Geography 38/42:376
GIS II

Topic 7: Surface Representation and Analysis

(Chang: Chapters 13 & 15)
DeMers: Chapter 10

What can we represent as a Surface?

- Surfaces can be used to represent:
  - Continuously distributed phenomena sampled at point locations
  - Areally discrete data represented at point locations
- Phenomena may be real or conceptual

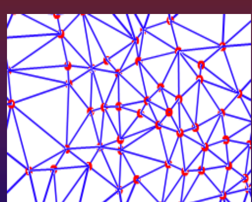
Common Surface Representations

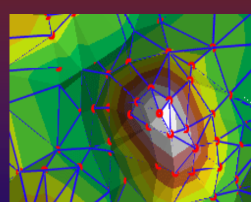
- Vector data model = TIN
  - Triangulated Irregular Network
- Raster data model = DEM
  - Digital Elevation Model

TINs - vector

- A network of interlocking triangles
- Corners “anchored” by control/sampled pts.
- Each facet has a slope and aspect

TINs

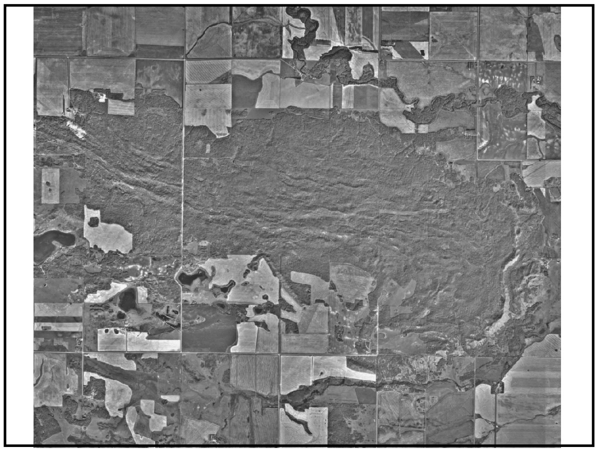


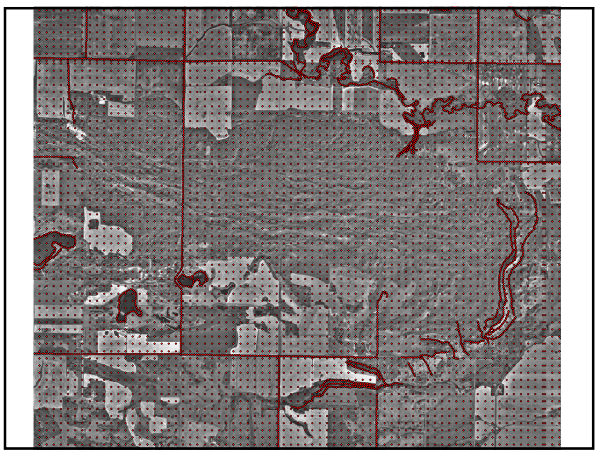


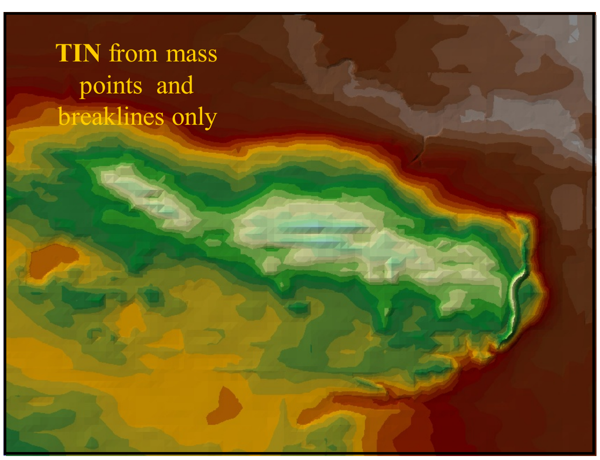
- **NOTE**: Level of detail determined by number/distribution of control pts. and resulting size and number of triangles

TINs

- Can be derived from multiple data sources
  - Point features (random or irregular)
    - Mass points
  - Line features (roads, railways, streams, ridges)
    - Mass Points (if 3D)
    - Breaklines
  - Polygons (study area, water body, drainage basin)
    - Breaklines
    - Mass Points (if 3D)
    - Erase
    - Replace
    - Clip



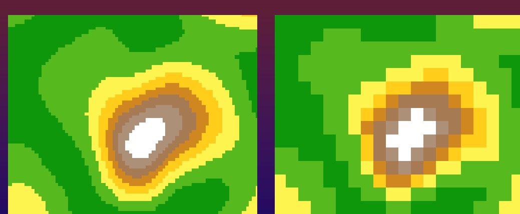



TIN from mass
points and
breaklines only

DEMs

- Array of discrete grid cells
- Cell value represents z-value (e.g. elevation amsl)
- Also improperly used to refer to a regular array of (x,y,z) pts; a lattice

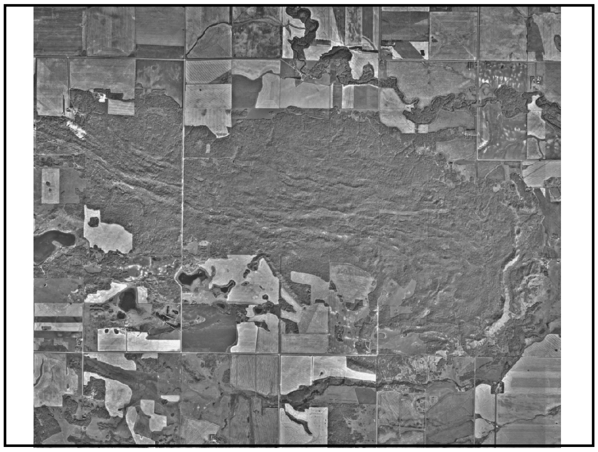
DEMs

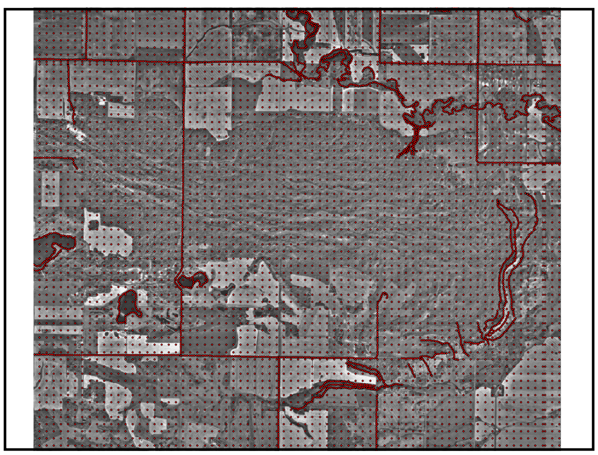


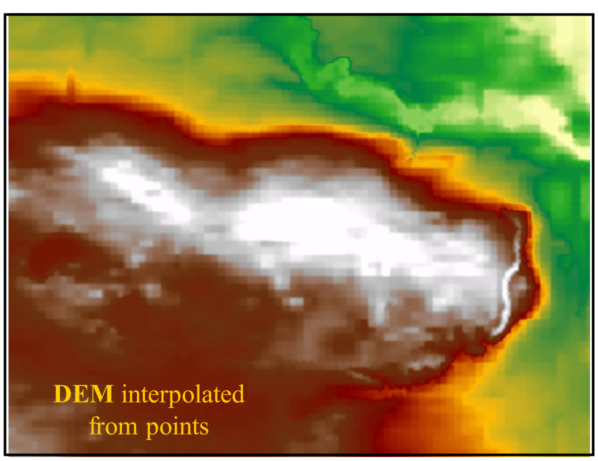
- Level of detail determined by spatial resolution of raster

DEMs

- Created through process of **spatial interpolation**
  - Definition?
- Limited data sources:
  - Point features
  - Line features
    - But only to define boundaries limiting neighbourhood






DEM interpolated
from points

TINs vs. DEMs

- TINs
  - Vector data model
  - Variety of data inputs
  - Variable level of detail
  - No data redundancy (when z-tolerance used) because triangle size varies
  - **Built not interpolated**
- DEMs
  - Raster data model
  - Limited data inputs
  - Fixed cell size/level detail
  - May contain significant data redundancy
  - **Interpolated not built**

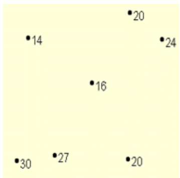
Spatial Interpolation

- Process of creating a continuous raster surface:
  - from cont. data sampled/measured at control pts.
  - or areally discrete data represented at point locations
- Each point consists of x, y and z-value
- Cell values interpolated based on control points
  - Based on first law of geography
- Variety of interpolation algorithms

Interpolation Algorithms

- User is required to:
  - choose interpolation algorithm
  - identify z-value
  - determine size/shape of neighbourhood (if applicable)
    - don't confuse with focal/neighbourhood operations
  - specify weight/power value (if applicable)
  - set output grid cell size (hmmm?)

Spatial Interpolation



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

What would be a reasonable output grid cell size?

Global vs. Local Interpolators

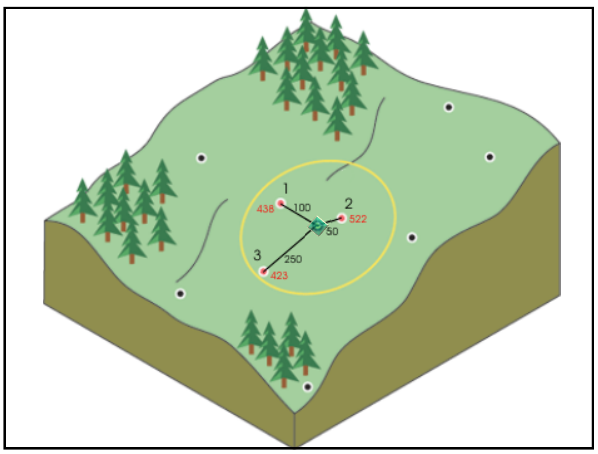
- Global interpolators consider value at all sampled locations when estimating each cell value
- Local interpolators consider sampled values within a defined neighbourhood
  - a specified number
  - a specified distance

Exact vs. Inexact Interpolators

- Exact interpolators result in surface that conforms exactly to the sampled locations
- Inexact interpolators result in surface that may not conform to sampled locations

Inverse Distance Weighting

- Interpolated value equal to mean of control points within defined neighbourhood
- Mean weighted inversely proportional to distance
- Sample points are identified:
  - By specifying an exact number of points
  - Defining a specific search radius



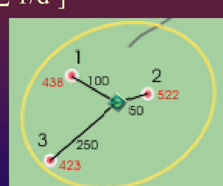
Inverse Distance Weighting

$$z_0 = [\sum z_i \ 1/d^k] / [\sum 1/d^k]$$

Where:

z_i = value at known location
d = distance
k = power of exponent

Power = 2		z_i $1/d^k$	$1/d^k$
z_1 = 438	d_1 = 100	0.0438	0.0001
z_2=522	d_2 = 50	0.2088	0.0004
z_3=423	d_3 = 250	0.006768	0.000016
	$\sum$ =	0.259368	0.000516



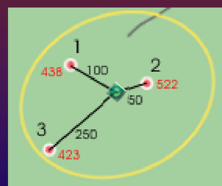
z_0 = 502.65

If Power = 6

$$z_0 = [\sum z_i \, 1/d^k] / [\sum 1/d^k]$$

Where:

- z_i = value at known location
- d = distance
- k = power of exponent



Power = 6		$z_i \, 1/d^k$	$1/d^k$
z_1 = 438	d_1 = 100	4.4×10^{-10}	1.0×10^{-12}
z_2 = 522	d_2 = 50	3.3×10^{-8}	6.4×10^{-11}
z_3 = 423	d_3 = 250	1.7×10^{-12}	4.1×10^{-15}
	$\sum$ =	3.3×10^{-8}	6.5×10^{-11}

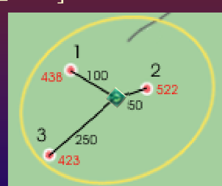
$z_0 = 520.73$

If Power = 1

$$z_0 = [\sum z_i \, 1/d^k] / [\sum 1/d^k]$$

Where:

- z_i = value at known location
- d = distance
- k = power of exponent



Power = 1		$z_i \, 1/d^k$	$1/d^k$
z_1 = 438	d_1 = 100	4.38	0.01
z_2 = 522	d_2 = 50	10.44	0.02
z_3 = 423	d_3 = 250	1.692	0.004
	$\sum$ =	16.512	0.034

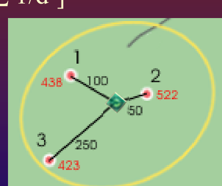
$z_0 = 485.65$

If Power = 0

$$z_0 = [\sum z_i \, 1/d^k] / [\sum 1/d^k]$$

Where:

- z_i = value at known location
- d = distance
- k = power of exponent



Power = 0		$z_i \, 1/d^k$	$1/d^k$
z_1 = 438	d_1 = 100	438	1
z_2 = 522	d_2 = 50	522	1
z_3 = 423	d_3 = 250	423	1
	$\sum$ =	1383	3

$z_0 = 461$

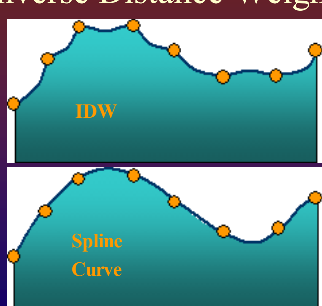
Inverse Distance Weighting

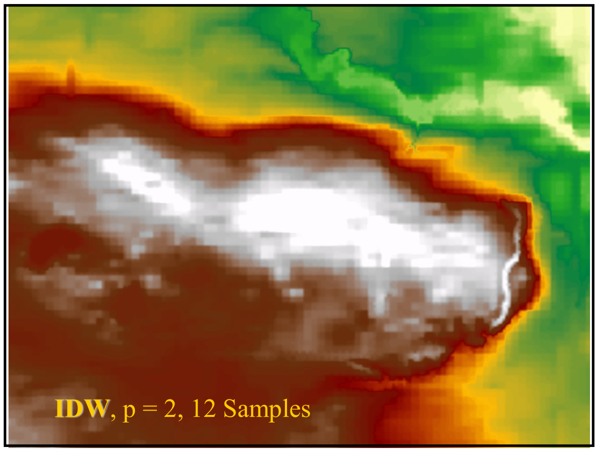
- How is an appropriate k value determined?
- As k ↑, influence of more distant points ↓
  - good if rate of change in z value is exponential
- When k = 1, linear relationship exists
  - good if rate of change is linear

Inverse Distance Weighting

- IDW is an exact interpolator
  - when d = 0 division error; defaults to sample value
- IDW is a local interpolator
  - neighbourhood is defined
- z_0 is limited to range of control points
  - results in “bulls-eye” effect at sample points

Inverse Distance Weighting

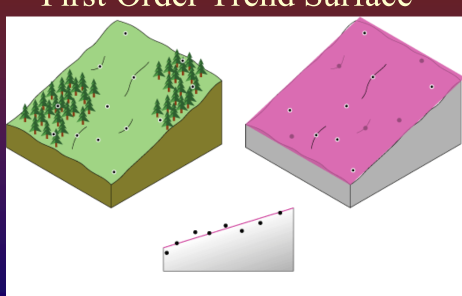




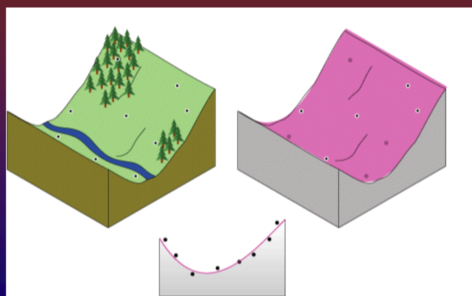
Trend Surface

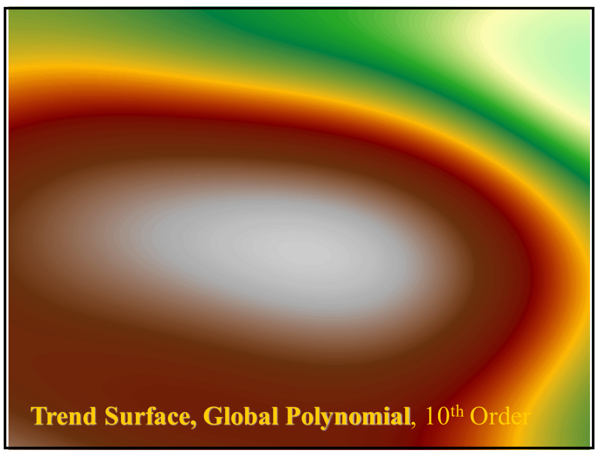
- Control points used to derive least-squares polynomial equation to represent <u>entire surface</u>
  - x, y are independent variables, z dependent variable
- Order of equation adjusted to “fit” sampled points

First Order Trend Surface

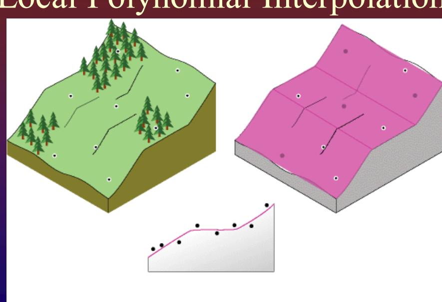


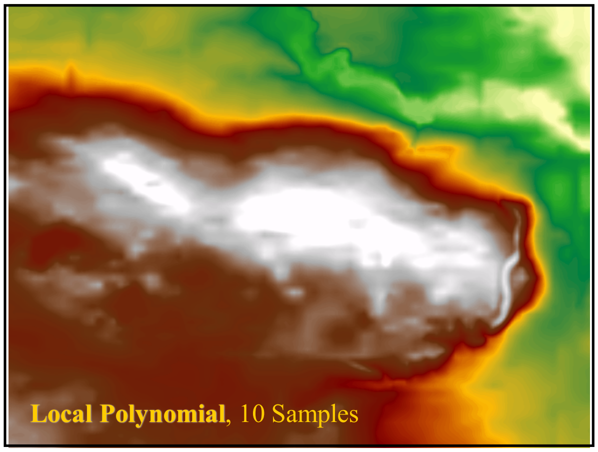
Second Order Trend Surface





Local Polynomial Interpolation

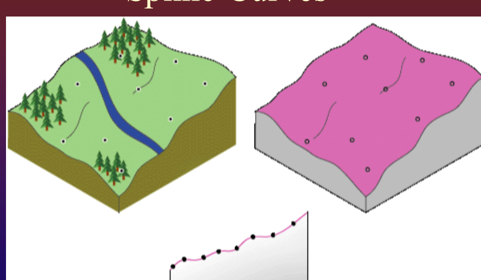


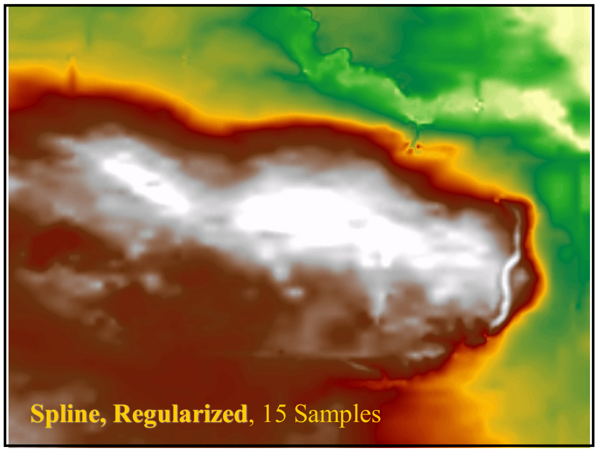


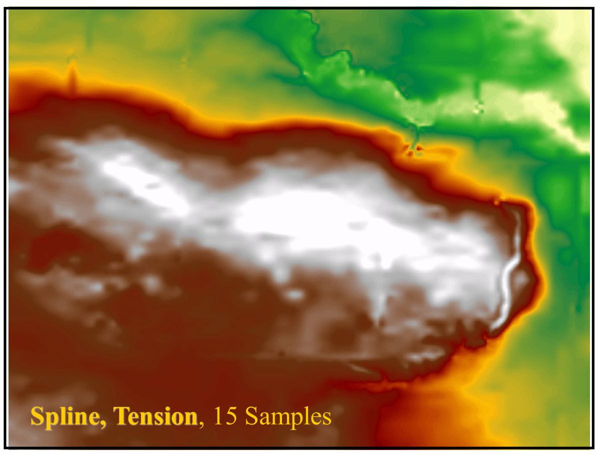
Spline Curves

- Creates a minimum curvature surface between control points
- Surface passes through control points
- Uses a specified number of neighbouring sample points
- Predicted values are not limited to range of selected sample values

Spline Curves







Kriging

- Limitations of <u>deterministic</u> methods:
  - Selection of sample points
  - Orientation/shape of neighbourhood
  - Best weight/power fcn to use
  - Identifying, measuring, or estimating errors

Kriging

- Kriging is a geostatistical method
- Recognizes naturally occurring surfaces cannot be modeled by a smooth mathematical function
- Stochastic method in that it recognizes random but spatially autocorrelated variability

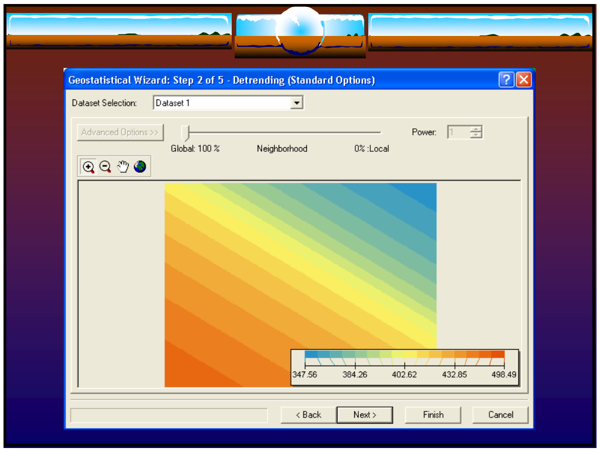
Kriging

- Regionalized variable theory states that spatial variation of any phenomena can be predicted by considering three components:
  1. Overall trend/drift modeled by a deterministic function
  2. Random but spatially autocorrelated component; regionalized variable
  3. Spatially uncorrelated, random, unpredictable noise

$$Z(x) = m(x) + e'(x) + e''(x)$$

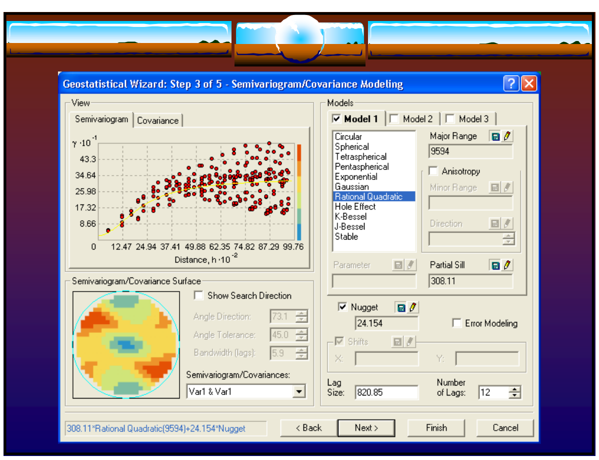
Kriging

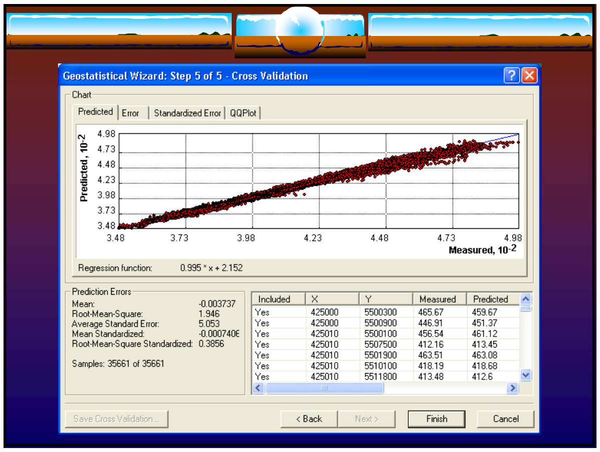
- Theoretically, once the overall trend is accounted for, the remaining deviations (minus random error) can be accounted for by spatially autocorrelated variations that are a function of distance

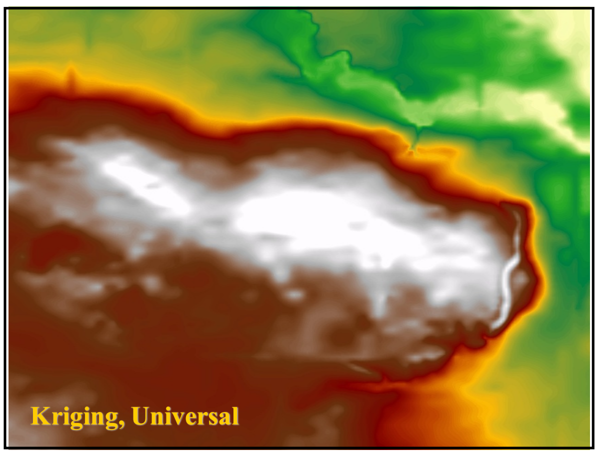


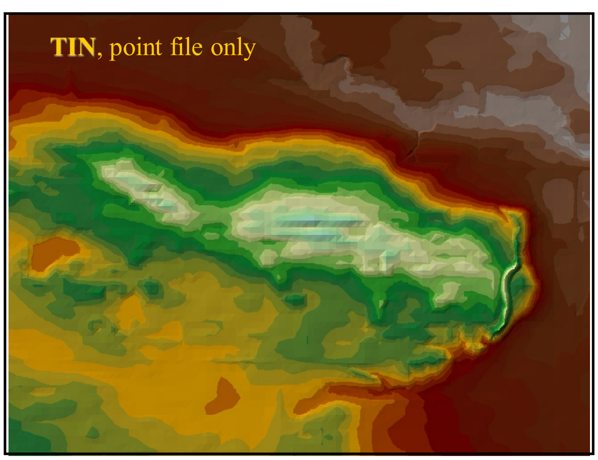
Kriging

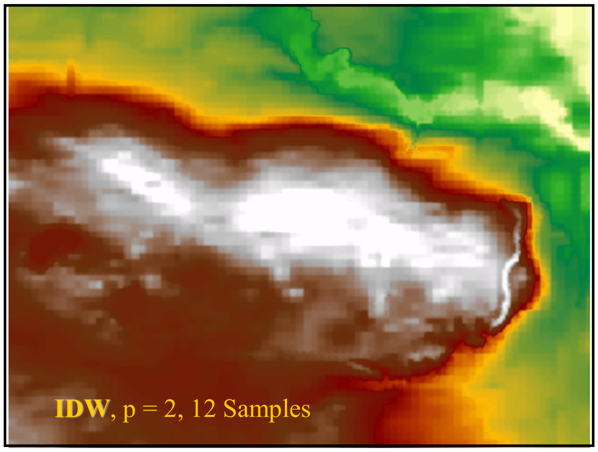
- Regional variable can be estimated from sample data by creating a plot of:
  - γ(h) = variance
  - against h = lag distance
- This is known as a semivariogram
  - Provides optimal information for interpolation:
    - the lag distance (i.e. size/shape of neighbourhood)
    - and an estimate of the unpredictable random component of the equation

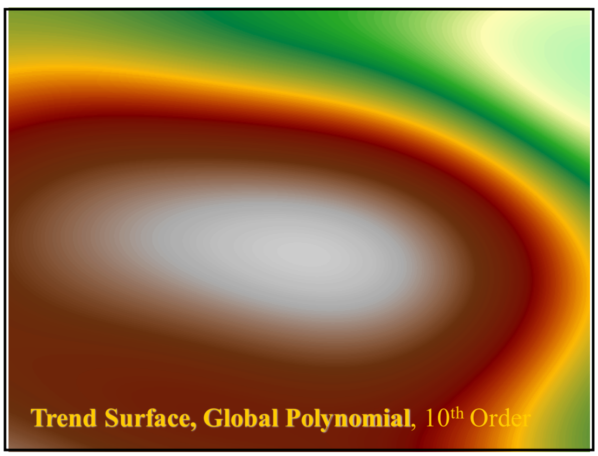


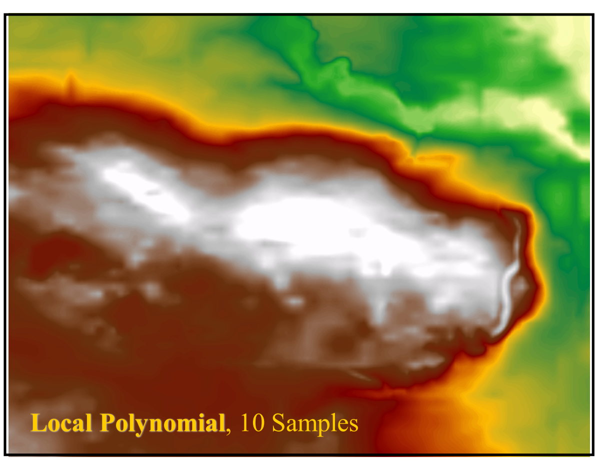

Geostatistical Wizard: Step 5 of 5 - Cross Validation
Chart
Predicted
Error
Standardized Error
QQPlot
Predicted, 10-2
4.98
4.73
4.48
4.23
3.98
3.73
3.48
3.48
3.73
3.98
4.23
4.48
4.73
4.98
Measured, 10-2
Regression function: 0.995 * x + 2.152
Prediction Errors
Mean: -0.003737
Root-Mean-Square: 1.946
Average Standard Error: 5.053
Mean Standardized: -0.000740E
Root-Mean-Square Standardized: 0.3856
Samples: 35661 of 35661
Included X Y Measured Predicted
Yes 425000 5500300 465.67 459.67
Yes 425000 5500900 446.91 451.37
Yes 425010 5500100 456.54 461.12
Yes 425010 5507500 412.16 413.45
Yes 425010 5501900 463.51 463.08
Yes 425010 5510100 418.19 418.68
Yes 425010 5511800 413.48 412.6
Save Cross Validation...
< Back
Next >
Finish
Cancel

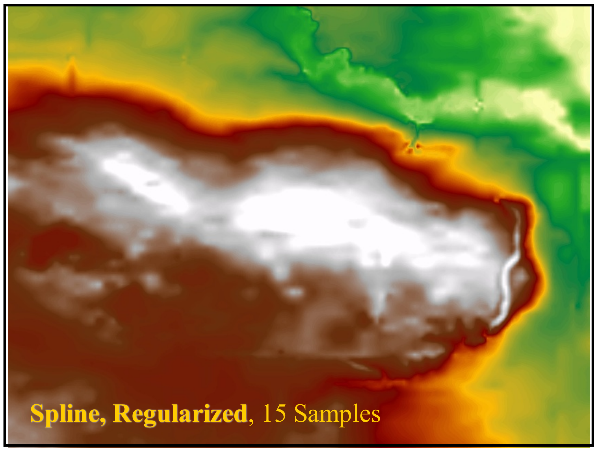

Kriging, Universal

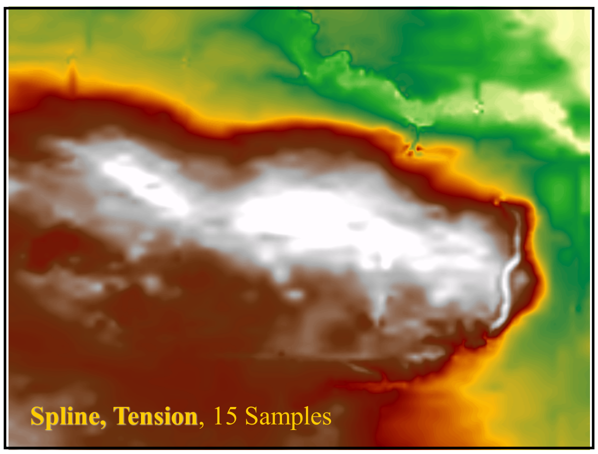

TIN, point file only

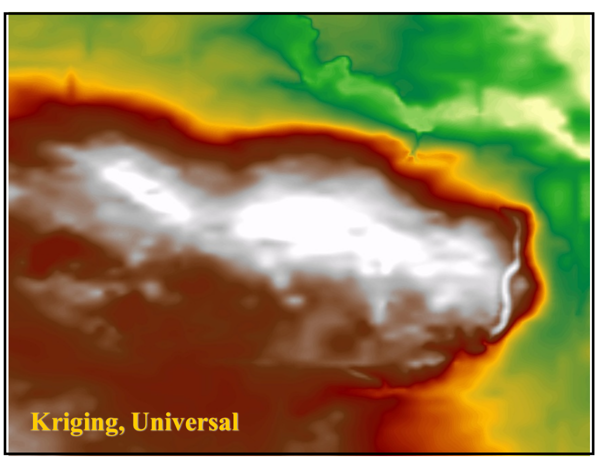

IDW, p = 2, 12 Samples

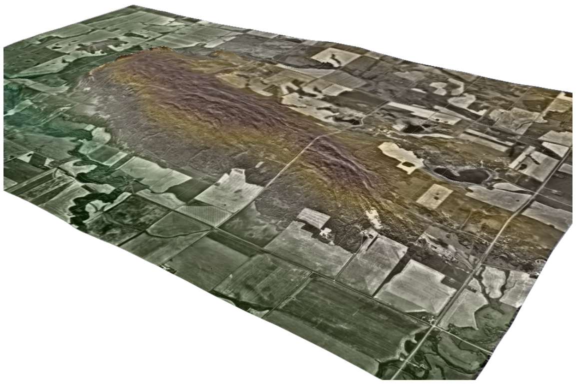

Trend Surface, Global Polynomial, 10th Order


Local Polynomial, 10 Samples


Spline, Regularized, 15 Samples


Spline, Tension, 15 Samples


Kriging, Universal



Other “Sufaces”: Thiessen Polygons

- Network of triangels constructed <u>around</u> points
- Also derived from Delaunay triangles
  - Thiessen polys created by drawing lines perpendicular to midpoint
- A built surface not interpolated
- Used to assign catchment or sevice areas

Other Surfaces: Density Surface

- Created from input point or line data
- Numeric field can be specified or count of features
- Results in raster output
  - Cell value is calculated density per unit area