IDL-based Geospatial Data Processor (IGDP)
: A new spatial allocator and raster data processor

Hyun Cheol Kim [1,2], Daewon W. Byun[1], Clint Harper[3], Fong Ngan[1,4], Daniel Tong[1,2], Pius Lee[1], and Ivanka Stajner[5]

[1]NOAA/ARL, Silver Spring, MD [2]ERT, Laurel, MD [3]UCAR, Boulder CO [4]TCEQ, Austin TX [5]Noblis, Inc., Falls Church, VA

Development of IGDP:

A new tool for fine resolution geospatial data processing has been developed. Fast and accurate Geographic Information System (GIS) data processing tools are essential in air quality studies; especially in preparing model emission inputs as finer resolution air quality simulations become more commonplace. An IDL-based Geospatial Data Processor (IGDP), written in Interactive Data Language (IDL by ITT Visual Information Solutions), has been created and can process GIS data both in vector format (e.g. ESRI shapefiles) and raster format (e.g. GEOTIFF and IMG) for any given domain. Processing speeds have been improved through the use of polygon-clipping routines and other algorithms optimized for particular applications. The raster tool utilizes a histogram reverse-indexing method, which enables easy access of grouped pixels, so it can generate statistics of pixel values within each grid cells, with improved speed and enhanced control of memory usage. IGDP supports map projection conversions between Lambert Conformal Conic (LCC), Rotated latitude-longitude (RLL, used in National Centers for Environmental Protection (NCEP) North American Model (NAM)), geographic latitude-longitude coordinates, and also is capable of easy expansion to all 40 projections that IDL supports.

Spatial allocator:

- In order to generate "surrogate" files for emission processing, spatial allocator is an essential tool.
- Spatial allocator requires huge computational power to calculate fractional weightings between GIS polygons and/or polylines and gridded cells, so an efficient polygon/polyline cutting algorithm is important. Some GIS data have more than millions of entities.
- A key for faster spatial allocator is to optimize computational iterations in both polygon clipping and map projection calculations.

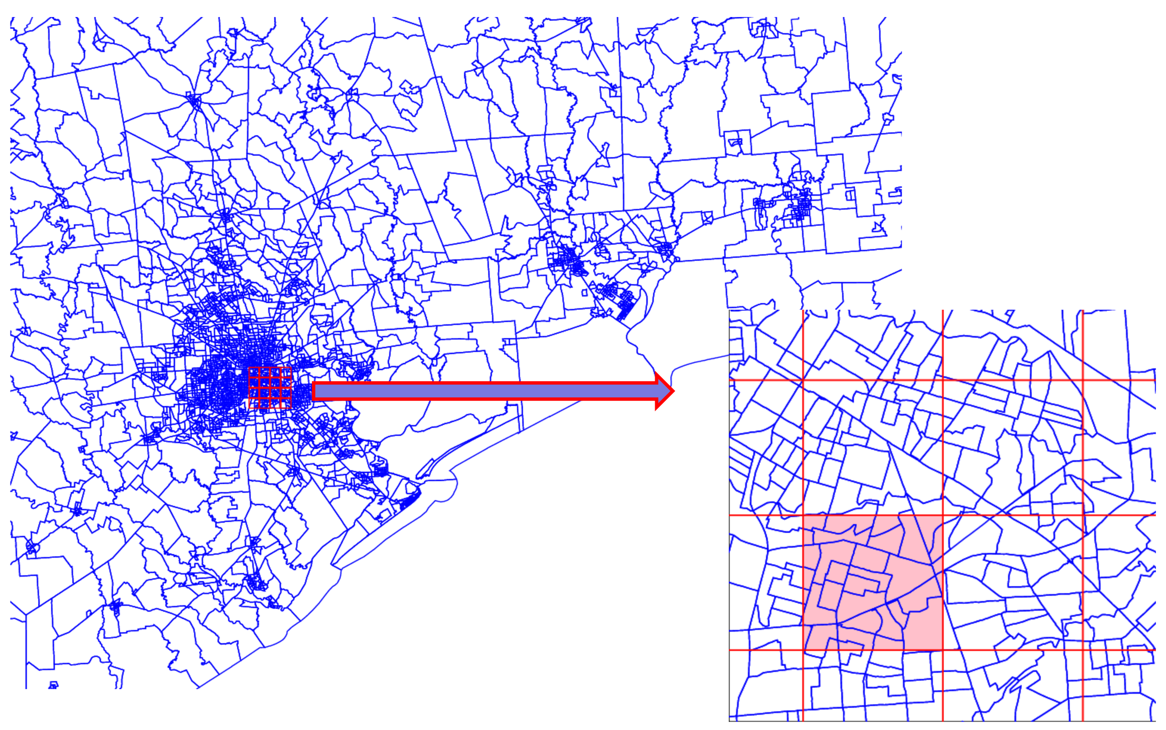


Figure 1. GIS data for population (census track) in Houston, with 4km grid cells (red line)

Polygon clipping algorithm:

The Sutherland-Hodgman algorithm is used for clipping polygons. It works by extending each line of the convex clip polygon in turn and selecting only vertices from the subject polygon that are on the visible side.

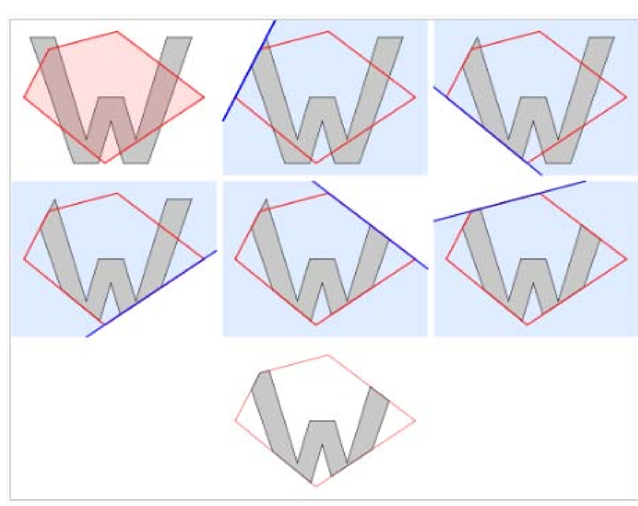


Figure 2. The Sutherland-Hodgman polygon cutting algorithm

Test of spatial allocator:

- We have tested new spatial allocator for various domains settings (Fig. 3)
- The tools supports both Lambert Conformal Conic (LCC) and Rotated Lat/Lon (RLL) map projection

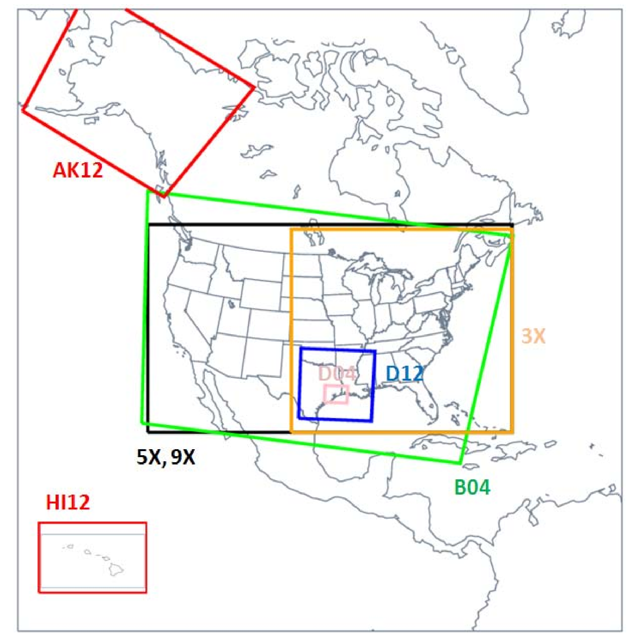


Figure 3. Testing domains

Name	5X	3X	D12	D04	HI12	AK12	B04	9X
Resolution	442X265 (12 km)	268X259 (12 km)	89X89 (12 km)	83X65 (4km)	80X52 (12 km)	199X163 (12 km)	1114X880 (4 km)	5200X3200 (1 km)
Projection	LCC	LCC	LCC	LCC	LCC	LCC	RLL	LCC
# of surrogates	130	130	130	57	62	104	130	130
Processing Time	13.4 hr	9.8 hr	5.5 hr	3.5 hr	1.2 hr	3.2 hr	25.8 hr	~1 week

Raster data processor:

The raster tool uses a histogram reverse-indexing method in IDL histogram function, and is capable of faster access of grouped pixels. For each grid cells, the raster tool provides histogram and statistics of pixels inside. Figure 4 shows an example of 30-m NLCD LULC data near Houston region.

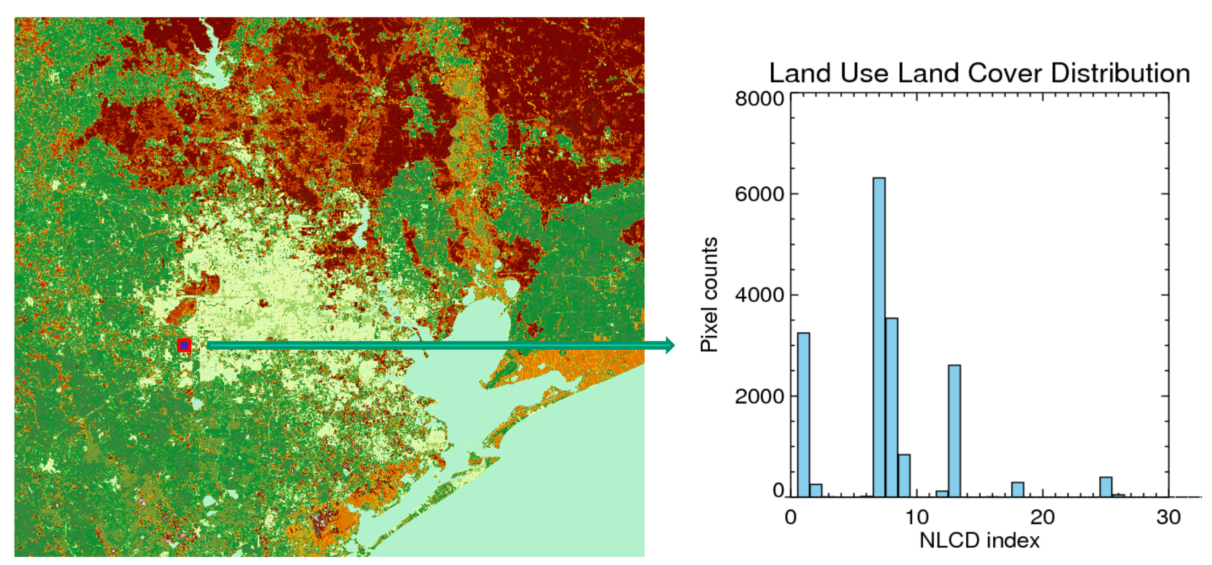
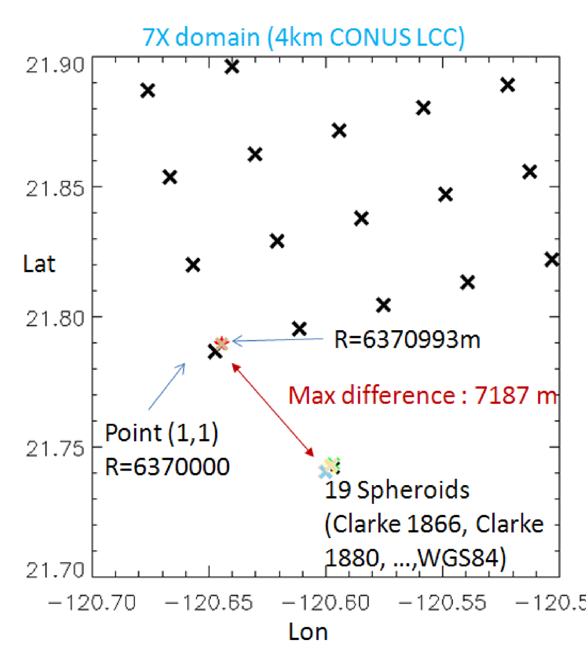


Figure 4. 30-m NLCD land us e land cover data set near Houston (left), and an example pixel distribution from a grid cell

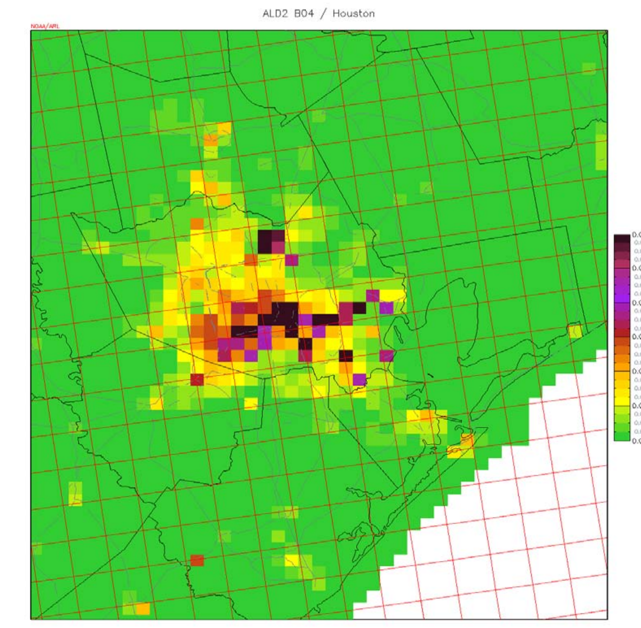
Common mistakes in Geospatial data processing:

- Most of Met. simulation assumes "ideal" spherical Earth shape, while most of GIS data are based on more "real" ellipsoidal earth shape. (e.g. Normal LCC (Sphere) and GCTP (General Cartographic Transformation Package), http://gcmd.nasa.gov/records/USGS-GCTP.html)
- Improper choice of geo-information (Earth radius and/or earth shape – e.g. "Datum") could cause significant error in placing emissions. (up to 7 km difference in 4km CONUS domain)



Applications : Art of spatial re-gridding:

- The new spatial allocator in IGDP can not only generate surrogate files for emission processing, but also has various applications in handling model outputs and GIS data.
- Regridding of model outputs for different map setting is crucial for inter-comparisons of modeled results, or for comparison to satellite observations.
- By using new spatial allocator, exact regridding with fractional weighting is possible. Regridding method can be chosen for total mass conservation (for emission), or for consistent concentration (for gas/matter density).

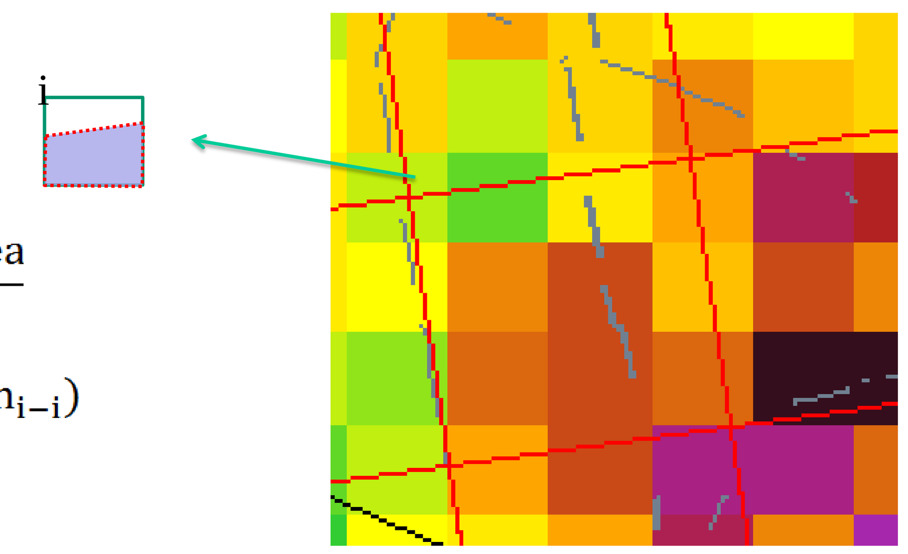


Fractional weighting (total mass)

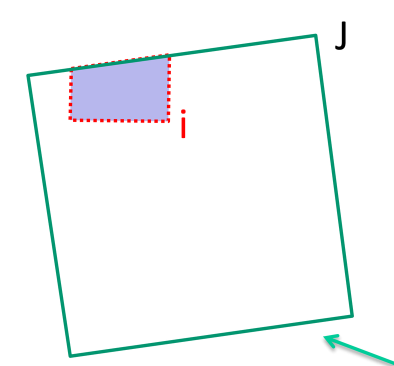
i : original grid cell indices
J : target grid cell indices

$$\text{Fraction}_{i-i} = \frac{\text{overlapped area}}{\text{area of cell i}}$$

$$M_J = \sum (M_i \cdot \text{Fraction}_{i-i})$$



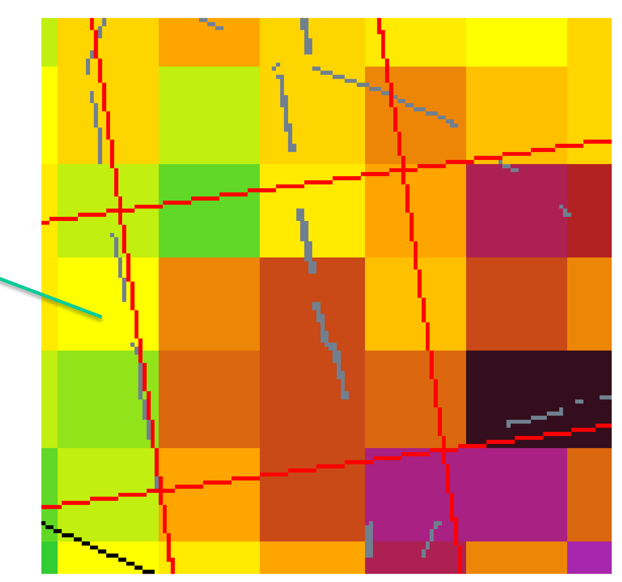
Fractional weighting (concentration)



i : original grid cell indices
J : target grid cell indices

$$\text{Fraction}_{i-J} = \frac{\text{overlapped area}}{\text{area of cell J}}$$

$$C_J = \frac{\sum (C_i . \text{Fraction}_{i-J})}{\sum \text{Fraction}_{i-J}}$$



Highlights:

- The IGDP is newly developed geospatial data processing tool, to handle GIS shapefiles and raster data, in generating emissions surrogates and land/ocean inputs, for air quality simulations.
- With smart optimization of polygon clipping algorithms and map projection capability, the IGDP shows a significant improvement in data processing speed. It shows 10~100 times faster data processing time, compared to current spatial allocators or GIS processing softwares.
- Application of IGDP spatial allocator enables exact regridding of spatial data, using fractional weighting method, which conserves total mass, or mean density.

CONTACT: Hyun Cheol Kim (hyun.kim@noaa.gov)