Geol 588

-

GIS for Geoscientists II

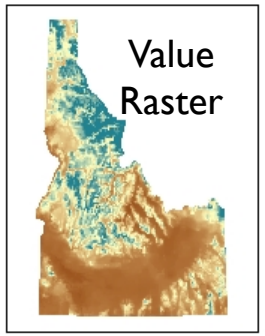
Feb 22, 2011

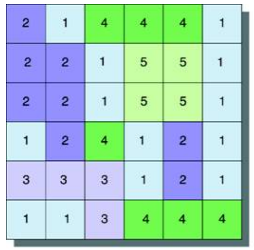
Zonal statistics

Interpolation

Zonal functions

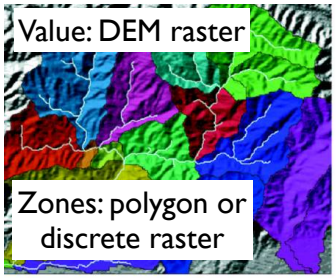
- Needs 2 inputs:
- a **zone data:**
  - discrete raster or
  - features (polygon, line points)
    - which attribute?
- a **value raster**
- Zones don't need to be continuous!



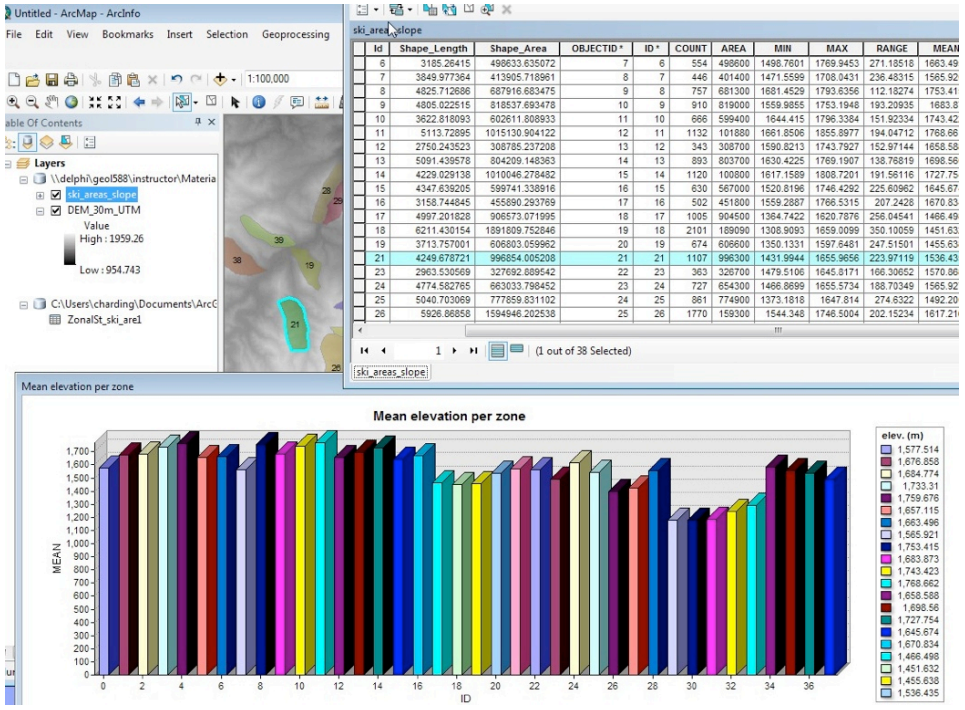


Zonal statistics
Sp. Analyst Tools - Zonal

- Choose correct attribute for zones (usually: must be unique ID for each zone)
- Choose which stat(s)
- for each zone:
  - get value raster cells that are inside the zone
  - calculate **stat(s)** of these cells
  - store stat(s) for this zone in a table
- Tool: Zonal Statistics **as Table**,
  - creates a standalone table that can be **joined** to zone layer
  - (ArcMap TOC: list by source to see standalone tables!)
- Tool: Zonal Statistics:
  - stores 1 type of stat in a new raster
  - new raster will have Nodata outside of the zones
  - inside a zone: all cells have the same value (e.g. the mean of the zone)



- ALL —All of the statistics will be calculated. This is the default.
- MEAN— Calculates the average of all cells in the value raster that belong to the same zone as the output cell.
- MAJORITY — Determines the value that occurs most often of all cells in the value raster that belong to the same zone as the output cell.
- MAXIMUM — Determines the largest value of all cells in the value raster that belong to the same zone as the output cell.
- MEDIAN — Determines the median value of all cells in the value raster that belong to the same zone as the output cell.
- MINIMUM — Determines the smallest value of all cells in the value raster that belong to the same zone as the output cell.
- MINORITY — Determines the value that occurs least often of all cells in the value raster that belong to the same zone as the output cell.
- RANGE — Calculates the difference between the largest and smallest value of all cells in the value raster that belong to the same zone as the output cell.
- STD — Calculates the standard deviation of all cells in the value raster that belong to the same zone as the output cell.
- SUM — Calculates the total value of all cells in the value raster that belong to the same zone as the output cell.
- VARIETY — Calculates the number of unique values for all cells in the value raster that belong to the same zone as the output cell.
- MIN_MAX—Both the Minimum and Maximum statistics are calculated.
- MEAN_STD—Both the Mean and STD statistics are calculated.
- MIN_MAX_MEAN—The Minimum, Maximum and Mean statistics are calculated.

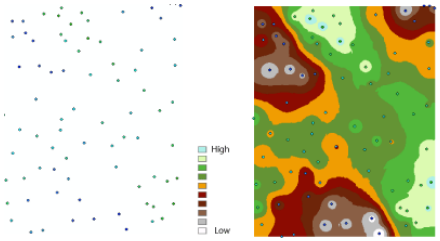


Spatial Interpolation

data\Interpolation Exercise

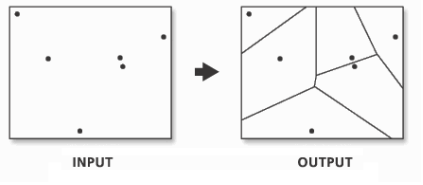
- point samples (x,y,"value")
- fill each cell (center) in raster with an "appropriate" value
- Principle: the *closer* together points are, the *more similar* their value (should be)
- (depends on many factors: type of underlying phenomenon, etc.)
- Examples?
- in general: distance **and** sample value matter

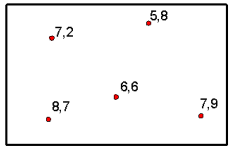
13	14	16	20	23
14	14	16	19	24
18	16	16	18	22
24	22	19	19	21
30	27	23	20	20

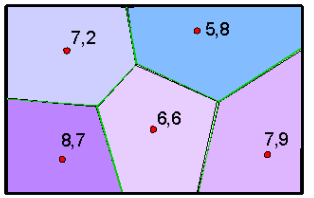


Thiessen (Voronoi) polygons

- Space division scheme
- Assign each sample "its fair share" of space around it
- Raster: fill this space (polygon) with **same** value
- Keep in mind for Natural Neighbor interpolation method

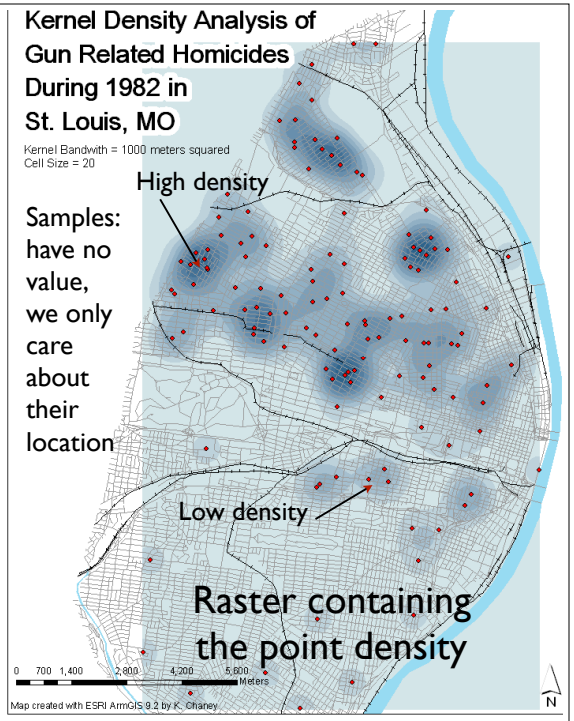


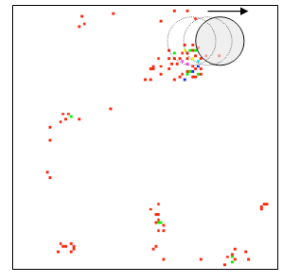


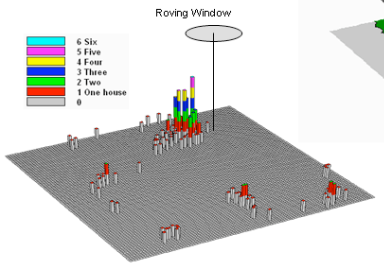


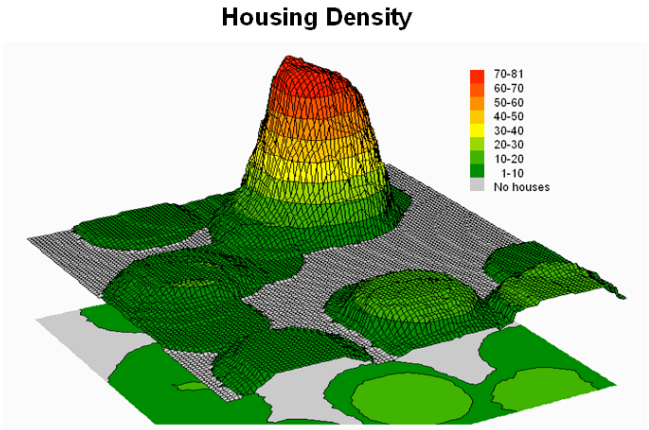
Raster of point density

- NOT a **value** interpolation scheme!
- decide on a circle (kernel) size (radius,"bandwidth")
- go through all cells ....
- count how many samples fall within a that circle
- divide the count by the circle's area
- assign this density value to the cell
- next cell
- Tools:
  Spatial Analyst - Density
  - Point density
  - Kernel Density





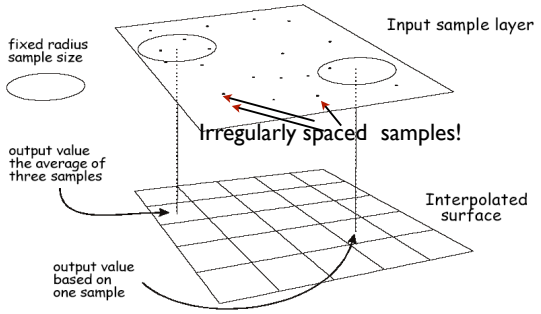


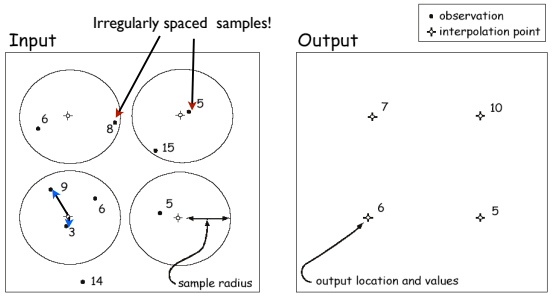


... the Point Density technique passes a "roving window" over a project area calculating the total number of houses within a specified distance

Interpolating values

- super simple: moving averages
- decide on a radius
- go through all cell centers
- grab all the point samples within the search radius
- Sum these points' values and divid by their number
- How does size of radius affect results?
- Does sample to center distance play any role
- How could we deal more fairly with farther away samples (9) compared to closer samples (3)?



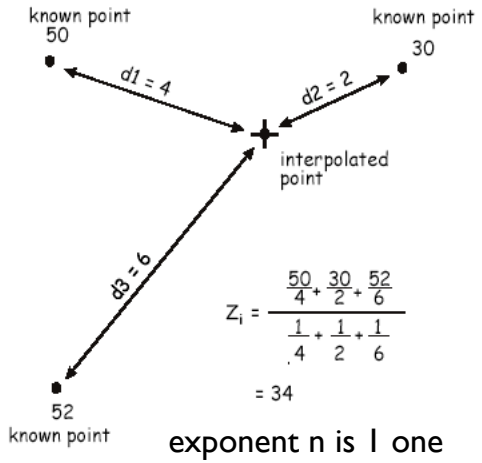


Inverse Distance Weighted (IDW) interpolation

Zi is value of known point
Dij is distance to known point
Zj is the cell value (unknown)
n is a user selected exponent or power (often 1,2 or 3)

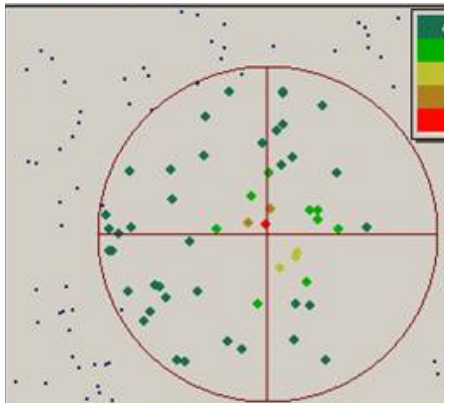
$$Z_j = \frac{\sum_i \frac{Z_i}{d^n_{ij}}}{\sum_i \frac{1}{d^n_{ij}}}$$

- grab all the point samples inside a radius
- When calculation the cell's value, take the sample distances into account ("weight")
- Weight of each sample point is an inverse proportion to its distance to the cell
- The further away the point, the less the weight it gets (contributes less)
- effect of exponent (1,2,3, ..)?

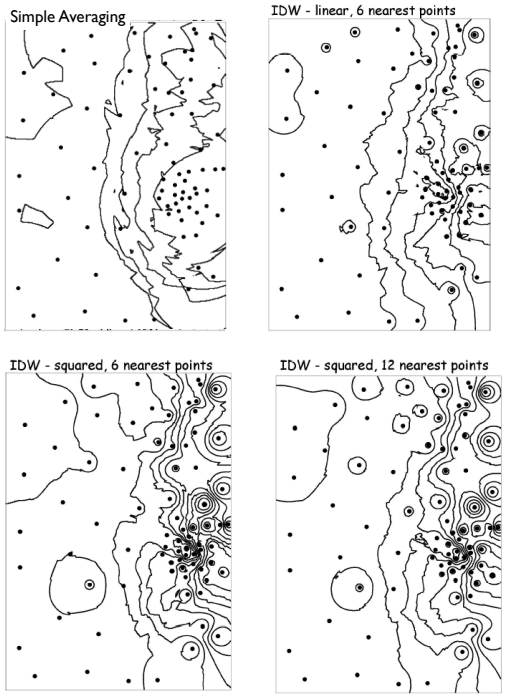


exponent n is 1 one

- Effect:
- samples that are closer to the cell (center) to be estimated, get a larger weight (are more important)
- The larger the sample to center distance, the less important the sample is



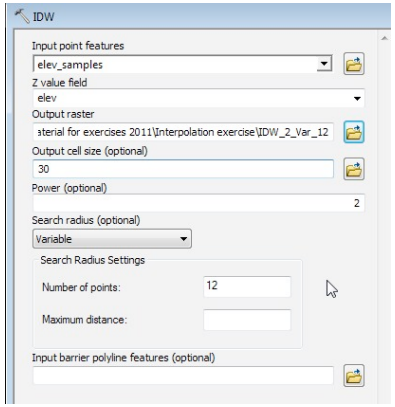
- (We could use all the samples for each cell's estimation
- Using a search radius is more efficient (why?)



- Higher exponents: less weight to distant points (point that are farther away)
- (closer point are more important)
- Use more samples: "smoother" distribution
- interpolated value at sample location?

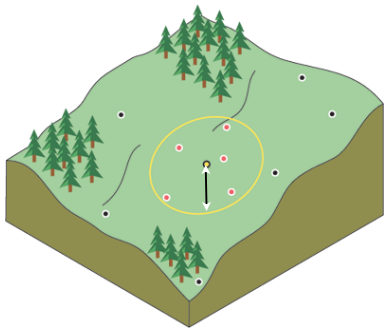
"linear": exponent = 1
"squared": exponent = 2

Tool: Spatial Analyst Interpolation - IDW

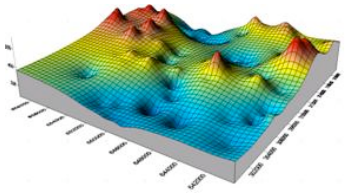


- point features (here: elevation samples)
- set correct attribute (here: elev)
- encode the parameters in your output raster's file name:
- IDW => interpolation method
- 2 => power
- Var_12 => Search parameters
- bar => used barriers

IDW: search parameters

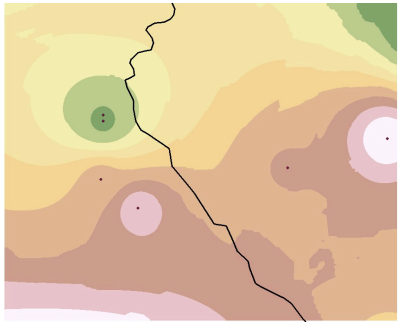


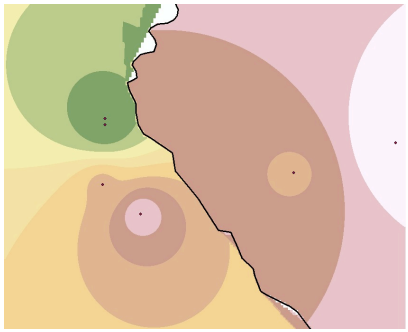
- **size** of search radius: multiples of cell size
- **variable** search radius:
  grab n (12) closest points, up to a distance of d (default 0, => no distance limit)
- **fixed** search radius:
  grab ALL points within a distance of (250) units, use at least n points (default 0, => use all)



IDW: using break lines

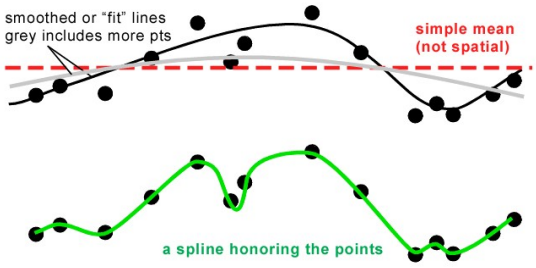
- line features to limit (block) point "grabbing"
- samples "from the other side" of the break line will NOT be used
- use for: ridges, faults, ...
- Caveats: much slower, some NoData values (extrapolation issues?)

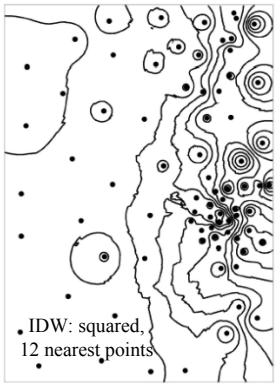


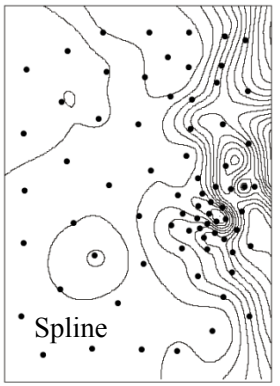


Splines

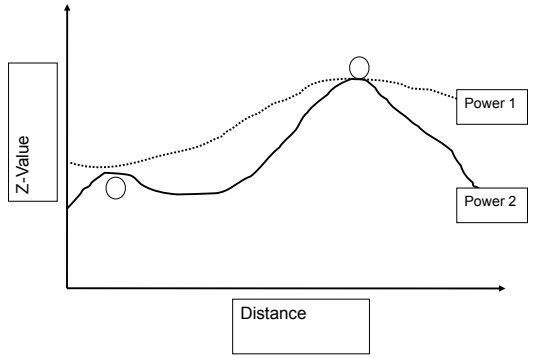
- smooth surface, non-exact
- based on minimizing curvature
- rubber sheet "bent" around samples
- can over-shoot / under-shoot
- ArgGIS: 2 types - regularized and tension
- ArcGIS:Weight - smoothness "tweak" factor
- gotcha: both have different meaning of weight
- tension (0.1 - 15):
  low = smooth, high = coarse)

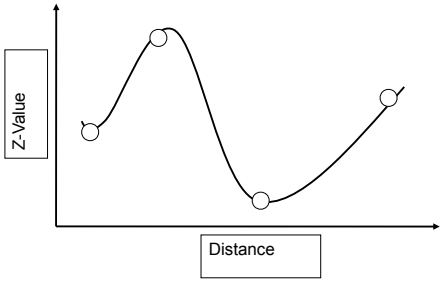






IDW:





Spline