Urban GIS for Health Metrics

Dajun Dai

Department of Geosciences, Georgia State University
Atlanta, Georgia, United States

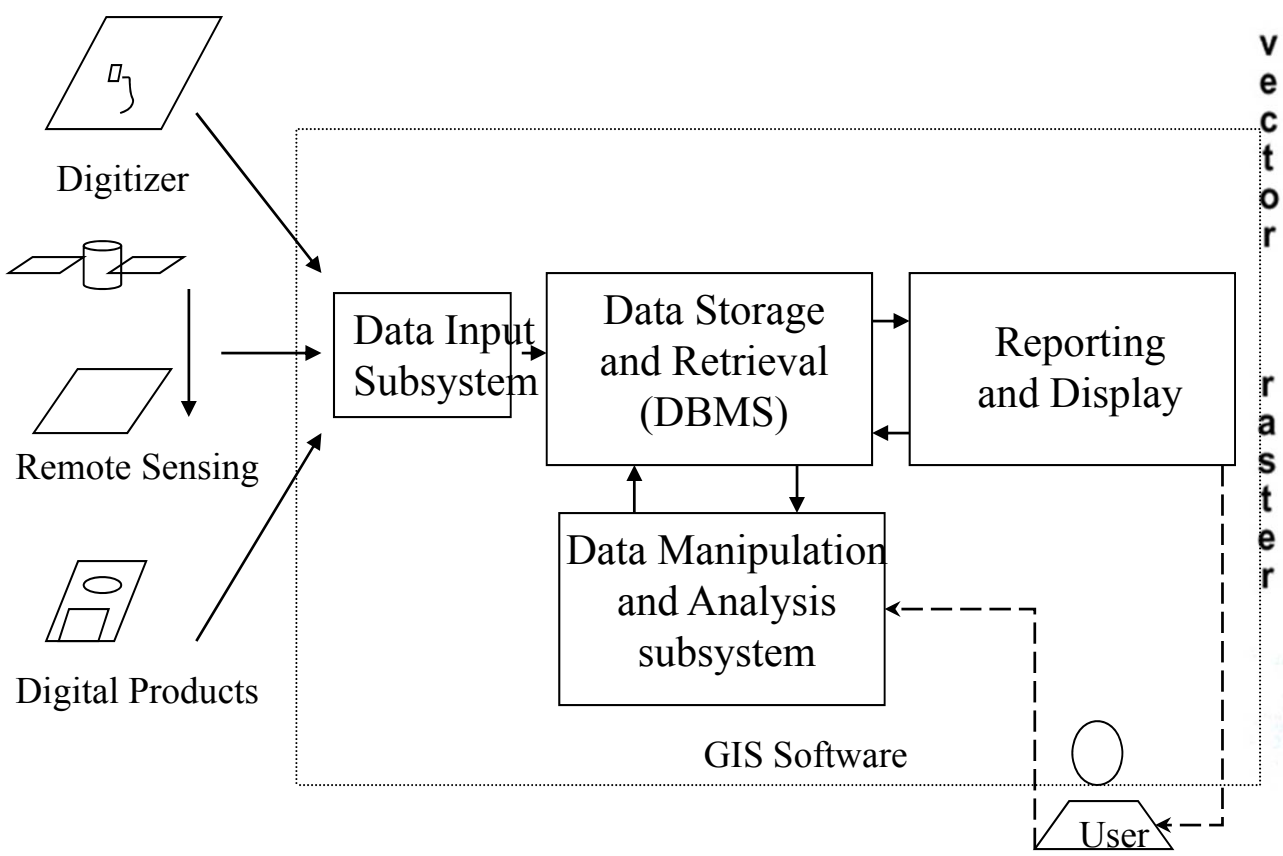
Presented at International Conference on Urban Health, March 5th, 2014

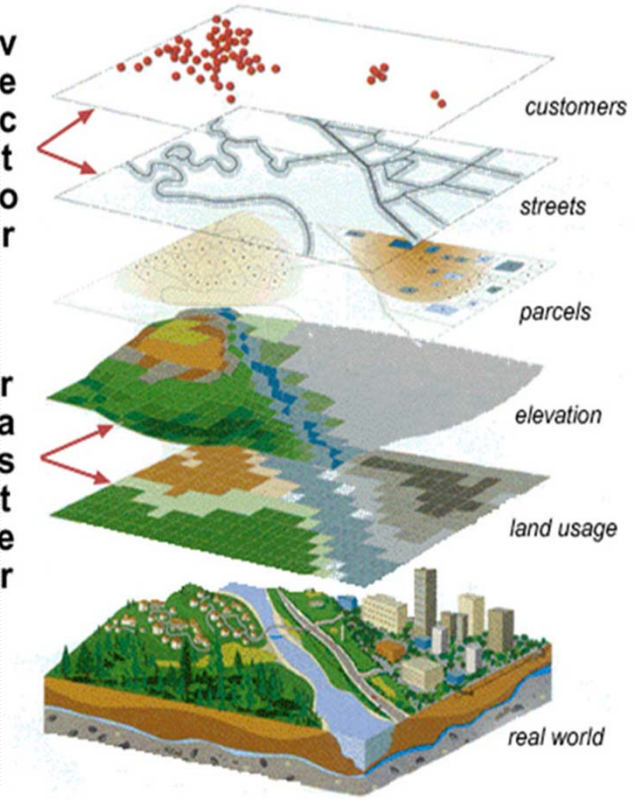
People, Place, and Health

- Location, location, location!!!
  - Almost everything that happens, happens somewhere
  - GIS keeps track not only events, activities, and things, but also where they happen
- Geography
  - Where (activity space & migration)
  - People affected by their environments (natural, built, social, economic, etc)
- Pubic health
  - Not simply the absence of disease
  - State of physical, social, and emotional well-being of residents

What is Geographic Information Systems (GIS)?

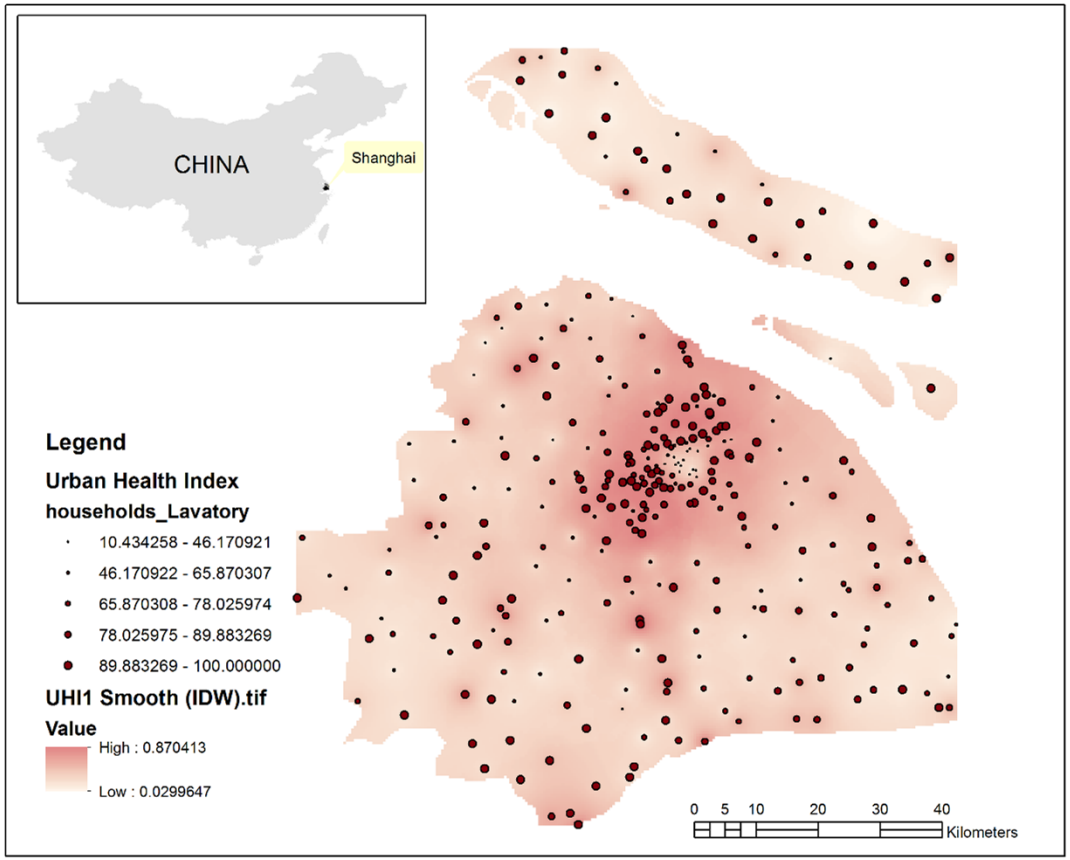
- Computer hardware & software for capturing, storing, retrieving, analyzing, and output spatial data.
  - Maps: a vital role in the analysis (visualization) and display components of GIS.



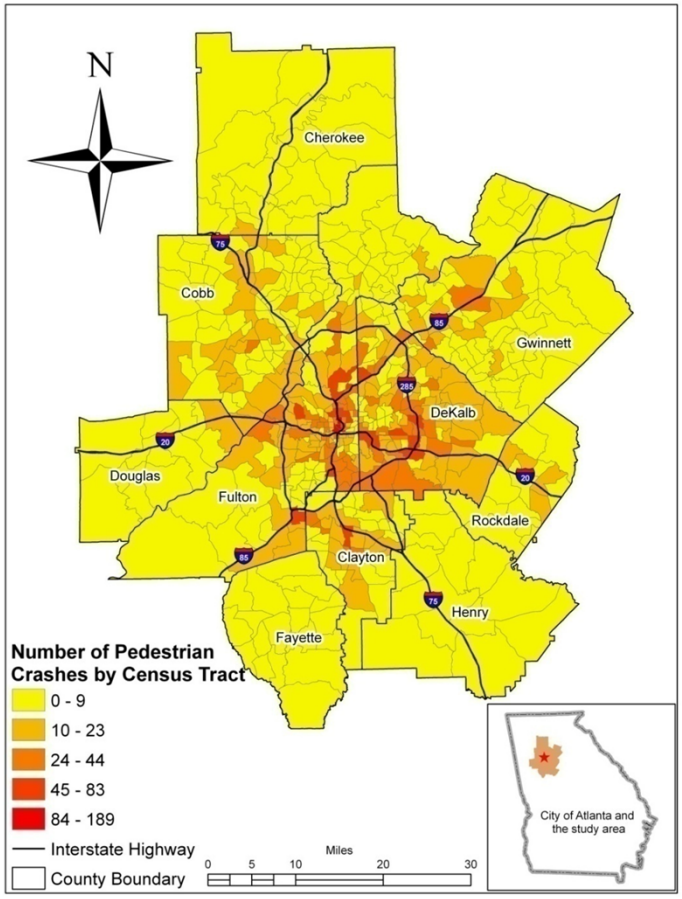


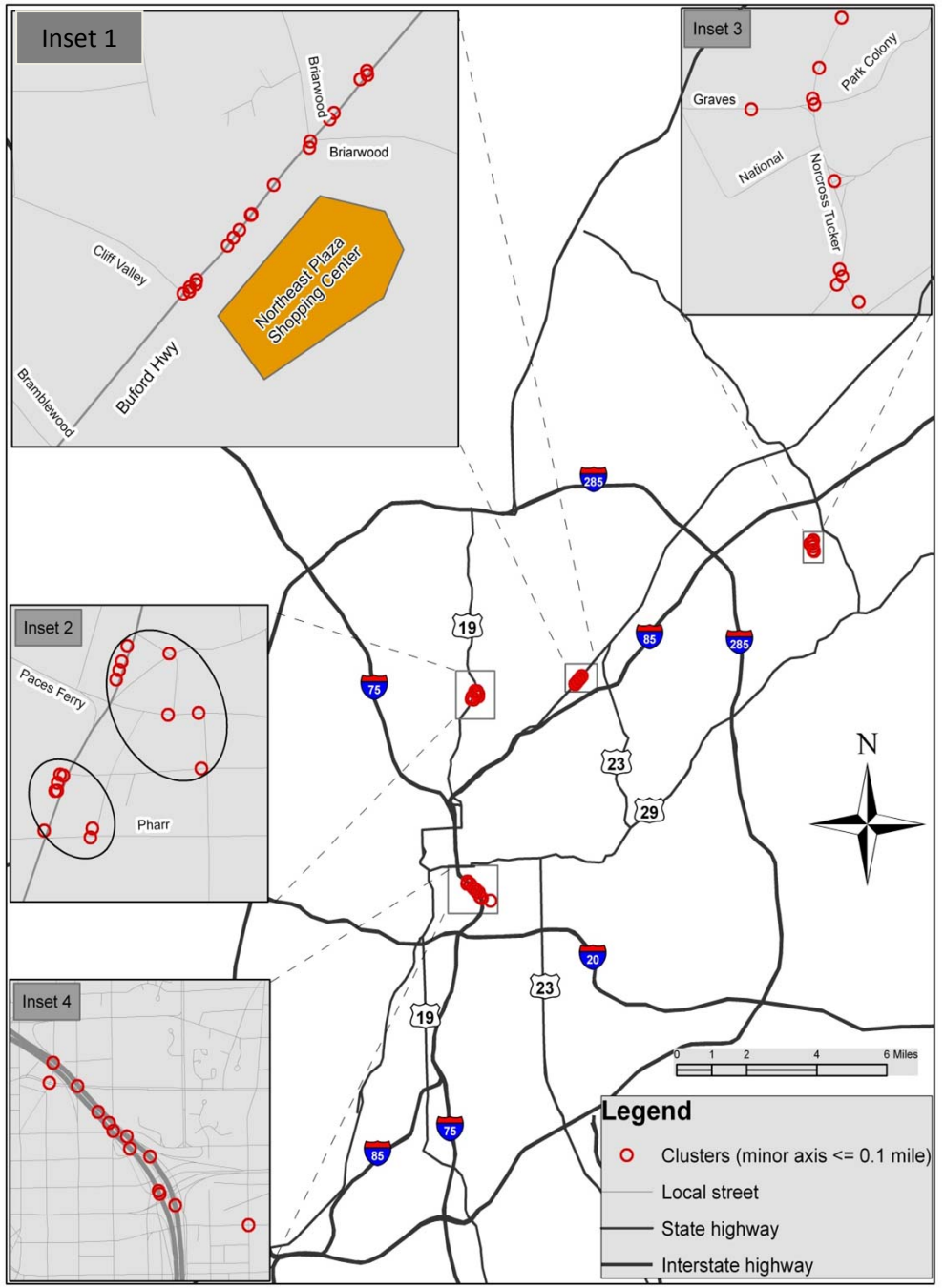
Mapping Health

- Spatio-temporal variation
  - Geocoding for individual cases
  - Chropoleth map for aggregated cases
  - Dot density/size map



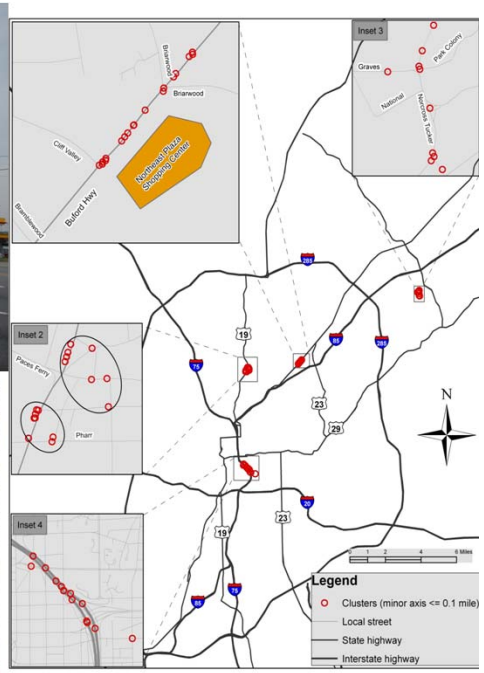
Pedestrian Crashes in Metro Atlanta

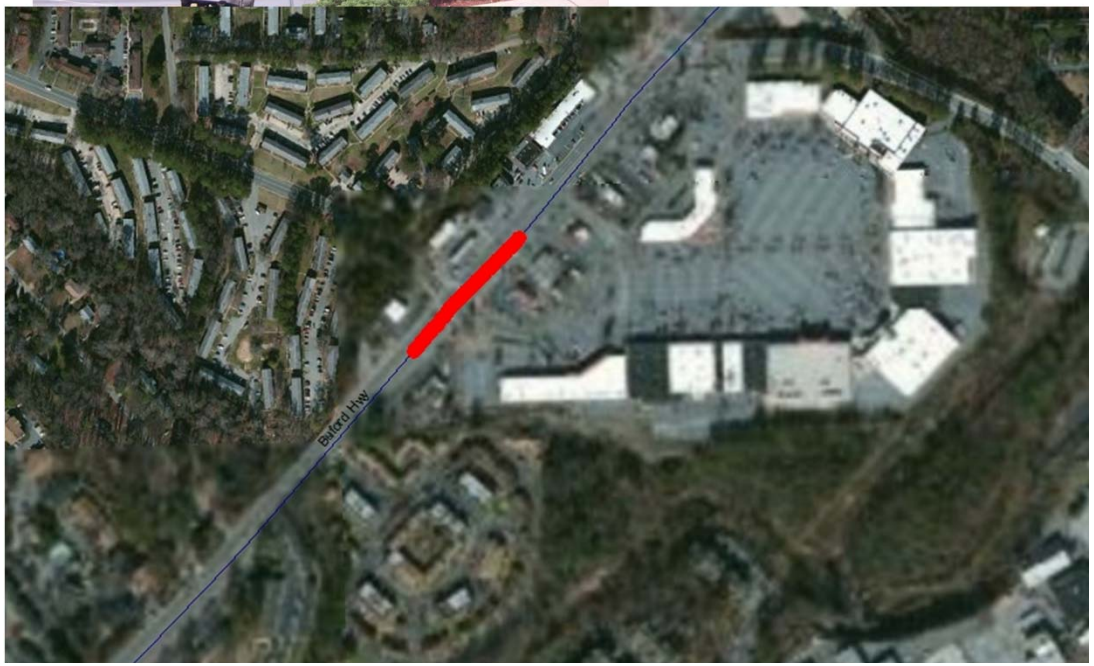




Built Envion't

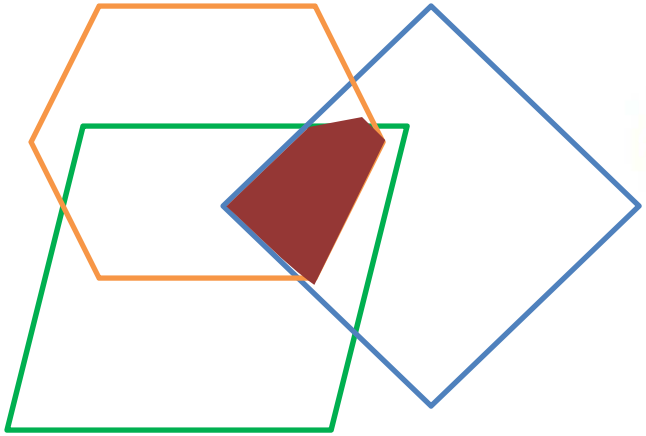
- Cluster 1
  - Statistically significant(P=0.01; 01-04)
- Sidewalks
  - discontinuous
  - only available on one side
  - Apartment complex bisected by busy roads
  - Mixed with commercial properties.

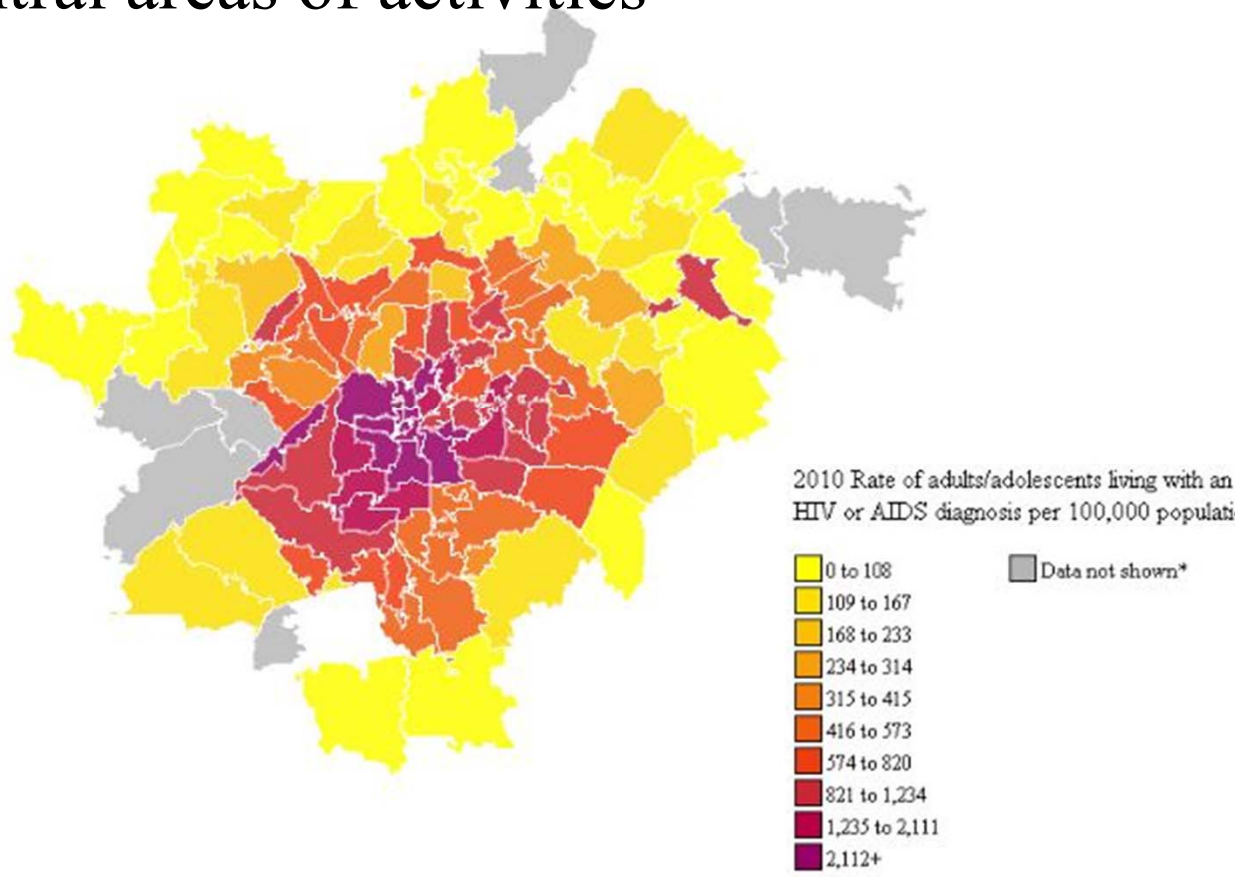


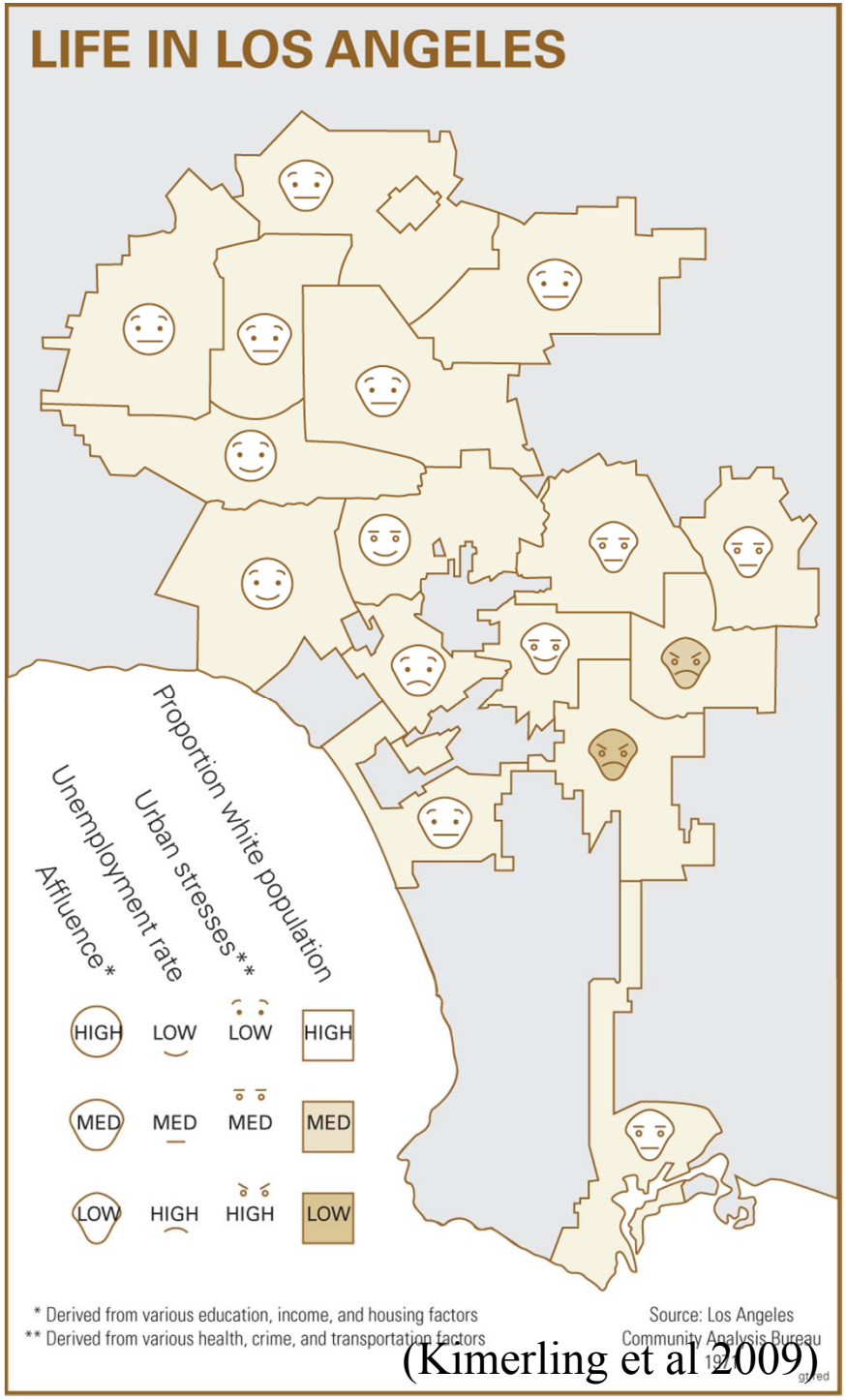


Spatial Surveillance

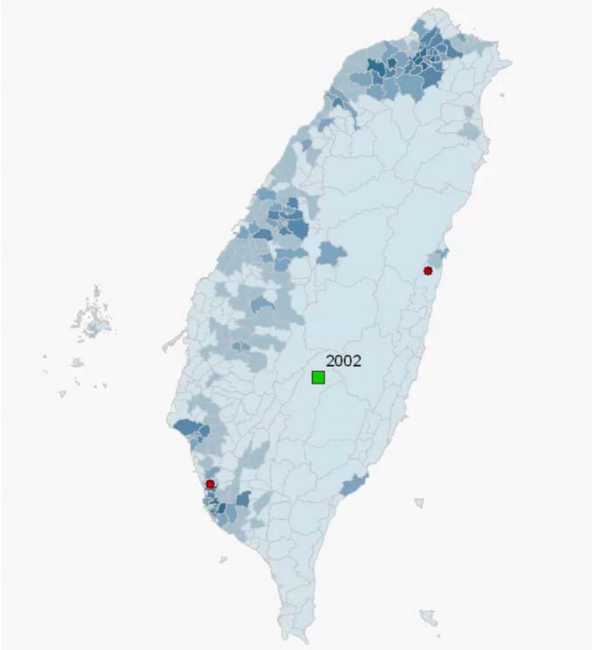
- Spatial clustering in GIS
  - E.g., Spatial Clustering of HIV in Atlanta (Hixson et al, 2011)
- Identify core central areas of activities



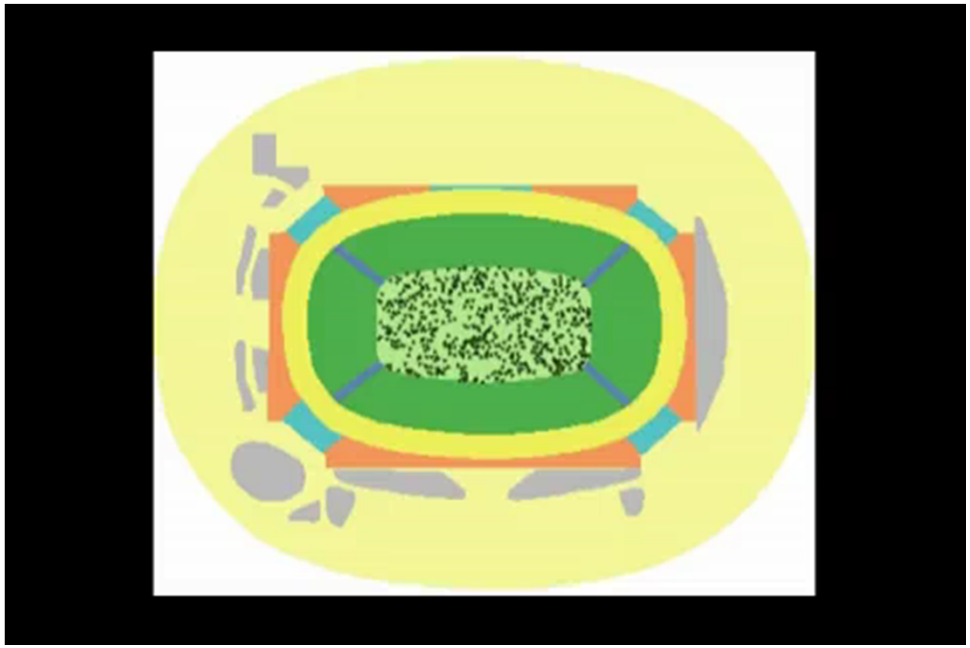


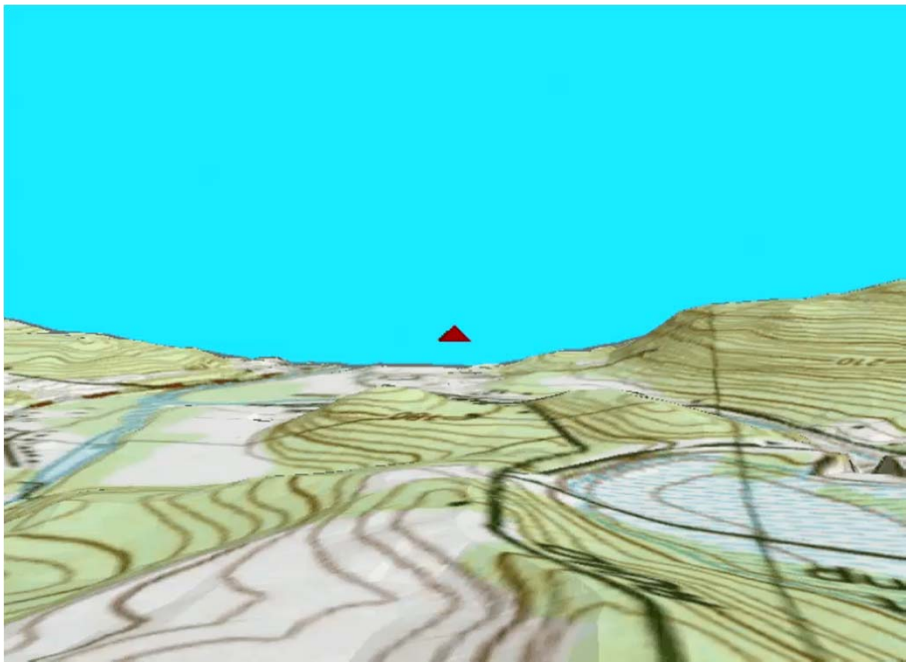


(Kimerling et al 2009)

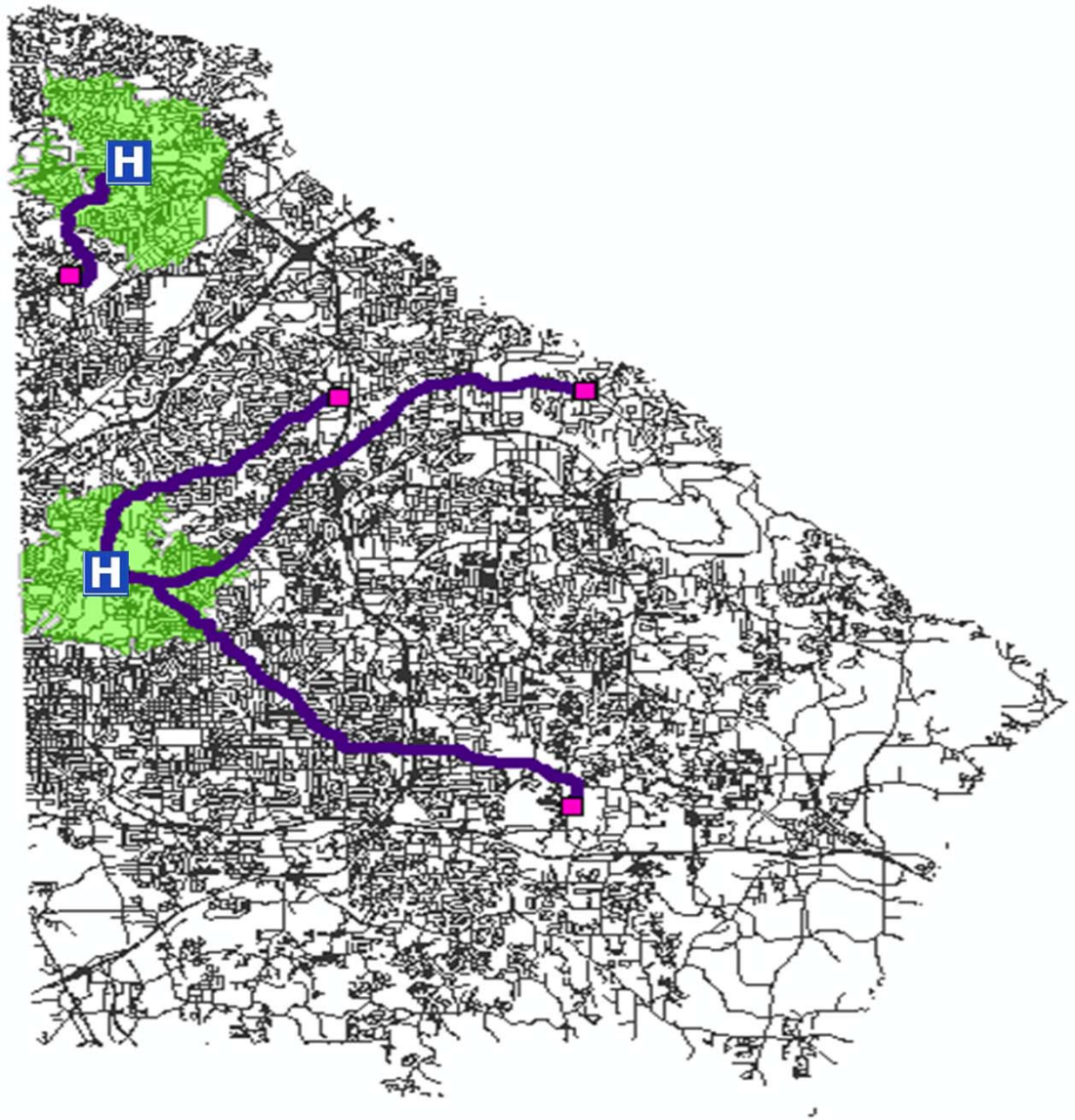


Space-Time Visualization



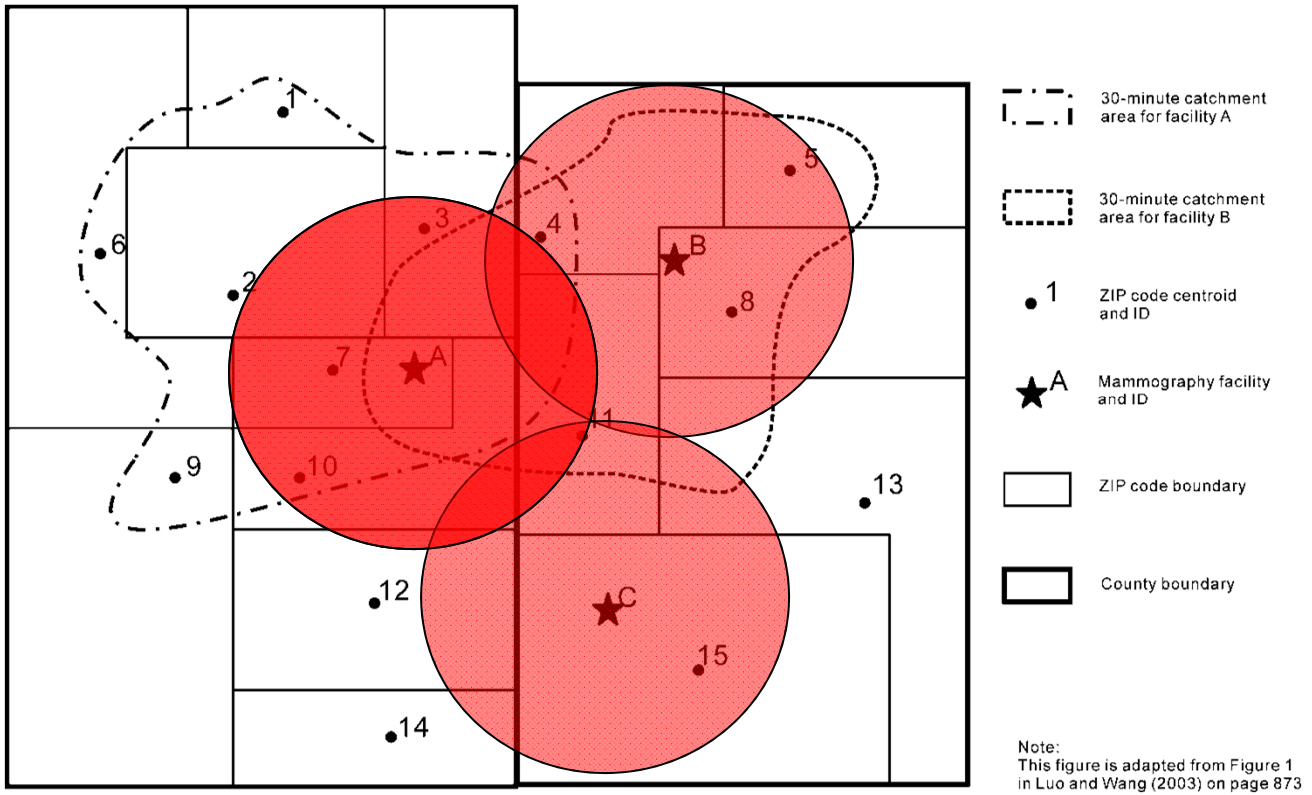


Healthcare Access: Routing & Coverage



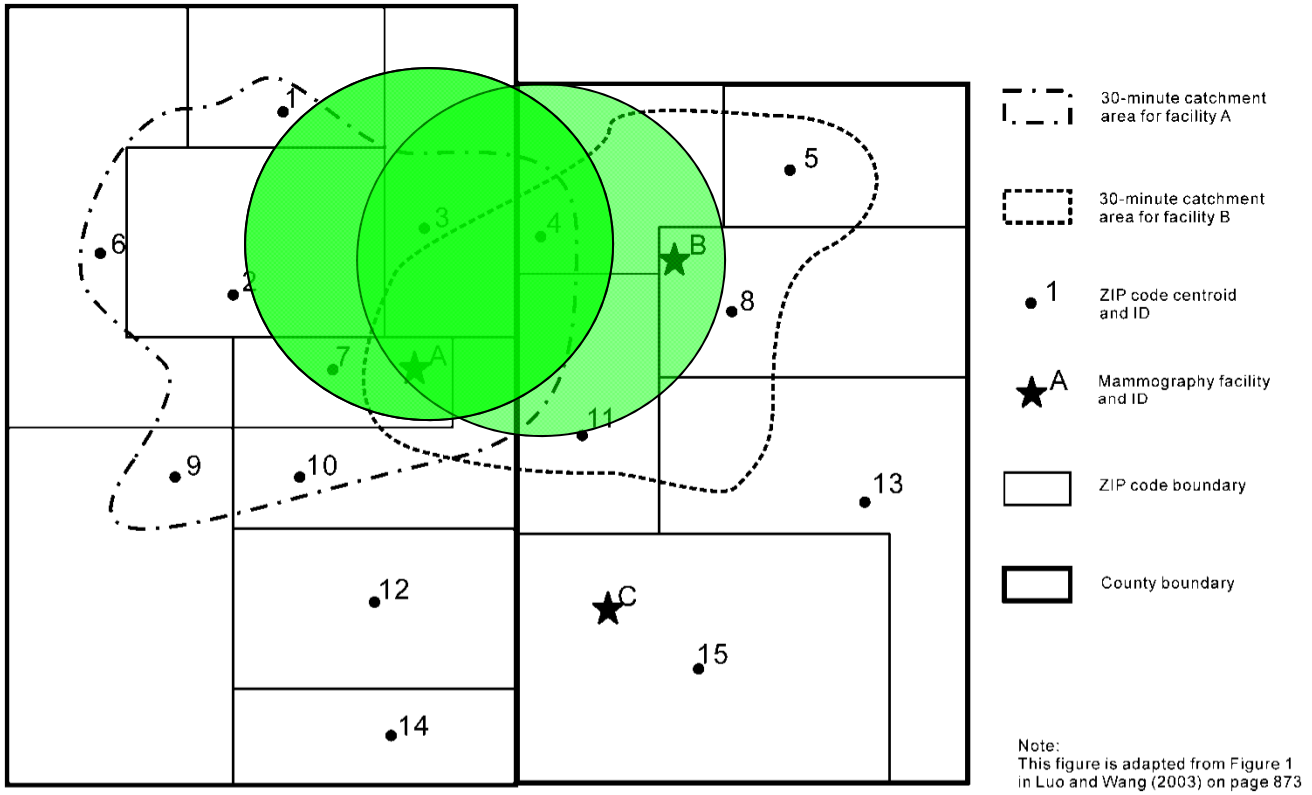
Mapping spatial access to Mammography

- Step 1: float a catchment on Mammography facilities
  - For each facility: (1) search all population locations that are within the catchment; (2) inverse the population to obtain facility-population ratio (v_j)



Measure spatial access (con't)

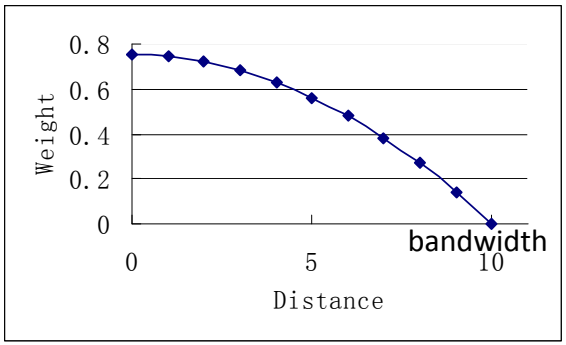
- Step 2: float a catchment on population locations
  - For each location: (1) search all facilities that are within the catchment; (2) weigh their facility-population ratios (v_j) using the kernel function; and (3) summarize the weighted ratios

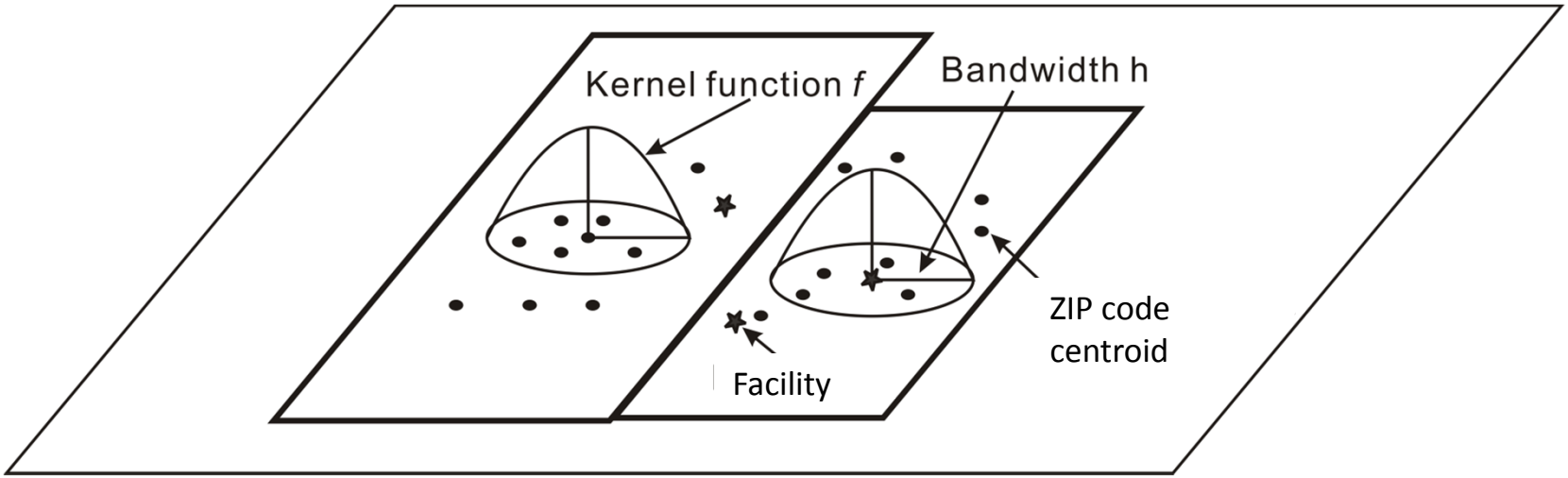


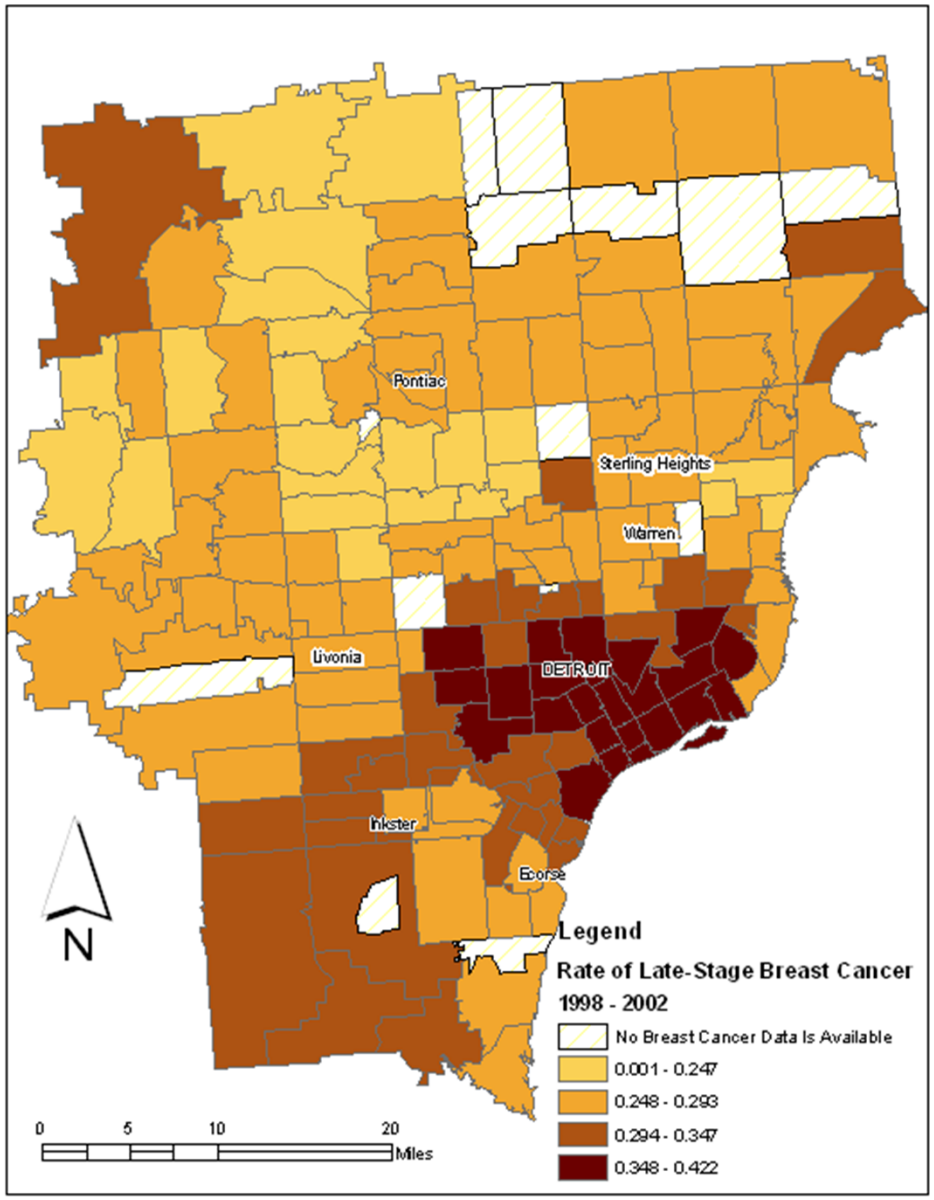
Measure spatial access to health care (con't)

- Integrating kernel function (Gaussian) to create weighted populations when computing the facility-population ratio
- Kernel bandwidth = catchment size

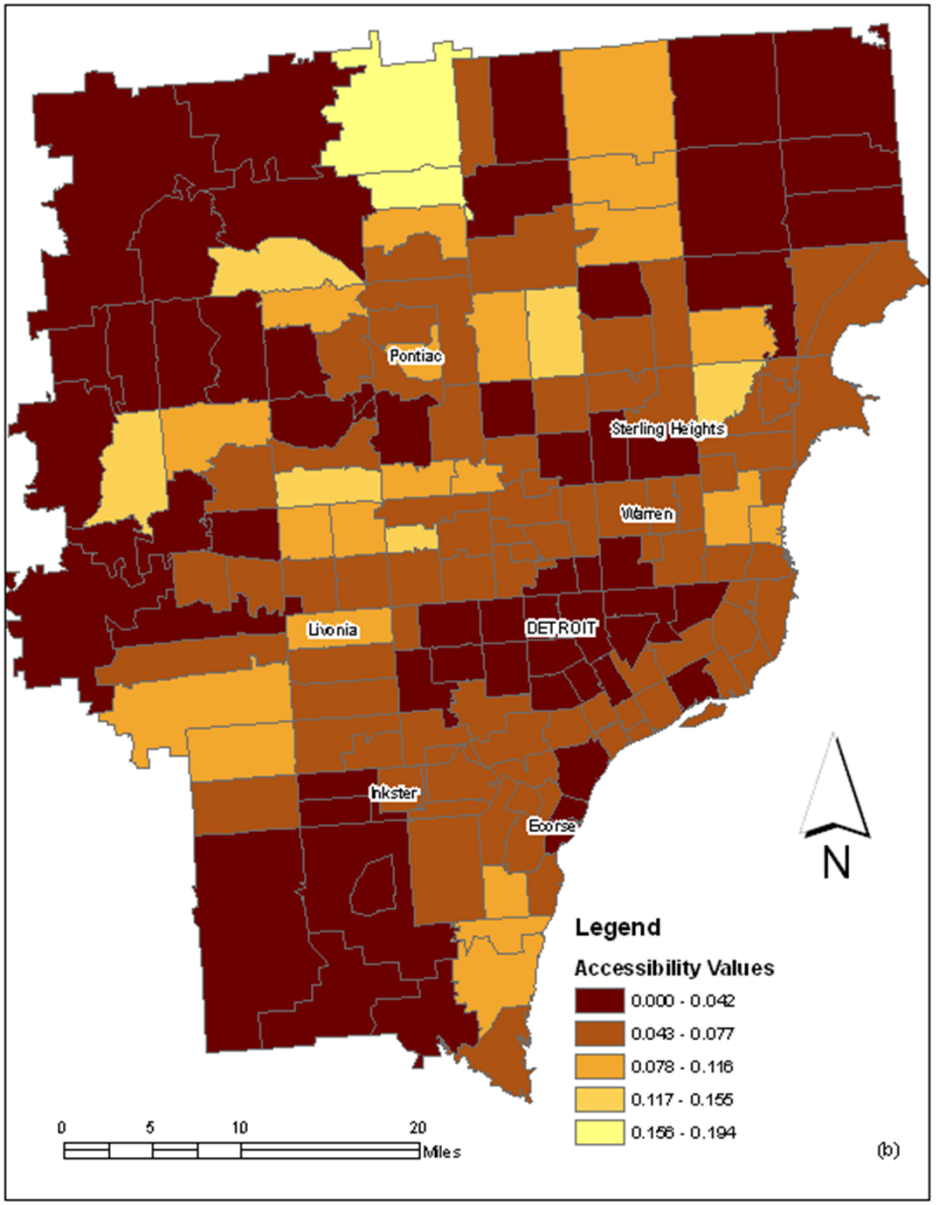
$$G(d_{ij}, d_0) = \begin{cases} \dfrac{e^{-1/2\times(d_{ij}/d_0)^2} - e^{-1/2}}{1 - e^{-1/2}} & \text{if } d_{ij} \le d_0 \\ 0 & \text{if } d_{ij} > d_0 \end{cases}$$







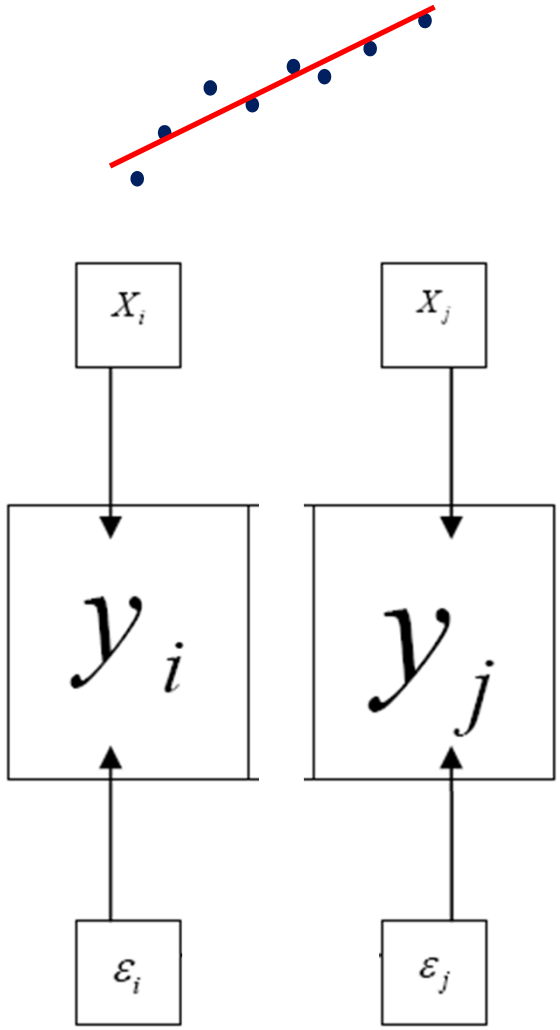
Late stage breast cancer%



Access to mammography facilities (d=10 min)

- Multiple linear regression model:
  - Regression: $Y=a+b_1x_1+b_2x_2+\ldots.b_nx_n+\varepsilon$
  - Housing Price = Sq.ft. + Age + Median Income + Dist_Marta + *error*

Something the model can't account for



- Assumptions
  - Random errors have a mean of zero
  - Random errors have a constant variance and are uncorrelated
  - Random errors have a normal distribution
- The assumptions may not be always satisfied in practice

Standard Linear Regression

Why is spatial regression?

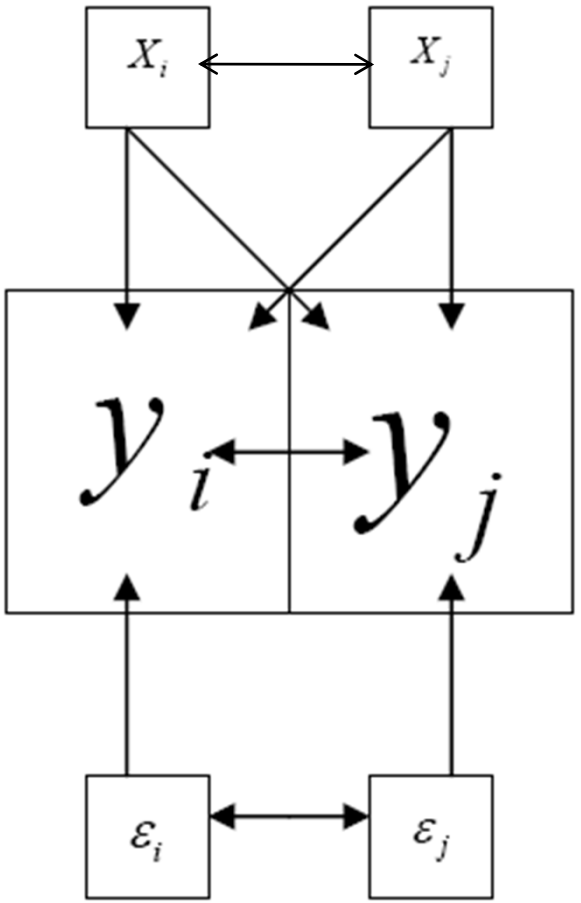
- First law of geography: values of a variable systemically related to geographic location

Price = Sq.ft. + Age + Median Income + Dist_Marta + error

Housing price is related to location nearby

Median income is related to location nearby

Housing price is related to median income nearby

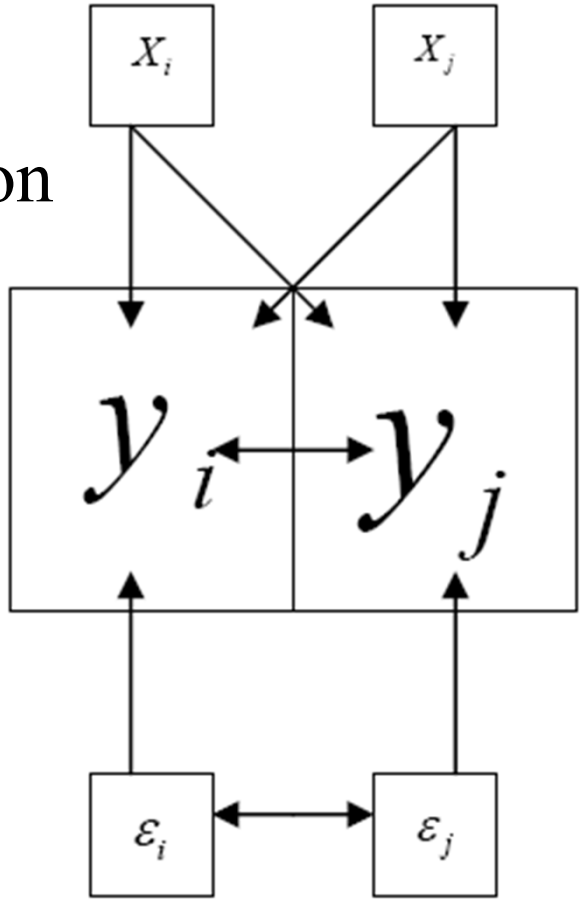


Assumptions in standard regression may not be satisfied

Spatial Regression Model

Spatial Lag Model: $Y=\rho WY+aX+\varepsilon$

- Account for the spatial autocorrelation

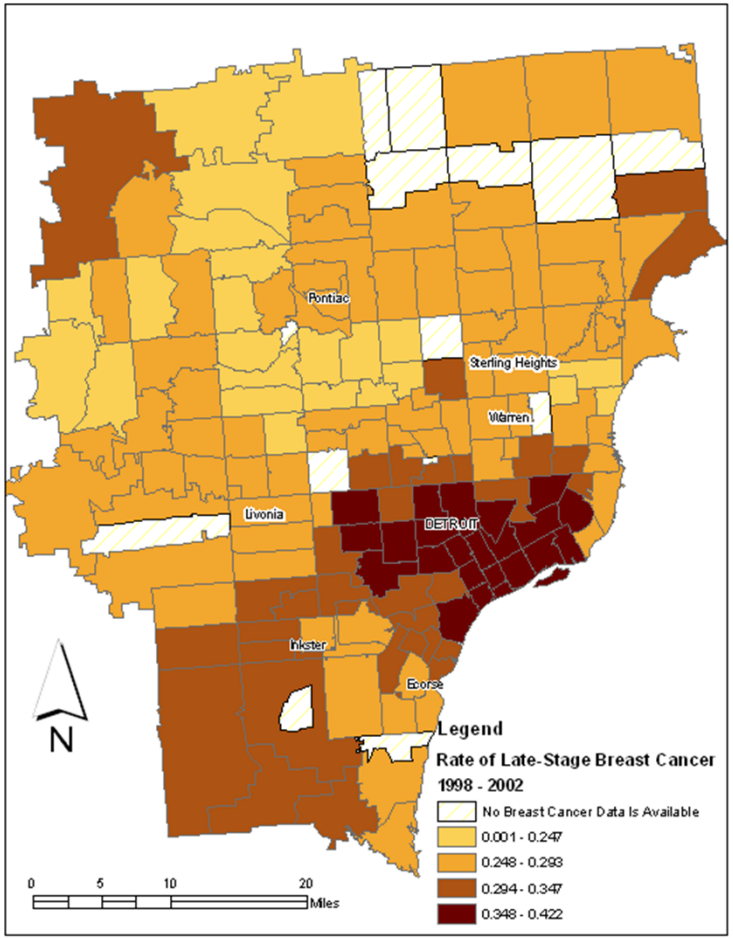


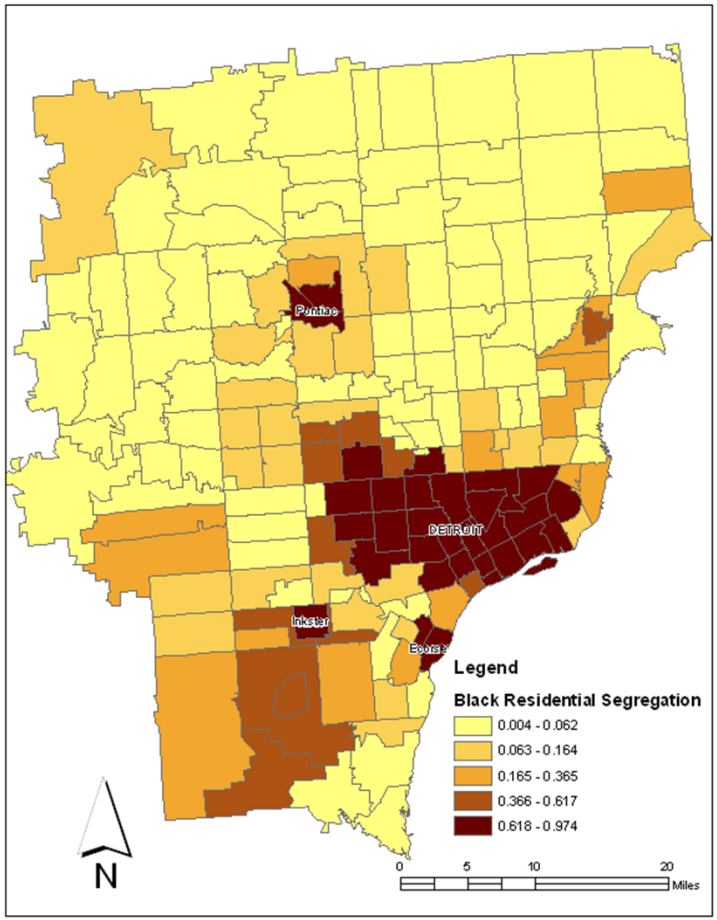
Spatial autocorrelation

Price = W*Price + Sq.ft. + Age + Median Income + Dist_Marta + error

Spatial Regression Model

Late-stage breast cancer and black residential segregation in City of Detroit and its 30-min buffer zone





A case study using spatial lag model

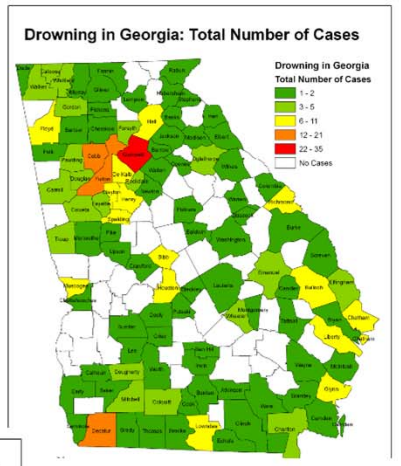
	St. Regression Model		Spatial Lag Model	
	Coefficients	*t* values	Coefficients	*t* values
Constant	0.261**	74.889	0.048**	3.843
Black Segregation	0.113**	13.513	0.046**	7.184
Spatial Lag			0.790**	17.344
R^2	0.544		0.817	

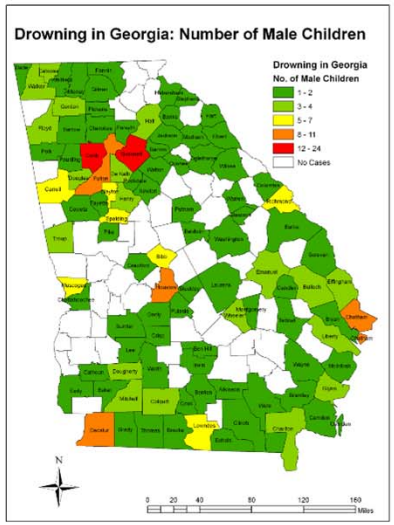
**significance at the 0.001 level

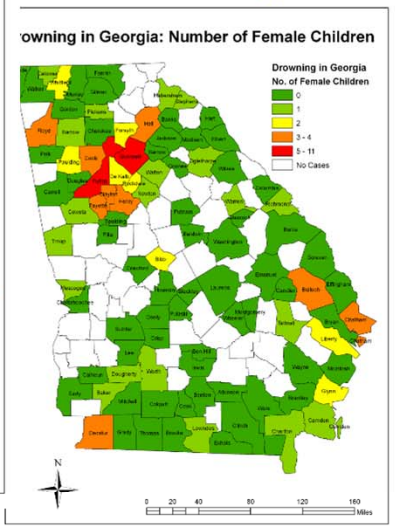
Standard Regression vs. Spatial Lag Model

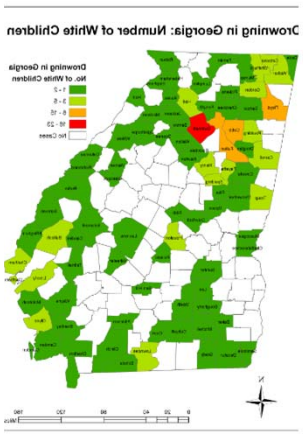
Mapping Childhood Drowning

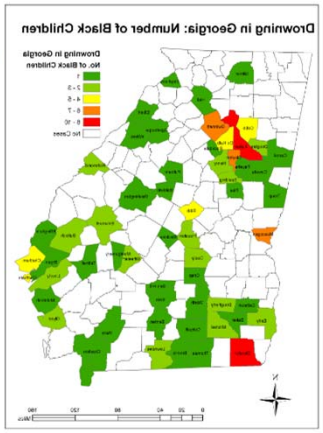
- 2002-2008 childhood drowning cases (N=276)
  - Residential address
  - Demographic info
  - Drowning place (descriptive)
  - Source: Georgia Office of the Child Advocate

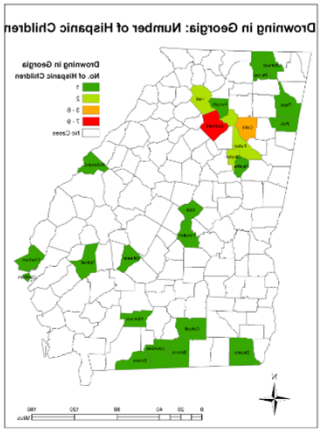


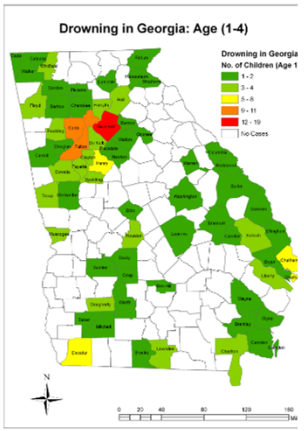


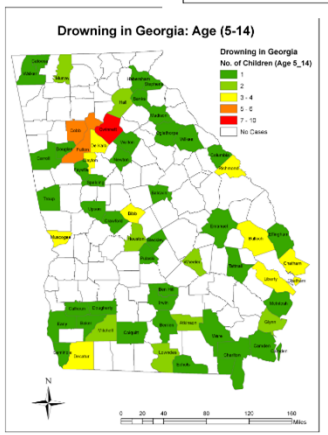


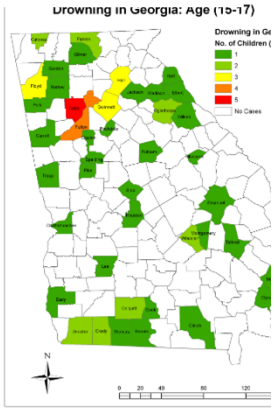




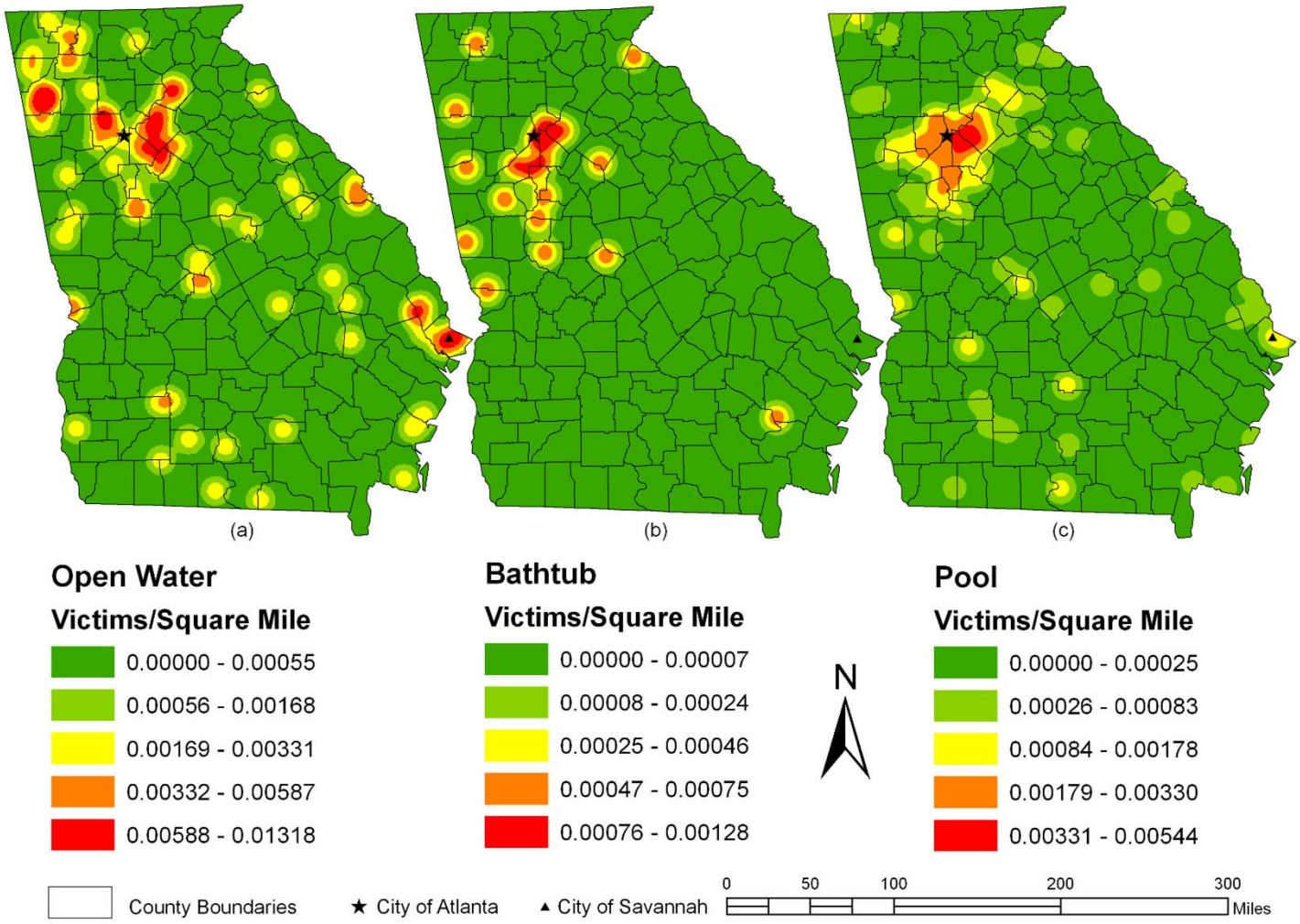






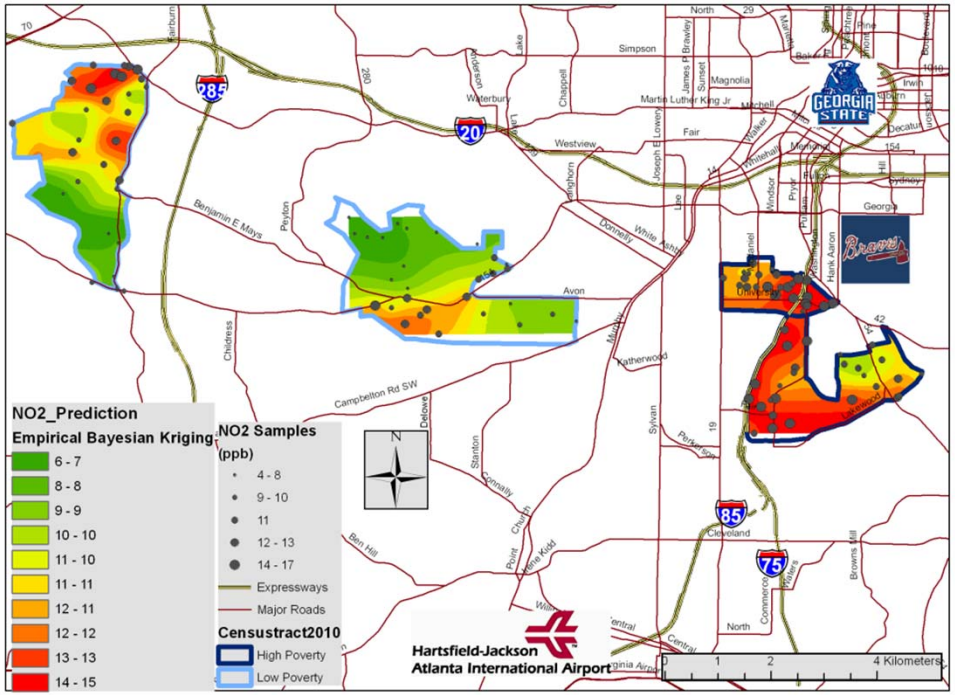


Spatial Smoothing Using Density



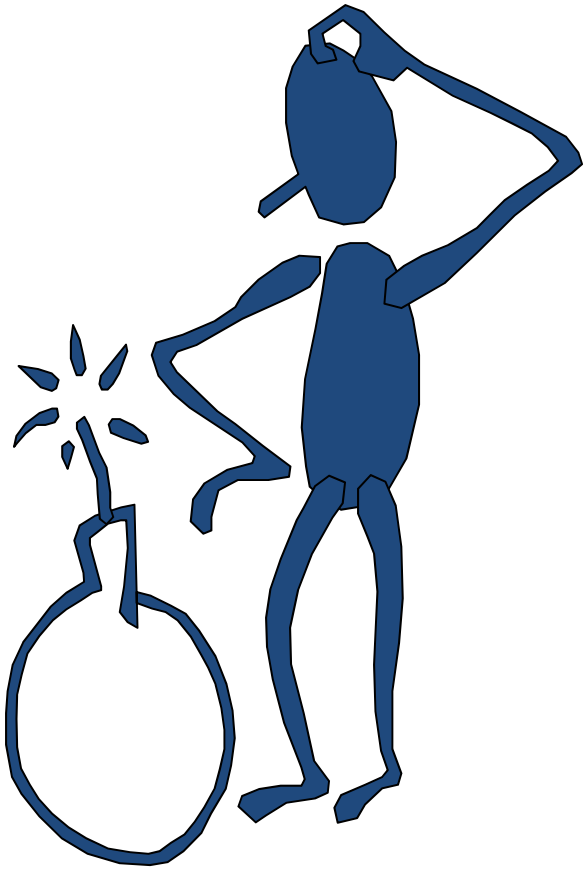
Spatial Interpolation

- Spatial Interpolation
  - Use points with known values to estimate values at other points
- Sample points: points with known values
  - The number and distribution determine the accuracy of spatial interpolation
- Basic assumption
  - 1st Law of Geography (Everything is related to everything else, but near things are more related than distant things)



Challenges

- Big Data
  - Location
- Privacy
- Accuracy
  - Latency
  - Migration



Thank you! Questions?

Dajun Dai

ddai@gsu.edu

ICUH 2014 Manchester