GEOG 3340: Introduction to Human Geography Research

Lecture 2: Representing Social Process in GIS

Guofeng Cao

www.myweb.ttu.edu/gucao

Department of Geosciences
Texas Tech University

guofeng.cao@ttu.edu

Fall 2015

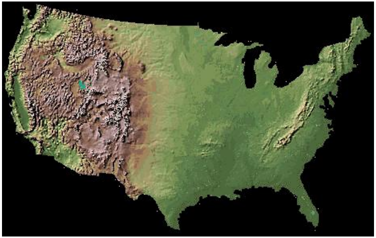
Reading for last week:

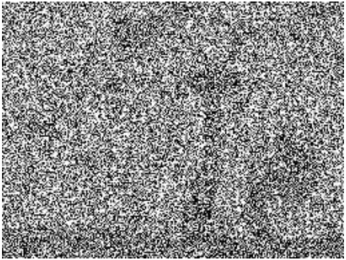
- Goodchild, M. F., Anselin, L., Appelbaum, R. P., & Harthorn, B. H. (2000). Toward spatially integrated social science. International Regional Science Review, 23(2), 139-159.
- http://www.csiss.org/

1. Spatial (and temporal) Context: "Everything is related to everything else, but near things are more related than distant things"
   - Waldo Toblers First Law (TFL) of geography
   - nearby things are more similar than distant things
   - phenomena vary slowly over the Earth's surface
   - Compare time series

Laws of GIScience

- Implication of Tobler's First Law (TFL)
  - We can do samplings and fill the gap using estimation procedures (e.g. weather stations)
  - Spatial patterns
  - Image a world without TFL:
    - White noise
    - No lines, polygons or geometry (how to draw a polygon on a white noise map?)

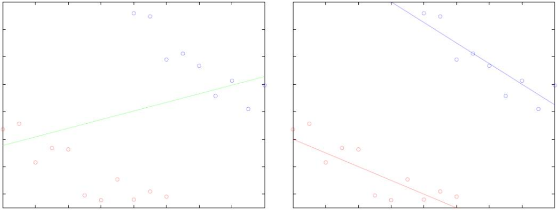




Laws of GIScience

2. Spatial heterogeneity

- "Second law of geography" (Goodchild, UCGIS 2003)
- Earths surface is non-stationary
- Laws of physical sciences remain constant, virtually everything else changes
  - Elevation,
  - Climate, temperatures
  - Social conditions
- Implications
  - Global model might be inconsistent with regional models
  - Spatial Simpsons Paradox (a special case of modified areal unit problem, which we will discuss more in the later of this class)



(a) Global Model   (b) Regional Models

Side note: example of Simpson's paradox

- Simpson's paradox usually fools us on tests of performance in real life
- The following is a real life example. Comparison of recovery rates between a new treatment and a traditional treatment for kidney stones.

	New Treatment	Traditional Treatment
Small Stones	93%(81/87)	87%(234/270)
Large Stones	73%(192/263)	69%(55/80)
All	78%(273/350)	83%(289/350)

- Comparison of batting average of two baseball players:

	1996	1997	Combined
Derek Jeter	25.0%(12/48)	31.4%(183/582)	31.0%(195/630)
David Justice	25.3%(104/411)	32.1%(45/140)	27.0%(149/551)

Laws of GIScience

- In a spatial settings, it is related to modified areal unit problem (MAUP) or omitted variable problem, which will discuss more in the later of this class

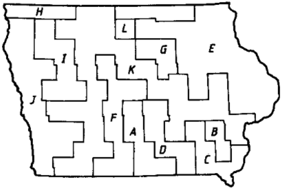


Figure 2a. Zoning system that minimises the regression slope coefficient (-24, r = -.25)

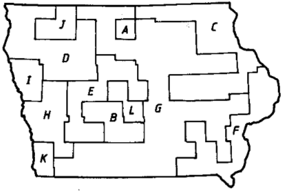


Figure 2b. Zoning system that maximises the regression slope coefficient (12, r = .87)

Figure : Image Courtesy of OpenShaw

3. Fractal behavior
   - What happens as scale of map changes?
   - Coast of Maine

- Implications
   - Scale is critical for the problem of study
   - Volume of geographic features tends to be underestimated
      - length of lines
      - area of polygons
   - Think of the difference of distances that an ant and elephant needed to travel from where I stand to the center of memorial circle

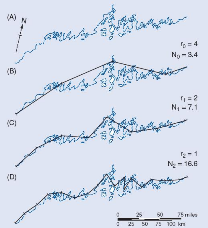
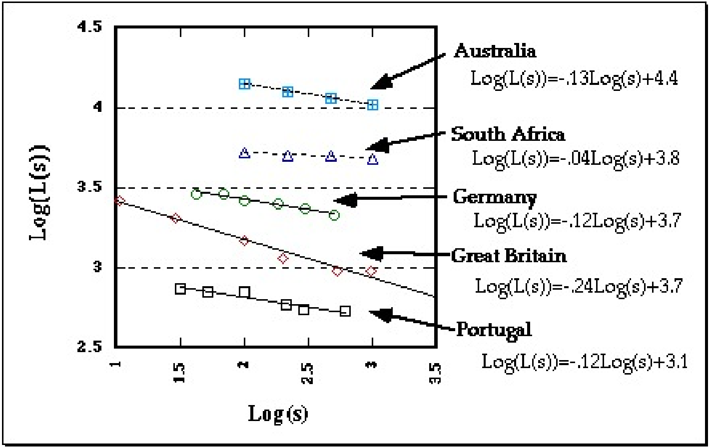


Figure 4.18 The coastline of Maine, at three levels of recursion: (A) the base curve of the coastline; (B) approximation using 100 km steps; (C) 50 km step approximation; and (D) 25 km step approximation

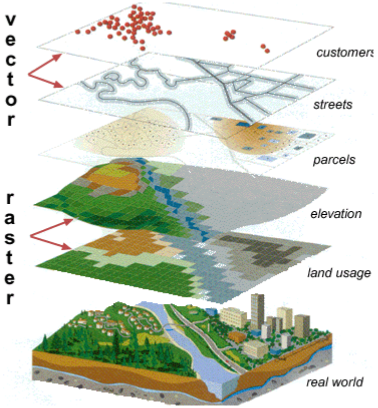
- Richardson Plot



Laws of GIScience

4. Objects and fields

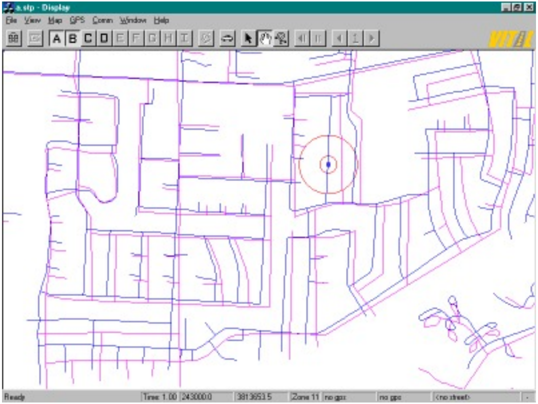
- Two ways of conceptualizing geographic variation
  - as a collection of discrete objects
  - as a collection of continuous fields, functions of location



Laws of GIScience

4. The uncertainty principle

- No representation of the Earth's surface can be perfect
  - no measurements of position can be perfect
  - a GIS will always leave doubt about the true nature of Earth's surface



Laws of GIScience

5. Derivative principle

- Principles that can be derived by combining fundamental ones
- TFL and the principle of uncertainty
  - errors will be spatially autocorrelated
  - relative accuracy will be better than absolute accuracy
  - a map whose absolute positional accuracy is no better than 50m will still show objects in their correct relative location
  - elevations that are accurate to no better than 7m can still be used to estimate slope

Summary

- Spatial context/spatial pattern/spatial structure/spatial dependence/spatial texture..
- Spatial heterogeneity/locality
- Fractal behaviors/scaling effects

Elements of Spatial Data

Elements

- Georeferenced measurements (point or area/region specific samples) Spatial arrangement: regular or irregular (gridded or scattered sampling locations)
- variables/attributes: continuous or discrete (e.g., chemical concentration, soil types, disease occurrences)
- auto- and cross-correlation endemic to spatial data (Toblers first law of Geography)

Types of spatial data

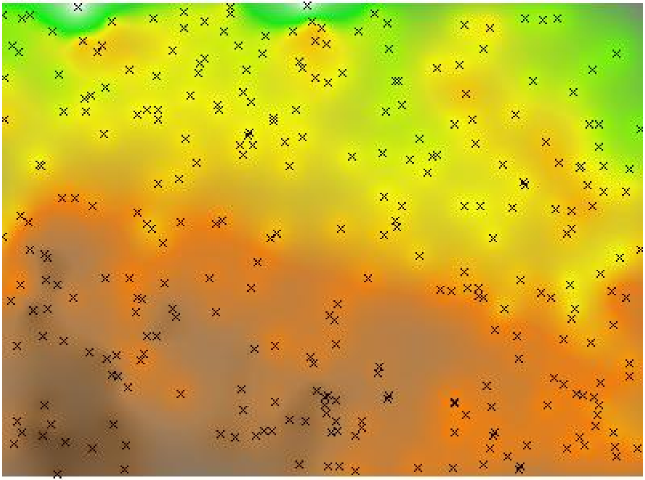
- Point pattern data
- Areal data
- Geostatisticla data
- Spatial interaction or network data

Types of Spatial Data: Geostatistical Data

Geostatistical data

- Attributes vary continuously in space, e.g., temperature, rainfall, elevation
- Measurements of nominal scale (e.g., soil types), or interval/ratio scale (e.g., depth of boreholes)
- Sampling only at fixed set of locations
- Occurs often in physical-related sciences

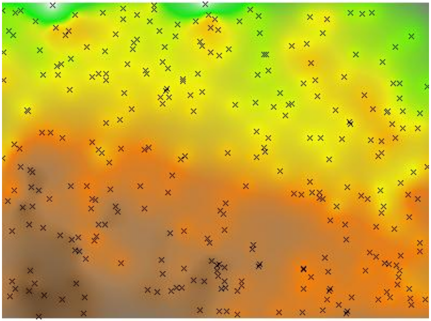
Example: 300 randomly placed points



Objective

- Mapping spatial variations of regional variables
- Make estimation at unsampled locations

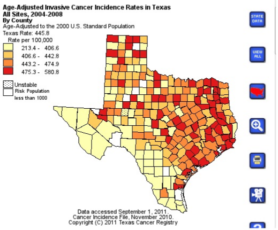
Example: *elevation surface generated from 300 points*

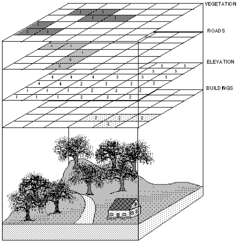


Areal (lattice) data

- attributes take values only at fixed set of areas or zones, e.g., administrative districts, pixels of satellite images
- Attributes distribute homogeneously within a region
- Lattice or uniform raster data could be taken as a special case of this type of data

Example:

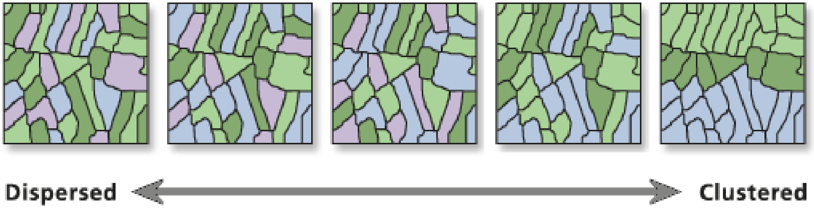




Objective

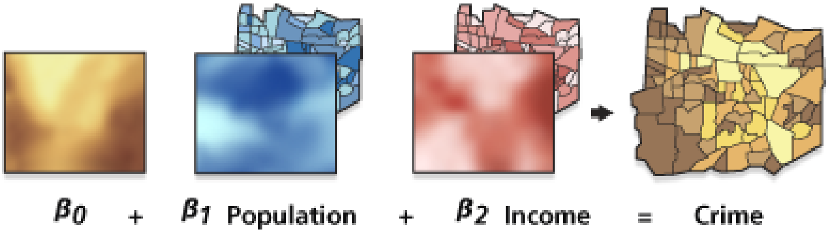
- Detect and model spatial patterns or trends in areal values
- Use covariates or relationships with adjacent areal values for inference (e.g., disease rates in light of socioeconomic variables)

Example:



Example 2: *find the correlation among maps*

- It is analog to the cases in traditional statistics, but each variable is (multidimensional) ‘maps’ instead of single 'numbers'



Types of Spatial Data: Point Pattern Data

Point pattern data

- series of point locations with recorded events, e.g., locations of trees, epic centers, disease or crime incidents
- attribute values also possible at same locations, e.g., tree diameter, magnitude of earthquakes (marked point pattern)

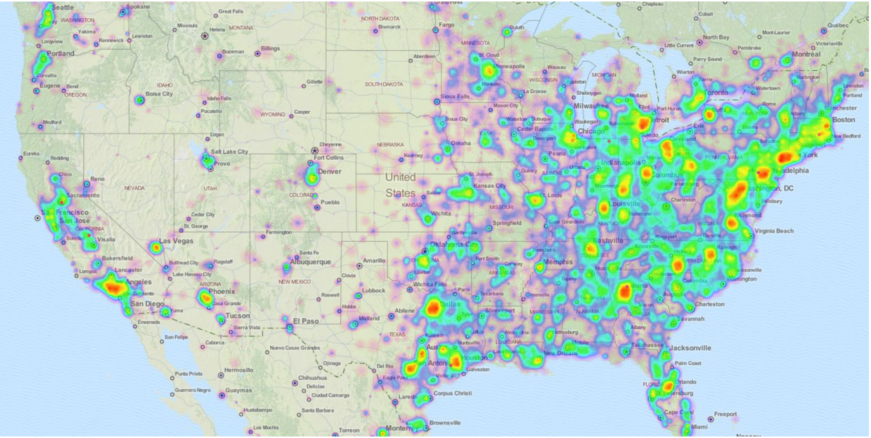
Example

Objective

- detect clustering or regularity, as opposed to complete randomness, of event locations (in space and time)
- If abnormal clustering detected, investigate possible relations with potential factors, e.g., density of disease occurrences with socio-economic status
- Difference with geostatistical point data

Types of Spatial Data: Point Pattern Data

Example:



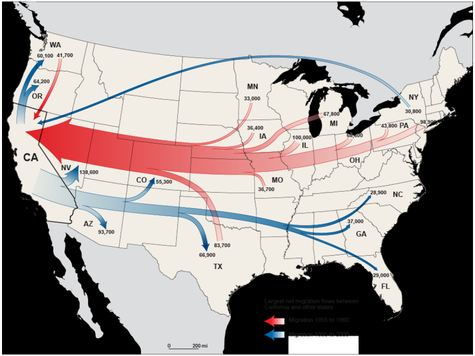
Spatial interaction or network data

- Topological space (not Euclidean space)
- Attributes relate to pairs of points or areas: flows from origins to destinations, e.g., population migrating from CA to TX
- Mostly interested in spatial patterns of aggregate interaction, rather than individuals themselves
- Not a major topic of this class

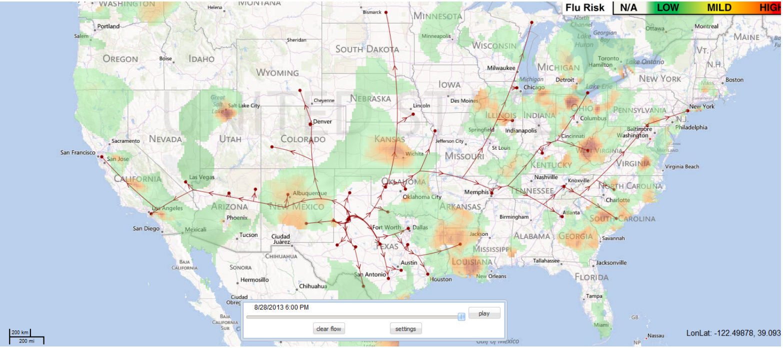
Objective

- Modeling of flow patterns
- Mostly interested in spatial patterns of aggregated interaction, rather than individual behaviors

Example



Example


Flu Risk
N/A
LOW
MILD
8/28/2013 6:00 PM
play
clear flow
settings
LonLat: -122.49878, 39.093
200 km
200 mi
Washington
Montana
Minnesota
South Dakota
Wisconsin
Oregon
Idaho
Wyoming
Michigan
New York
Maine
Iowa
Nebraska
Nevada
Utah
Colorado
Kansas
Illinois
Indiana
Ohio
Pennsylvania
Missouri
Kentucky
Virginia
California
Arizona
New Mexico
Oklahoma
Tennessee
North Carolina
Arkansas
South Carolina
Texas
Alabama
Georgia
Mississippi
Louisiana
Florida
Portland
Salem
Boise
Helena
Bismarck
Sheridan
Minneapolis
Milwaukee
Chicago
Detroit
Toronto
Hamilton
Ottawa
Montreal
Boston
New York
Philadelphia
Baltimore
Washington
Salt Lake City
Cheyenne
Denver
Lincoln
Des Moines
Springfield
Indianapolis
Columbus
Cincinnati
Jefferson City
St Louis
Sacramento
San Francisco
San Jose
Wichita
Las Vegas
Los Angeles
Albuquerque
Oklahoma City
Memphis
Nashville
Knoxville
Charlotte
Virginia Beach
San Diego
Mexicali
Phoenix
Tucson
Fort Worth
Dallas
Atlanta
Birmingham
Jackson
Savannah
Austin
San Antonio
Houston
New Orleans
Tallahassee
Jacksonville
Tampa
Miami
Hermosillo
Chihuahua
Sonora

Types of Spatial Data: Summary

Summary

- Geostatistical data
- Spatial point pattern
- Areal (lattice) data
- Spatial interaction/network data

Readings of this week

- Tobler, W.R. (1970) A computer movie simulating urban growth in the Detroit re gion. Economic Geography 46: 234-240
- Sui, D.Z. (2004) Toblers First Law of Geography: A big id ea for a small world? Annals of the Association of American Geographers 94(2): 269 - 277