THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

ENVIRON

New Tools to Generate Spatial Surrogate and Speciation Profile Inputs to SMOKE

Alison Eyth, Limei Ran, R. Partheepan
Carolina Environmental Program
University of North Carolina at Chapel Hill

Greg Yarwood, Michele Jimenez, Sunja Rao
ENVIRON International Corporation

OAQPS Motivation for Tool Development

- Make it easier to produce, maintain and track SMOKE ancillary files
- Provide a transparent and self-documenting approach to modeling
- Integrate data creation tools with Emissions Modeling Framework (EMF) Data Management System
- Provide needed new options for integrated HAP and CAP modeling

Surrogate Tool

- Creates dozens of spatial surrogates in one run for multiple regions
  - 66 surrogates for U.S., plus Mexico
- Spatial Surrogates:
  - Used to spatially allocate county-based emissions to the grid cells (or census tracts) used by air quality models
  - Stored in files with rows formatted as srg code, county id, row, col, fraction e.g., 100 37001 10 36 0.3
  - Fractions sum to 1 for each county

Surrogate Tool

- Creates surrogates from Shapefile DBF attributes or functions of attributes
  - e.g., Urban Population, Industrial Land
  - Uses Spatial Allocator’s srgcreate program
- Merges surrogates
  - 0.75*Total Road Miles + 0.25*Population
- Gapfills surrogates to prevent dropping of emissions - up to four levels deep
  - e.g., Population gapfills Urban Population

Surrogate Tool Input Files

- .CSV files that can be managed by the EMF
  - **Global control variables file**
  - **Shapefile catalog**
  - **Surrogate specification file**
  - **Generation control file**
  - **Surrogate Code file**
  - **GRIDDESC (not .CSV but is ASCII)**
- Input files provide all information needed to regenerate the surrogates

Some Global Control Variables

OUTPUT_FILE_TYPE	RegularGrid
OUTPUT_GRID_NAME	US36KM_148x112
GRIDDESC	/srgtool/GRIDDESC
OUTPUT DIRECTORY	/data/ US36KM_148x112
OUTPUT SRGDESC FILE	SRGDESC.txt
OVERWRITE OUTPUT FILES	YES

Shapefile Catalog File

- Table, with required and optional columns that describe information for each Shapefile
- Required columns:
  - Shapefile name: pophu2k
  - Directory: /data/shapefiles/us
  - Ellipsoid: Sphere (others available)
  - Projection: lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97 (PROJ.4 syntax)
  - Shape type: Polygon (Line or Point)

Surrogate Specification File

- Table with columns to describe how to create each surrogate
- Required columns:
  - REGION: USA
  - SURROGATE: Water
  - SURROGATE CODE: 350
  - DATA SHAPEFILE: us_county_water02
  - DATA ATTRIBUTE: FIPSSTCO
  - WEIGHT SHAPEFILE: uslw1k
  - WEIGHT ATTRIBUTE: AREA

Surrogate Specification File

- Remaining Columns:
  - WEIGHT FUNCTION: COM1+COM2
  - FILTER FUNCTION: H20_CODE !=2
  - MERGE FUNCTION: 0.75*Total Road Miles+ 0.25*Population
  - SECONDARY SURROGATE: Navigable Waterway Activity
  - TERTIARY SURROGATE: Navigable Waterway Miles
  - QUARTERNARY SURROGATE: Land

Surrogate Generation Control File

REGION	SURROGATE	SRG CODE	GENERATE	QUALITY ASSURE
USA	Population	100	YES	YES
CAN	Population	100	NO	YES
USA	Urban Population	120	YES	YES
USA	Total Road Miles	240	NO	YES

Running the Surrogate Tool

- Open source software
- Runs on Linux, Unix, and Windows
- Requires Java 1.5
- Uses one command line argument:
    - Global control variables file (which refers to all other input files)

```
java gov.epa.surrogate.SurrogateTool control_variables_file
```

Surrogate Tool Output Files

- Surrogate files: One file for each surrogate and region
- Srgcreate scripts
- Srgmerge input files for gapfilling, merging
- Log file – with summary table at end
- Gridded Shapefiles of surrogate numerators from srgcreate for QA
- SRGDESC File

Example SRGDESC File

#GRID US36KM_148X112 *more grid info*

USA,100,”Population”,USA_100_NOFILL.txt

USA,120,”Urban Population”, USA_120_FILL.txt

USA,240,”Total Road Miles”, USA_240_FILL_NORM.txt

CAN,240,”Total Road Miles”, /ext_srgs/can/road_miles.txt

Related Java Tools

- In same package as Surrogate Tool
- Normalization tool
- QA Tool: produces several .CSV summaries by county
  - Gapfilling report
  - Counties with sum not equal to 1
  - Counties with no data
- Java implementation of merging
- Java implementation of gapfilling

Speciation Modeling Data Tool

- Generates chemical speciation profiles for emission models (e.g., GSPRO, GSCNV)
- Supports any defined chemical mechanism
- Supports both VOC and PM species
- Supports integrated and non-integrated HAPS
- Supports both active and tracer toxics

Speciation Tool
Software Requirements

- Free, open source software
  - PostgreSQL: www.postgresql.org
  - Perl: www.perl.com
  - Perl DBI (Database Interface module)

Speciation Tool Inputs

- Initialization Files
  - CBIV and SAPRC99 chemical mechanisms
  - SPECIATE 4.0 profile weights for VOCs and simplified PM2.5
  - SPECIATE 4.0 species data
  - Pre-computed speciation profiles

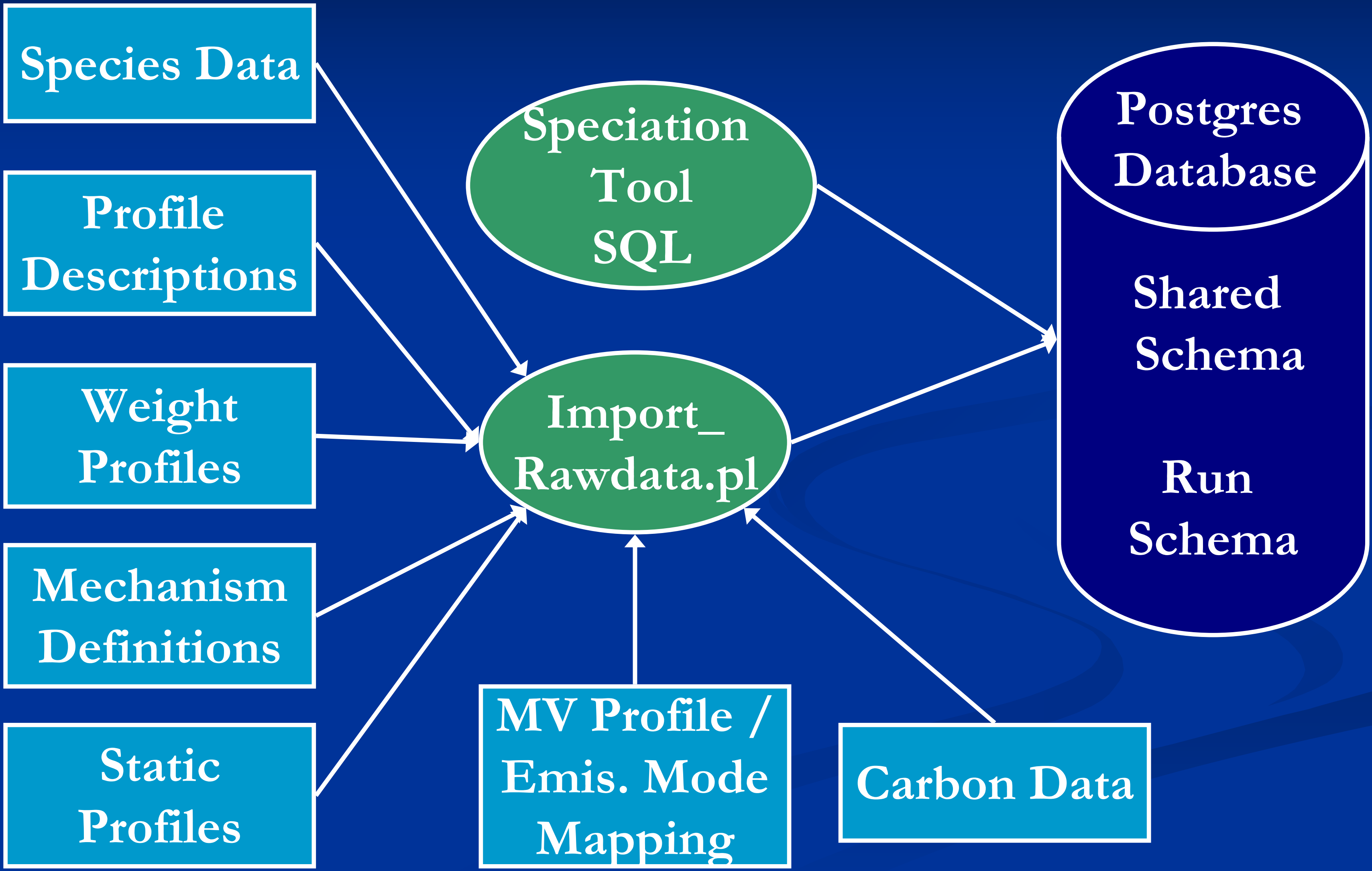
Speciation Tool Inputs – continued

- User Defined Inputs
  - Run Control File
    - Run options
    - Optional input file names
    - Output file names
  - Explicit HAPS treatment (optional)
  - Active/Tracer compounds (optional)
  - Profile Weights (optional)

Speciation Tool Initialization

- Creates tool database *sptoolv1*
- Creates *shared* schema within *sptoolv1*
  - *Shared* schema stores data shared by all runs
- Loads the Speciation Tool SQL modules into *shared* schema
- Loads default data into *shared* schema

Initialize Database



Species Data

Profile Descriptions

Weight Profiles

Mechanism Definitions

Static Profiles

Speciation Tool SQL

Import_ Rawdata.pl

MV Profile / Emis. Mode Mapping

Carbon Data

Postgres Database

Shared Schema

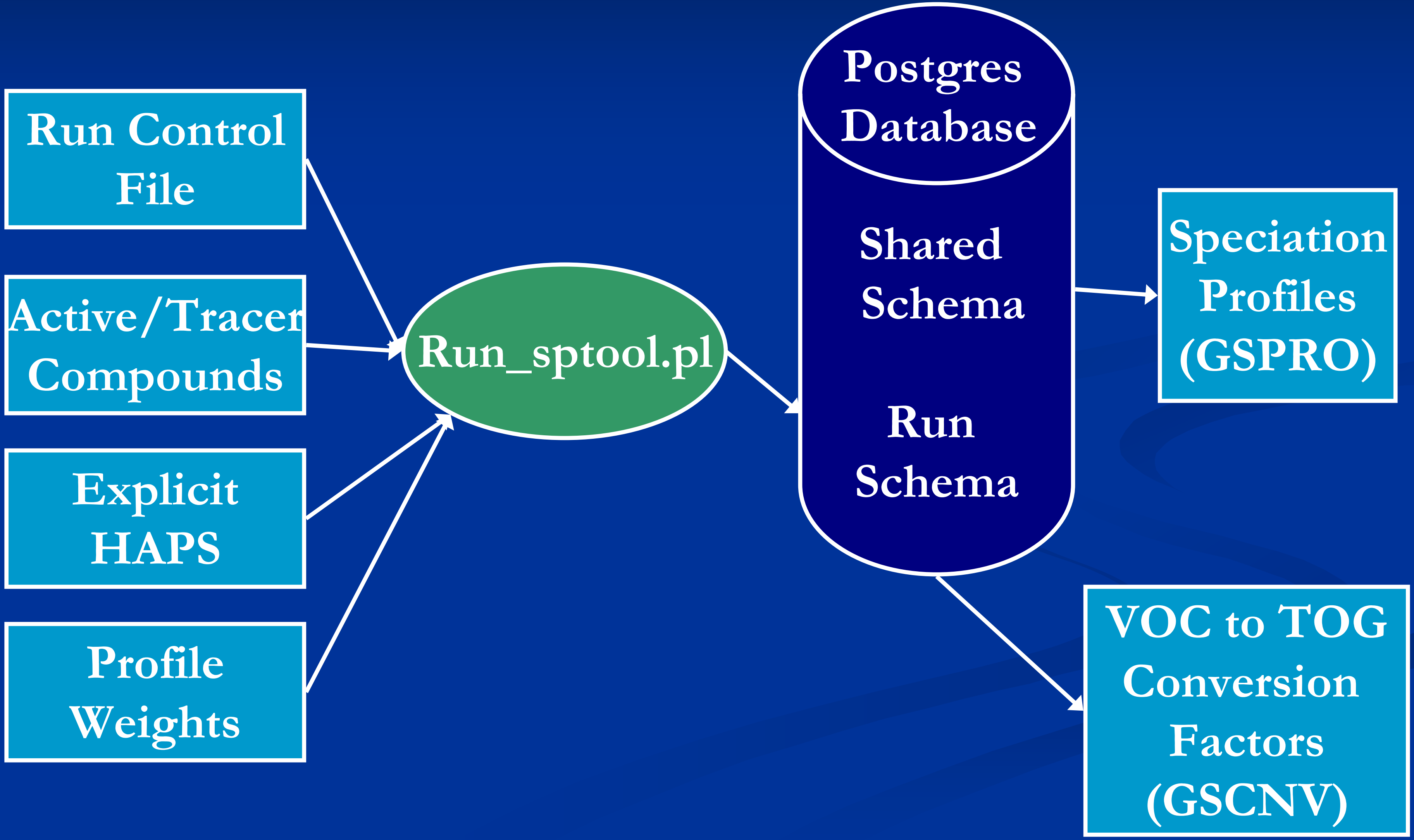
Run Schema

Speciation Tool Model Run

```
perl run_sptool.pl sptoolv1 RunName control_file
```

- Creates RunName schema within initialized *sptoolv1* tool database
- Performs calculations
- Generates outputs as files

Speciation Tool Run



Run Control File

Active/Tracer Compounds

Explicit HAPS

Profile Weights

Run_sptool.pl

Postgres Database

Shared Schema

Run Schema

Speciation Profiles (GSPRO)

VOC to TOG Conversion Factors (GSCNV)

Speciation Tool Run Types

- ALLCRITERIA
  - All VOC model species, including toxic species, are computed from criteria VOC.
- INTEGRATE
  - Specified HAPS are to be integrated from a separate inventory. Part of VOC mass is replaced with HAP VOC mass.
- NOINTEGRATE
  - Both criteria VOC mass and HAP VOC mass are treated separately.

Speciation Tool
Calculation Methodology

- Calculate moles of model species per gram of emissions
- Calculate mole fractions
- Calculate average molecular weight of each species in profile, and AQM lumped compound
- Calculate splits values
- Calculate conversion factors

Speciation Tool Outputs

- Speciation Profiles – GSPRO format
- VOC-to-TOG Conversion Factors – in an updated GSCNV format
- Informative run progress messages
- Intermediate calculations stored within the Run schema of the database
- SMOKE files now, but could be modified to output data to be used by other emission models

Speciation Tool Output Examples

GSPRO:

```
0000      BENZENE     BENZ      1.00000027  78.11184000    1.00000027
0000      ETHBENZ     ETHBENZ   0.99999998 106.16500000    0.99999998
0000      ETHBENZ     PAR       0.12500000  13.27062500    0.12500000
0000      ETHBENZ     TOL       0.87499999  92.89437500    0.87499999
MJ2540    NONHAPTOG   ALD2      0.01704131  32.99254327    0.01704131
MJ2540    NONHAPTOG   ETH       0.07337444  28.05316000    0.07337444
MJ2540    NONHAPTOG   FORM      0.00004998  16.49627164    0.00004998
MJ2540    NONHAPTOG   OLE       0.03663635  30.43771380    0.03663635
MJ2540    NONHAPTOG   PAR       0.50270839  15.62038150    0.50270839
MJ2540    NONHAPTOG   TOL       0.09527041  98.80567974    0.09527041
MJ2540    NONHAPTOG   UNR       0.17751876  15.62097230    0.17751876
MJ2540    NONHAPTOG   XYL       0.09739918 115.53876835    0.09739918
```

GSCNV:

```
VOC       TOG       MJ2540        1.01882178
```

Applications of Tools

- Surrogate Tool
  - Compared to GIS generated surrogates
  - Generated 36km and 12km surrogates for OAQPS 2002 modeling platform
  - Generated surrogates on rotated lat-lon grid for NOAA (WRF-nmm model)
- Speciation Tool
  - Will be used early this summer to generate data for 3 cases for OAQPS

References

- Spatial Allocator Release web site: http://www.cep.unc.edu/empd/projects/mims/spatial/
- Surrogate Tool Release web site: http://www.cep.unc.edu/empd/projects/mims/spatial/srgtool
- Surrogate Tool will be updated 5/31
- Speciation Tool will be released later this FY after its initial application