BY PROFILE
VOC              TOG          2402          1.03316458
VOC              TOG          3150          1.08471635
```

GSPRO:

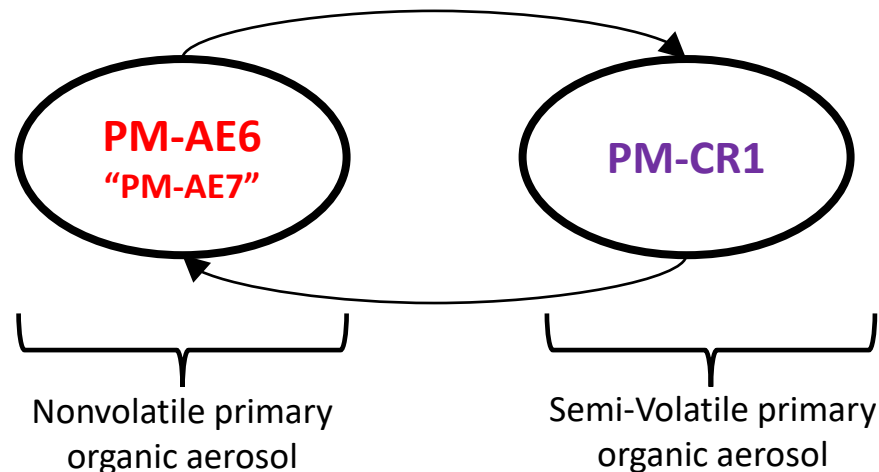
```
#S2S_AQM          CMAQ
#S2S_CAMX_FCFS   ./input/camx_fcfs.profile.csv
#S2S_MW           ./input/mechanism_mw.csv
#S2S_MECH_BASIS  CB6R3_AE7
#S2S_RUN_TYPE     CRITERIA
#S2S_RUN_DATE     2023-08-09
#S2S_TBL_TOX     Not Applicable
2402             TOG          ACET          3.210000E-02  5.810000E+01  3.210000E-02
2402             TOG          ALDX          1.160839E-03  5.810000E+01  1.160839E-03
2402             TOG          ETOH          2.280000E-02  4.610000E+01  2.280000E-02
2402             TOG          IVOC          1.170000E-02  2.405000E+02  1.170000E-02
2402             TOG          KET           6.038634E-02  2.900000E+01  6.038634E-02
2402             TOG          PAR           4.552528E-01  1.400000E+01  4.552528E-01
2402             TOG          TOL           1.296000E-01  9.210000E+01  1.296000E-01
2402             TOG          XYLMN        2.870000E-01  1.062000E+02  2.870000E-01
2402             TOG          NMOG          1.000000E+00  1.000000E+00  1.000000E+00
```

- The contents of the GSCNV and GSPRO (ancillary file inputs to SMOKE) vary depending on:
 - The HAPs incorporated from the inventory.
 - HAP treatment methods (e.g., Integrate, HAP-use; No-Integrate, HAP-use; Criteria).
 - Photochemical model and chemical mechanism selected by modeler.

Semi-Volatile Organic Aerosol

- Organic PM_{2.5} can partition between the aerosol and gas-phase. Older chemical mechanisms exclusively treated primary organic aerosol as non-volatile and do not allow partitioning.
- S2S-Tool enables all PM_{2.5} profiles in SPECIATE to be cross compatible with different chemical mechanisms and treatments of primary organic aerosol.
 - The POA_VolatilityBins input file in S2S-Tool allows users to specify source-specific organic aerosol volatility profiles.

PM treatments in CMAQ:



Semi-Volatile Organic Aerosol

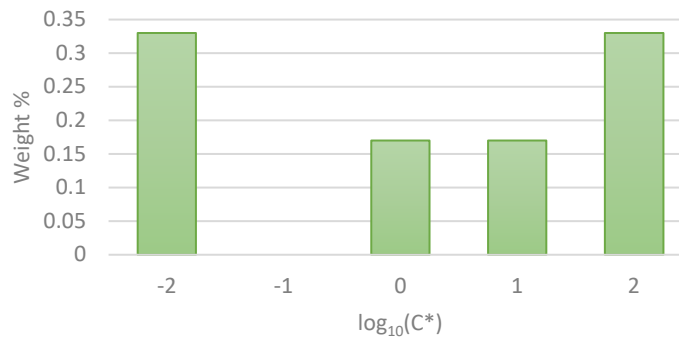
- **Example:** Profile 95873 - Catalytic Woodstove Burn Phase and Wood Composite
 - Entered into SPECIATE as a PM-AE6 profile:

PM-AE6:

PROFILE_CODE	SPECIES_NAME	WEIGHT_PERCENT
95873	PAL	1.49E-06
95873	PCA	3.21E-04
95873	PCL	3.97E-03
95873	PEC	5.63E-02
95873	PFE	5.95E-05
95873	PH2O	3.02E-03
95873	PK	9.88E-03
95873	PMG	8.77E-05
95873	PMN	6.15E-06
95873	PMOTHR	3.37E-03
95873	PNA	1.39E-04
95873	PSI	5.95E-04
95873	PSO4	1.26E-02
95873	PTI	2.00E-06
95873	PNCOM	4.38E-01
95873	POC	4.72E-01

S2S-Tool translates this profile into a PM-CR1 profile using POA_VolatilityBins:

May et al., 2013



PNCOM + POC is summed and distributed using this volatility profile.

PM-CR1:

PROFILE_CODE	SPECIES_NAME	WEIGHT_PERCENT
95873	PAL	1.49E-06
95873	PCA	3.21E-04
95873	PCL	3.97E-03
95873	PEC	5.63E-02
95873	PFE	5.95E-05
95873	PH2O	3.02E-03
95873	PK	9.88E-03
95873	PMG	8.77E-05
95873	PMN	6.15E-06
95873	PMOTHR	3.37E-03
95873	PNA	1.39E-04
95873	PSI	5.95E-04
95873	PSO4	1.26E-02
95873	PTI	2.00E-06
95873	ROC2ALK	3.00E-01
95873	ROC0ALK	1.55E-01
95873	ROC1ALK	1.55E-01
95873	ROC2ALK	3.00E-01

Mechanism Mappers

- “Mechanism mapping” refers to the process of mapping individual compounds within SPECIATE to the appropriate chemical mechanism-specific specie(s). This mapping file is a key component of the S2S-Tool’s input.
- Historically, this process has often been manually performed by mechanism developers. EPA has developed automated scripts for this process to make it more transparent.
- Available for CB6R3_AE7, CB6R5_AE7, CB7_AE7, CRACMMv1.0, SAPRC07TC_AE7, CB6R4_CF2, CB7_CF2, SAPRC07_CF2.

mechanism_forImport File:

```
Mechanism,SPECIES_ID,Species,Moles  
CB6R3_AE7,1,TOL,1  
CB6R3_AE7,1,PAR,3  
CB6R3_AE7,2,PAR,11  
CB6R3_AE7,3,XYLMN,1  
CB6R3_AE7,3,PAR,2
```

For example, (2-methylbutyl)cyclohexane is mapped to 11 moles of PAR in the CB6R3_AE7 mechanism.

Conclusions

- VOC and PM_{2.5} speciation are important steps in the compilation of our emissions inventories and in air quality modeling.
- Various methods are employed at the EPA to incorporate HAPs from a base inventory into the modeling platforms.
- NEI VOC-HAP emissions are largely generated using “HAP-augmentation” in EIS (nonpoint, point), provided by MOVES (mobile), or submitted (point).
- Speciating emissions from a base inventory for a modeling platform requires many databases and tools:
 - SPECIATE, HAP-Augmentation, S2S-Tool, SMOKE
- We are continually looking to make updates and improvements to our tools and datasets.
Please reach out with any questions, comments, or concerns regarding speciation!

Thank you – any questions?
seltzer.karl@epa.gov