EPA Spatial Allocator
User Guide

Prepared for:

William Benjey, Thomas Pierce, and Dan Loughlin
United States Environmental Protection Agency
National Exposure Research Laboratory
Office of Research and Development
Research Triangle Park, North Carolina

Prepared by:

Science Applications International Corporation
615 Oberlin Road, Suite 100
Raleigh, North Carolina 27605

Contract GS-35F-4461G
Task Order GST0403BF0247

December 2004

Table Of Contents

Introduction to the Spatial Allocator

The MIMS Spatial Allocator was originally designed (and has since been expanded) as a tool to help prepare emission inventory information without the use of commercial Geographic Information Systems (GIS). Emissions inventories are generally created based on political boundaries or attached to specific locations (e.g., railways), but most models require emissions to be located within specific grid cell boundaries. In conjunction with the Sparse Matrix Operating Kernel Emissions (SMOKE) system, the Spatial Allocator compares model grid boundaries with the geographic boundaries of inventories to distribute activity and emissions data properly over the modeling grid. The Spatial Allocator was designed to prepare the AGPRO, MGPRO, and BGPRO surrogate inputs required by the SMOKE system.

To perform the allocating operations, the Spatial Allocator uses:

- a definition of the modeling grid (which may be created through the Grid Family Window in the MIMS Framework),
- boundary GIS files showing the extents of the emission inventory (e.g., county boundaries for county level emissions files), and
- surrogate GIS files that specify how to distribute the emissions (e.g., census tract information showing population counts).

The files that the Spatial Allocator creates act as cross-references between an emissions inventory and the modeling grid. For example, the line below indicates that emissions from county 47003 in the inventory should be allocated to the modeling grid cell in the sixth column and second row.

    2   47003   6   2   0.625

The 0.625 number indicates that only a fraction of the emissions from this county should be put in that grid cell. The remaining 0.375 fraction will be allocated to other grid cells.

In addition to creating spatial surrogates, the Spatial Allocator was expanded to complete other GIS functions:

1. Aggregate or average data from smaller spatial units to larger ones (e.g., converting county data to state information);
2. Create a copy of a GIS Shape file with a new mapping projection; and
3. Convert gridded biogenic emission files to a different grid.

EPA Spatial Allocator

Unlike the MIMS Framework, the Spatial Allocator does not provide a graphical user interface (GUI). Instead users must create a script file that specifies the required parameters and file locations and then run an executable program. The program was written in the C language. The Spatial Allocator has been tested on several UNIX platforms, Windows NT/2000 and LINUX, and identical results were obtained on all platforms.

The five sections below will provide details on installing the Spatial Allocator, illustrations of spatial surrogate calculations, instruction on operating the executable files, and examples of the basic operations:

Installing the Spatial Allocator

To install the Spatial Allocator and run the test case, follow these steps:

1. Download the executable files, source code, and scripts for the appropriate platform into a single directory:

   Windows executables and source code and scripts

   Solaris executables and source code and scripts

   Red Hat Linux executables and source code and scripts

   IRIX executables and source code and scripts

   Check these files against those posted at the Spatial Allocator download site (go to http://www.cep.unc.edu/empd/projects/mims/spatial/ and click on Installing the Software and Running the Test Case) to ensure that your files are not outdated.

2. Download the input data files and the documentation into the same directory as Step 1. If you prefer mapping with a latitude/longitude projection, download the latitude/longitude input data files.
3. Uncompress the downloaded files onto a disk containing at least 40 MB of free space. As the files uncompress, the data files will be placed in the appropriate subdirectories.
4. In the scripts directory, run the generate_surrogates.bat executable (Windows platforms) or the generate_surrogates.csh executable (UNIX platforms). This operation will create the surrogate values for the test case. Users who wish to direct the output to alternative directories should first edit the BASDIR environment variable in the generate_surrogates file.
5. In the scripts directory, run the compare_surrogates.bat executable (Windows platforms) or the compare_surrogates.csh executable (UNIX platforms). This operation compares the surrogates generated in the previous step with a set of reference values. The statement "Test suite successful!" should appear at the end of the program.
6. The output directory now contains the file SRG_M_08_99NASH.TXT. This file can be referenced by SMOKE as both the AGPRO and MGPRO file. After the first line in the file (which describes the grid), this file contains the following columns:

- Numeric indicator for the surrogate type (the user may assign positive integers to any surrogate type. The sample file uses 2=airports, airstrips, and helipads; 3=land area; 4=ports; 5=navigable water; 6=highways; 7=households; and 8=population)
- County FIPS code
- Grid column number
- Grid row number
- Commented columns showing the numerator and denominator of the fraction, followed by a cumulative fractional tally for the county (information following the explanation point is commented)

The output lines shown below indicate that airport emissions from county 47003 in the inventory should be allocated to the modeling grid cell in the sixth column and second row.

2	47003	6	2	0.625	!	5	8	0.625
2	47003	6	3	0.375	!	3	8	1

The 0.625 number indicates that only a fraction of the airport emissions from this county should be put in that grid cell. The remaining 0.375 fraction will be allocated to the neighboring grid cell (sixth column and third row).

The remaining GIS files and databases can be used to plot the gridded information with mapping software.

If you have a C compiler, you may also build your own executable from the source code. The code has been compiled with both gcc and native C compilers and has been developed for operation with most standard C compilers. See the README.TXT file in the src directory for further instructions.

How Spatial Surrogates, Aggregates, and Averages Are Calculated

The map below shows the area surrounding Nashville, Tennessee. Sample grid boundaries are show in navy blue, county boundaries in black, and census tract boundaries in gray. The red aircraft represent airports, airstrips, and helipads. The central county (shown in yellow) is Davidson County. Two of the Davidson County census tracts show purple cross-hatching. The surrogate discussion will share examples based on this diagram. Users interested in the mathematical formulae for surrogates are referred to http://www.cep.unc.edu/empd/projects/mims/spatial/. Recall that spatial surrogates represent a fractional relationship between two overlapping areas.

The sample files contain script files (generate_surrogates.*) that show the proper coding for these options.
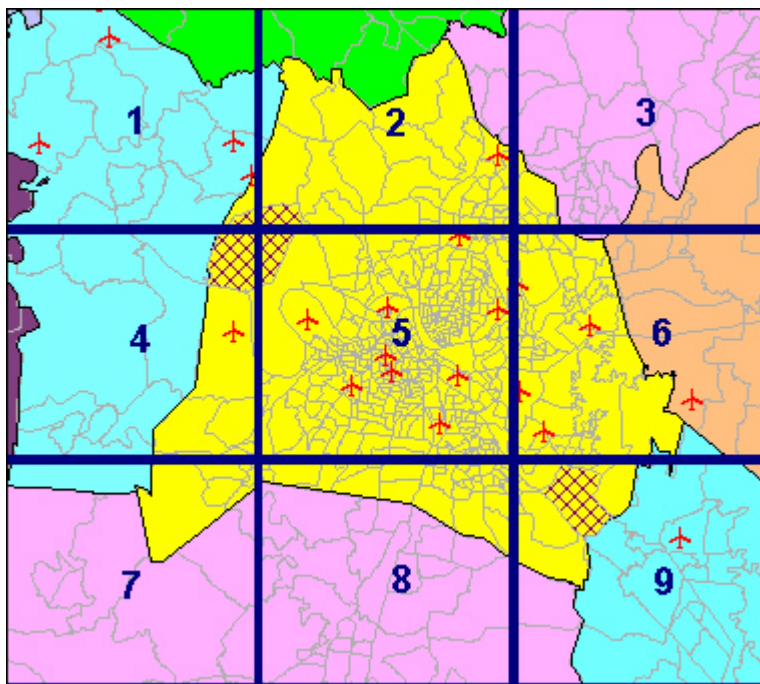
Area-Weighted Surrogates

An area-weighted surrogate approach distributes emissions to the grid cells based on the geographic areas of overlap between the emissions units (e.g., counties) and the

grid cells. If the geographic boundaries of an emissions unit are entirely contained within a single grid cell, the area-weighted surrogate value would be 1.0.

To determine an area-weighted surrogate value for a grid cell as it overlaps a county, determine the area of overlap and divide this number by the total county area. For example, Grid Cell 5 represents 173 square miles, and Davidson County covers 525 square miles. Therefore, the land area-based surrogate for Grid Cell 5 within Davidson County (FIPS Code 47037) would be 173/525, or 0.33.

Grid Cell 8 would have two area-weighted surrogate values assigned to it, one for Davidson County and a second for Williamson County in the central and southern sections. The overlap of Grid Cell 8 and Davidson County is just 32 miles, so the first surrogate for Grid Cell 8 would be 32/525, or 0.06.

The remaining grid cells in the diagram would have three or more surrogate values since they intersect more than two counties.

Line-Weighted Surrogates

Similar to area-weighted surrogates, a line-weighted surrogate approach distributes emissions to the grid cells based on the geographic areas of overlap between the lines defining emissions units (e.g., roads or railways) and the grid cells. If the line object is entirely contained within a single grid cell, the line-weighted surrogate value would be 1.0. Line-weighted surrogates are not described in this example.

Count-Weighted Surrogates

A count-weighted surrogate approach distributes emissions to the grid cells based on the number of objects from the emissions units (e.g., counties) located within the grid cells. If all of the objects for an emission unit are entirely contained within a single grid cell, the count-weighted surrogate value would be 1.0 for that cell and 0.0 for all others.

In the example above, fifteen airports, airstrips, and helipads are located in Davidson County. One of these Davidson County airports (Oakley Airport) is located in Grid Cell 4. Therefore, the surrogate value for Grid Cell 4 within Davidson County would be 1/15, or 0.07. Grid Cell 2 would also have a surrogate value within Davidson County of 0.07. Grid Cell 5 would have a surrogate value of 9/15, or 0.60. Grid Cell 6 would have a surrogate value for Davidson County of the remaining 4/15, or 0.27 (Grid Cell 6 would have an additional surrogate value for Wilson County to the east). Note that the fractional surrogate values for Davidson County total 1 (although totals for counties which are not completely covered by a grid will be less than 1).

Attribute-Weighted Surrogates

An attribute-weighted surrogate approach distributes emissions to the grid cells based on a weight associated with each object (e.g., census tract) from the emissions units

6

(e.g., counties) located within the grid cells. It also distributes that weight based on the area within the particular grid cell. If the entire emission unit is contained within a single grid cell, the attribute-weighted surrogate value would be 1.0 for that cell.

The 2000 U.S. Census data is stored at the census tract level (in the POP2000 field), and census tract boundaries do not cross county lines. The census tracts from Davidson County show a total population near 570,000. The cross-hatched census tract in Grid Cell 9 recorded a population of 3,955. Therefore, this census tract would contribute 3,955/570,000 (0.0069) to the surrogate value for Grid Cell 9 within Davidson County. This process would be repeated for all of the census tracts that lie in both Grid Cell 9 and Davidson County, then the values would be summed to get the final surrogate value for Grid Cell 9 within Davidson County.

However, some census tracts lie in more than one grid cell (e.g., the cross-hatched census tract lying at the intersection of Grid Cells 1, 2, 4, and 5). That tract recorded 787 persons in 2000 and covers 16.7 square miles. The portion of this census tract that lies in Grid Cell 4 is 7.6 square miles. The land area is used to apportion the census tract to the four grid cells. Therefore, the contribution of this census tract to the surrogate value for Grid Cell 4 within Davidson County would be calculated as (787/570,000)x(7.6 sq mi/16.7 sq mi), or 0.00063.

This sample attribute-weighted surrogate is based on polygonal areas, but they may also be computed based on lines or point objects.

Spatial Aggregates

Spatial aggregation is the summation of data from smaller objects into total values for larger area (polygonal) objects. One example would be to sum all of the population values for counties up to the state level. As a spatial aggregate, a state map would be placed over the counties map. The population numbers for counties falling within that state's boundaries would then be summed to give a state total.

In Aggregate processing mode, the Spatial Allocator can also determine the total population within the county boundaries, based on the census tract information. The Spatial Allocator determines which census tracts fall within the individual counties and then totals the population.

Spatial Averages

Spatial averaging is similar to spatial aggregation. Spatial averaging determines the mean for data from smaller objects contained within larger area (polygonal) objects.

In Average processing mode, the Spatial Allocator can determine the average population of the census tracts within the county boundaries, based on the census tract information. The Spatial Allocator determines which census tracts fall within the individual counties and then averages the population per census tract for all within each

county. In this instance, the POP2000 field would be used as the attribute weight (ATTR_WEIGHT variable).

Operating the Spatial Allocator

The Spatial Allocator is a tool with the ability to perform geographic operations on user-provided information. However, the only national information provided in the example data files is for airports and navigable waters. In general, the users must:

- Collect the necessary data;
- Convert the input data into the proper format;
- Choose the appropriate environment variables;
- Compose and run a script; and
- Recognize the output format.

Collecting the necessary data

Before using the Spatial Allocator, users must locate the following information:

- Map projection, ellipsoid, and projection parameters
- Grid specifications (e.g., number of rows, number of columns, starting location, and size of grid cells)
- ESRI Shape files to describe the surrogate areas
- Surrogate weighting information (population by census tract, port locations, etc.) associated with each emissions area

The air quality modelers will specify the map projection parameters based on the desired modeling domain. These specifications can be used in the MIMS Framework's Grid Family Window to create the appropriate grid specifications. Within the Grid Family Window, choose the Export GRIDDESC option from the file menu to create a file specifying the grid.

Currently the polygons describing the surrogate areas must be in ESRI's Shape file format. Many other GIS packages (e.g., MapInfo, Intergraph, and AutoCAD) provide translators that can convert mapping files into this format.

At least one of the surrogate files must link entries to the emission inventory files (e.g., county FIPS codes) and to the polygons in the Shape file.

Input Data Format

The map projection and grid specifications are detailed in the GRIDDESC file, and the surrogate areas and weighting information are described under Shape files and associated data files.

GRIDDESC file

The Spatial Allocator specifies the map and grid parameters in the GRIDDESC file. The format for the comma-delimited GRIDDESC file appears below:

Line	Columns	Description
1	A	Name of first projection in single quotation marks
2	A	Integer specifying projection type for *1st* projection (1=Latitude-longitude, 2=Lambert, 3=Mercator, 4=Stereographic, 5=UTM, 6=Polar, 7=Equatorial Mercator, and 8=Transverse Mercator)
	B,C,D	Projection alpha, beta, and gamma parameters for *1st* projection (see Projection Information in Supplementary Information)
	E,F	Central x and y points for the *1st* projection (see Projection Information in Supplementary Information)
3	A	Name of *2nd* projection in single quotation marks
...	...	...
2n-1	A	Name of *nth* projection in single quotation marks
2n	A	Integer specifying projection type for *nth* projection
	B,C,D	Projection alpha, beta, and gamma parameters for *nth* projection
	E,F	Central x and y points for the *nth* projection
2n+1	A	Single quotation marks around a space
2n+2	A	Name of *1st* grid in single quotation marks
2n+3	A	Name of projection associated with *1st* grid in single quotation marks
	B,C	X and y origins for the *1st* grid (lower left corner)
	D,E	Size of each grid cell in the x and y directions
	F,G	Integer number of grid cells in the x and y directions
	H	Number of grid cells that the grid should be expanded beyond each boundary
2n+4	A	Name of *2nd* grid in single quotation marks
...	...	Specifications associated with subsequent grids
last line	A	Single quotation marks around a space

A sample grid description file appears below for a single latitude-longitude projection with two grids named M_08_99NASH and nest.

```
'1'

1,  0.0D0, 0.0D0, 0,0D0, -100D0, 0.0D0

' '

'M_08_99NASH'

'1', -90.61333D0, 33.53333D0, 792.666D-3, 1.075112D0, 10, 5, 1

'nest'

'1', -89.027998D0, 35.683554D0, 198.1665D-3, 268.778D-3, 12, 12, 1

' '
```

A GRIDDESC file may be constructed through the Grid Family Window of the MIMS Framework, and no further modifications would be required to the file. The Grid Family window acts as a tool to create the GRIDDESC file in the format that the Spatial Allocator and the MIMS Framework recognize.

Shape Files and Associated Data

The Shape files are necessary to determine the positions of the surrogate features. Each information set should have at least three files of the same name with extensions .shp, .shx, and .dbf. The dbf file represents the database of information about the polygons. The dbf file should have at least one field that corresponds directly with a field in the emissions inventory (e.g., FIPS code); see the description of the ATTR_DATA_ID environment variable below. The dbf file should also have a field that corresponds to the weighting each area is to be assigned (e.g., number of housing units); see the description of the ATTR_WEIGHT environment variable below.

Environment Variables

The Spatial Allocator recognizes the following case-sensitive environment variables:

General Processing Variables

> GRIDDESC - directory and file name for the file containing the grid specifications.

> GRID_NAME - name of the output grid, as specified in the GRIDDESC file.

> MIMS_PROCESSING - variable that controls the operation that the program performs. Valid choices are SURROGATE, CONVERT_SHAPE, AGGREGATE, AVERAGE, and CONVERT_BELD.

USE_CURVED_LINES - set to YES to compute using Great Circle distances. The default is NO and ignores the Earth's curvature.

USE_DW_FILE - Directory and file name for the intermediate file (overlay of weight polygons on data polygons) that will initialize the intersection of data and weight polygons. If this variable is not set or set to NONE, no intermediate file will be used.

Input Shape File Variables

ATTR_DATA_ID - Field attribute name that associates data polygons with a unique polygon in the inventory (e.g., FIPS_CODE for a county-level emissions file). The country/state/county code (6-digit integer YSSCCC) format should be used if the output is to be used in SMOKE.

ATTR_WEIGHT - attribute field to use for the weighting of the surrogate (floating point or integer). Use NONE if length, area, or count should be used instead of a field. If ATTR_WEIGHT is assigned to a string field, then a count of the rows will be used instead of the field. To create multiple surrogates based on different attributes in the same file (e.g., housing and population), separate the ATTR_WEIGHT fields with commas.

POLY_DATA - directory and file name for file containing data polygons (do not include the .shp extension in the variable declaration for shape files).

POLY_DATA_TYPE - File type containing data polygons. Valid values are currently ShapeFile or FractionalVegetationFile (used in CONVERT_BELD mode).

POLY_WEIGHT - directory and file name for file containing weight polygons (do not include the .shp extension in the variable declaration for shape files). If no weights are desired, specify NONE.

POLY_WEIGHT_TYPE - File type containing weight polygons. The only currently valid value is ShapeFile.

Map Projection Variables

DATA_POLY_ELLIPSOID - optional [PROJ.4](#) ellipsoid for the data polygons. The PROJ.4 program provides tools to convert geographic coordinates into other projections. Specify this variable to assign a non-spherical surface to the projection for the data files. The default setting is SPHERE.

DATA_POLY_MAP_PRJN - optional [PROJ.4](#) map projection parameters for the data polygons. The default projection is latitude/longitude and may be specified as LATLON.

OUTPUT_POLY_ELLIPSOID - optional PROJ.4 ellipsoid for the output polygons. Specify this variable to assign a non-spherical surface to the projection for the output files. The default setting is SPHERE. This variable affects CONVERT_SHAPE and AGGREGATE processing but does not affect output from SURROGATE processing.

OUTPUT_POLY_MAP_PRJN - optional PROJ.4 map projection parameters for the data polygons. The default projection is latitude/longitude and may be specified as LATLON. This variable affects CONVERT_SHAPE and AGGREGATE processing but does not affect output from SURROGATE processing.

WEIGHT_POLY_ELLIPSOID - optional PROJ.4 ellipsoid for the weight polygons. Specify this variable to assign a non-spherical surface to the projection for the weight files. The default setting is SPHERE.

WEIGHT_POLY_MAP_PRJN - optional PROJ.4 map projection parameters for the weight polygons. The default projection is latitude/longitude and may be specified as LATLON.

Output Variables

CATEGORY_WEIGHT - spatial surrogates code (integer) to indicate the surrogate method. This variable corresponds to the entries in the first column of the output file. When creating multiple surrogates based on different attributes in the same file (e.g., housing and population), separate the CATEGORY_WEIGHT integers with commas.

MIMS_HEADER - When set to YES, the first line of the surrogate file will show the information related to the grid and map projection. The SMOKE preprocessor looks for this information. The default setting for the MIMS_HEADER is NO. Note that the command "mims_spatial.exe -header" will also produce a header line.

MIMS_QASUM - When performing SURROGATE processing and this variable is set to YES, a new column appears in the surrogate file after the comment indicator. The values in this column reflect the cumulative sum of the surrogate fractions up to that line in the file (for a given inventory area) and should never total more than one. The default setting is NO.

OUTPUT_SRG_DENOMINATOR - When performing SURROGATE processing and this variable is set to YES, a new column will appear in the surrogate file after the comment indicator. The values in this column reflect the numerator for the surrogate fraction. The default setting is NO.

OUTPUT_SRG_NUMERATOR - When performing SURROGATE processing and this variable is set to YES, a new column will appear in the surrogate file after the

comment indicator. The values in this column reflect the denominator for the surrogate fraction. The default variable is NO.

POLY_OUT_NAME - directory and file name for the output shape file. When this variable is specified, an output shape file is created in SURROGATE, CONVERT_SHAPE, and AGGREGATE processing.

POLY_OUT_TYPE - file type for the output shape file. The only value currently recognized is RegularGrid.

SAVE_DW_FILE - Directory and file name for the intermediate file (overlay of weight polygons on data polygons). This file is independent of the grid. If this variable is not set or set to NONE, no intermediate file will be saved.

SURROGATE_FILE - directory, file name, and extension for the output surrogate file. In CONVERT_BELD processing, this variable specifies the directory, file name, and extension for the output file.

Compose and Run a Script

The basic scripts to run the Spatial Allocator can be constructed with a few commands:

copy - *duplicate a file's contents into a new file*

del - *delete a file*

echo - *print a message to the console window*

mims_spatial.exe - *run the Spatial Allocator*

mkdir - *create a directory*

more - *pipe file contents to the console window or to another file*

set or setenv - *assign a value to an environment variable*

:: - *indicate a comment line*

The *generate_surrogates.bat* file shows the appropriate command formats for DOS/Windows-based operating systems, and the *generate_surrogates.csh* file shows the command formats for UNIX/Linux-based operating systems. These and other basic scripts for the Spatial Allocator are found in the scripts directory.

To run a script, type its name at the command prompt and press the Enter key.

Output Data Format

The Spatial Allocator can produce a spatial surrogate file, a comma-delimited file describing the gridded data, and the files associated with a Shape file (.shp, .shx, and .dbf extensions) to present the gridded data with GIS tools.

The Spatial Allocator's SURROGATE_FILE environment variable specifies the directory and file name for the spatial surrogates file. The spatial surrogates file may contain the spatial allocation factors for all area sources and non-link mobile sources. This file name may be specified as the AGPRO or MGPRO file by the SMOKE preprocessor.
 The file description appears below:

Line	Columns	Description
1	A	#GRID - line 1 only appears if MIMS_HEADER environment variable is set to YES (default) or if "-header" is part of the mims_spatial.exe command line
	B	GRID_NAME environment variable
	C	X origin in units of the projection
	D	Y origin in units of the projection
	E	X direction cell length in units of the projection
	F	Y direction cell length in units of the projection
	G	Number of columns
	H	Number of rows
	I	Number of boundary cells
	J	Projection types: Latitude-Longitude: "LAT-LON" or "LATGRD3" Lambert Conformal: "LAMBERT" or "LAMGRD3" Universal Transverse Mercator: "UTM" or "UTMGRD3"
	K	Projection units
	L-N	Projection alpha, beta, and gamma values
	O	X-dir projection center in units of the projection
	P	Y-dir projection center in units of the projection
2+	A	Area source spatial surrogates code, mobile source county feature/roadway type, or CATEGORY_WEIGHT environment variable (integer)
	B	Country/State/County Code (integer) or ATTR_DATA_ID environment variable. The country/state/county code (6-digit integer YSSCCC) format should be used if the output is to be used in SMOKE.
	C	Grid column number (integer)
	D	Grid row number (integer)
	E	Spatial surrogate ratio for area sources or fraction of county feature in cell for mobile sources (real number)

Line	Columns	Description
	F	! - indicating that the information following is treated as a comment
	G-I	Numerator, denominator, and cumulative surrogate sum if OUTPUT_SRG_NUMERATOR, OUTPUT_SRG_DENOMINATOR, and MIMS_QASUM environment variables are set to YES

The first four lines of the spatial surrogates file might appear as:

#GRID M_08_99NASH   -90.613330   33.533330   0.792666   1.075112   10   5   1 ...

2   47003   6   2   0.625   !   5   8   0.625

2   47003   6   3   0.375   !   3   8   1

2   47005   4   3   1   !   2   2   1

The comma-delimited file (which is assigned the name of POLY_OUT_NAME and the extension .csv) lists only the grid column, grid row, and numerator (OUTPUT_SRG_NUMERATOR) fields. Similarly, the database file (.dbf extension) holds the column, row, and numerator. An example of the first four lines of the .csv file appears below:

COL,ROW,NUMERSUM

1,1,0.000000

1,2,0.180799

1,3,0.008979

The Shape files (.shp and .shx extensions following the POLY_OUT_NAME) allow the grids to be plotted with GIS software but contain no more information about the surrogates. Instead, their NUMERSUM field is populated by the sum of all numerators for the grid cell (OUTPUT_SRG_NUMERATOR column).

Spatial Allocator Examples

The Spatial Allocator is designed to perform five functions:

Create spatial surrogates

Create spatial averages

Create spatial aggregates

Convert Shape files to new map projections

Convert biogenic data to a new grid

Three sample applications (for both Windows and UNIX-based systems) of these functions are provided during installation in the SCRIPTS directory and are briefly described below.

Spatial Surrogates Example

The sample *generate_surrogates.bat* and *generate_surrogates.csh* scripts will create surrogate files that can reference spatial information about the number of airports/airstrips/helipads, land area, berths at ports, navigable waters, highways, population, and number of households. The sample script operates on a limited data set centered near Nashville, Tennessee. While creating the surrogate file, the Spatial Allocator will also create separate Shape files of the grid for each surrogate variable (except households).

The sample *generate_surrogates* scripts are organized to accomplish these ordered tasks:

- Set general environment variables
- Write header line to create the surrogate file
- Run Spatial Allocator for airports/airstrips/helipads and add to surrogate file
- Run Spatial Allocator for land area and add to surrogate file
- Run Spatial Allocator for ports and add to surrogate file
- Run Spatial Allocator for navigable waterways and add to surrogate file
- Run Spatial Allocator for highways and add to surrogate file
- Run Spatial Allocator for households and population and add to surrogate file

To run the sample script, type the script name and the Enter key.

The environment variables are assigned the following values in the Windows version:

General Processing Variables

    BASDIR = .. (one directory lower than the scripts directory)

    WORK_DIR = ..\output

    DATA = ..\data

    MIMSDIR = ..\src

    GRIDDESC = ..\data\GRIDDESC.txt

    MIMS_PROCESSING = SURROGATE

    EXE = mims_spatial.exe

    USE_CURVED_LINES = not set

    USE_DW_FILE = not set

    TIME = "time /t"

Map Projection Variables

    DATA_POLY_ELLIPSOID = SPHERE

    DATA_POLY_MAP_PRJN = LATLON

    OUTPUT_POLY_ELLIPSOID = not set

    OUTPUT_POLY_MAP_PRJN = not set

    WEIGHT_POLY_ELLIPSOID = SPHERE

    WEIGHT_POLY_MAP_PRJN = +proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

Output Variables

    MIMS_HEADER = NO

    MIMS_QASUM = YES

    OUTPUT_SRG_DENOMINATOR = YES

OUTPUT_SRG_NUMERATOR = YES

SAVE_DW_FILE = not set

SURROGATE_FILE = ..\output\tmp_srg.M_08_99NASH.txt (temporarily generated during each allocation run for the various category weights)

SRG_FILE = ..\output\srg_M_08_99NASH.txt (represents the merge of all of the SURROGATE_FILE iterations)

Input Shape File Variables

CATEGORY_WEIGHT and Surrogate	ATTR_WEIGHT	POLY_WEIGHT	POLY_OUT_NAME
2 Airport	NONE	..\data\us_air-pt	..\output\grid_airpt_M_08_99NASH
3 Land area	NONE	NONE	..\output\grid_area_M_08_99NASH
4 Ports	BERTHS	..\data\tn_ports	..\output\grid_ports_M_08_99NASH
5 Navigable waters	LENGTH	..\data\us_nav_h20	..\output\grid_navig_M_08_99NASH
6 Highways	NONE	..\data\tn_roads	..\output\grid_highway_M_08_99NASH
7 Households	HOUSEHOLDS	..\data\tn_pophous	..\output\grid_pop_M_08_99NASH
8 Population	POP2000	..\data\tn_pophous	..\output\grid_pop_M_08_99NASH

ATTR_DATA_ID = FIPS_CODE

GRID_NAME = M_08_99NASH

POLY_DATA = ..\data\cnty_tn

POLY_DATA_TYPE = ShapeFile

POLY_OUT_TYPE = RegularGrid

POLY_WEIGHT_TYPE = ShapeFile

After running the sample *generate_surrogates* script, users have the opportunity to check that the *generate_surrogates* script ran correctly by comparing the results with the sample reference file srg_nash_ref.txt. Run the sample script *compare_surrogates.bat* (Windows/DOS) or *compare_surrogates.csh* (UNIX) to check that the results are within a tolerance of 2E-5. An error message will be displayed if the results do not match.

Spatial Aggregates Example

The sample scripts *aggregate.bat* (Windows/DOS) and *aggregate.csh* (UNIX) operate on a Shape file describing the Tennessee counties. The *aggregate* scripts sum the

values for the census tracts to the county FIPS code level, based on the county within which each census tract is located. This processing produces a Shape file (agg_pophou.shp) and associated database with three fields (FIPS_CODE, POP2000, and HOUSEHOLDS).

The scripts are organized to accomplish these ordered tasks:

- Set environment variables
- Run Spatial Allocator for Aggregate processing

To run the script, type the script name and the Enter key.

The environment variables are assigned the following values in the Windows version:

General Processing Variables

 MIMSDIR = .. (one directory lower than the scripts directory)

 DATADIR = ..\data

 OUTPUT = ..\output

 GRIDDESC = ..\data\GRIDDESC.txt

 MIMS_PROCESSING = AGGREGATE

 EXE = ..\src\mims_spatial.exe

Map Projection Variables

 DATA_POLY_ELLIPSOID = SPHERE

 DATA_POLY_MAP_PRJN = LATLON

 OUTPUT_POLY_ELLIPSOID = SPHERE

 OUTPUT_POLY_MAP_PRJN = LATLON

 WEIGHT_POLY_ELLIPSOID = SPHERE

 WEIGHT_POLY_MAP_PRJN = +proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

Input Shape File Variables

 ATTR_DATA_ID = FIPS_CODE

 ATTR_WEIGHT = POP2000, HOUSEHOLDS (creates 2 fields in the output)

 POLY_DATA = ..\data\cnty_tn

POLY_DATA_TYPE = ShapeFile

POLY_OUT_NAME = ..\output\agg_pophous

POLY_WEIGHT = ..\data\tn_pophous

POLY_WEIGHT_TYPE = ShapeFile

With a GIS tool, the Shape file *agg_pophou.shp* will display a representation of the counties.

Converting Shape Files Example

The sample scripts *convert_shape.bat* (Windows/DOS) and *convert_shape.csh* (UNIX) operate on a Shape file that the user specifies. The *convert_shape* scripts transform the polygons in the original files to a new set of files with an alternate map projection.

The scripts are organized to accomplish these ordered tasks:

- Set environment variables
- Run Spatial Allocator for Convert_Shape processing
- Copy the associated original dbf file to a file with the new name

To run the script, type the script name, the path and name of the original file (without extension), the path and name of the new file (without extension), and the Enter key. For example, type "convert_shape.bat ..\output\agg_pophou ..\output\agg_pophou_latlon" and Enter to convert the agg_pophou file to a latitude/longitude format.

The environment variables are assigned the following values in the Windows version:

General Processing Variables

MIMS_PROCESSING = CONVERT_SHAPE

MIMS_EXE = ..\src\mims_spatial.exe

Map Projection Variables

DATA_POLY_ELLIPSOID = SPHERE

DATA_POLY_MAP_PRJN=+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

OUTPUT_POLY_ELLIPSOID = SPHERE

OUTPUT_POLY_MAP_PRJN = LATLON

Input Shape File Variables

POLY_DATA = specified by user in command line

POLY_DATA_TYPE = ShapeFile

POLY_OUT_NAME = specified by user in command line

POLY_OUT_TYPE = RegularGrid

Supplementary Information on the Spatial Allocator

The following topics are covered on this page:

GRIDDESC Projection Information

PROJ.4 Projections and Ellipsoids

Licenses and External Software

Creating a New Writer

How to Implement a Customized Reader

GRIDDESC Projection Information

PROJECTION	TYPE	PARAMETERS
lat-lon	1	*Latitude/Longitude* PROJ_ALPHA, PROJ_BETA, PROJ_GAMMA unused. Coordinate units are degrees, -180.0 < X,Y <= 180.0 Note that Western hemisphere longitudes and Southern Hemisphere latitudes are taken to be negative.
Lambert	2	*Lambert Conformal* PROJ_ALPHA is less than or equal to PROJ_BETA, and these two latitudes determine the projection cone; PROJ_GAMMA is the central meridian. (X_CENT,Y_CENT) are the lat-lon coordinates for the center (0,0) of the Cartesian coordinate system. Coordinate units are meters.

PROJECTION	TYPE	PARAMETERS
Mercator	3	*General Mercator*

PROJ_ALPHA and PROJ_BETA are the latitude and longitude of the coordinate origin (within the tangent circle)
PROJ_GAMMA is the angle between the cylinder axis and the North polar axis.
(X_CENT,Y_CENT) are the lat-lon coordinates for the center (0,0) of the Cartesian coordinate system.
Coordinate units are meters.

Stereographic	4	*General Tangent Stereographic*

PROJ_ALPHA is the "true latitude," the latitude at which the stereographic plane is secant to the Earth (or 90, if the projection is a North Polar tangent stereographic).
PROJ_BETA is the angle of rotation of the Y-axis relative to the Greenwich meridian, i.e., the longitude meridian which becomes the negative Y axis.
PROJ_GAMMA is unused.
(X_CENT,Y_CENT) are the lat-lon coordinates for the center (0,0) of the Cartesian coordinate system.
Coordinate units are meters.

UTM	5	*Universal Transverse Mercator*

PROJ_ALPHA is the UTM zone, considered as a DOUBLE.
PROJ_BETA and PROJ_GAMMA are unused.
(X_CENT,Y_CENT) are the UTM coordinates of the origin for offset UTM coordinates (or are (0,0) for Equator-based UTM coords).
Coordinate units are meters.

Polar	6	*Polar Secant Stereographic*

PROJ_ALPHA is 1.0 for North Polar, -1.0 for South Polar.
PROJ_BETA is the secant latitude (latitude of true scale).
PROJ_GAMMA is the central meridian.
(X_CENT,Y_CENT) are the lat-lon coordinates for the center (0,0) of the Cartesian coordinate system.
Coordinate units are meters.

Equatorial Mercator	7	*Equatorial Mercator*

PROJ_ALPHA is the latitude of true scale.
PROJ_BETA is unused.
PROJ_GAMMA is the longitude of the central meridian.
Coordinate units are meters.

PROJECTION	TYPE	PARAMETERS

Transverse Mercator   8   *Transverse Mercator*

PROJ_ALPHA is the latitude of the origin.
PROJ_BETA is the scale factor at the central meridian.
PROJ_GAMMA is the longitude of the central meridian.
Coordinate units are meters.

Note that the standard names for UTM coordinate systems are given in the form *UTM_<nn>*, where *nn* is the number for the UTM zone. Much of the Eastern US would use coordinate system UTM_17, for example.

PROJ.4 Projections and Ellipsoids

(from http://www.cep.unc.edu/empd/projects/mims/spatial/)

The PROJ.4 program provides tools to convert geographic coordinates into other projections. The PROJ.4 library (http://www.remotesensing.org/proj) is used by the Spatial Allocator to specify map projections and ellipsoids. Map projections can be specified independently for data, weight, and output polygons using the DATA/WEIGHT/OUTPUT_POLY_MAP_PRJN environment variables. If no map projection is specified, or the projection is specified as LATLON, latitude-longitude is assumed. Information on specifying map projection parameters for PROJ.4 can be found at the bottom of the web pages linked from (http://www.remotesensing.org/geotiff/proj_list/). The projection parameters should start with "+proj=" and the parameters should be in a comma-delimited list without white space. For example, to use a Lambert conformal conic projection with standard parallels of 33 and 45 degrees, a central latitude of 40 degrees, and a central longitude of -97 degrees as the projection for data polygons, use the following syntax:

   setenv DATA_POLY_MAP_PRJN "+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97"

To specify a UTM projection for zone 17 for the weight polygons, use the following syntax:

   setenv WEIGHT_POLY_MAP_PRJN "+proj=utm,+zone=17"

The Spatial Allocator allows the user to set the shape of the Earth independently for data, weight, and output polygons with the DATA/WEIGHT/OUTPUT_POLY_ELLIPSOID environment variables. If these variables are not specified, or they are specified as SPHERE, the default ellipsoid is the MM5 sphere where R=6370977.0. To specify an ellipsoid, users must specify either the name of a standard Ellipsoid, or the values that traditionally describe the shape of the Earth including a (Ellipsoidal Semi-Major Axis), b (Ellipsoidal Semi-Minor Axis), R (radius), and rf. A plus sign must precede either the name of the Earth Shape or each ellipsoidal parameter. When supplying multiple parameters, the list must be comma-separated without any white space between the parameters. For example, you may specify the MERIT spheroid using the following syntax:

       setenv WEIGHT_POLY_ELLIPSOID=+MERIT

Or you may specify the ellipsoid parameters as follows:

setenv WEIGHT_POLY_ELLIPSOID=+a=6378137.0,+rf=298.257

Valid Earth Shape Specifications:

MERIT	a=6378137.0	rf=298.257	MERIT 1983
SGS85	a=6378136.0	rf=298.257	Soviet Geodetic System 85
GRS80	a=6378137.0	rf=298.257222101	GRS 1980(IUGG, 1980)
IAU76	a=6378140.0	rf=298.257	IAU 1976
airy	a=6377563.396	b=6356256.910	Airy 1830
APL4.9	a=6378137.0	rf=298.25	Appl. Physics. 1965
NWL9D	a=6378145.0	rf=298.25	Naval Weapons Lab., 1965
mod_airy	a=6377340.189	b=6356034.446	Modified Airy
andrae	a=6377104.43	rf=300.0	Andrae 1876 (Den., Iclnd.)
aust_SA	a=6378160.0	rf=298.25	Australian Natl & S. Amer. 1969
GRS67	a=6378160.0	rf=298.2471674270	GRS 67(IUGG 1967)
bessel	a=6377397.155	rf=299.1528128	Bessel 1841
bess_nam	a=6377483.865	rf=299.1528128	Bessel 1841 (Namibia)
clrk66	a=6378206.4	b=6356583.8	Clarke 1866
clrk80	a=6378249.145	rf=293.4663	Clarke 1880 mod.
CPM	a=6375738.7	rf=334.29	Comm. des Poids et Mesures 1799
delmbr	a=6376428.	rf=311.5	Delambre 1810 (Belgium)
engelis	a=6378136.05	rf=298.2566	Engelis 1985
evrst30	a=6377276.345	rf=300.8017	Everest 1830
evrst48	a=6377304.063	rf=300.8017	Everest 1948
evrst56	a=6377301.243	rf=300.8017	Everest 1956
evrst69	a=6377295.664	rf=300.8017	Everest 1969
evrstSS	a=6377298.556	rf=300.8017	Everest (Sabah & Sarawak)
fschr60	a=6378166	rf=298.3	Fischer (Mercury Datum) 1960
fschr60m	a=6378155.	rf=298.3	Modified Fischer 1960
fschr68	a=6378150.	rf=298.3	Fischer 1968
helmert	a=6378200.	rf=298.3	Helmert 1906
hough	a=6378270.0	rf=297.	Hough
intl	a=6378388.0	rf=297.	International 1909 (Hayford)
krass	a=6378245.0	rf=298.3	Krassovsky, 1942
kaula	a=6378163.	rf=298.24	Kaula 1961
lerch	a=6378139.	rf=298.257	Lerch 1979
mprts	a=6397300.	rf=191.	Maupertius 1738
new_intl	a=6378157.5	b=6356772.2	New International 1967
plessis	a=6376523.	b=6355863.	Plessis 1817 (France)
SEasia	a=6378155.0	b=6356773.3205	Southeast Asia
walbeck	a=6376896.0	b=6355834.8467	Walbeck
WGS60	a=6378165.0	rf=298.3	WGS 60
WGS66	a=6378145.0	rf=298.25	WGS 66
WGS72	a=6378135.0	rf=298.26	WGS 72
WGS84	a=6378137.0	rf=298.257223563	WGS 84

Licenses and External Software

The spatial analysis tool was created using multiple packages (PROJ.4, Shapelib, and Generic Polygon Clipper). License information is found at the following links:

http://www.cep.unc.edu/empd/projects/mims/spatial/license.txt - PROJ.4, Shapelib, and Generic Polygon Clipper

http://www.cep.unc.edu/empd/projects/mims/spatial/LICENSE.LGPL - GNU General Public License required by PROJ.4 and Shapelib

Creating a New Writer / Adding New Operations

(based on http://www.cep.unc.edu/empd/projects/mims/spatial/programmers_guide.html)

Currently the Spatial Allocator is organized into two main operations: creating surrogates (implemented in *reportSurrogate*) and performing spatial aggregation (*reportSum*). These operations each take a PolyObject * (a list of Shapes), perform a computation based on the attributes, and write the results to the file formats applicable to that operation. The *reportSurrogate* operation uses the format of the surrogate file embedded in it. The *reportSum* operation calls *polyShapeWrite* to write the .shp portion of the shape file, manipulates the attributes as appropriate (summing or averaging), and writes the new values with calls to DBFWriteDoubleAttribute.

Three methods support the surrogate generation and the spatial aggregation: sum1Poly, sum2Poly, and avg1Poly. If a developer adds an interpolation method to fit into the current aggregation framework, a new method (e.g., interp1Poly) could be developed to interpolate the values of the desired attributes. If the new method would not fit into this framework, a new branch to the main program should be added.

How to Implement a Customized Reader

(based on http://www.cep.unc.edu/empd/projects/mims/spatial/programmers_guide.html)

A. The developer must change the PolyReader() method. The second argument to this method is a character string that points to the environment variable that specifies the type of the reader. In addition to adding customized code for the new reader, the developer must insert a section similar to the following segment:

```
else if (!strcmp(type,"ShapeFile"))

{

    poly PolyShapeReader(ename);
```

B. To implement the customized reader (i.e., the substitute for PolyShapeReader), the developer must provide a reader method that returns a PolyObject and takes a character string argument pointing to the environment variable that specifies the file name. This method should read all the data and fill in an appropriate PolyObject struct. A pointer to the struct should be obtained by a call to getNewPoly(nObjects). The reader code should specify the type of polygon (polygon, line, or point), the map information struct, and the bounding box (see documentation on PolyShapeReader and RegularGridReader for complete examples).

```
poly->nSHPType = SHPT_POLYGON;
poly->map = map;
poly->bb = newBBox(xmin, ymin, xmax, ymax)
```

Then, in a loop over the objects, the developer can do the following:

- Get a pointer to a PolyShape struct:

```
ps = getNewPolyShape(number_of_parts);
ps->num_contours = 0;
```

- Loop over the parts (i.e., contours or polylines) and get a pointer to a Shape struct:

```
shp = getNewShape(number_of_vertices);
```

In a loop over the vertices, fill in the x and y for each vertex.

```
shp->vertex[i].x = xcoord;
shp->vertex[i].y = ycoord;
```

Close vertices loop.

Add a contour to the PolyShape through:

```
gpc_add_contour(ps,shp,SOLID_POLYGON);
```

Close the loop over the parts.

Add the PolyShape struct to the PolyShape Object.

```
polyShapeIncl(&(poly->plist),ps,NULL);
```

C. A replacement for the attachAttribute method must also be developed for the new file format. This routine must properly associate the attribute information with the polygons.

Appendix A. Sample Spatial Allocator Files for Generating Surrogates

Sample Files

Script file = generate_surrogates.bat

Grid description file = griddesc.txt

Output file = srg_M_08_99NASH.txt

The generate_surrogates.bat file also uses shape files, such as cnty_tn.shp and us_air-pt.shp, but their geospatial data are not presented in this appendix.

Script file = generate_surrogates.bat

```
::******************** Generate Surrogates Run Script
************************

:: This script generates surrogates for the MIMS spatial tool test case

:: (8km over Tennessee).

::

:: Script created by : Alison Eyth, Carolina Environmental Program

:: Last edited : March 2003

::

::*****************************************************************

@echo off

:: Set installation directory

set BASDIR=..

:: Set directory for output surrogate and shape files

set WORK_DIR=%BASDIR%\output
```

```
:: Set Location of shapefiles

set DATA=%BASDIR%\data

mkdir %WORK_DIR%

:: Set Grid settings

:: used to read "set GRIDDESC=%DATA%\GRIDDESC.txt

set GRIDDESC=%BASDIR%\input\GRIDDESC.txt

set GRID_NAME=M_08_99NASH

:: Set Location of executable

set MIMSDIR=%BASDIR%\src

set EXE=mims_spatial.exe

:: Set name and path to temporary surrogate file

set SURROGATE_FILE=%WORK_DIR%\tmp_srg.%GRID_NAME%.txt

:: Set name and path to final merged surrogate file from spatial tool

set SRG_FILE=%WORK_DIR%\srg_%GRID_NAME%.txt

:: mode of operation for program

set MIMS_PROCESSING=SURROGATE

:: MIMS_QASUM=YES prints surrogate sums by county in file

:: OUTPUT_SRG_NUMERATOR=YES writes surrogate numerator as comment in file

:: OUTPUT_SRG_DENOMINATOR=YES writes denminator (county totals) for srg
weight

set MIMS_QASUM=YES

set OUTPUT_SRG_NUMERATOR=YES

set OUTPUT_SRG_DENOMINATOR=YES

:: Print header info

set MIMS_HEADER=YES

:: Specify type of data files to use
```

```
set POLY_OUT_TYPE=RegularGrid

set POLY_DATA_TYPE=ShapeFile

set POLY_WEIGHT_TYPE=ShapeFile

:: The data polygons should be the shape file containing polygons for the
grid

set POLY_DATA=%DATA%\cnty_tn

set ATTR_DATA_ID=FIPS_CODE

set DATA_POLY_MAP_PRJN=LATLON

set DATA_POLY_ELLIPSOID=SPHERE

:: Set weight projection to that of the EPA files

set WEIGHT_POLY_MAP_PRJN=+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

set WEIGHT_POLY_ELLIPSOID=SPHERE

set TIME="time /t"

:: Generate surrogate header line

echo Writing header

%MIMSDIR%\%EXE% -header > %SRG_FILE%

if %errorlevel% NEQ 0 (

    goto error

)

:: Generate surrogates for each category

:: For each one, specify:

:: CATEGORY_WEIGHT (integer ID for type of surrogate)

:: POLY_WEIGHT (shapefile for weights),

:: ATTR_WEIGHT (attribute to weight by (or NONE for area/length/count)

:: POLY_OUT_NAME (optional - name of shapefile of gridded weights to output)

echo Generating airport surrogate
```

```
set CATEGORY_WEIGHT=2

set POLY_WEIGHT=%DATA%\us_air-pt

set ATTR_WEIGHT=NONE

set POLY_OUT_NAME=%WORK_DIR%\grid_airpt_%GRID_NAME%

%TIME%

%MIMSDIR%\%EXE%

if %errorlevel% == 0 (

   more %SURROGATE_FILE% >> %SRG_FILE%

   del %SURROGATE_FILE%

)

echo Generating land area surrogate

set CATEGORY_WEIGHT=3

set POLY_WEIGHT=NONE

set ATTR_WEIGHT=NONE

set POLY_OUT_NAME=%WORK_DIR%\grid_area_%GRID_NAME%

%TIME%

%MIMSDIR%\%EXE%

if %errorlevel% == 0 (

   more %SURROGATE_FILE% >> %SRG_FILE%

   del %SURROGATE_FILE%

)

...
```

Grid description file = griddesc.txt

```
! Grid family:l

'1'

1,  0.0D0, 0.0D0, 0,0D0, -100D0, 0.0D0
```

32

' '

'M_08_99NASH'

'1', -90.61333D0, 33.53333D0, 792.666D-3, 1.075112D0, 10, 5, 1

'nest'

'1', -89.027998D0, 35.683554D0, 198.1665D-3, 268.778D-3, 12, 12, 1

' '

Output file = srg_M_08_99NASH.txt

```
#GRID M_08_99NASH -90.613330 33.533330 0.792666 1.075112 10 5 1 LAT-LON
meters ...

#GRID   M_08_99NASH    -90.613330      33.533330       0.792666
        1.075112        10 ...

   2   47003       6       2       0.625       !           5
              8       0.625

   2   47003       6       3       0.375       !           3
              8       1

   2   47005       4       3       1       !               2
              2       1

   2   47007       8       3       1       !               1
              1       1

   2   47009       9       3       1       !               5
              5       1

   2   47011       8       2       1       !               1
              1       1

   2   47013       9       3       1       !               3
              3       1

   2   47017       3       3       1       !               4
              4       1

   2   47021       5       3       1       !               5
              5       1

   2   47025       9       3       1       !               1
              1       1

   2   47027       7       3       1       !               1
              1       1
```

```
2    47029       10        3              1       !              1
                  1        1

2    47031        6        2              1       !              6
                  6        1

2    47035        7        3            0.5       !              2
                  4      0.5

2    47035        8        3            0.5       !              2
                  4        1

2    47037        5        3       0.933333       !             14
                 15   0.93333

2    47037        6        3      0.0666667       !              1
                 15        1

2    47039        4        2              1       !              2
                  2        1

2    47041        7        3              1       !              1
                  1        1

2    47043        5        3              1       !              3
                  3        1

2    47045        2        3              1       !              1
                  1        1

2    47047        2        2              1       !              7
                  7        1

2    47049        8        3              1       !              1
                  1        1

2    47051        6        2              1       !              6
                  6        1

2    47053        3        3              1       !              3
                  3        1

2    47055        5        2              1       !              1
                  1        1

2    47059       10        3       0.666667       !              2
                  3   0.66667
```

...

```
2    47001        8        3           0.25       !              1
                  4      0.25
```

```
2   47001        9       3          0.75       !                    3
              4       1


#GRID   M_08_99NASH     -90.613330      33.533330      0.792666
        1.075112        10 ...

3   47003        5       2     0.0128527    !        0.00156971
      0.122131      0.012853

3   47003        6       2     0.972641     !        0.118789
      0.122131      0.98549

3   47003        6       3     0.0145062    !        0.00177166
      0.122131      1

3   47005        4       3          1       !        0.113114
      0.113114      1

3   47007        7       2     0.721973     !        0.0757648
      0.104941      0.72197

3   47007        7       3     0.165214     !        0.0173377
      0.104941      0.88719

3   47007        8       2     0.0305539    !        0.00320637
      0.104941      0.91774

3   47007        8       3     0.0822595    !        0.00863242
      0.104941      1

...
```

Appendix B. Sample Spatial Allocator Files for Generating Aggregates

Sample File

Script file = aggregate.bat

The aggregate.bat file also uses the shape files cnty_tn.shp and tn_pophous.shp, but their geospatial data are not presented in this appendix. Similarly the output file agg_pophou.shp is not presented in the appendix.

Script file = aggregate.bat

```
:: ****************** Aggregate Shapefiles Run Script
************************

:: This script aggregates population and housing data from NC tracts to NC

:: counties.

::

:: Script created by : K. Hanisak, Carolina Environmental Program, UNC-CH

:: Last edited : March 2003

::

::
**************************************************************************

:: Set Input Directories

@echo off

set MIMSDIR=..

set DATADIR=%MIMSDIR%\data

set OUTPUT=%MIMSDIR%\output

:: Set executable

set EXE=%MIMSDIR%\src\mims_spatial.exe
```

```
:: Select method of spatial analysis

set MIMS_PROCESSING=AGGREGATE

:: Data polygon shapefile (state counties) -- level of aggregation

:: Set name and path of shapefile

set POLY_DATA=%MIMSDIR%\data_latlon\cnty_tn

set POLY_DATA_TYPE=ShapeFile

set DATA_POLY_MAP_PRJN=LATLON

set DATA_POLY_ELLIPSOID=SPHERE

:: Set Shapefile from which to aggregate data (tracts)

set POLY_WEIGHT=%DATADIR%\tn_pophous

set POLY_WEIGHT_TYPE=ShapeFile

set WEIGHT_POLY_MAP_PRJN=+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

set WEIGHT_POLY_ELLIPSOID=SPHERE

:: set output file name

set POLY_OUT_NAME=%OUTPUT%\agg_pophou

set OUTPUT_POLY_MAP_PRJN=LATLON

set OUTPUT_POLY_ELLIPSOID=SPHERE

:: Set attribute to aggregate by (county)

set ATTR_DATA_ID=FIPS_CODE

:: Set attributes to aggregate to county level

set ATTR_WEIGHT=POP2000,HOUSEHOLDS

echo Aggregating tracts to county for population and housing

%EXE%
```

Appendix C. Sample Spatial Allocator Files for Converting Shape Files

---

Sample File

Script file = convert_shape.bat

The convert_shape.bat file requires names for input and output shape files, but their geospatial data are not presented in this appendix. Users may refer to the agg_pophou.shp and agg_pophou_latlon.shp files for examples.

Script file = convert_shape.bat

```
::****************** Convert Shape Run Script *************************

:: This script converts shape file from one map projection to another

:: Note: the .proj component of the shape file is not created.

::

::*************************************************************************
*

@echo off

:: Location of executable

set MIMS_EXE=..\src\mims_spatial.exe

set MIMS_PROCESSING=CONVERT_SHAPE

set POLY_OUT_TYPE=RegularGrid

set POLY_DATA_TYPE=ShapeFile

set POLY_DATA=%1

set POLY_OUT_NAME=%2

::set input projection for nashville grid to convert mims_spatial outputs to
ll

::set DATA_POLY_MAP_PRJN +proj=lcc,+lat_1=30,+lat_2=60,+lat_0=40,+lon_0=-100
```

```
::set input projection to EPA Lambert to convert surrogate input files to ll

set DATA_POLY_MAP_PRJN=+proj=lcc,+lat_1=33,+lat_2=45,+lat_0=40,+lon_0=-97

set DATA_POLY_ELLIPSOID=SPHERE

set OUTPUT_POLY_MAP_PRJN=LATLON

set OUTPUT_POLY_ELLIPSOID=SPHERE

echo Converting from %DATA_POLY_MAP_PRJN% to %OUTPUT_POLY_MAP_PRJN%

echo Input file = %POLY_DATA

echo Output file = %POLY_OUT_NAME

%MIMS_EXE%

copy %POLY_DATA%.dbf %POLY_OUT_NAME%.dbf
```