The MIMS Spatial Allocator: A Tool for Generating Emission Surrogates without a Geographic Information System

Alison M. Eyth, Kimberly Hanisak

Carolina Environmental Program

University of North Carolina at Chapel Hill

May 1, 2003

Multimedia Integrated Modeling System (MIMS)

- Program sponsored by EPA Office of Research and Development
- Successor to Models-3 Computer Framework
- Components:
  - Java-based computer framework for running multimedia (and other complex) models
  - Spatial allocator, Analysis engine, etc.
  - Framework contains tool for designing/visualizing air quality model grids (also developed by MCNC)
- Free to users
  - mostly open source components, but some proprietary (but free) libraries
- http://www.epa.gov/asmdnerl/mims

MIMS Spatial Allocator

- Initially developed at MCNC
  - Most of MCNC EMC is now at UNC Chapel Hill
  - New version provided in March with optimizations, new features, and bug fixes
- Operating modes:
  - Generates spatial surrogates that can be input to SMOKE
  - Change map projection of Shapefiles
  - Performs other types of spatial allocation
    - Aggregate county data to state data
    - Convert between county data and gridded data

Benefits of Spatial Allocator

- **Zero cost makes surrogate generation accessible to more people**
- **Focused purpose of software makes it easier to use to create surrogates than a GIS**
- **Runs on UNIX and Windows**
- **Input data in commonly used ESRI Shapefile format**
- **Output Surrogates**
  - **Computed for regular grid (but could support adaptive grids)**
  - **Written in SMOKE format (but other formats could be added)**
  - **Supported map projections include Universal Transverse Mercator (UTM), Lambert Conformal, and Latitude-Longitude**

What are Spatial Surrogates?

- **Used to map county level emission inventory data into the rectangular grid cells used by air quality models**
  - **For example, dry cleaning emissions values may be available for a county, but CMAQ requires them by grid cell**
- **Surrogate value is the fraction of the emissions for a county that should be apportioned to a grid cell**
- *emis(GC) = srg(Cty,GC) * emis(Cty)*
- **Surrogates allow for more spatial accuracy in emissions distribution than assuming a uniform spread over the county**
  - **E.g. Applying a population surrogate causes higher levels emissions to be placed in the grid cells that cover more densely populated parts of the county**

Computing Spatial Surrogates

$$srg(Cty, GC) = \frac{Wt(Cty \cap GC)}{Wt(Cty)}$$

- Surrogates are computed using a fraction
- Numerator = the value of a weight attribute in the area of intersection between the grid cell and county
- Denominator = the value of a weight attribute in the entire county
- Sum of surrogates values for each county within the grid should be 1
- Weight attributes can be based on objects that are points, lines, or polygons (e.g. port berths, railroads, population)
- Sometimes use number of points, length, or area for weight

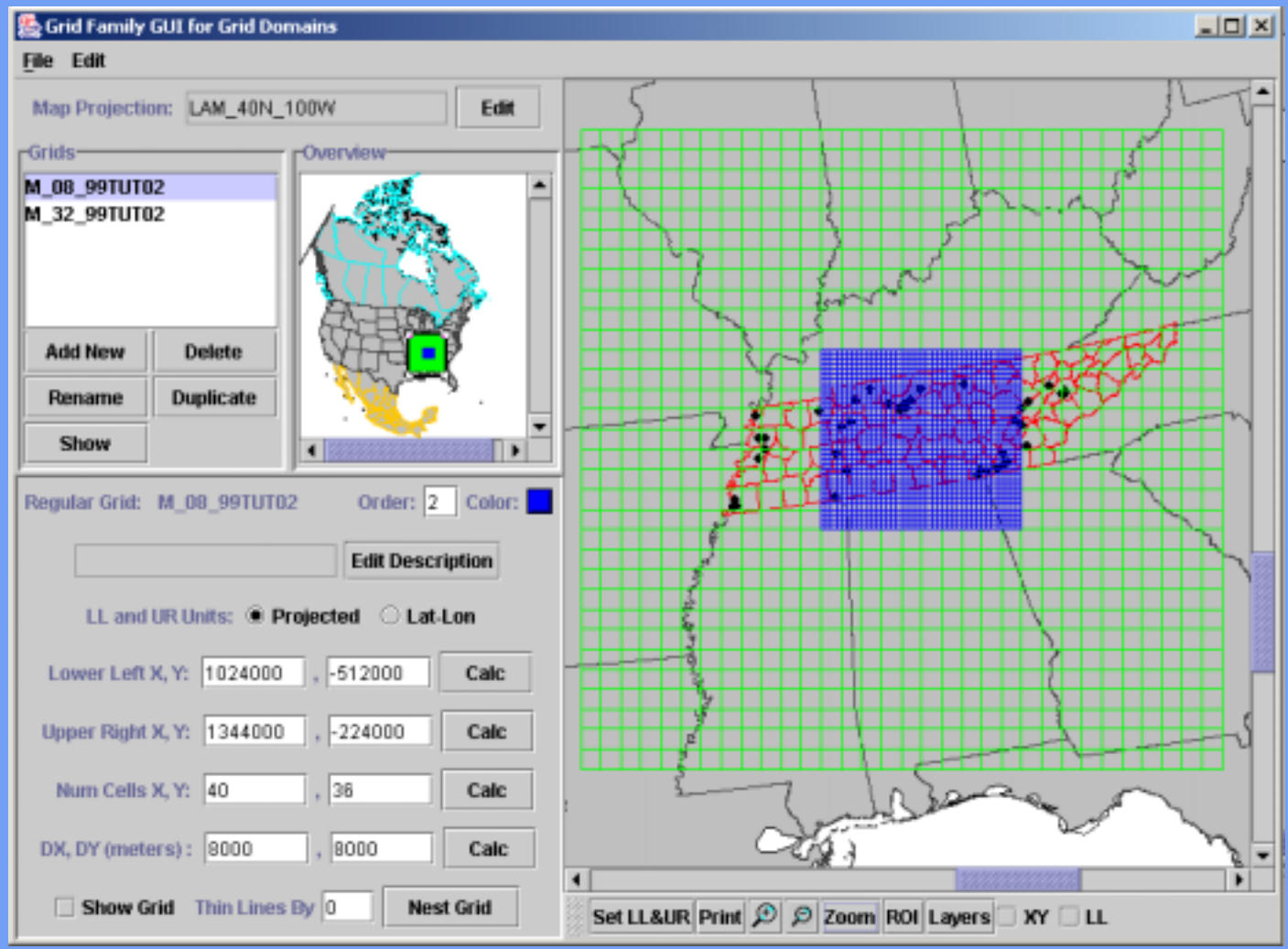
Impact of Weight Attribute on Surrogate Values

Grid Cell	# Ports in gc i, cty C	# Berths in gc i, cty C	Surrogate wt=Count	Surrogate wt=Berths
1	0	0	0	0
2	1	6	1 / 4 = 0.25	6 / 12 = 0.5
3	2	6	2 / 4 = 0.5	6 / 12 = 0.5
4	1	0	1 / 4 = 0.25	0
Total	4	12	1	1

Using the GRIDDESC File to Specify the Output Grid

```
!  coords --line:  name; type, P-alpha, P-beta, P-gamma, xcent, ycent
'LAT_LON'
1, 0.0D0, 0.0D0, 0.0D0, 0.0D0, 0.0D0
'UTM_10'
5, 10.0D0, 0.0D0, 0.0D0, 0.0D0, 0.0D0
'LAM_40N105W'
2, 30.0D0, 60.D0,-105.D0,-105.D0, 40.D0
' '  !  end coords.  Grid name;xorig,yorig,xcell,ycell,ncols,nrows,nthik
'EPAW36_56X78'
'LAT_LON' , -127.0D0, 26.0D0, 0.5000D0, 0.33333D0, 56, 78, 1
'NEW_YORK'
'UTM_18', 480.0D3, 4440.0D3, 5.0D3, 5.0D3, 58, 46, 1'
'DENVER8_34X45'
'LAM_40N105W', -116.D3, -188.D3, 8.D3, 8.D3, 34, 45, 1
```

Visualizing Grids with MIMS Grid Family GUI



Example Windows .bat for Ports Surrogate

```
set MIMS_PROCESSING=SURROGATE
set POLY_OUT_TYPE=RegularGrid
set DATA=C:\surrogates\inputs
set GRIDDESC=%DATA%\GRIDDESC.txt
set GRID=M_08_99NASH

set POLY_DATA_TYPE=ShapeFile
set POLY_DATA=%DATA%\cnty_tn
set ATTR_DATA_ID=FIPS_CODE
set POLY_WEIGHT_TYPE=ShapeFile
set POLY_WEIGHT=%DATA%\tn_ports
set ATTR_WEIGHT=BERTHS
set CATEGORY_WEIGHT=4

set SURROGATE_FILE=C:\surrogates\output\srg_ports.%GRID%.txt
C:\surrogates\bin\mims_spatial.exe
```

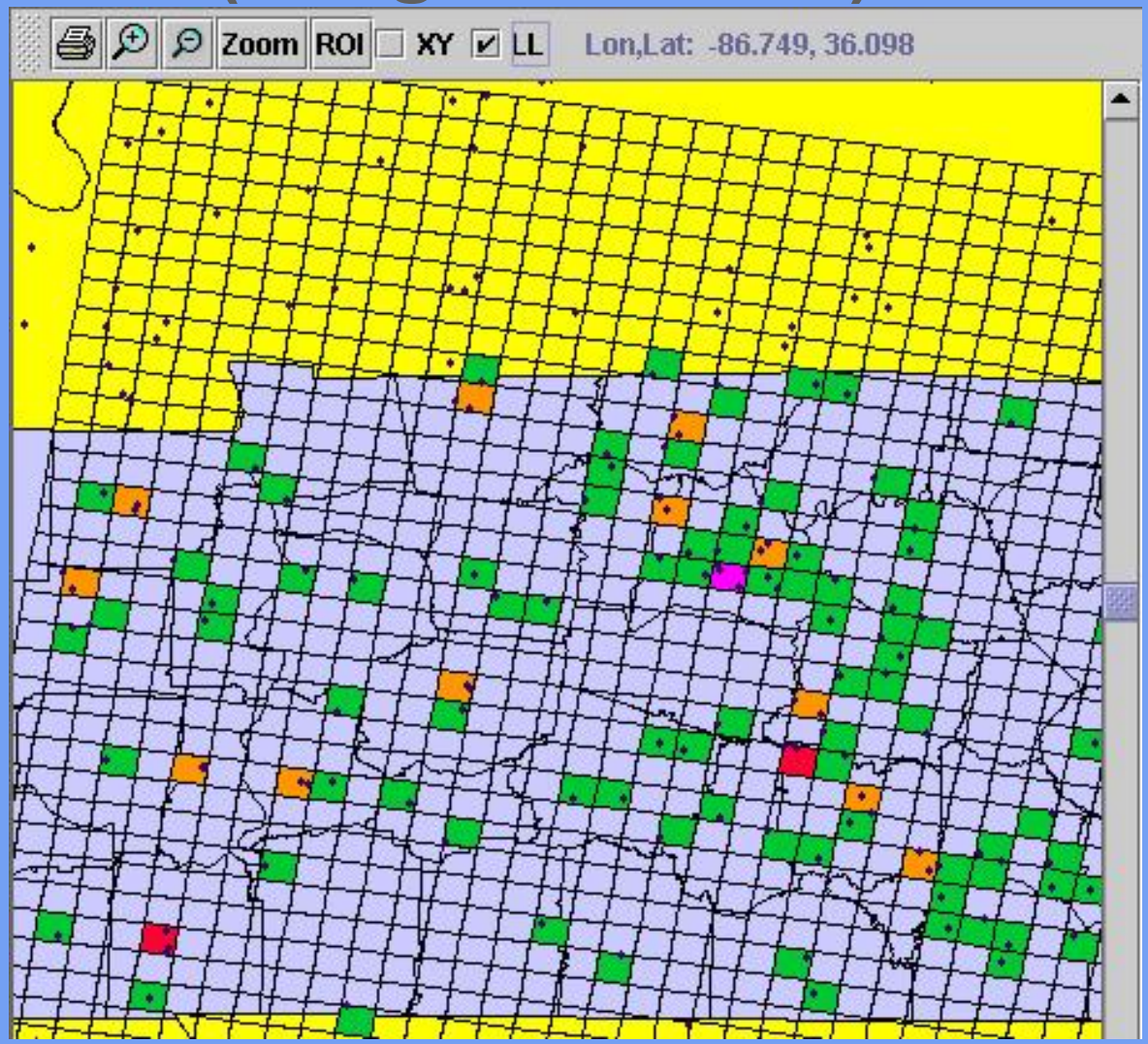
Quality Assurance Options for Surrogates

- **OUTPUT_SRG_NUMERATOR: writes surrogate numerator as a comment in output file**
- **OUTPUT_SRG_DENOMINATOR: writes surrogate denominator as a comment in output file**
- **MIMS_QASUM: writes a running sum of the surrogate values for each county in output file (should sum to 1)**
- **POLY_OUT_NAME: creates a shape file (and .csv file) that contain sums of the surrogate numerators for each grid cell (gridded version of weight attribute**

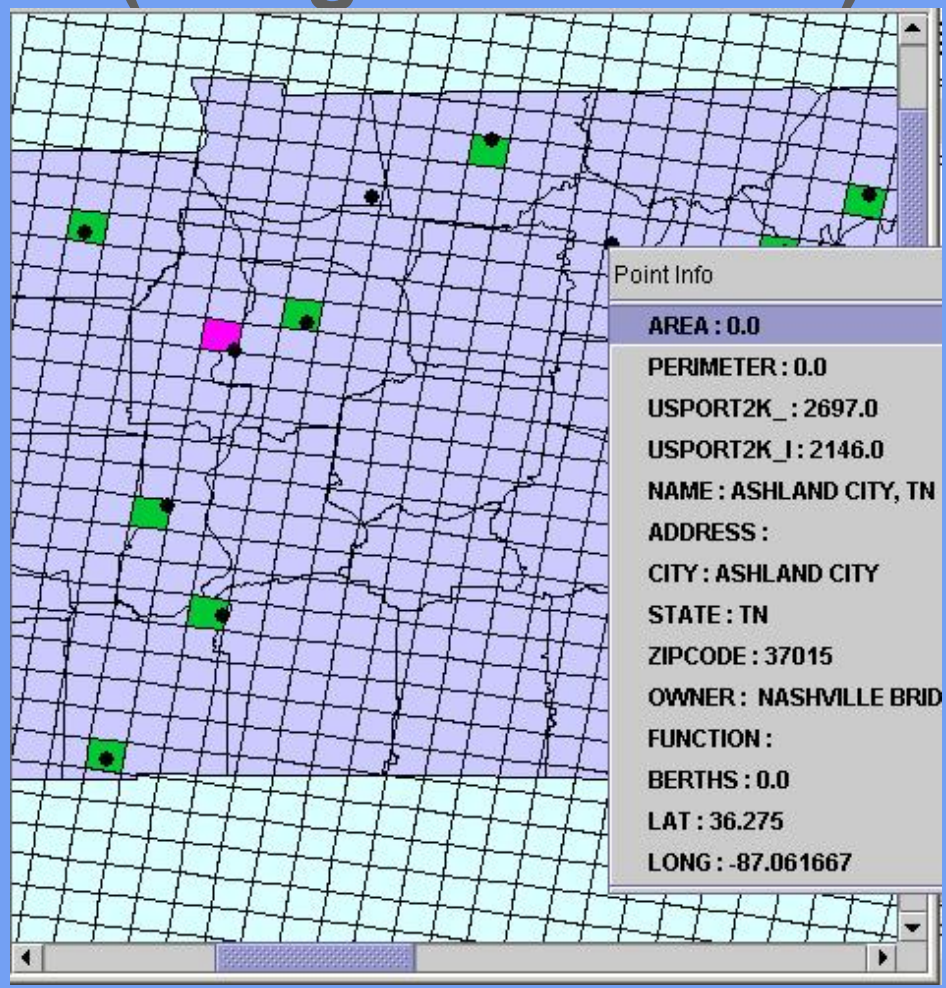
Excerpt of Surrogate Output with Quality Assurance Options On

Cat	County	Col	Row	Srg value		Numerator	Denominator	QA Sum
4	47011	44	20	1	!	8	8	1
4	47037	19	28	0.2	!	2	10	0.2
4	47037	20	27	0.3	!	3	10	0.5
4	47037	20	28	0.3	!	3	10	0.8
4	47037	21	29	0.2	!	2	10	1
4	47039	6	18	1	!	3	3	1
5	47027	32	34	0.491036	!	12917.80	26307.3	0.49104
5	47027	32	35	0.005787	!	152.24	26307.3	0.49682
5	47027	33	35	0.338548	!	8906.28	26307.3	0.83537
5	47027	33	36	0.164629	!	4330.96	26307.3	1

GIST Visualization of Airport Surrogate (Weight = Count)



GIST Visualization of Port Surrogate (Weight = Berths)



Specifying Map Projections and Ellipsoids

- **PROJ.4 library is used (http://www.remotesensiong.org/proj)**
- **Supports most map projections**
- **Lambert conformal example:**
  setenv DATA_POLY_MAP_PRJN
  "+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97“

- **UTM example:**
  setenv WEIGHT_POLY_MAP_PRJN "+proj=utm,+zone=17“
- **Ellipsoid examples:**
  - **setenv WEIGHT_POLY_ELLIPSOID=+GRS80**
  - **setenv WEIGHT_POLY_ELLIPSOID=+a=6378137.0,+rf=298.2572**

Converting the Map Projection of Shapefiles

```
#!/bin/csh -f
setenv MIMS_PROCESSING CONVERT_SHAPE
setenv POLY_DATA_TYPE   ShapeFile
setenv POLY_DATA    $argv[1]      # no extension
setenv POLY_OUT_TYPE    ShapeFile
setenv POLY_OUT_NAME   $argv[2]   # no extension

setenv DATA_POLY_MAP_PRJN
   "+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97"
setenv DATA_POLY_ELLIPSOID +WGS84
setenv OUTPUT_POLY_MAP_PRJN   LATLON
setenv OUTPUT_POLY_ELLIPSOID  SPHERE
/apps/mims_spatial/bin/mims_spatial.exe
```

Software Implementation

- Software is written in C
- Blocks of code to perform specific tasks that are reused for different operating modes
  - First read data (then weight) polygons & convert map projection to output projection
  - Next compute intersection of weight and data polygons (then of weight-data polygons with grid polygons)
  - Then compute surrogates
    - Numerator = sum of weight for each county and gc
    - Denominator = sum of weight for each county
- Public domain libraries used
  - PROJ.4 for map projection conversions
  - Shapelib for reading / writing shapefiles
  - Generic Polygon Clipper for polygon intersection

Limitations

- **Currently only SMOKE-ready output**
- **Output assumed to be on sphere**
- **Line-based inputs must be "dissolved" at the county boundaries**

Future Directions

- **Produce biogenic inputs for SMOKE**
- **Easier conversion between county and gridded data**
- **Generalize spatial allocation to support more forms of regridding**
- **Create surrogates for adaptive and other non-regular grids**
- **Further reduce memory usage to support use of larger data sets (~ 1 GB)**

Jason Eyth Cobb
3/30/2003
6 lbs 14 oz (1.97 kg)

Overall Surrogate equation

$$Wt(Cty \cap GC) = \sum_{wp} Wt(wp) * \frac{area(wp \cap Cty \cap GC)}{area(wp)}$$

$$Wt(cty) = \sum_{wp} Wt(wp) * \frac{area(wp \cap Cty)}{area(wp)}$$

Where Wt(x) = value of weight attribute for x, Cty = County, GC = grid cell, wp = weight polygon