

# **New Features in Version 3 of the MIMS Spatial Allocator**

**Alison Eyth  
Carolina Environmental Program  
University of North Carolina at Chapel Hill**

**September 27, 2005**

# Background

- Spatial Allocator developed as part of the Multimedia Integrated Modeling System
- Free open-source software
- Targeted software for spatial functions but does not require a GIS
- Shapefiles are primary input format
  - Also I/O API files and ASCII "PointFiles"
- Controlled using environment variables
  - Driven from scripts

# FY05 Project - Phase 1: More Complex Surrogates

- Posted January 31, 2005
- Included new functions to create:
  - Filtered surrogates and Shapefiles
  - Surrogates based on weight functions
  - Merging surrogates using functions
  - Gap filling surrogates
- Executables provided were mims\_spatial, srgmerge, diffsurr for Windows, Linux, and AIX
- See <http://www.cep.unc.edu/empd/projects/mims/spatial/>

# Phase 2 Components

## Posted July 9, 2005

- **allocator**: performs several types of spatial allocation
- **beld3smk**: creates biogenic inputs to SMOKE based on BELD3 land use
- **diffioapi**: differences I/O API files
- **dbf2asc**: creates a .csv file from a .dbf file
- **srgcreate**: creates spatial surrogates
- **srgmerge**: merges and gapfills surrogates
- **diffsurr**: differences surrogates
- **libspatial**: library shared by applications

# Modes of Allocator Program

- CONVERT\_SHAPE: create a copy of a Shapefile with a new map projection
- FILTER\_SHAPE: Apply a filter to a Shapefile to create a new Shapefile that is a subset of the original
- OVERLAY: Print data values of an input file that are overlap a shape (e.g., a grid)
- ALLOCATE: Allocate data from one geospatial unit (e.g., grid cells) to another (e.g., counties or another grid)

# Example Uses of ALLOCATE

- Allocate gridded data onto a different grid
  - Input and output grids can have different map projections or resolutions
  - Formats include I/O API Files and Shapefiles
- Sum point sources in a PointFile by census tract (or county or state)
- Average point observations from monitors into grid cells in an IoapiFile or RegularGrid
- Sum population data in census tracts to create a Shapefile of state level population
- Assign a dominant time zone or land use category to each grid cell in a domain

# OVERLAY Mode

- Prints attribute values for shapes in the input file that are overlapped by a specified region
  - Writes to standard output or a delimited file
- OVERLAY\_TYPE variable specifies form of region
  - BoundingBox, RegularGrid, Shapefile, PolygonFile
  - PolygonFile - ASCII file with lines of coordinate pairs, each with the format: x1 y1
- Input data file formats:  
PointFile, Shapefile, IoapiFile
- OVERLAY\_ATTRS variable lists names of attributes to output (or ALL)

# Uses for OVERLAY Mode

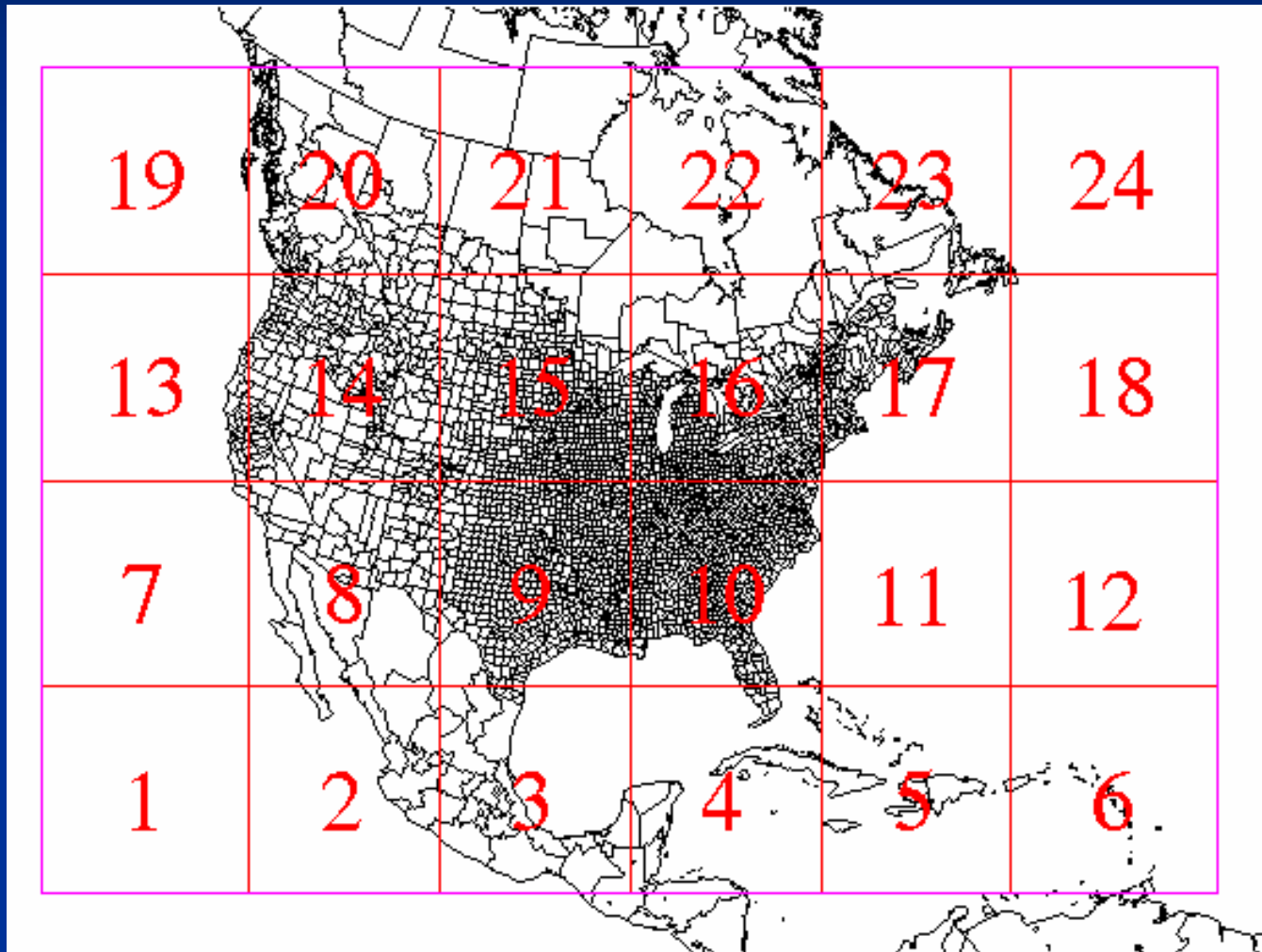
- Print abbreviations of states that overlap a bounding box (or grid) to the screen
  - Turn header and debug output off to drive a script
- Save the coordinates and values for observation stations that overlap a grid to a file
- Print names of counties and states that overlap a national park (specified in a PolygonFile)
- Print row and column numbers of grid cells that overlap polygons in a Shapefile (e.g., counties in NAA)
- Save modeled ozone values from an IoapiFile for grid cells that cover a Class 1 Region



# Creation of SMOKE-BEIS3 Inputs

- SMOKE-BEIS3 requires land use data for the modeled grid at the resolution of the grid as I/O API files: A, B, and TOT
- Input database is Biogenic Emissions Land use Data Version 3
  - 24 tiles over North America for 230 land use types at 1 km resolution

# BELD3 Tiles for 108 km MM5 Domain



# beld3smk

- Input files
  - 1km A, B, and Tot files in I/O API format (from BELD3 data)
  - A Shapefile of the 24 BELD3 tiles
  - GRIDDESC file containing the 24 BELD3 tile grids
- Use OVERLAY mode to find the tiles that overlap the air quality model grid
- Use ALLOCATE mode to allocate the 1km I/O API BELD 3 files onto the AQM grid for each overlapping tile
  - Handles any resolution or map projection changes
- Sum allocated files for each tile to create SMOKE-BEIS3 input files on the desired grid

# Surrogate Tool

- Part of Emissions Modeling Framework project
- Runs Spatial Allocator to create a set of surrogates for a modeling study
- Users can work with .csv input files and do not need to know scripting
  - Shapefile catalog, global control file, surrogate specification file, generation control file
- Runs on Windows and Linux (Java)

# Surrogate Tool (ctd)

- Supports multiple regions (e.g., countries)
- Can specify weight functions, merging, gap filling
- Creates surrogates for both old and new versions of SMOKE
- Will work with EPA's posted surrogate input files
- Will support grid-based and local scale models
- Will be released this fall

# Potential Enhancements

- Better allocation of data in PointFiles
- Read/write .prj component of Shapefiles
- Better support for “chunking” of input files
- In srgmerge, do gap-filling in conjunction with functions of surrogates
- Read an ASCII LineFile format
- Support I/O API on Windows
- Allow selection of layers and timesteps when allocating I/O API files to Shapefiles
- Output Shapefiles and I/O API files when using OVERLAY mode

# Outline

- Background
- Project Goals
- Filtered Shapefiles
- Surrogate Creation Enhancements
- Generic spatial allocation
- Creating SMOKE-BEIS3 inputs
- Other new features
- Potential enhancements

# FY 05 Project

- Goal 1: Support creation of more complex surrogates
  - Over sixty surrogates are now in use by EPA
  - Generated using sophisticated techniques
  - Completed January 1, 2005
- Goal 2: Improve flexibility for generic spatial allocation
- Goal 3: Create inputs to SMOKE-BEIS3 biogenic emissions model
- Phase 2 completed July 9, 2005
- Software available from <http://www.cep.unc.edu/empd/projects/mims/spatial/>



# December 2003 Spatial Allocator

- Starting point for this project
- Generated basic spatial surrogates for emissions modeling (e.g. ports, population, railroads)
- Converted map projection of Shapefiles
- Some other spatial allocation (e.g. aggregation)

# Filtered Surrogates and Shapefiles

- December 2003 Spatial Allocator created surrogates using all shapes in a Shapefile
- Often it is desirable to use only some shapes
  - Interstates from file with multiple road types
  - Water polygons from file with land and water
- Allocator and srgcreate accept a “filter file” to specify a subset of shapes to use
  - Can create a new Shapefile or surrogate
  - Can specify shapes to include and exclude based on attribute values

# State and County Name Filter

ATTRIBUTE\_NAME=STATE\_NAME

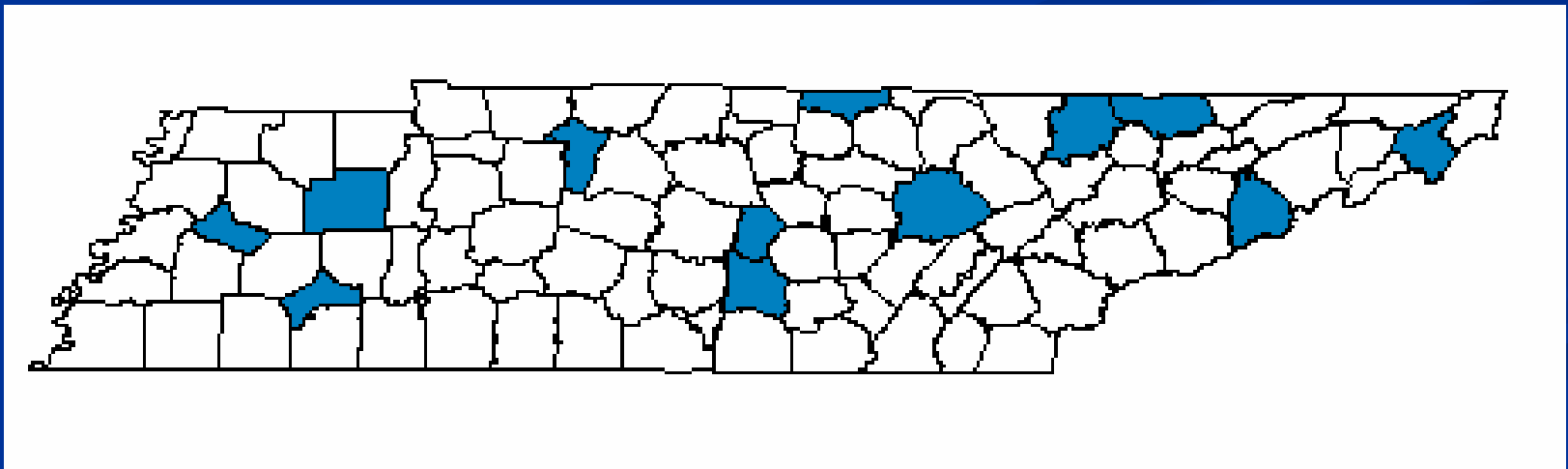
ATTRIBUTE\_TYPE =DISCRETE

INCLUDE\_VALUES=TENNESSEE

ATTRIBUTE\_NAME=COUNTY\_NAME

ATTRIBUTE\_TYPE =DISCRETE

INCLUDE\_VALUES=C\* # example of a regular expression  
(Input file was US Counties Shapefile)



# Surrogates based on Weight Functions

- Surrogates can be created based on a function of multiple attributes
  - e.g. Industrial space surrogate based on sum of multiple kinds of industrial space
  - $IND1 + IND2 + IND3 + IND4 + IND5$

# Srgmerge

- Creates surrogates that are functions of other surrogates
  - $0.5 * \text{Population} + 0.5 * \text{Housing}$
  - $0.25 * \text{Population} + 0.75 * \text{Total Road Miles}$
- Fills in “gaps” in surrogates using more detailed surrogates
  - When detailed surrogate is missing counties
  - Example: Primary=Class 2 Railroads,  
Secondary=Total Railroads,  
Tertiary=Population

# ALLOCATE Mode

- Input data types:
  - **Formats:** Shapefile, IoapiFile, PointFile (ASCII file w/ X,Y, and attributes)
  - **Shapes:** Points, lines, polygons, or a regular grid
- Output data types:
  - **Formats:** Shapefile, IoapiFile
  - **Shapes:** Polygons or a regular grid
- ALLOCATE\_ATTRS variable specifies which attributes of input file to allocate (or ALL)
  - Numeric attributes can be averaged or summed
  - Discrete attributes can be allocated based on max overlap or centroid

# beld3smk Input Data Files

- 1km A, B, and Tot files in I/O API format from BELD3 data
  - each contains only the non-zero land use types
- A Shapefile of the 24 BELD3 tiles
- GRIDDESC file containing the 24 BELD3 tile grids