GIS Data Analysis

Methods for Generating Secondary Information

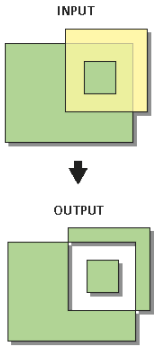
- Buffering – Generating AOI
- Overlay – Dasymetric Mapping (areal interpolation)
- Distance Measurement – Determining AOI
  - Thiessen polygons
- Pattern Analysis – Determining AOI
  - Nearest neighbor analysis
- Map Manipulation
  - Dissolve, clip, append, select, eliminate, update, erase
- Spatial interpolation
- Raster calculation

Buffering

- Proximity, edge effect, spatial interaction
- Point, line, polygon features
- Variations of buffering
- Irregular buffering (e.g., stream reaches)
- Applications
  - Protection zone
  - Neutral zone
  - Inclusion zone
  - Sampling scheme

Map overlay

- Overlay feature types
  - Point-in-polygon
  - Line-in-polygon
  - Polygon-on-polygon
- Overlay methods
  - Union
  - Intersect
  - Symmetrical Difference
  - Identity



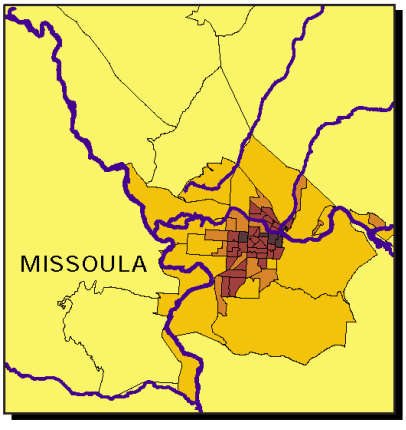
Overlay Procedures

- Determine the spatial reference for processing. All the input feature classes are projected (on the fly) into this spatial reference.
- Crack and **cluster** the features.
- Discover geometric relationships (overlap) between the input features and the overlap features.
- Assign attributes based on the type of overlay.
- Remove features based on the combinations of attributes and overlay types.

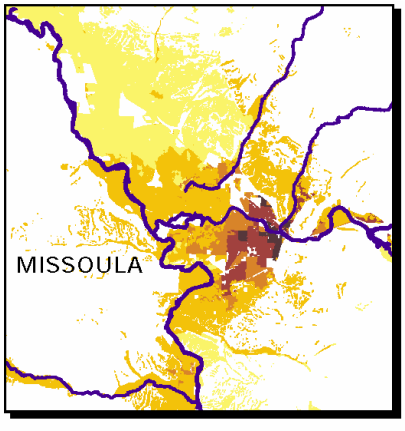
Map overlay (cont.)

- Considerations
  - Georeferencing, registration
  - Slivers (fuzzy tolerance, cluster tolerance, minimum mapping unit
  - Error propagation (The expected accuracy of a composite map cannot be better than the least accurate individual input map.)
- Applications
  - Site analysis
  - Areal interpolation (dasymetric mapping)

Choropleth vs Dasymetric Maps

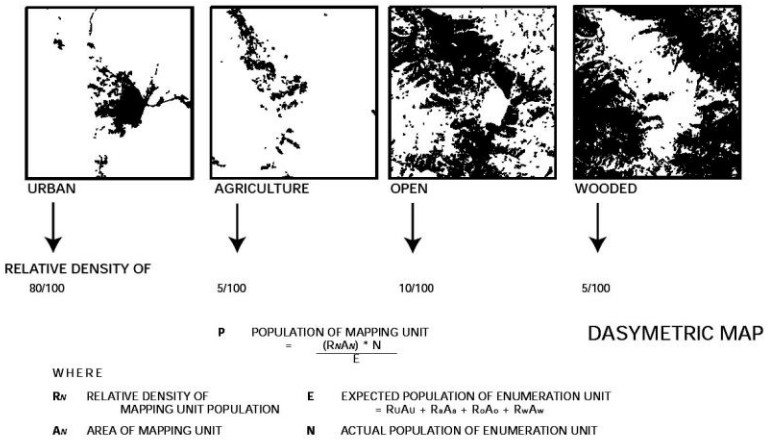


A

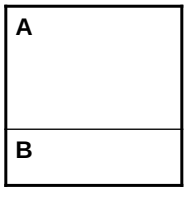


B

Dasymetric mapping



Dasymetric Mapping with Ancillary Info

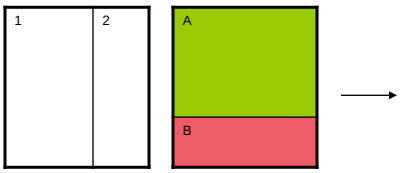


Example:
Total population N: 100
Total area of enumeration unit: 10 (unit^2)
Area of A A_A : 7 (unit^2)
Area of B A_B : 3 (unit^2)
B's population is twice as dense as A's (i.e., relative density A : B = 1 : 2)

Actual density of A and B: 1 x K, 2 x K (K is a constant)
Actual population of A and B: 1K x 7, 2K x 3
Total population 100 = 1K x 7 + 2K x 3
K = 100 / 13
Population of A = 1K x 7 = 100 / 13 x 7 = 53.85
Population of B = 2K x 3 = 2 x 100 /13 x 3 = 46.15

$K = N / (R_A x A_A + R_B x A_B)$
$P_A = R_A x A_A x K$, $P_B = R_B x A_B x K$
$P_{mu} = (R_A x A_A) x N / E$
$E = R_A x A_A + R_B x A_B$

Dasymetric Mapping (cont.)



$$P_{mu} = (R_A \times A_A) \times N / E$$

E_ID	Area	N
1	60	100
2	40	25

L_ID	Area	RD
A	70	0.1
B	30	0.9

O_ID	Area	E_ID	L_ID	N	RD	A*RD	E	P=A*RD*N/E	P
a	42	1	A	100	0.1	4.2	20.4	4.2 x 100 / 20.4	20.59
b	28	2	A	25	0.1	2.8	13.6	2.8 x 25 / 13.6	5.15
c	18	1	B	100	0.9	16.2	20.4	16.2 x 100 / 20.4	79.41
d	12	2	B	25	0.9	10.8	13.6	10.8 x 25 / 13.6	19.85

E_ID	E=Sum(A*RD)
1	20.4
2	13.6

See http://web.pdx.edu/~jduh/courses/geog492s10/lab04.htm for exercise.

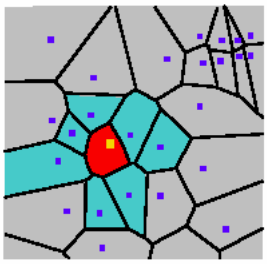
Distance measurement

- For features (i.e., points, lines) that cannot perform overlay analysis
- A method to link attributes between features that are not colocated
- For example, the NEAR and POINT-DISTANCE tools in Arctoolbox

Voronoi Map (Thiessen Polygons)

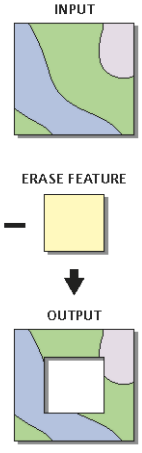
- Why use a Voronoi map?

- Assigning values to polygons
  - Simple
  - Mean
  - Cluster
  - Standard deviation
  - …



Map manipulation

- Dissolve
- Eliminate
- Append
- Erase/Clip
- Merge/Split



Raster Data Analysis

- Local (cell-by-cell) operations – multi-criteria analysis
- Neighborhood (moving-window) operations – “noise” reduction, raster map generalization, terrain analysis…
- Zonal operations – zonal statistics
- Raster distance measure operations - similar to buffering

Local Operations

- Cell-by-cell based
- Creates a new raster from either a single or multiple input rasters
- Includes Reclassification and Map Algebra

Neighborhood Operations

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

1.56	2.00	2.22
1.67	2.11	2.44
1.67	2.11	2.44

Neighborhood Means

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

2	2	2
2	2	3
2	2	3

Neighborhood Majority

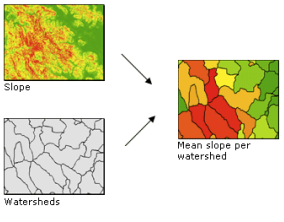
Zonal Operations

- Uses groups of cells that have the same value or like features
- Can be contiguous or noncontiguous
- For single rasters zonal operations measure the geometry of each zone (area, perimeter, thickness, centroid)
- For two rasters (an input raster and a zonal raster) a summary of values for the input values in each zone of the zonal raster is generated in an output raster (summary statistics and measures)

1	2	2	1
1	4	5	1
2	3	7	6
1	3	4	4

1	1	2	2
1	1	2	2
1	1	3	3
3	3	3	3

2.17	2.17	2.25	2.25
2.17	2.17	2.25	2.25
2.17	2.17	4.17	4.17
4.17	4.17	4.17	4.17



Raster to Vector / Vector to Raster