



Center for Spatially Integrated Social Science

Traditional Spatial Analysis Some Examples

Luc Anselin

Spatial Analysis Laboratory

Dept. Agricultural and Consumer Economics

University of Illinois, Urbana-Champaign

<http://sal.agecon.uiuc.edu>

Outline

- Address Matching
- Spatial Interpolation

Address Matching

Geocoding

➤ Problem

- addresses available in descriptive form
 - 101 High Street, Urbana IL 61801
- no geography
 - requires polygon boundaries or point coordinates

➤ Solution

- match address to data base of street segments
 - line segments, from-to
 - U.S. **Census Tiger files**: address range and topology
- assign location lat-lon by interpolation
 - if no exact match, interpolate on line segment

Address Matching

- locate addresses in geographical space
 - data base with address list
 - coverage with street network, zip codes,...
- automatic matching
 - standardize addresses
 - search spatial data base
 - typically less than perfect match
 - manual tuning

StreetMap: Geocode Addresses

Reference Theme: StreetMap USA

Address Table: pithom.dbf

Address Fields:

- Address: Address
- City: City
- State: State
- ZIP: <None>

Optional:

- Display Field: <None>
- Offset Distance: 0 ft

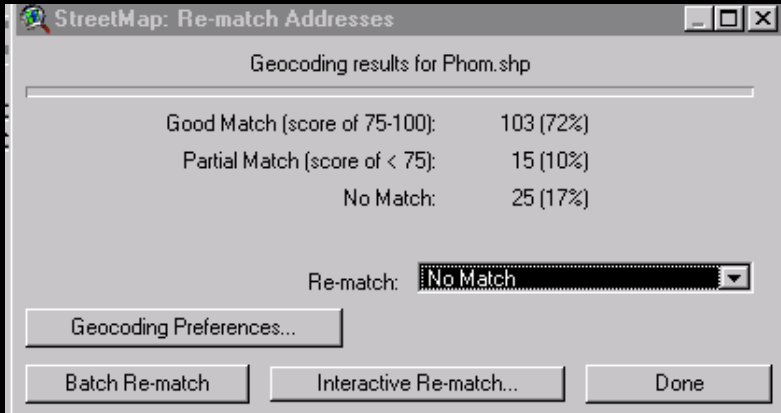
Alias Table: <None>

Geocoded Theme: c:\temp\geocd1.shp

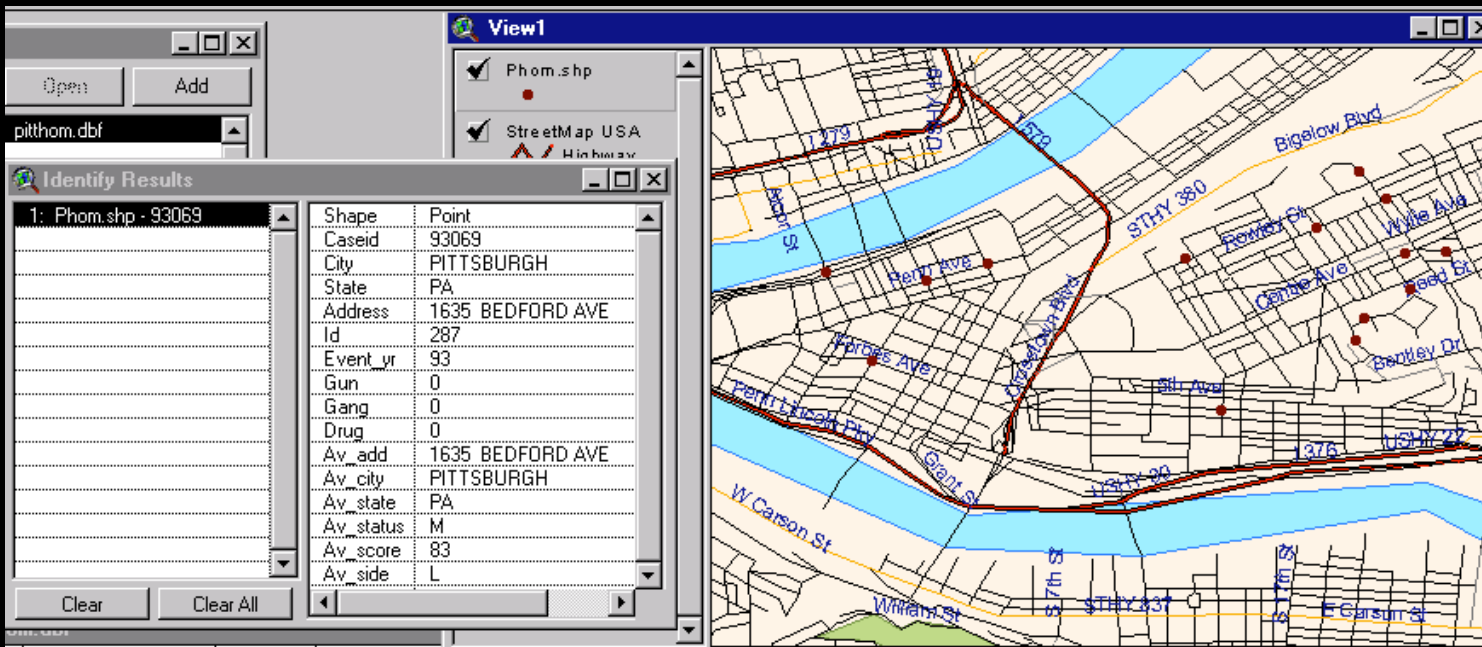
Buttons: Preferences..., Batch Match, Interactive Match..., Cancel

Caseid	City	State	Address	
93002	PITTSBURGH	PA	7904 INGLENOOK PLACE	790
93003	PITTSBURGH	PA	1203 N Lang Avenue	120
93004	PITTSBURGH	PA	5100 Columbo Street	520
93005	PITTSBURGH	PA	17 Pride Street	17
93006	PITTSBURGH	PA	34 Elmore Street	334
93007	PITTSBURGH	PA	MCCANDLESS ST	525
93008	PITTSBURGH	PA	900 PENN AVE	900
93009	PITTSBURGH	PA	5338 Mossfield Street	533
93010	PITTSBURGH	PA	2809 Bedford Avenue	280
93011	PITTSBURGH	PA	371 Penfort Street	371
93012	PITTSBURGH	PA	3245 Arlington Avenue	324
93013	PITTSBURGH	PA	1210 California Avenue	121

geocoding example using ESRI Streetmaps street files (lines) and data base of homicide address locations (Pittsburgh)



less than perfect match

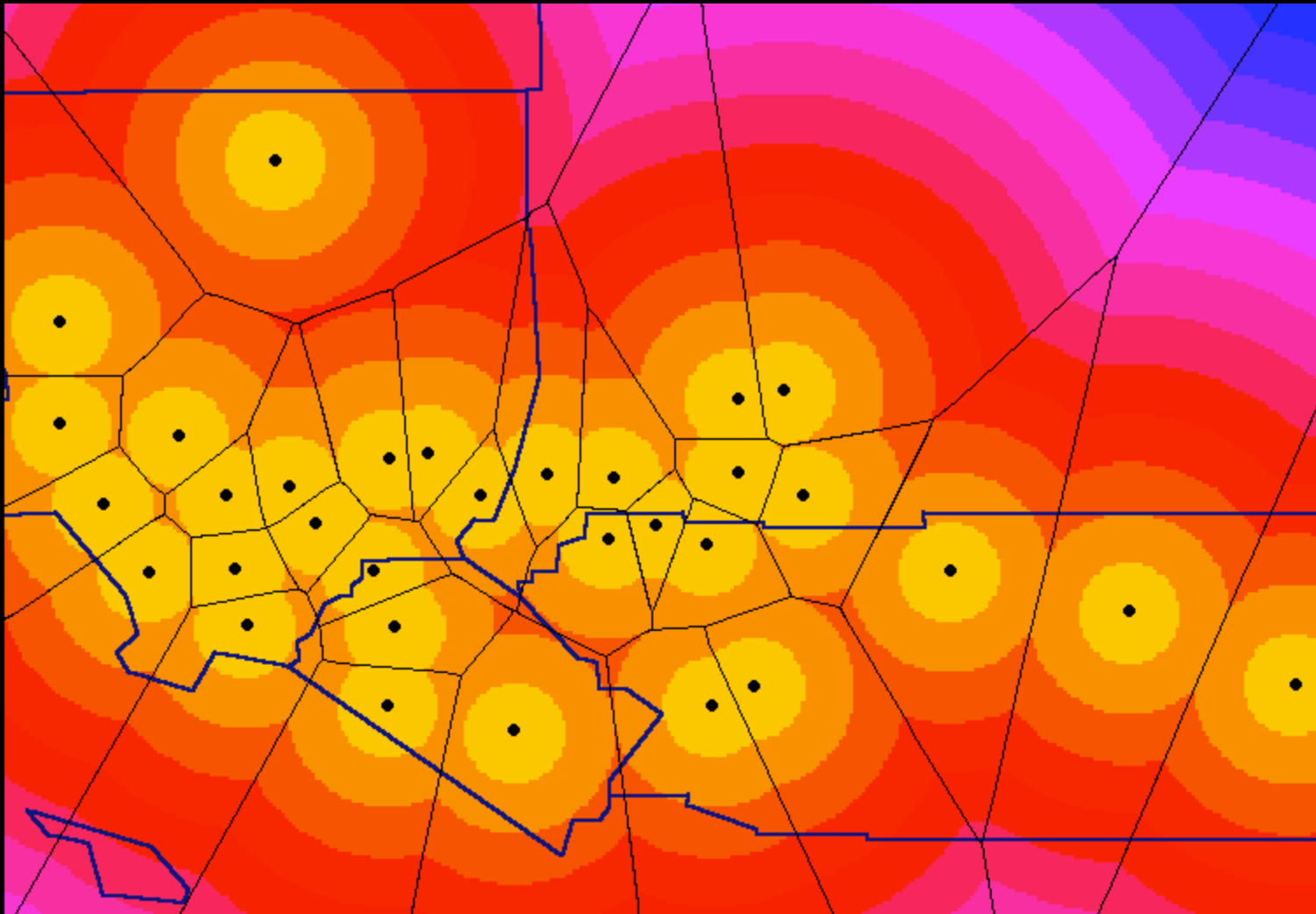


points (addresses) placed on street map

Spatial Interpolation

Spatial Interpolation

- Data Integration
 - merging spatially incompatible data sets
 - household survey data with census tract data
- Predicting Values for Locations with no Observations
 - use information on (spatial) pattern at one set of locations/scale
 - to predict values at different locations/scale
- Interpolation and Extrapolation
 - extrapolation: outside range observed values



Locations of air quality monitors, Thiessen polygons and IDW interpolation from monitor, LA Basin

Change of Support Problem

➤ COSP

- statistical approach

➤ Support

- in statistical terminology = range of values a random variable can take
 - Example: Normal r.v. $-\infty$ to $+\infty$
- spatial support = size, volume of variable and size, shape, geometry, spatial orientation of measurement "region"
 - Example: point, area

Interpolating Spatial Objects

➤ Point to Point

- from observed points to unobserved points
 - e.g. monitoring stations to house locations

➤ Point to Area

- from observed points to areal units
 - e.g. from monitoring stations to census tracts

➤ Area to Area

- from one set of observed areas to a different (overlapping) set of areas = areal interpolation
 - e.g. from counties to water management zones

Types of Interpolation

➤ By Data Structure

- fields
 - interpolating **surfaces**
- objects
 - **not interpolation**, but creating a field **representation** from discrete objects
 - density estimation = how many **objects/area**

➤ By Variable Type

- **spatially extensive**
 - **counts**, X ; example: tract population size
- **spatially intensive**
 - **per areal unit**: X/A ; example: tract population density

Interpolation Methods

- By Scope
 - **global**: use all the data points
 - **local**: use limited set (local) of data points
- By Fit
 - **exact**: observed data points predicted exactly
 - **approximate**: even observed data points predicted with error
- By Model
 - **deterministic**: mathematical model
 - **stochastic**: probabilistic model (random variable)

Point to Point Interpolation

➤ Problem:

- Data Observed at Points, Prediction for Other Points
 - e.g. assessed house value

➤ Techniques:

- Inverse Distance Weighting
- Thiessen Polygons
- Trend Surface
- Regression Models
- Splines
- Kriging

Inverse Distance Weighting

➤ Tobler's Law

- in space, everything is related to everything else, but closer locations more so
- use distance decay

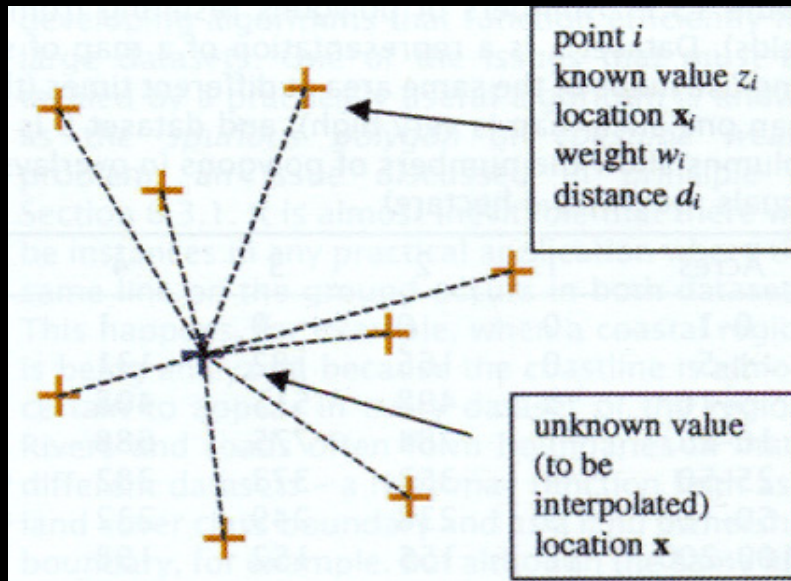
➤ IDW

- $z_x = \sum_i \lambda_i z_i$ with $\sum_i \lambda_i = 1$ or $z_x = \sum_i w_i z_i / \sum_i w_i$
- $w_i = \text{weight}$
 - inverse distance function: $1/d_{ix}^2$

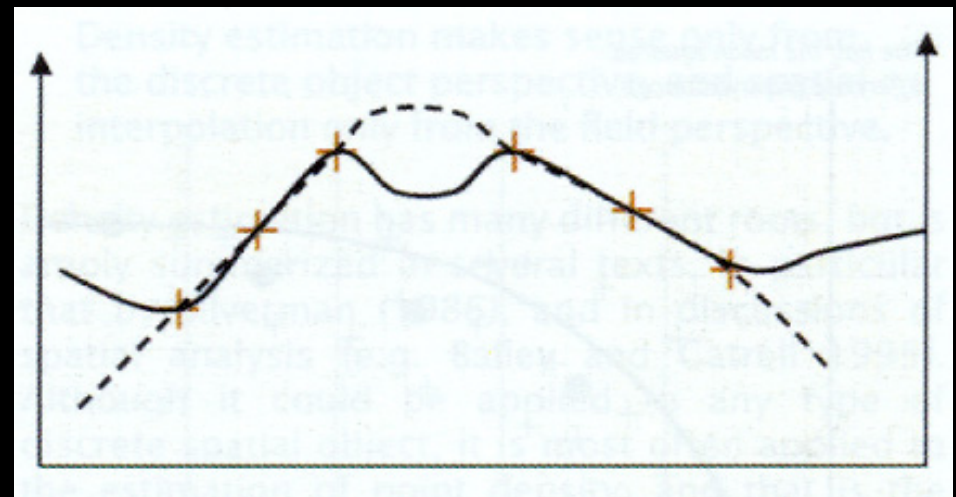
➤ Properties

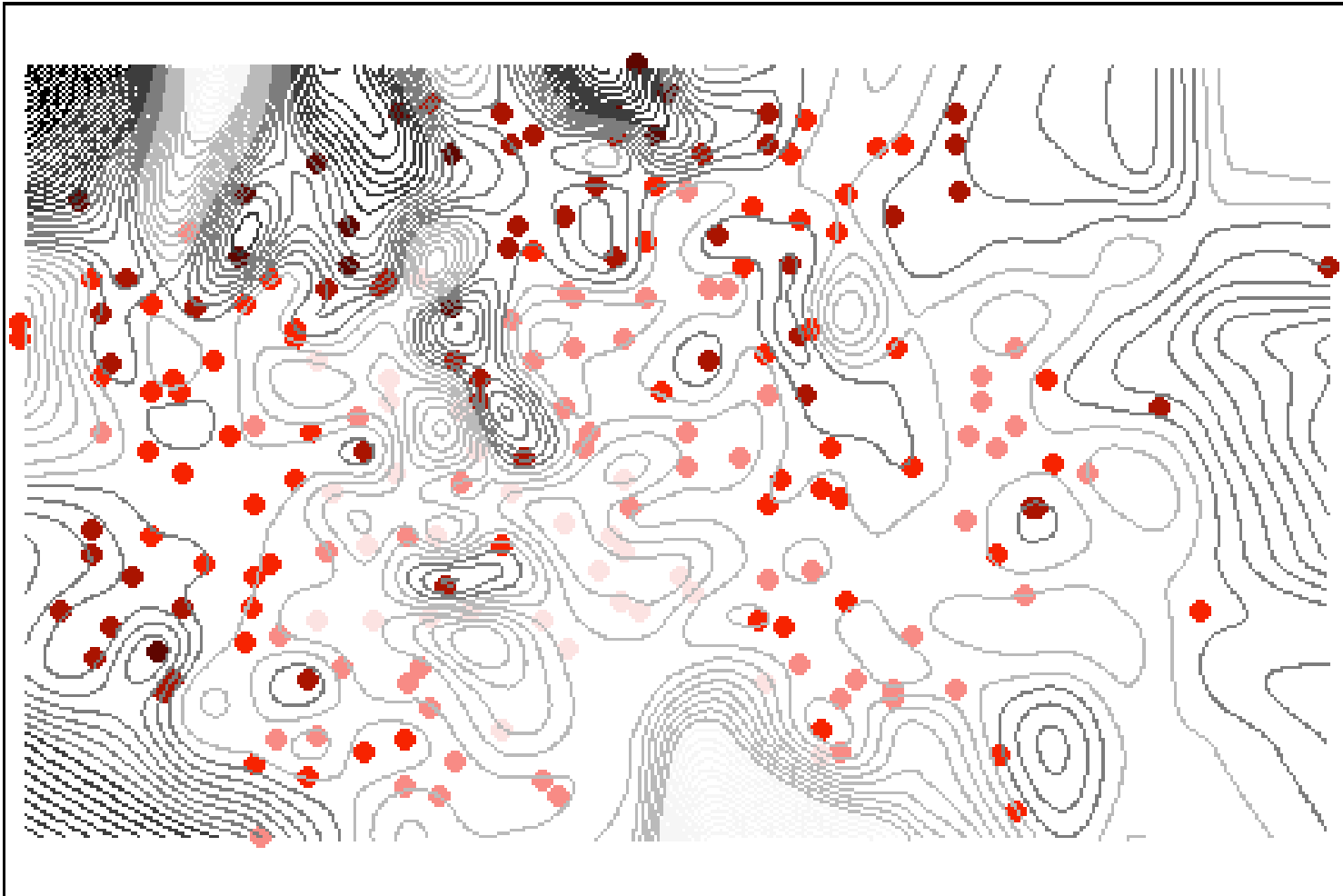
- local, exact, deterministic
- interpolated values are always within range

inverse distance weighting



potential problem:
no out of range predictions





Residential Sales Price, Baltimore MD (1980)
sample points (darker is higher) and IDW contours

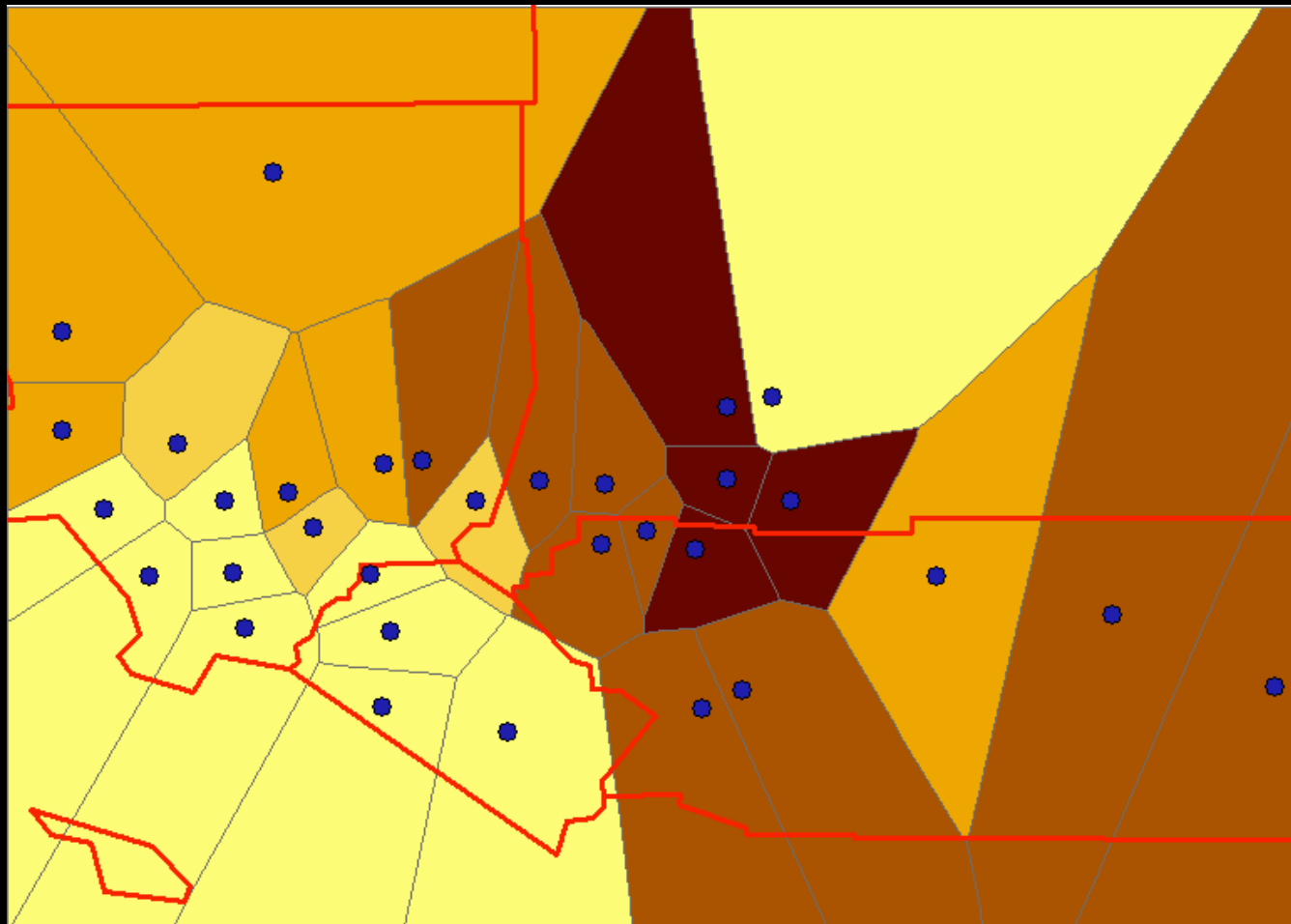
Tessellations

➤ Thiessen Polygons

- collection of all points in a plane that are closer to the central point than to any other point
- notion of minimum transport cost market area
- all points in the polygon are assigned the **same value as the central point**

➤ Properties

- local, exact, deterministic
- interpolated values underestimate heterogeneity



Thiessen Polygon Interpolation - LA Basin Ozone

