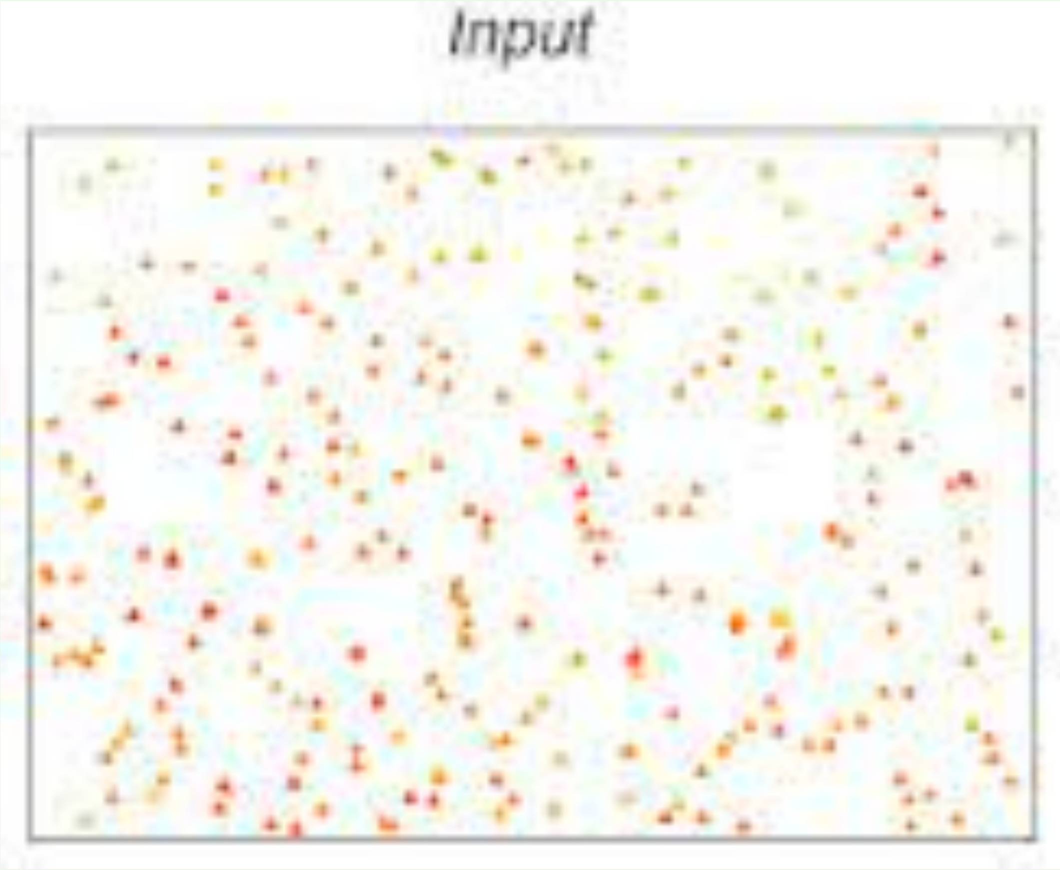
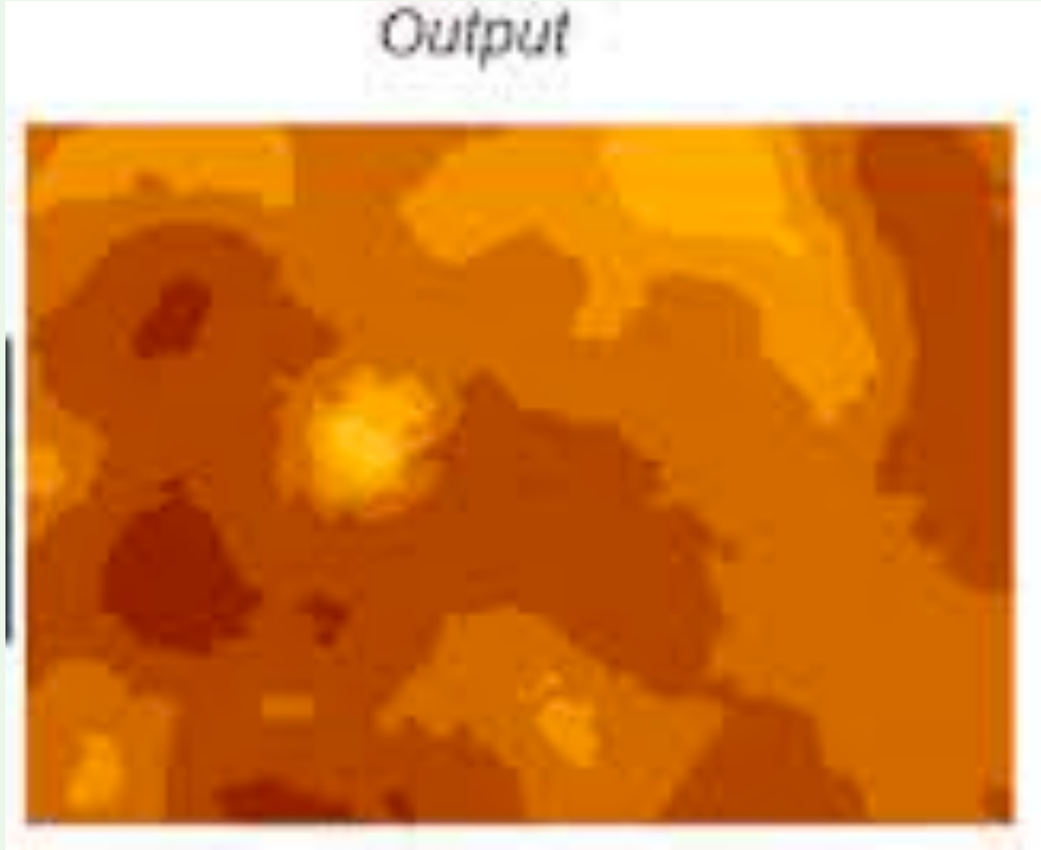
GEOG 358: Introduction to Geographic Information Systems

Spatial Interpolation



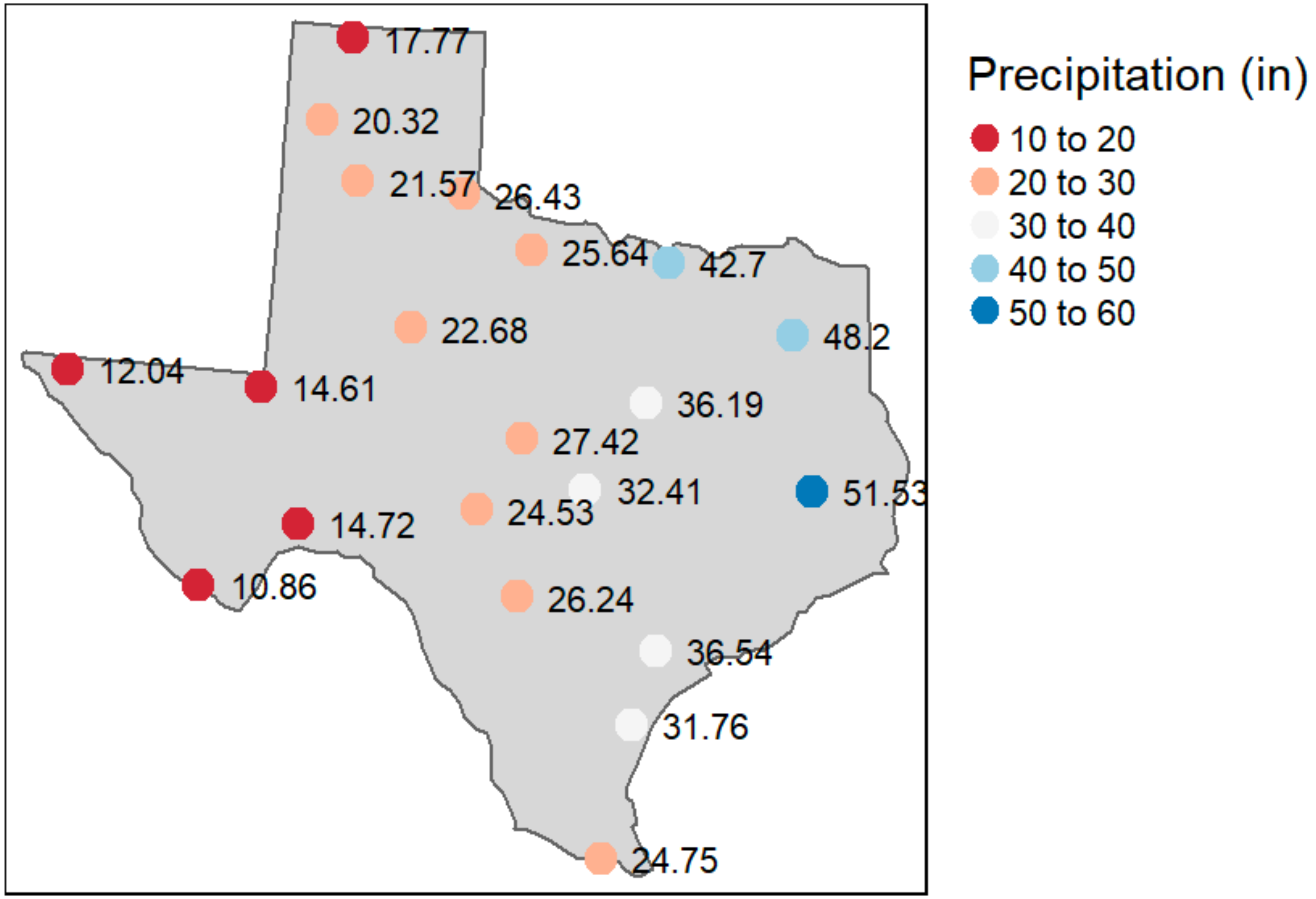


Topics

- Global spatial interpolation methods
  - Trend surface
- Local spatial interpolation methods
  - Nearest neighbor
  - IDW

What Is Spatial Interpolation?

- Estimate values at unsampled locations
  - No distinguishing between spatial interpolation and prediction
- Generate fields/surfaces

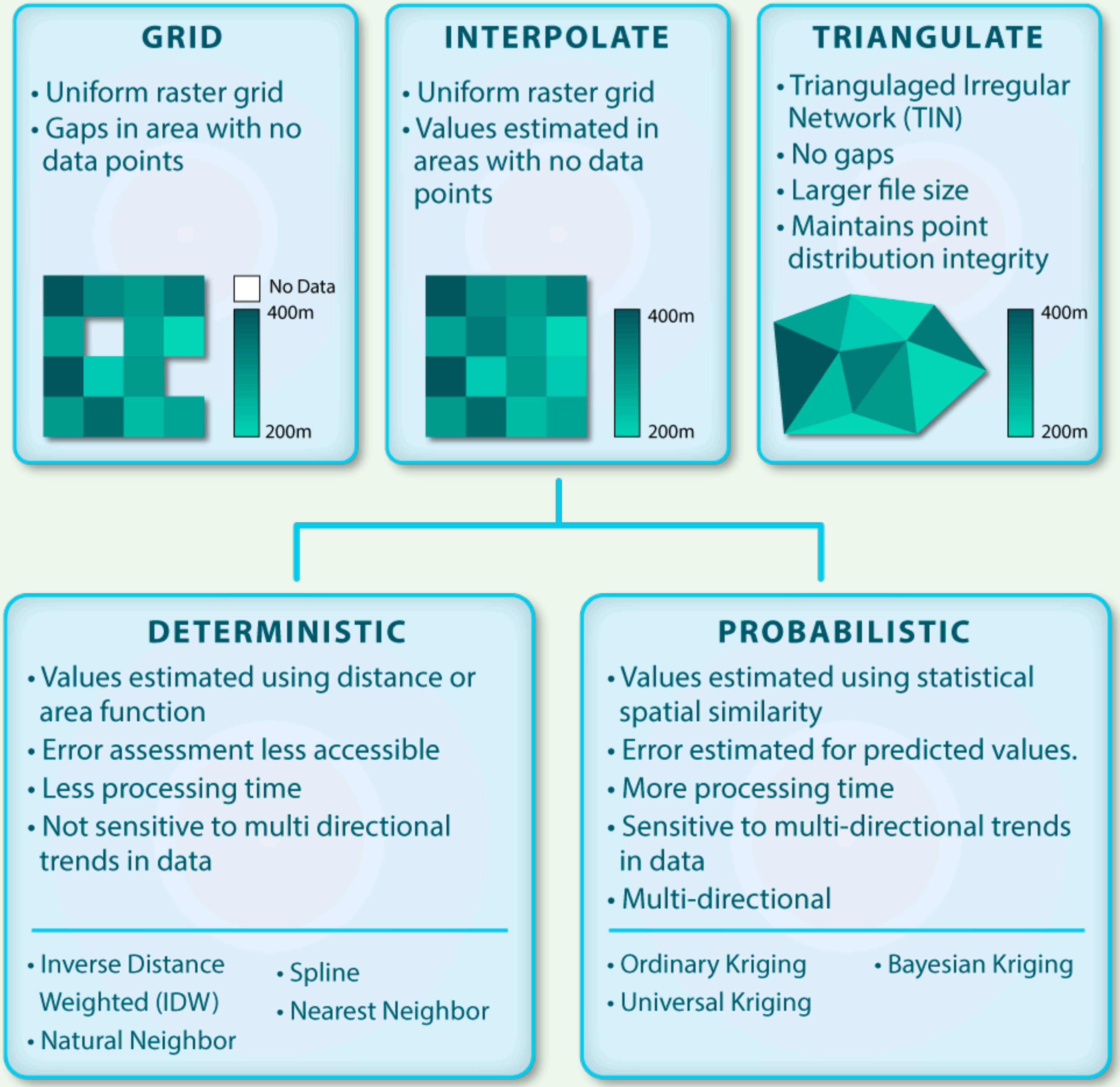


Why Interpolation?

- Can not measure at every location
  - Time
  - Money
  - Impossible (physical or legal)
- Estimate past values or predict future values

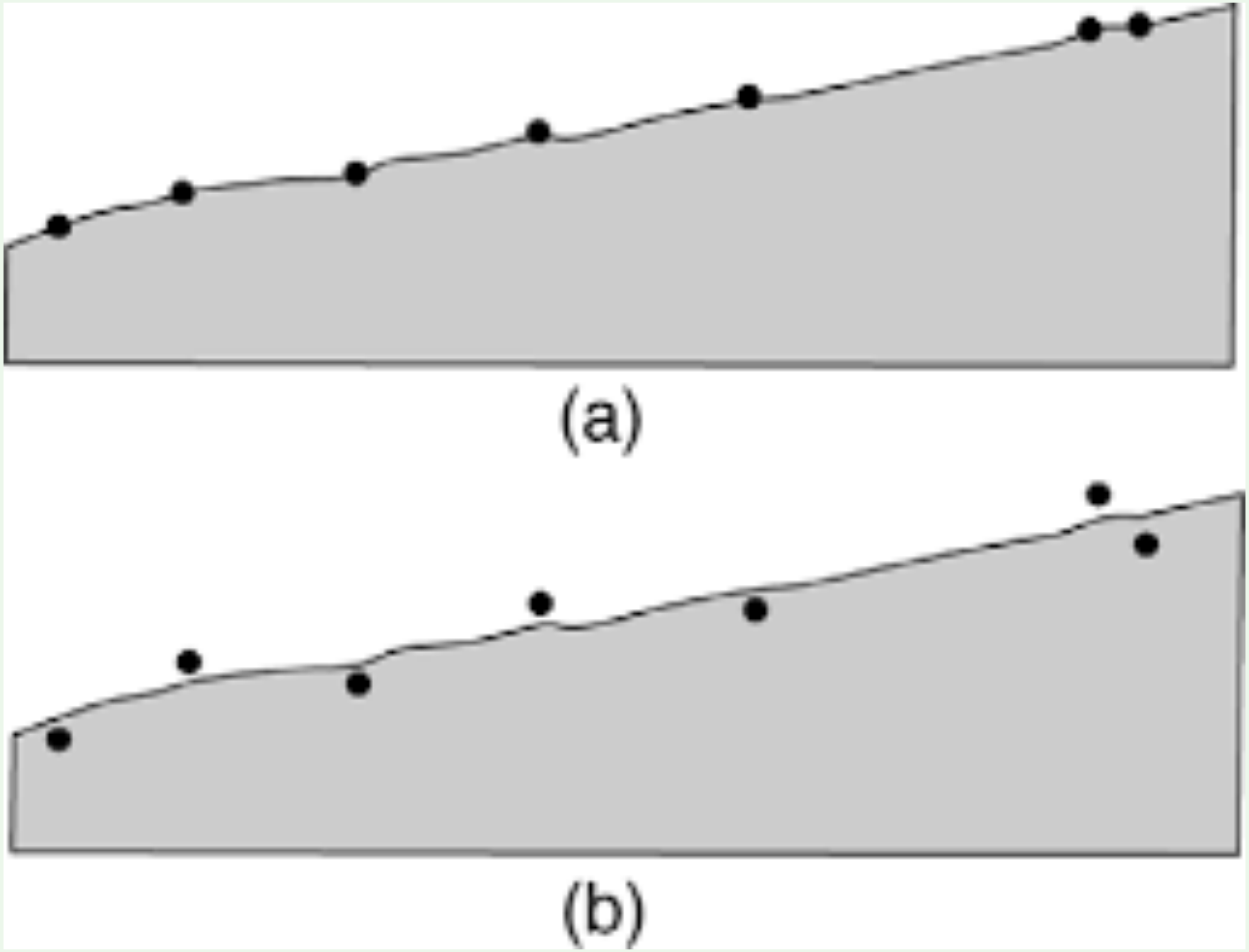
Interpolation Methods

- First law of geography (Tobler in the 60s)
  - Every thing is related to everything else. Near things are more related than distant things
- Local spatial interpolation
  - use nearby samples to estimate
- Global interpolation
  - use all sampled locations
- How to select nearby samples?
- How to use nearby sample values?


GRID
• Uniform raster grid
• Gaps in area with no data points
No Data
400m
200m
INTERPOLATE
• Uniform raster grid
• Values estimated in areas with no data points
400m
200m
TRIANGULATE
• Triangulaged Irregular Network (TIN)
• No gaps
• Larger file size
• Maintains point distribution integrity
400m
200m
DETERMINISTIC
• Values estimated using distance or area function
• Error assessment less accessible
• Less processing time
• Not sensitive to multi directional trends in data
• Inverse Distance Weighted (IDW)
• Natural Neighbor
• Spline
• Nearest Neighbor
PROBABILISTIC
• Values estimated using statistical spatial similarity
• Error estimated for predicted values.
• More processing time
• Sensitive to multi-directional trends in data
• Multi-directional
• Ordinary Kriging
• Universal Kriging
• Bayesian Kriging

	Predict values outside of sample point	Inherent Error Predicition	Good for variable point spacing	Good for dense point spacing	Surface passes through sample points
TIN				✓	✓
IDW			✓	✓	✓
Spline	✓				✓
Natural Neighbor			✓		✓
Ordinary Kriging	✓	✓	✓	✓	
Universal Kriging	✓	✓	✓	✓	
Bayesian Kriging	✓	✓	✓	✓	

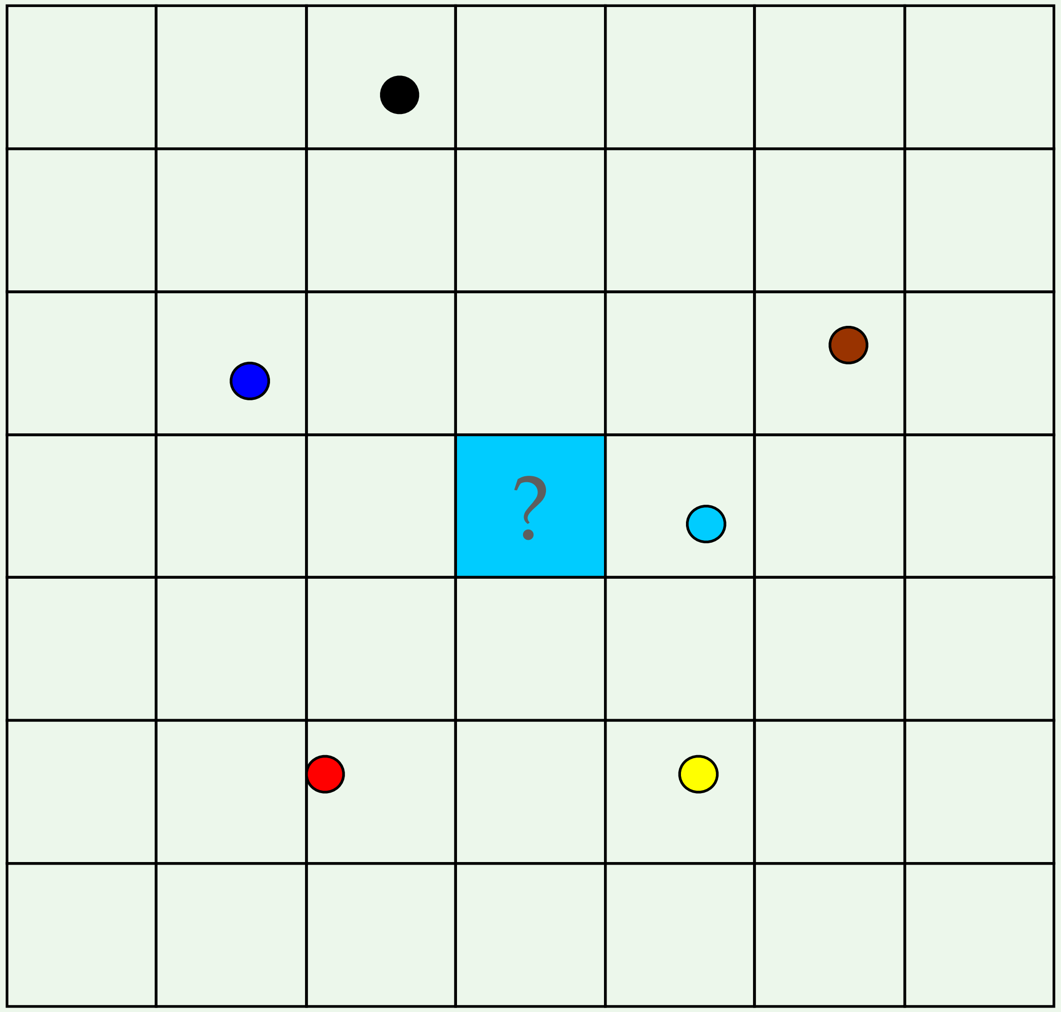
neon

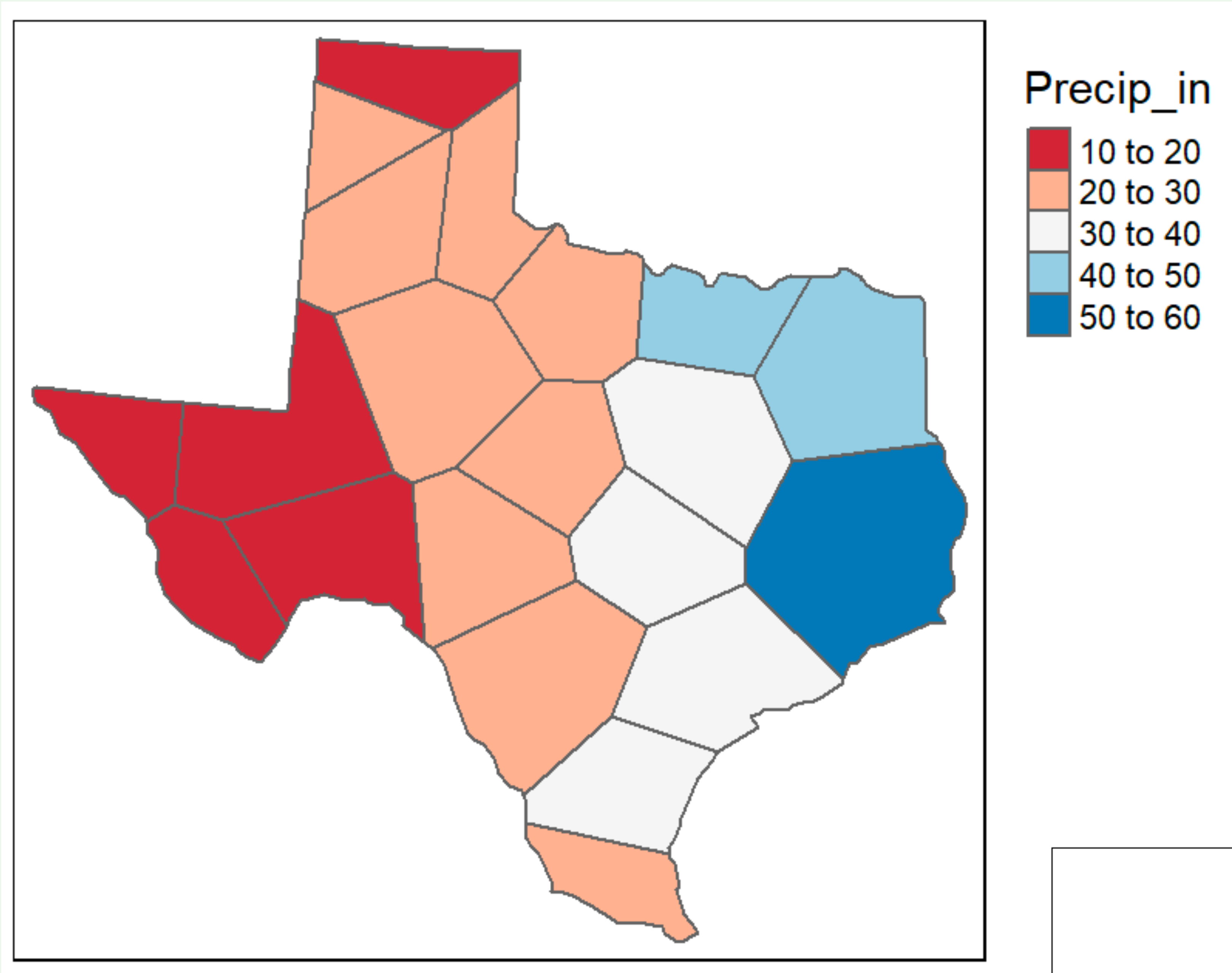

(a)
(b)

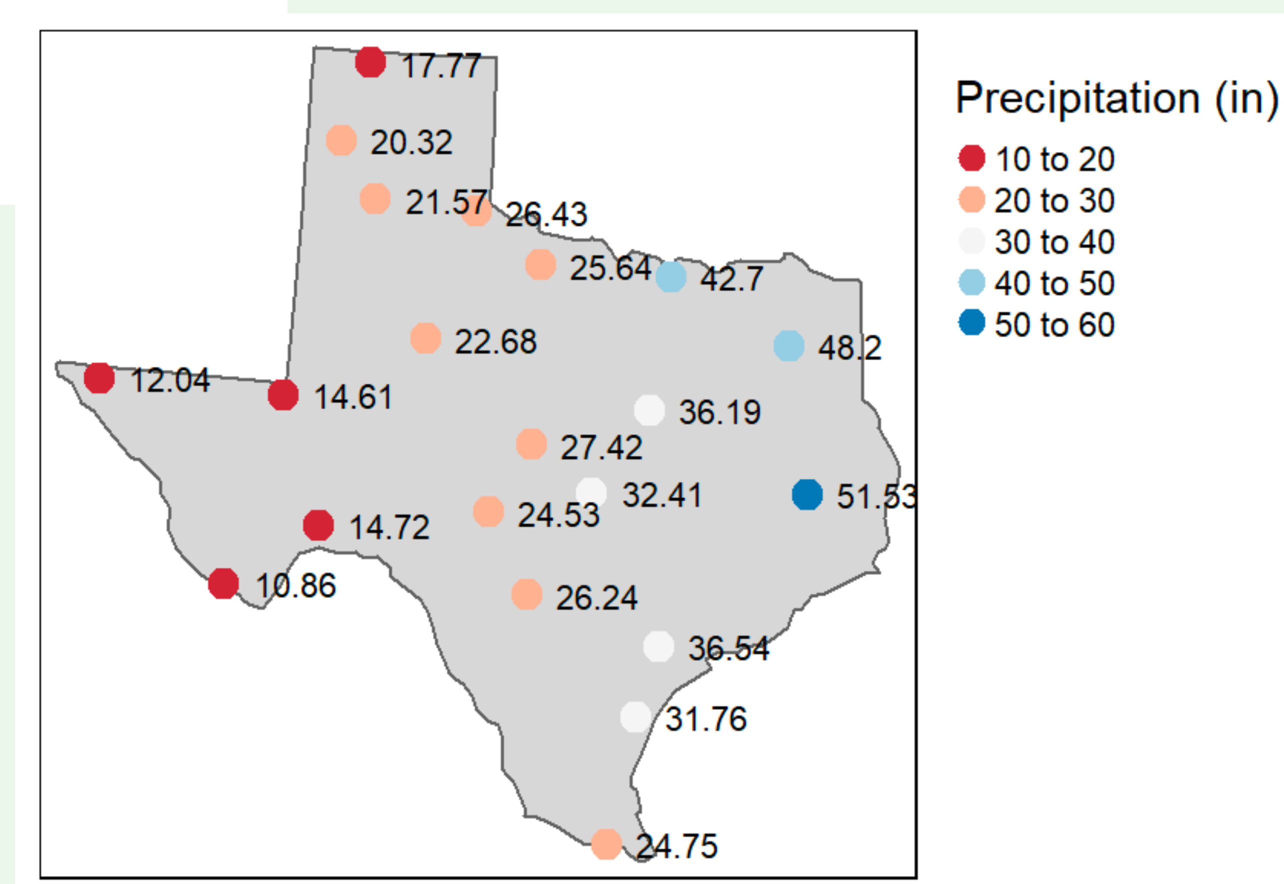
Deterministic Methods

Nearest Neighbor Interpolation

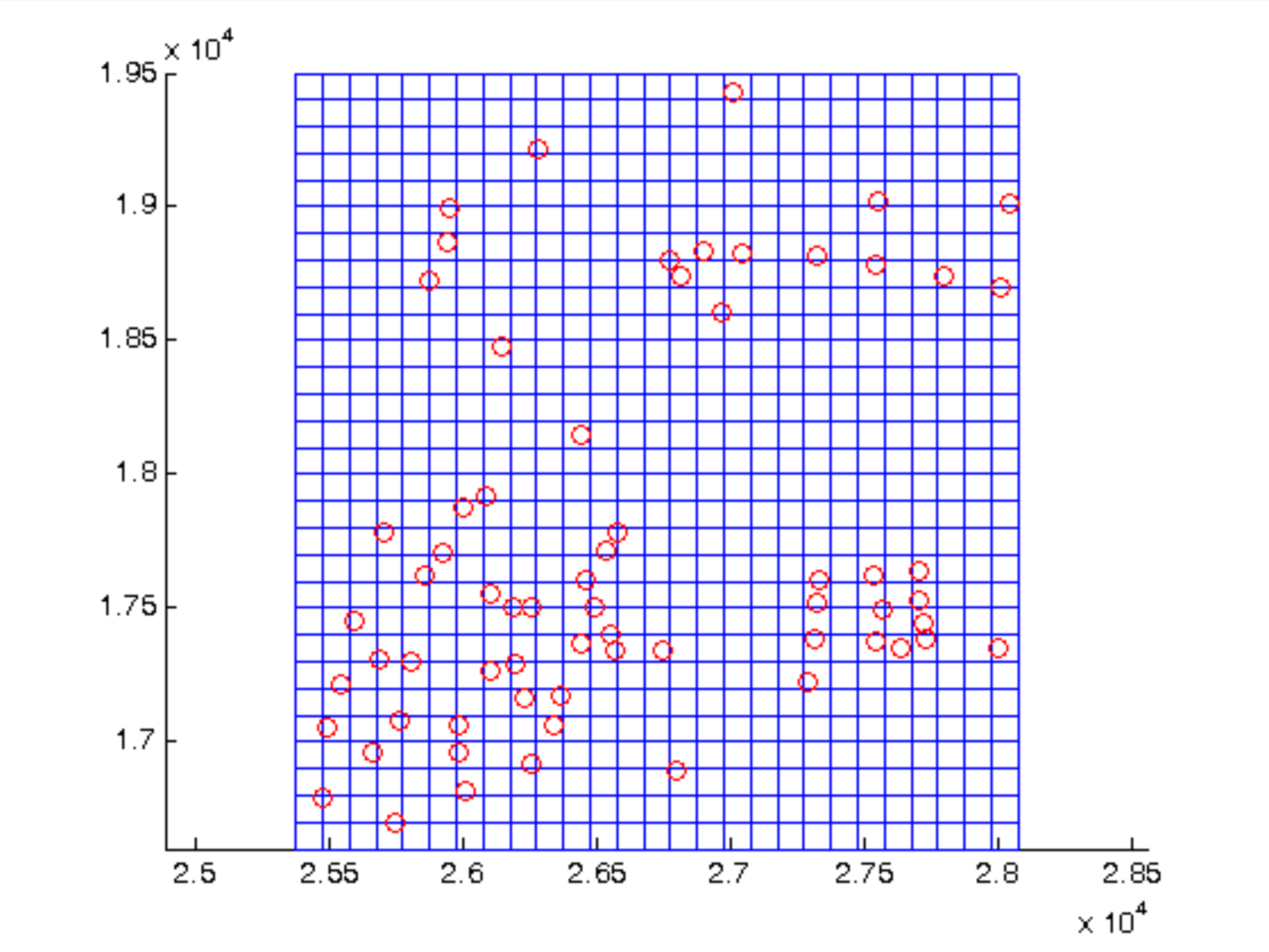
- Each cell is assigned to the value of its nearest sampled point.



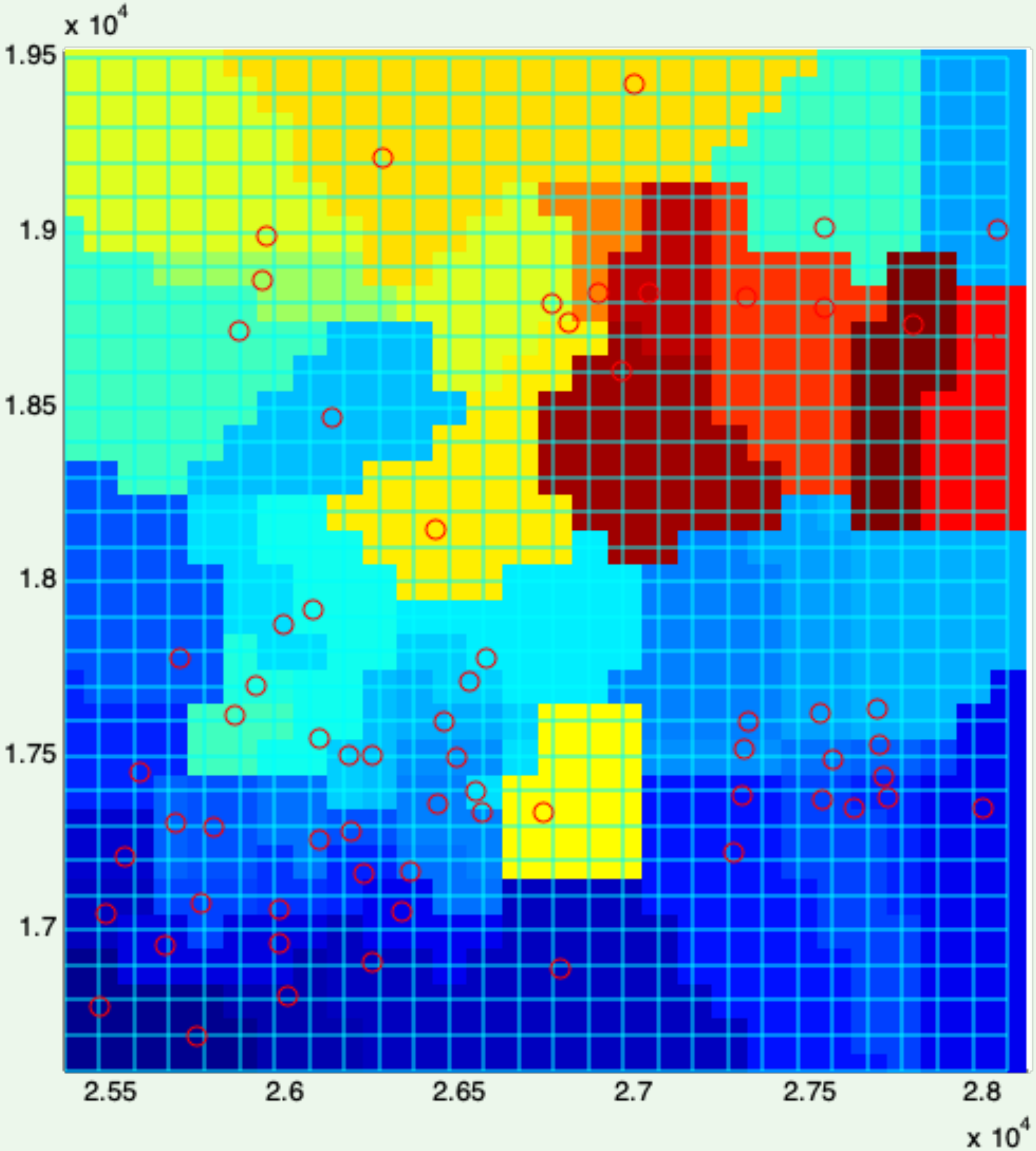

Precip_in
10 to 20
20 to 30
30 to 40
40 to 50
50 to 60


Precipitation (in)
10 to 20
20 to 30
30 to 40
40 to 50
50 to 60
17.77
20.32
21.57
26.43
25.64
42.7
22.68
48.2
12.04
14.61
36.19
27.42
32.41
51.53
24.53
14.72
10.86
26.24
36.54
31.76
24.75

Interpolate Ground Water Table Elevation



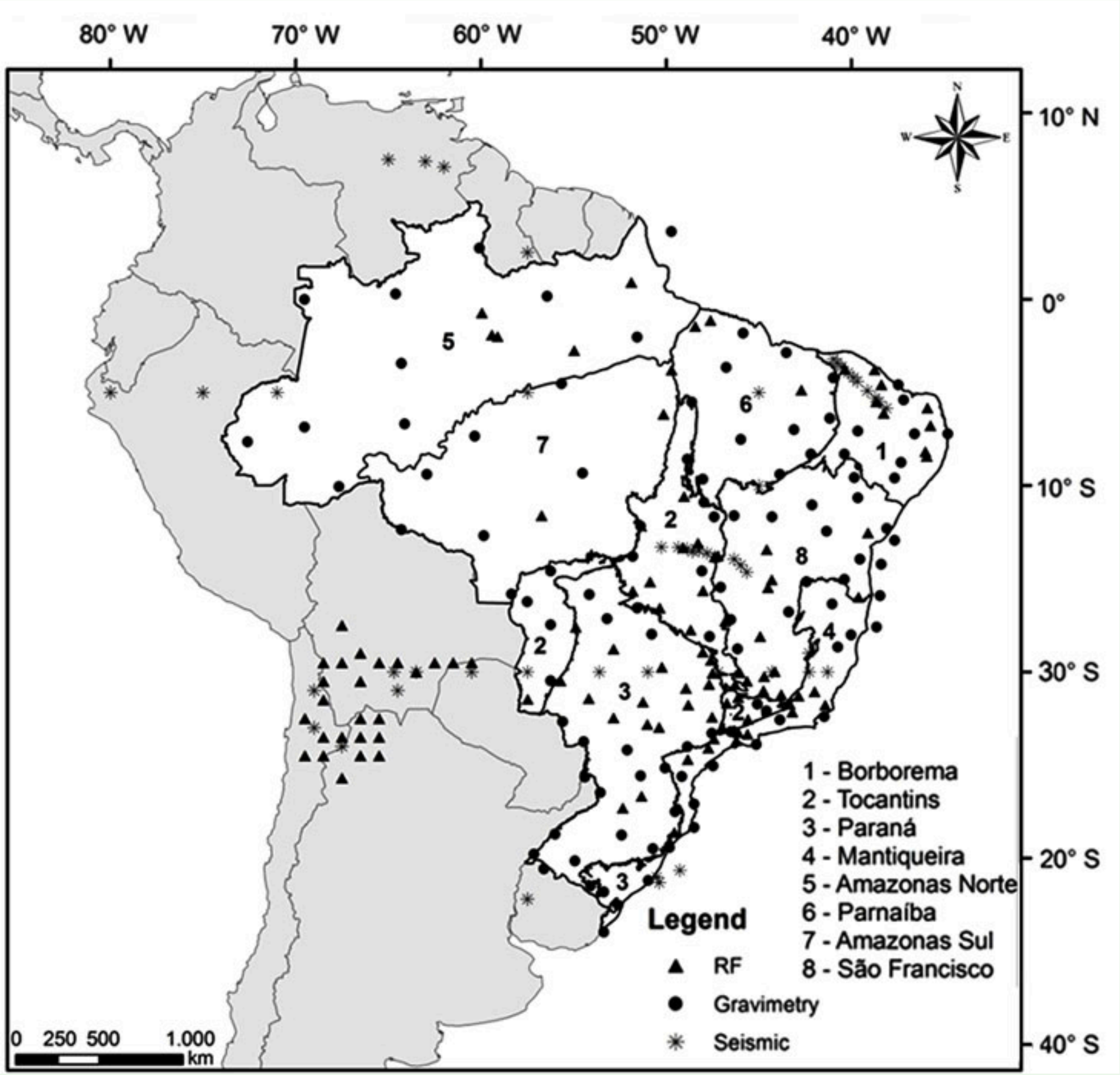
Nearest Neighbor Interpolation

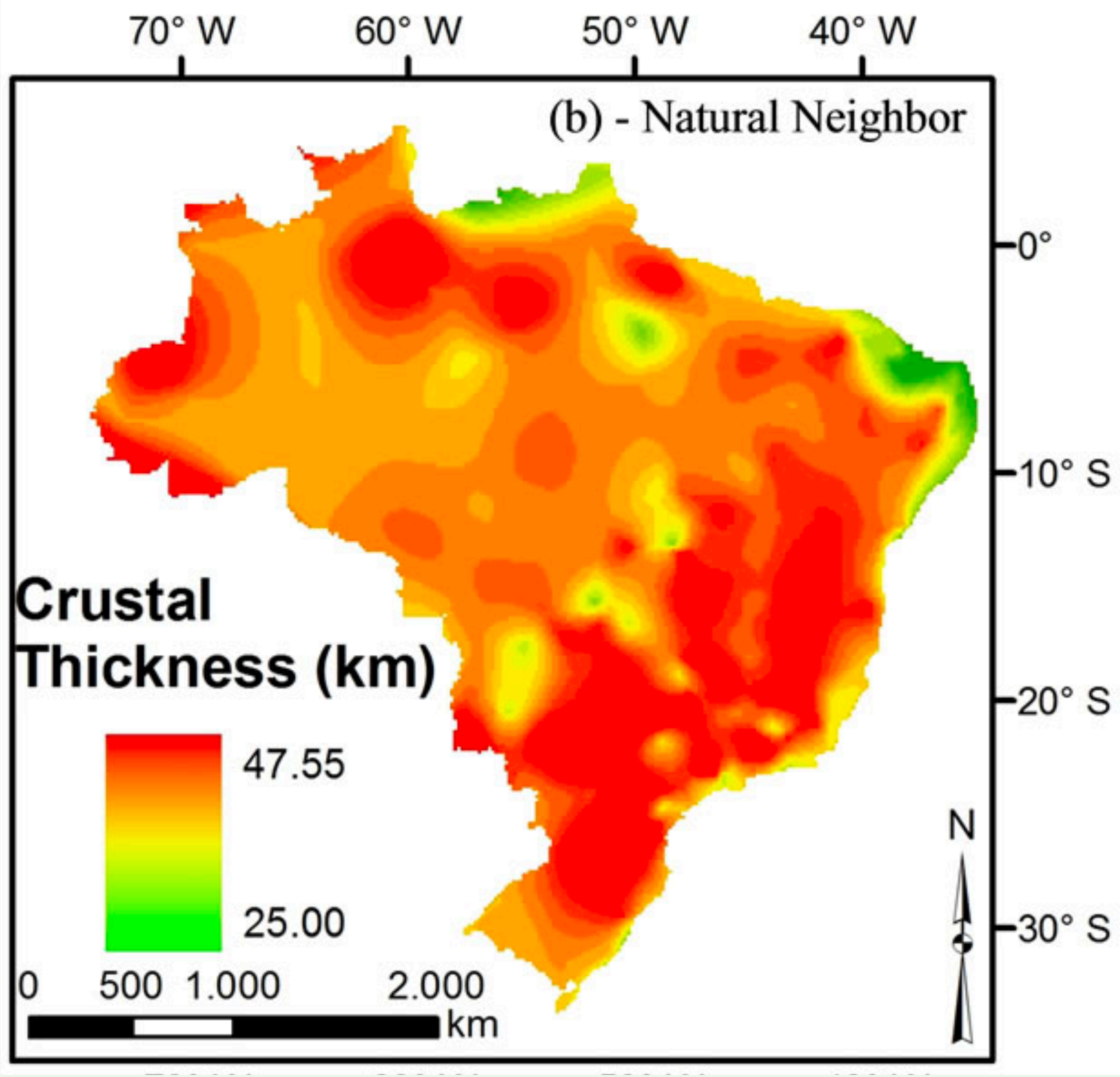


Thiessen's polygons

Proximal regions

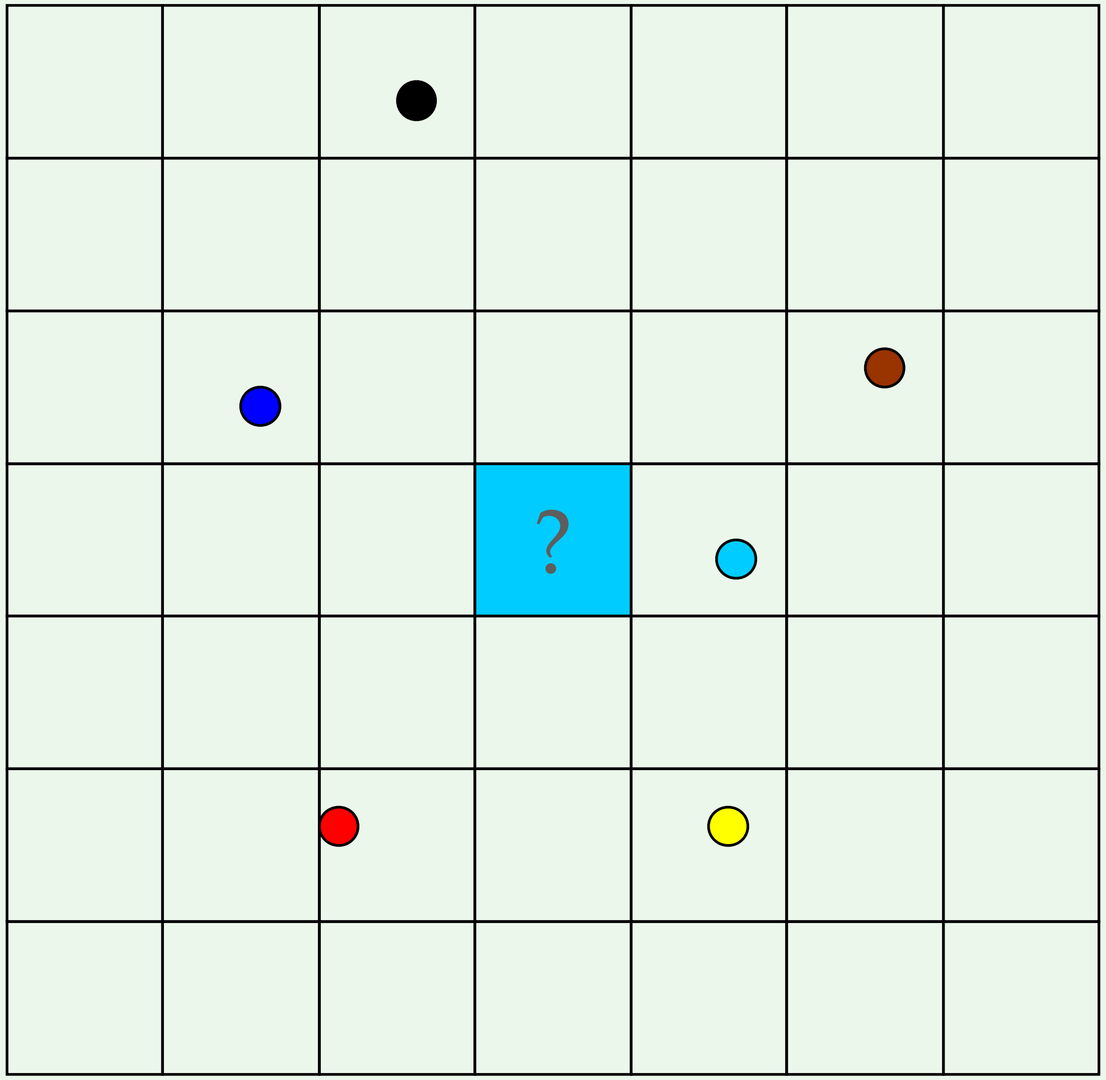
- Simple calculation
- Output has the same value range as the samples do
- Resulting surface is not continuous


80° W
70° W
60° W
50° W
40° W
10° N
0°
10° S
20° S
30° S
40° S
N
W
E
S
1 - Borborema
2 - Tocantins
3 - Paraná
4 - Mantiqueira
5 - Amazonas Norte
6 - Parnaíba
7 - Amazonas Sul
8 - São Francisco
Legend
RF
Gravimetry
Seismic
0 250 500 1.000
km


70° W
60° W
50° W
40° W
(b) - Natural Neighbor
0°
10° S
20° S
30° S
Crustal Thickness (km)
47.55
25.00
0
500
1.000
2.000
km
N

Using More Than One Nearby Samples

- How to select nearby samples?
  - Fixed distance
  - Fixed number of neighbors
- How to combine nearby values
  - Averaging
  - Weighted averaging



Fixed Radius Search

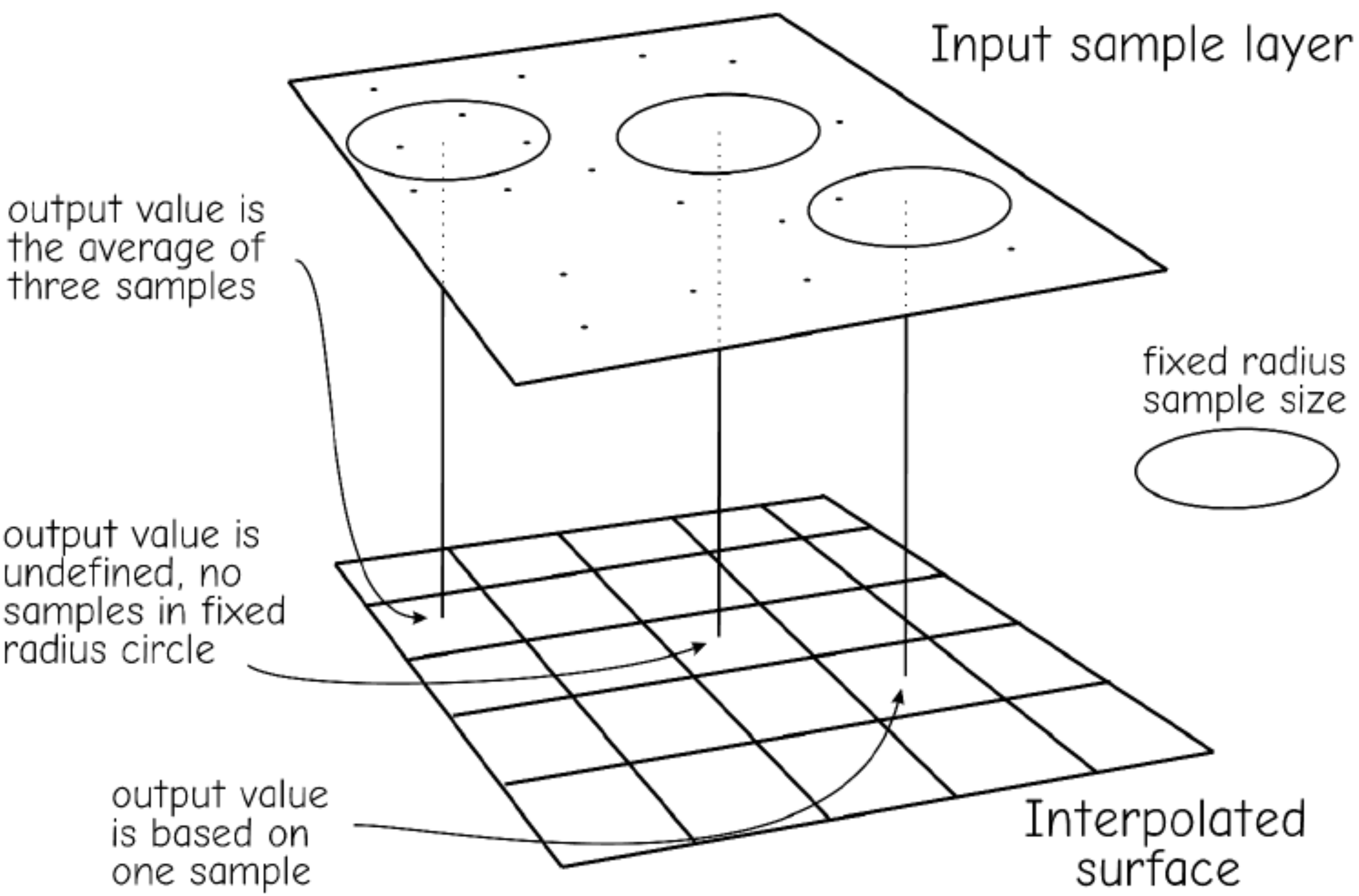
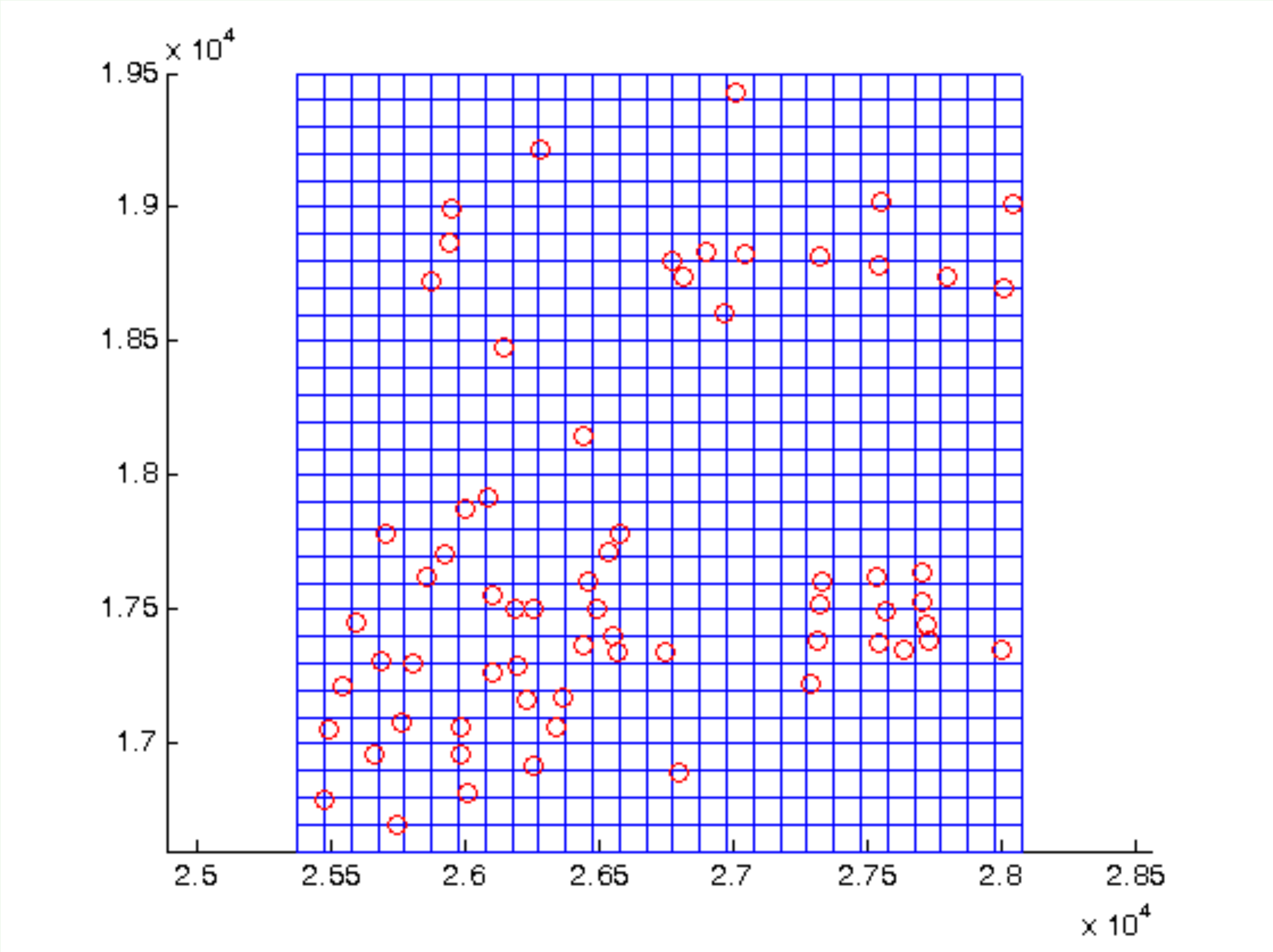
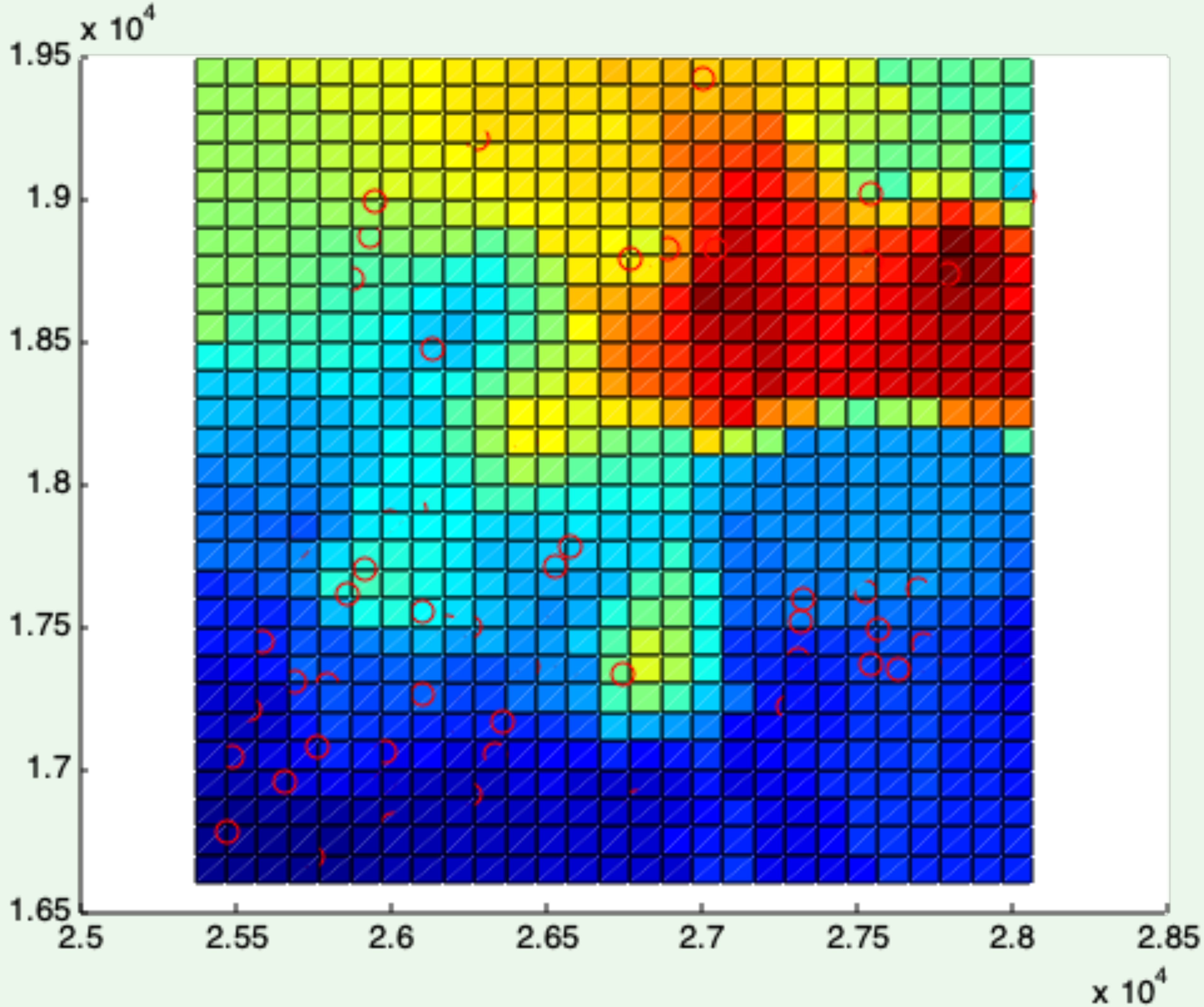


Figure 12-7: A perspective diagram of fixed radius sampling. A circle is centered on each raster cell location. Samples within the circle contribute to the value assigned to each corresponding raster cell (adapted from Mitchell, 1999).

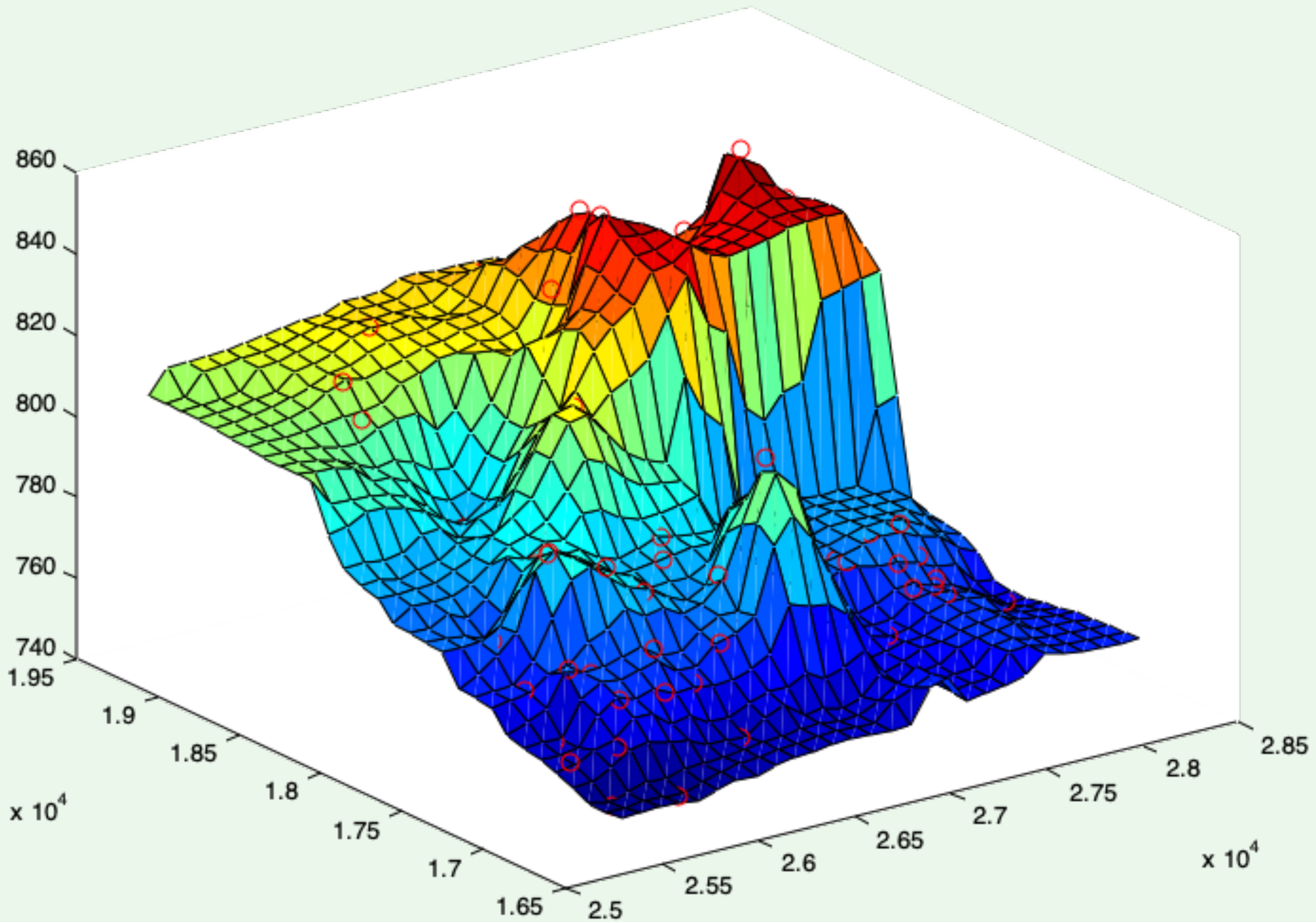
Interpolate Ground Water Table Elevation



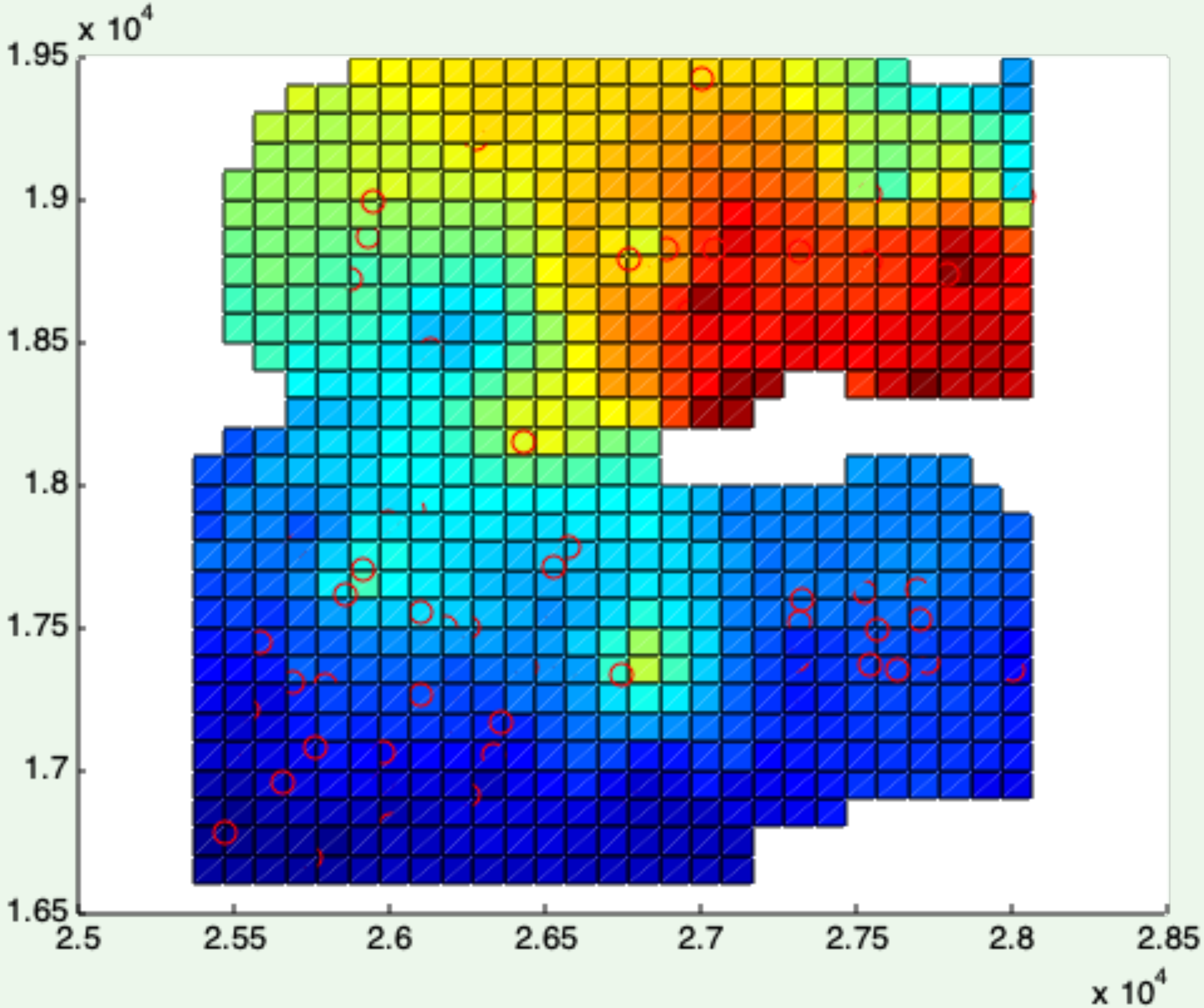
Fixed Number (3 nearest neighbors)



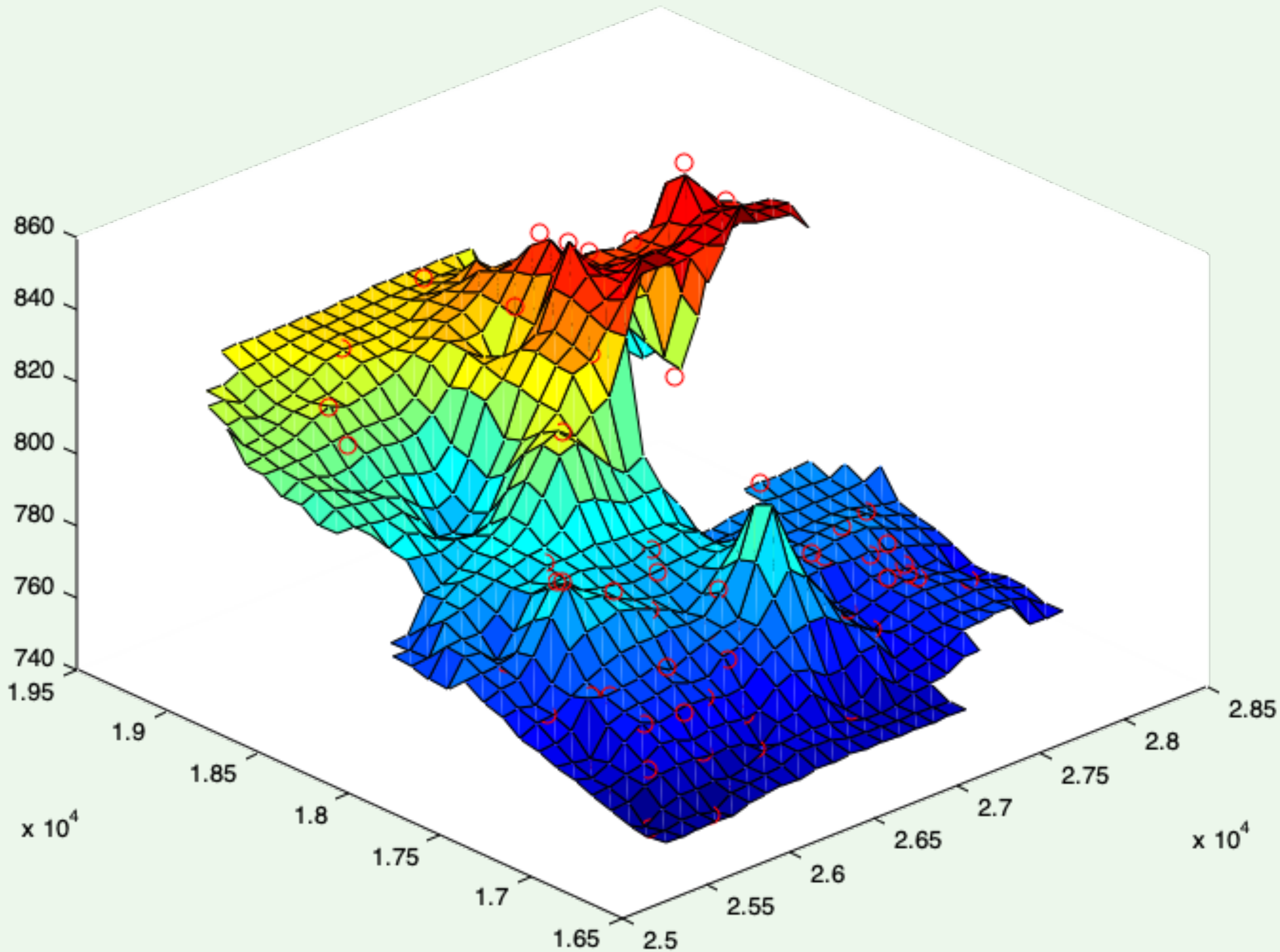
Fixed Number (3 nearest neighbors)



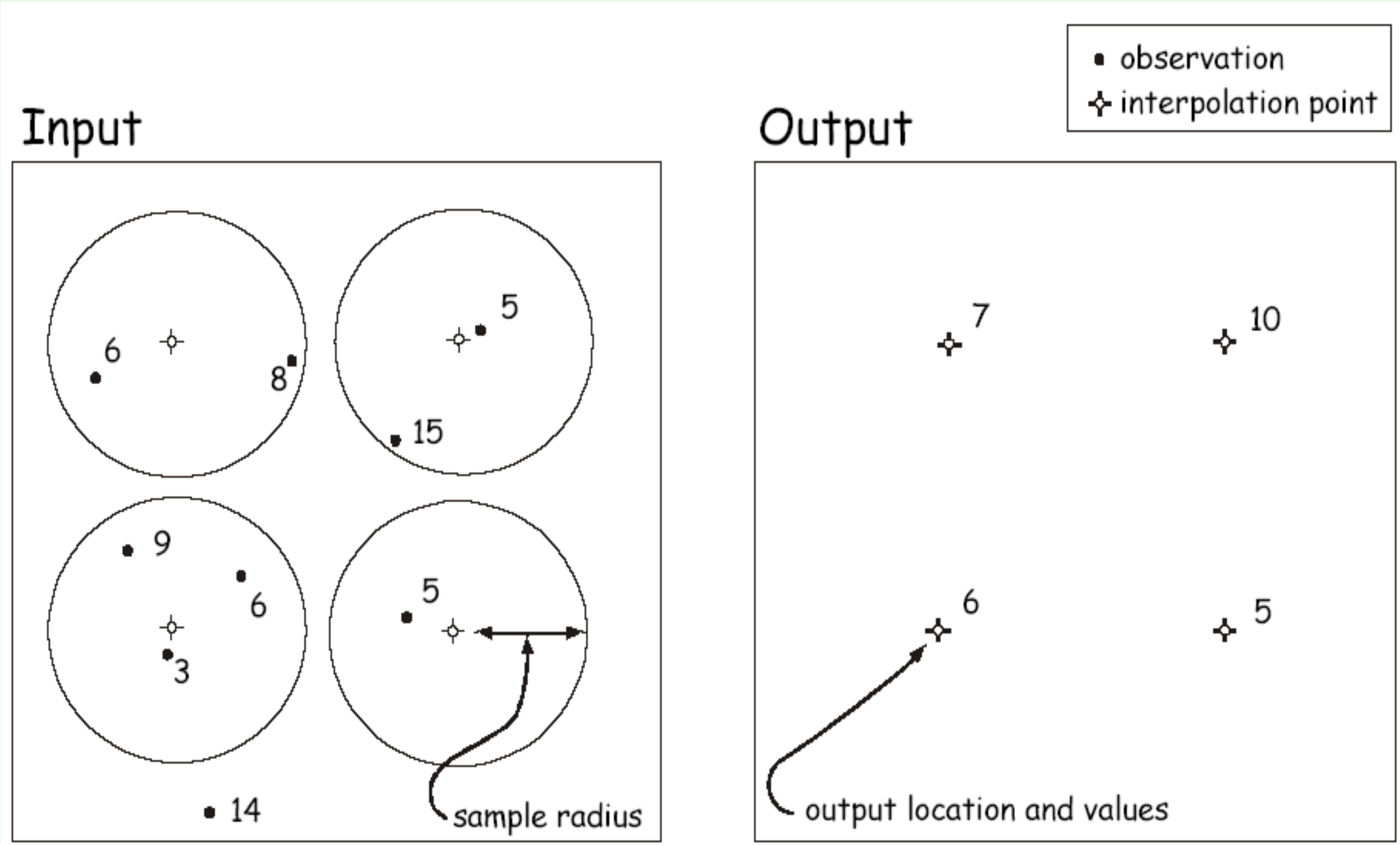
Fixed Distance (500 ft)



Fixed Distance (500 ft)

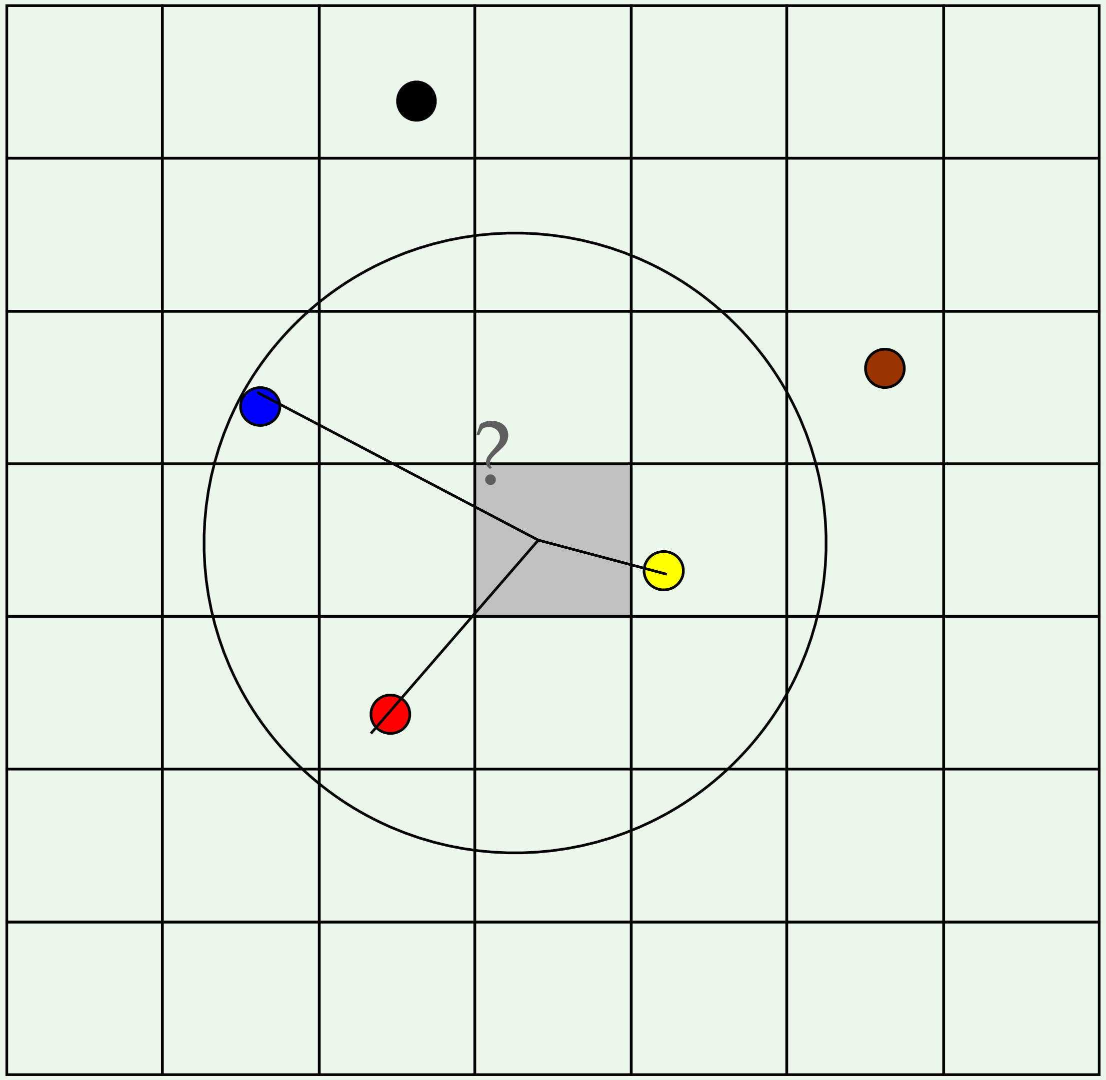


Averaging Nearby Sample Values



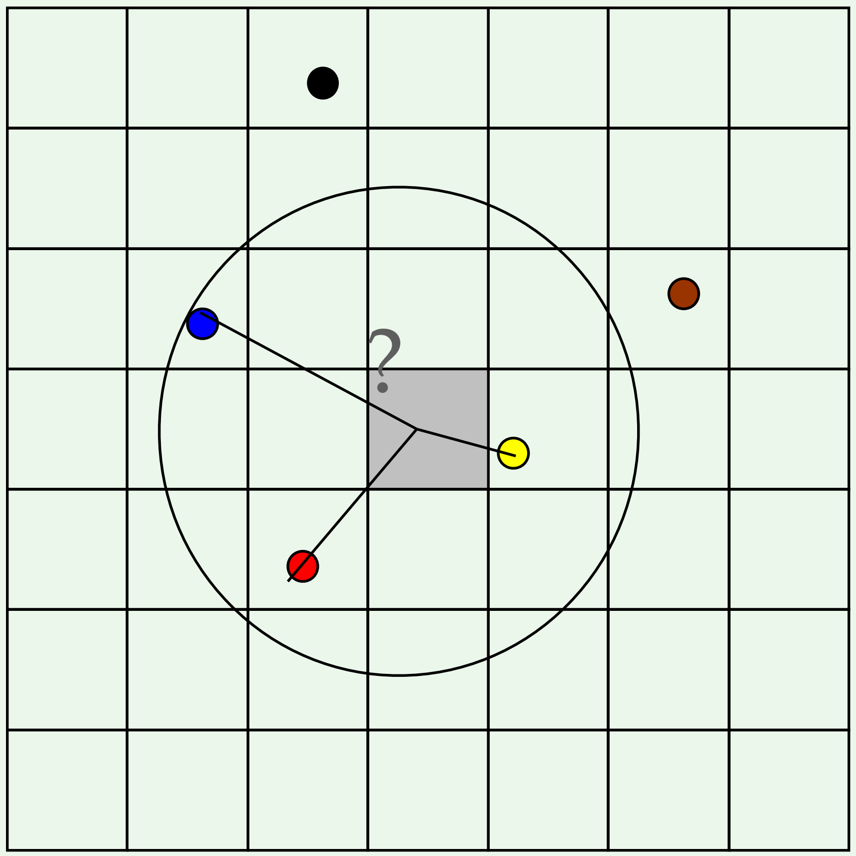
Weighted Combination

- If the first law of geography is true, how should we weight nearby samples?
- Weights should be inversely proportional to the distance to an unsampled location
- Inverse distance weighted (IDW)



Inverse Distance Weighted

$$z = \frac{\sum_{i=1}^{n} z_i \frac{1}{d_i^q}}{\sum_{i=1}^{n} \frac{1}{d_i^q}}$$



z_i = sample value

n = # of nearby sample points to be used

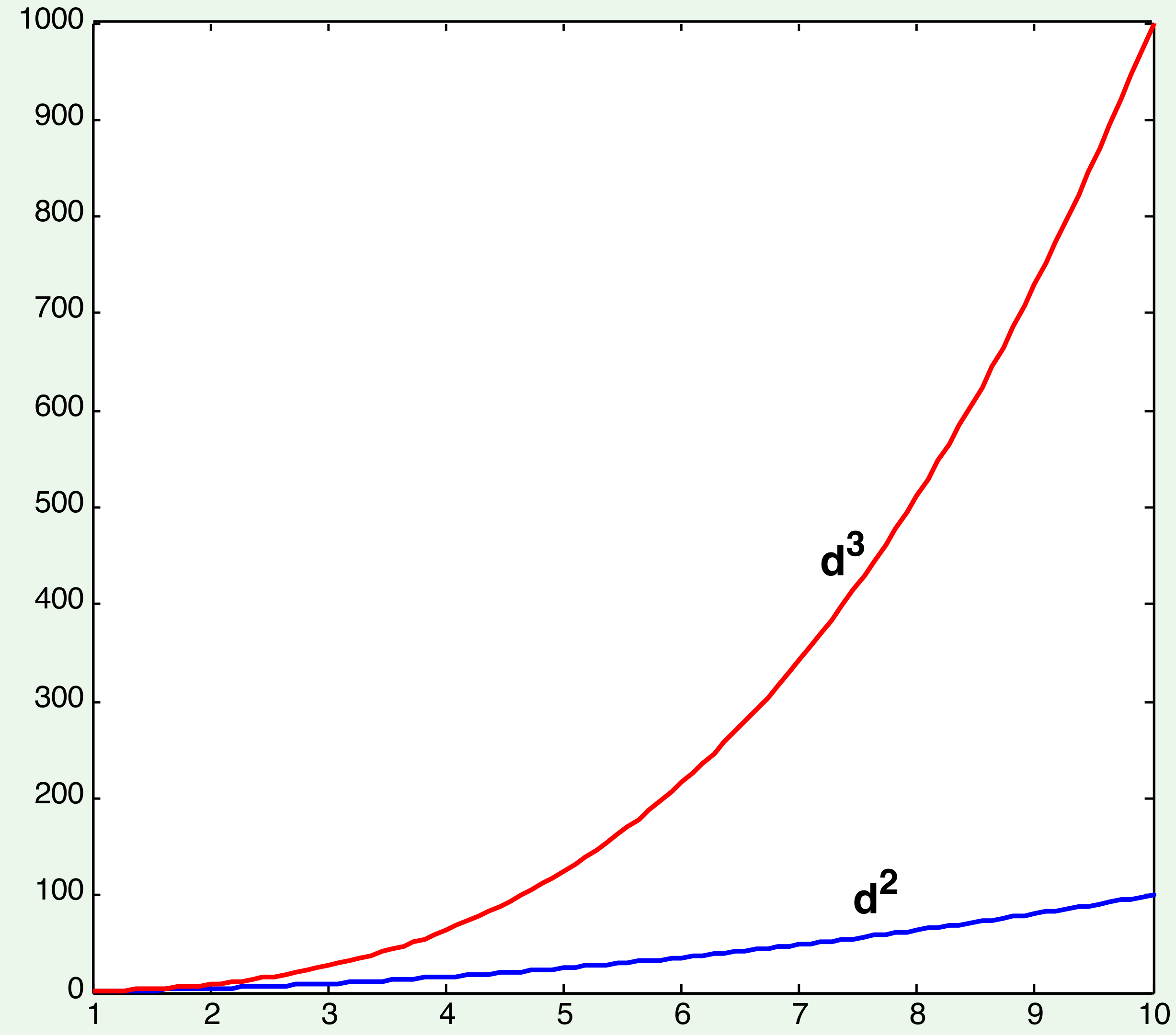
q = exponent of the distance

d_i = distance between a cell and sample point i.

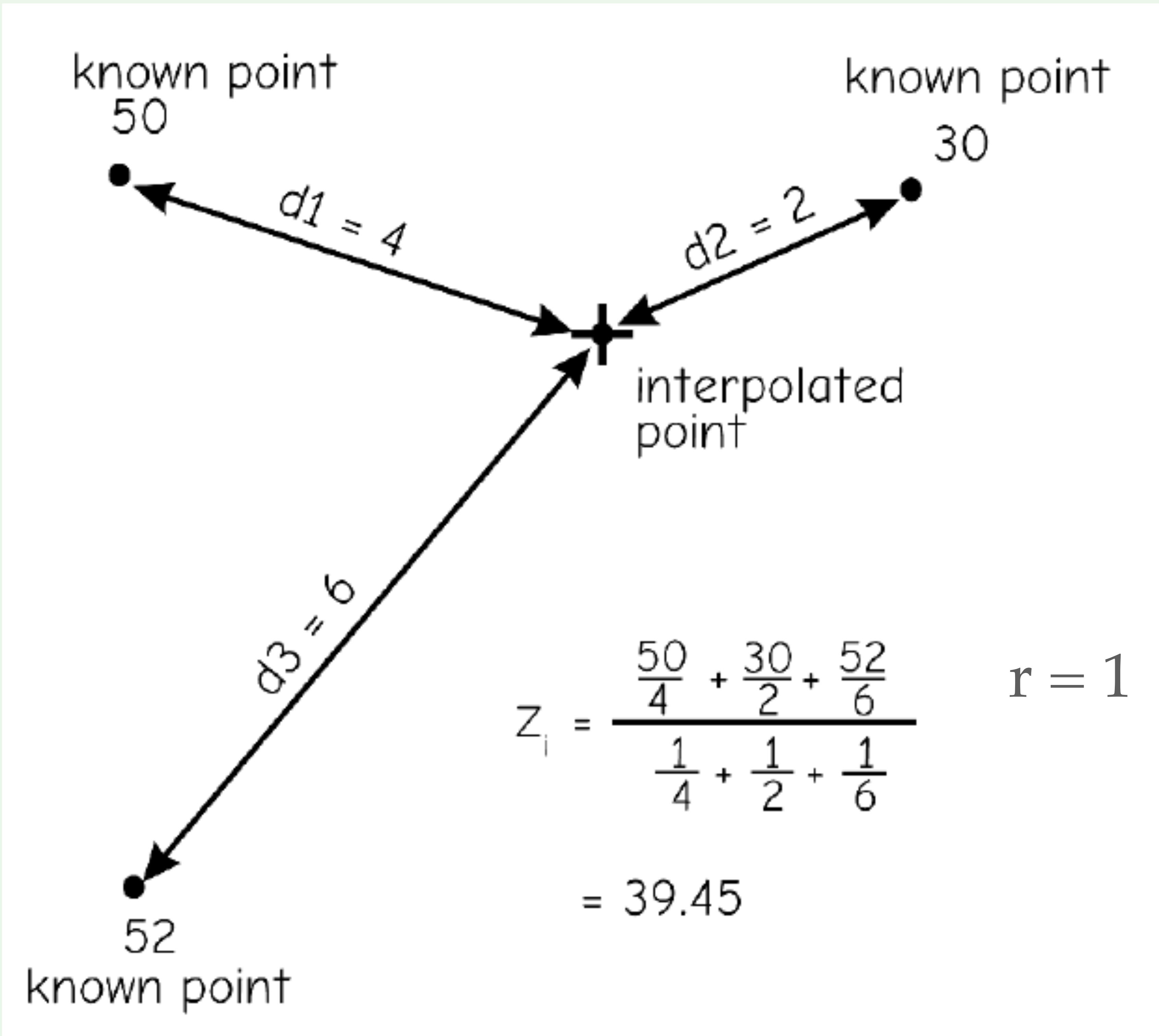
Inverse Distance Weighted--Exponent

$$z = \frac{\sum_{i=1}^{n} z_i \frac{1}{d_i^q}}{\sum_{i=1}^{n} \frac{1}{d_i^q}}$$

A higher exponent gives larger weights to closer neighbors



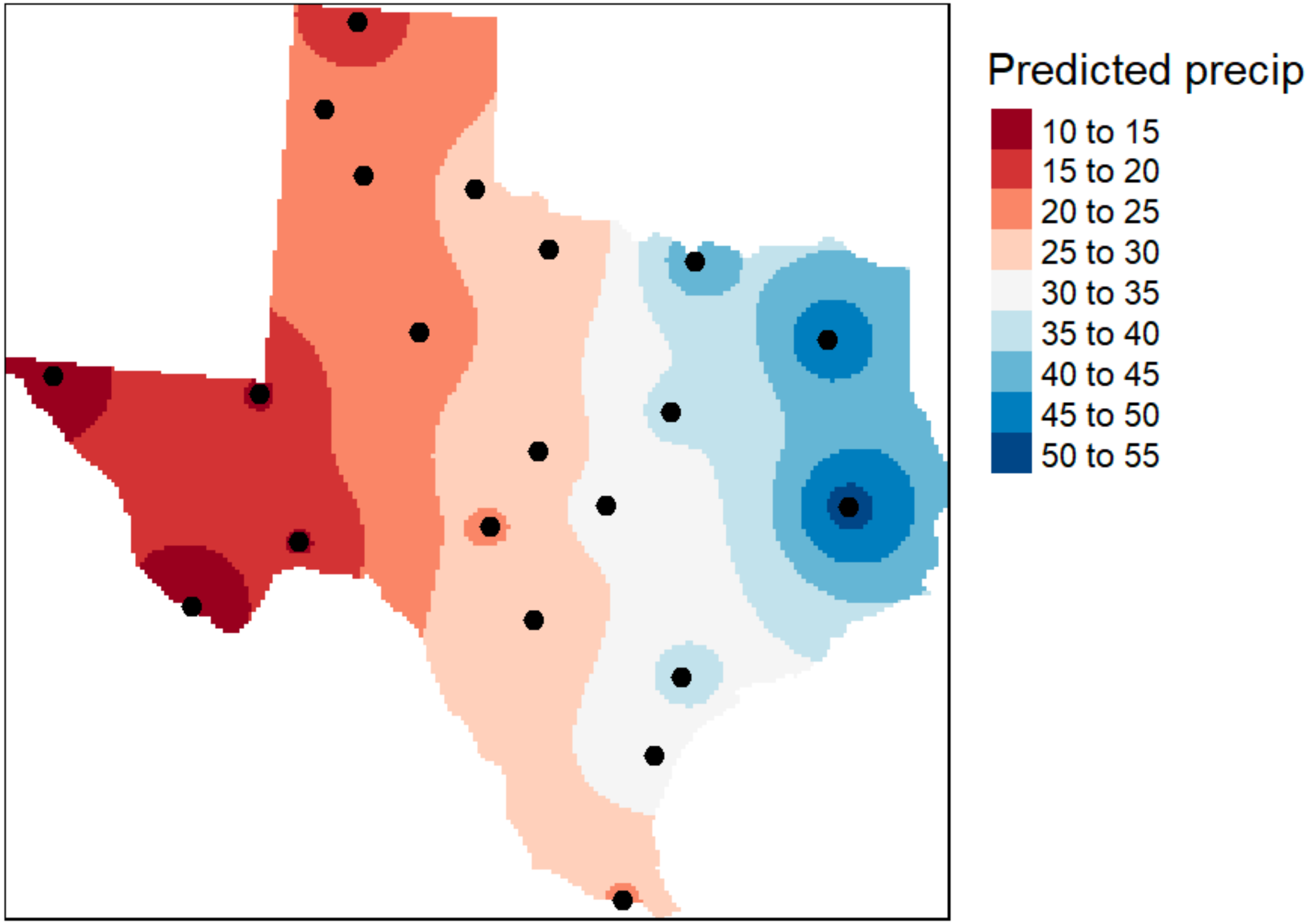
IDW Calculation



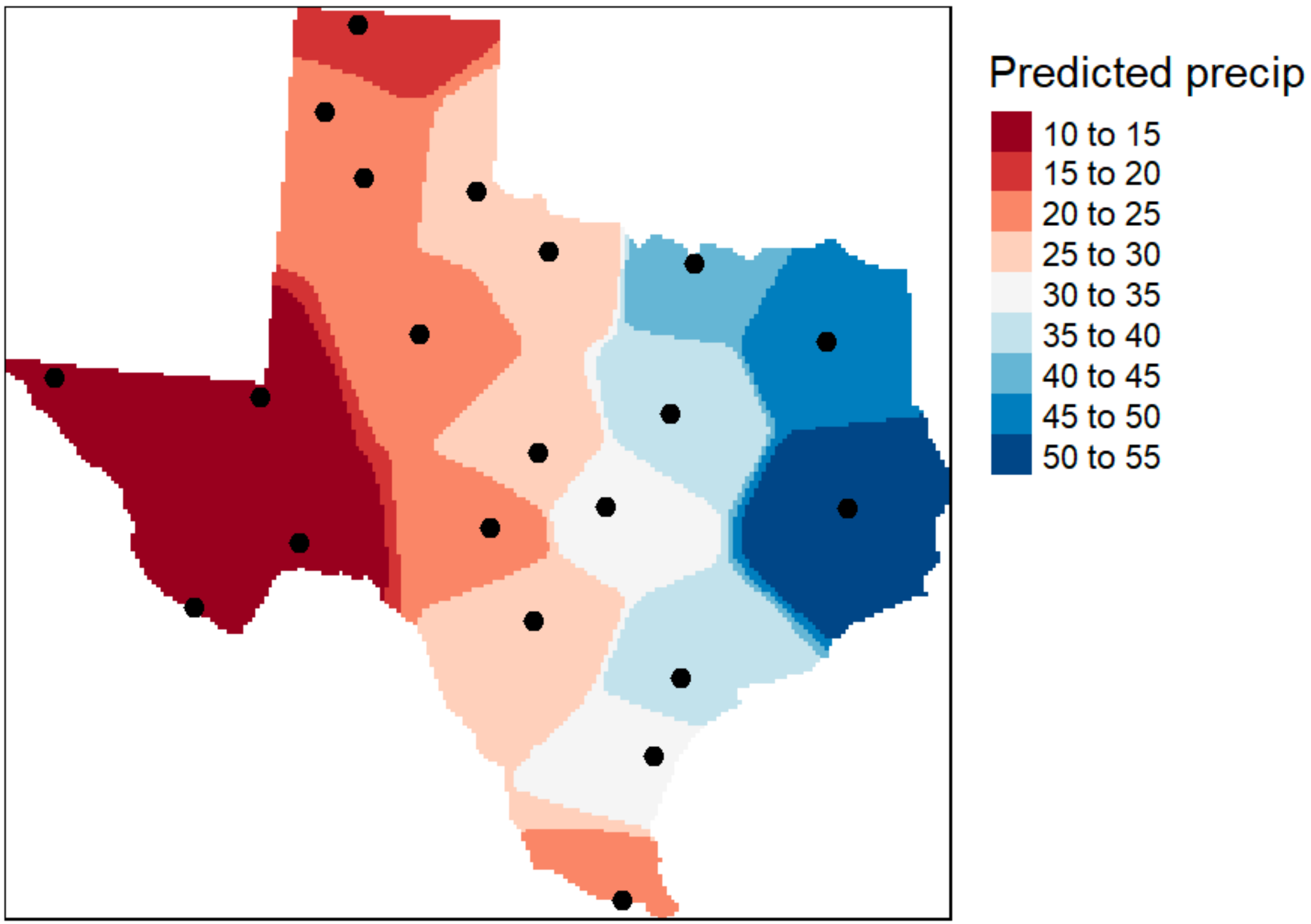
$r = 1$

IDW

- Results in smooth and continuous surface
- Output surface has the same range as the samples
- Requires subjective selection of parameters
    - Search radius
    - Exponent


Predicted precip
10 to 15
15 to 20
20 to 25
25 to 30
30 to 35
35 to 40
40 to 45
45 to 50
50 to 55

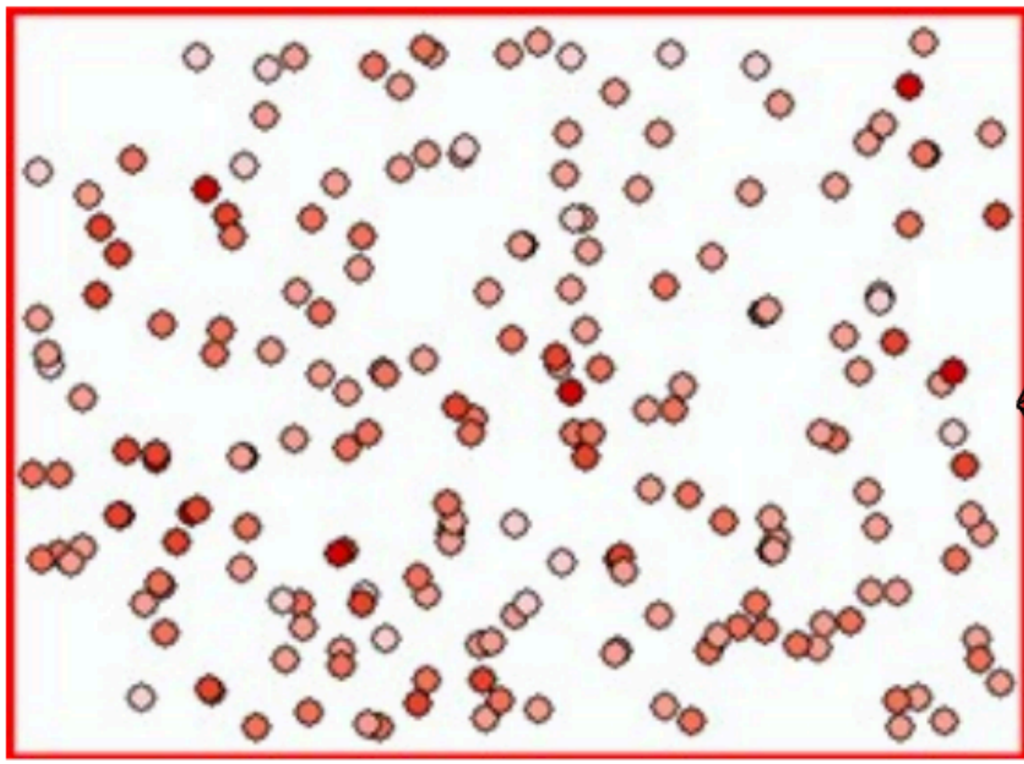
$$q = 2$$

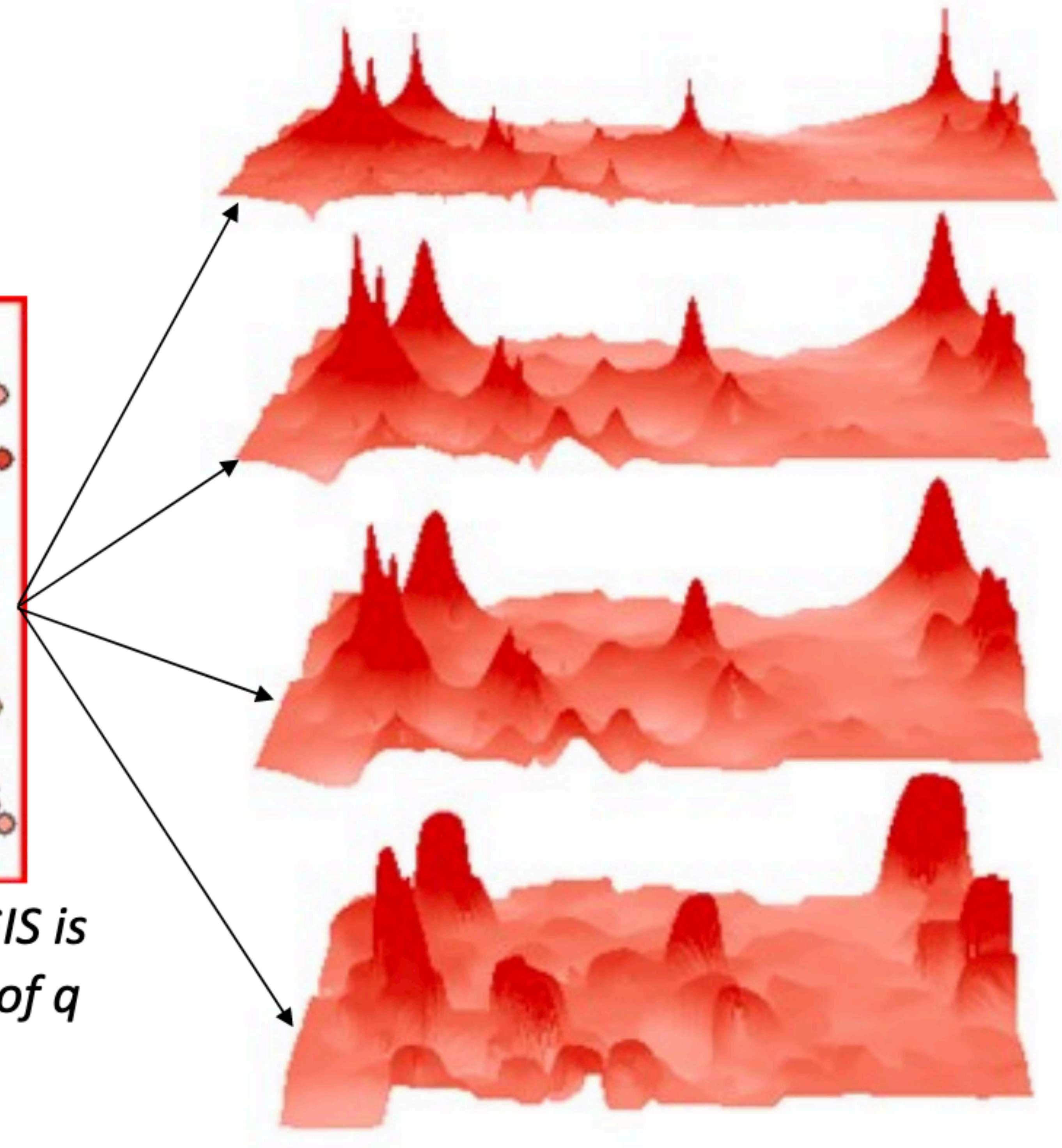

Predicted precip
10 to 15
15 to 20
20 to 25
25 to 30
30 to 35
35 to 40
40 to 45
45 to 50
50 to 55

$$q = 15$$

Examples of IDW with Different *q*'s

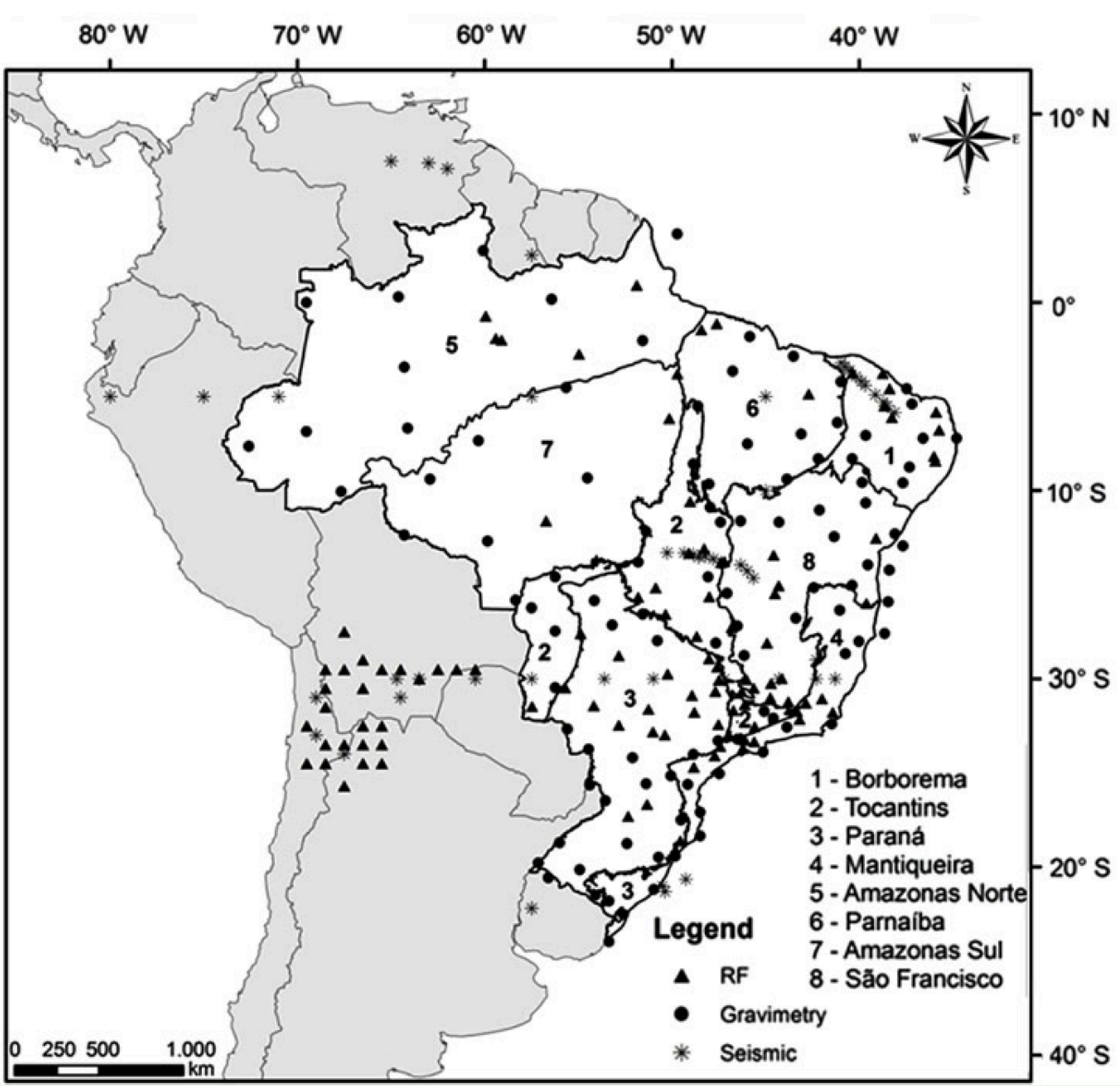
Gold concentrations at locations in western PA

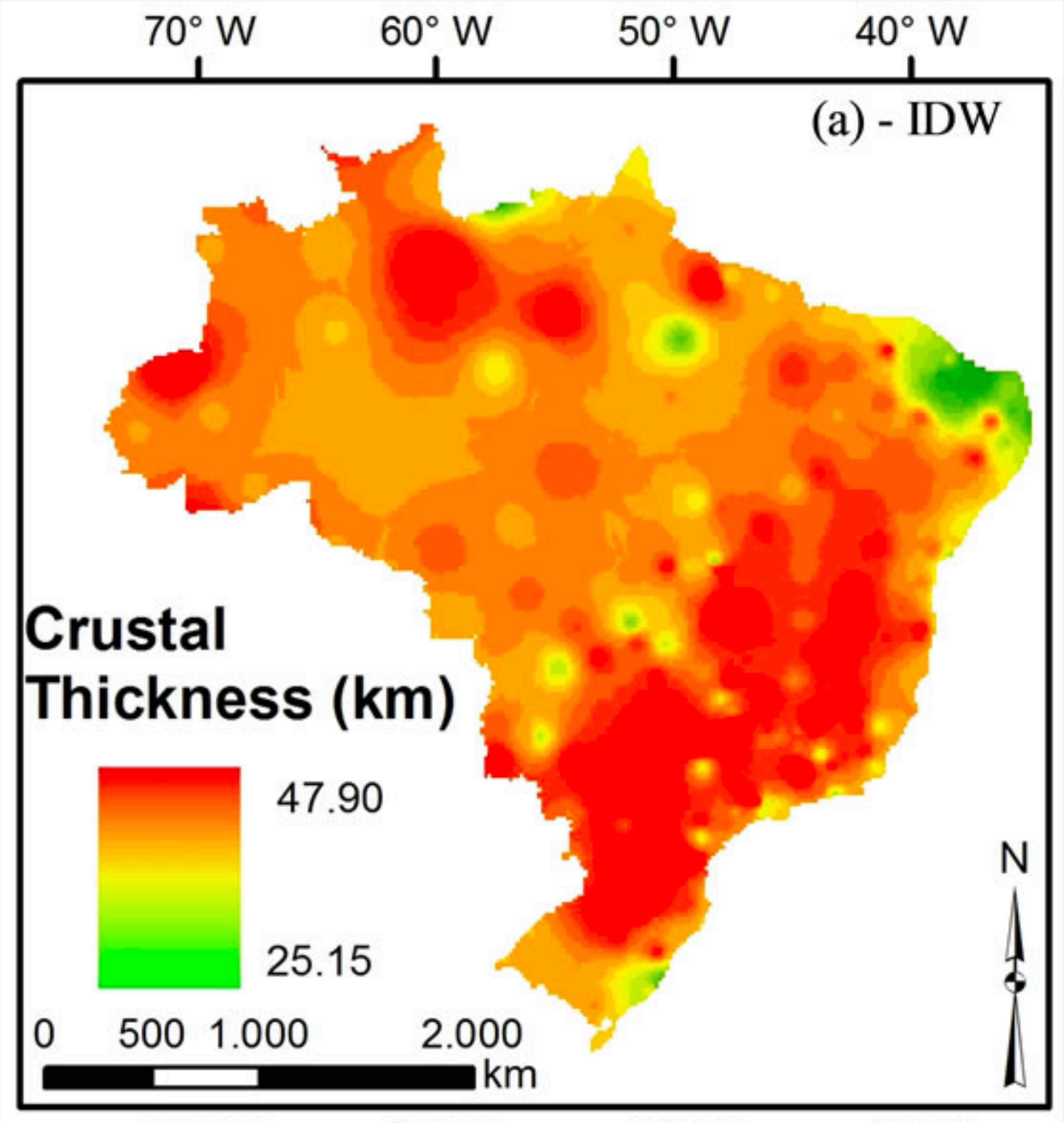




The Geostatistical Analyst of ArcGIS is able to tell you the optimal value of q by seeing which one yields the minimum RMSE. (Here, it is q=1).

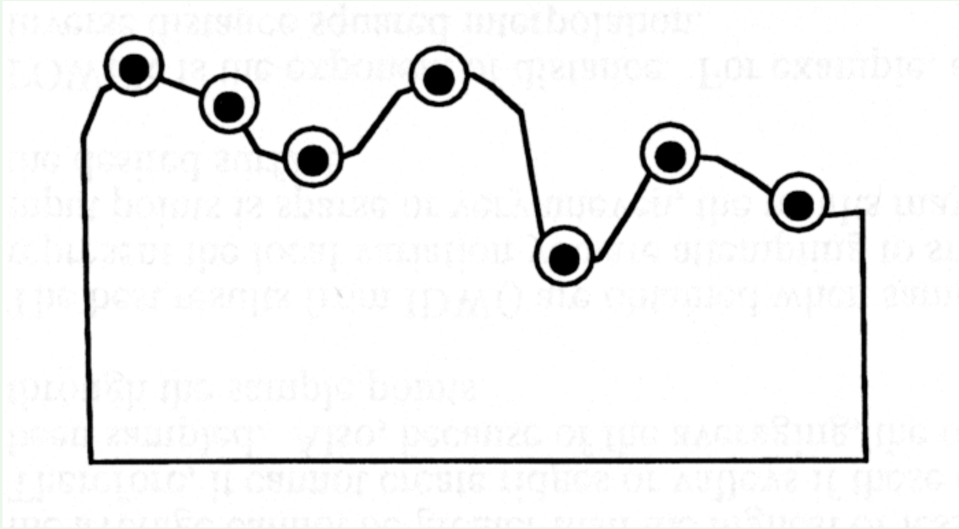
- Larger *q*'s (i.e., power to which distance is raised) yield smoother surfaces
- Food for thought: What happens when *q* is set to 0?


80° W
70° W
60° W
50° W
40° W
10° N
0°
10° S
20° S
30° S
40° S
N
S
E
W
5
7
6
1
2
8
2
3
4
2
3
1 - Borborema
2 - Tocantins
3 - Paraná
4 - Mantiqueira
5 - Amazonas Norte
6 - Parnaíba
7 - Amazonas Sul
8 - São Francisco
Legend
RF
Gravimetry
Seismic
0 250 500 1.000
km

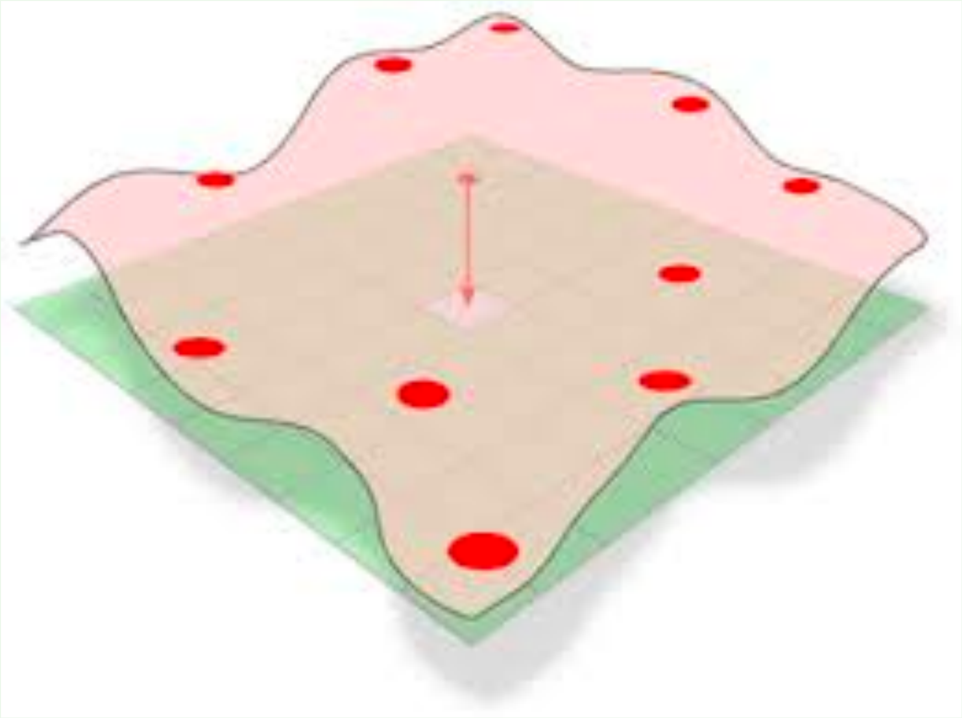

70° W
60° W
50° W
40° W
(a) - IDW
Crustal Thickness (km)
47.90
25.15
0 500 1.000 2.000
km
N

Spline

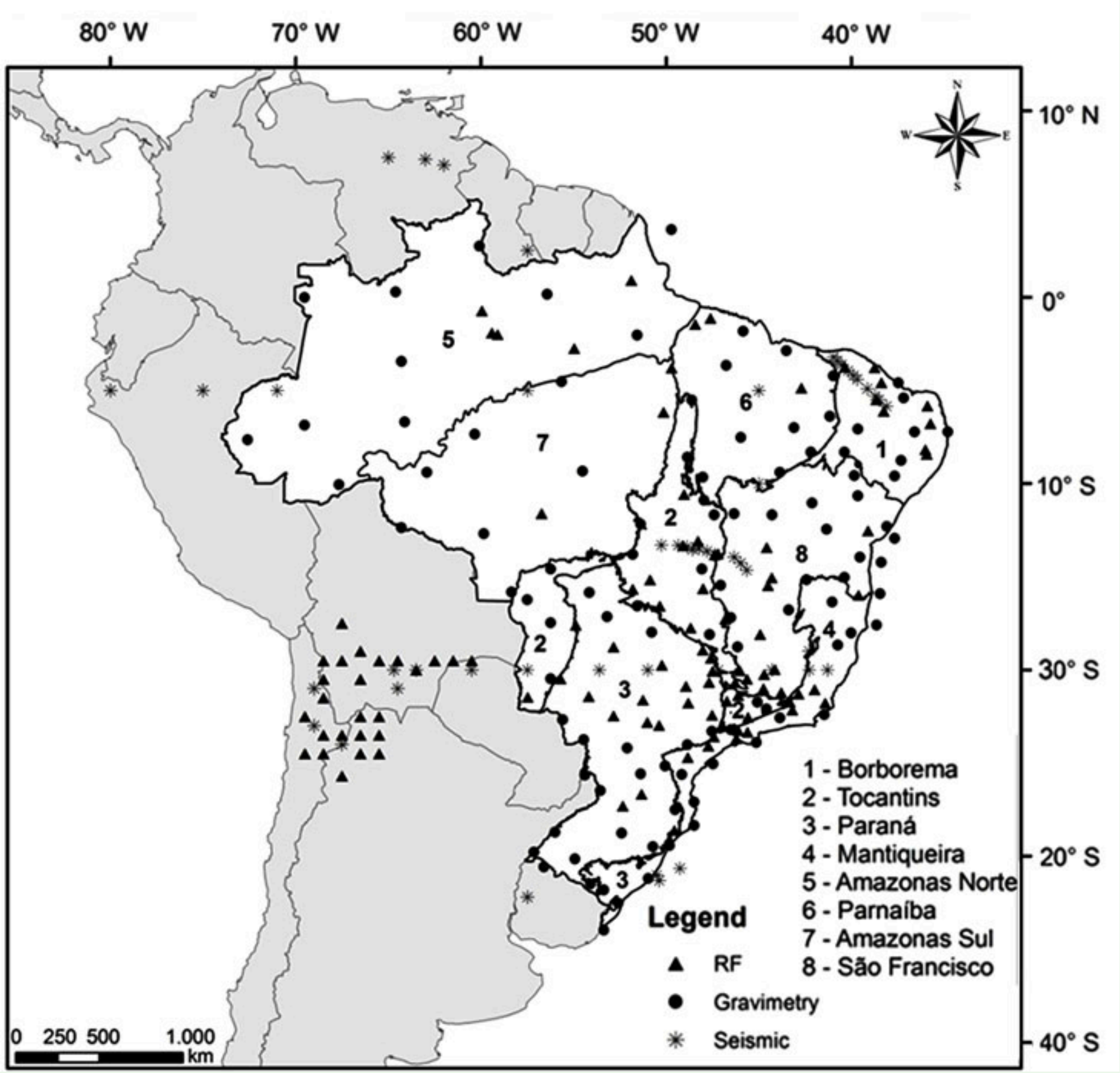
- The spline interpolation fits a minimum-curvature surface **through** a specified number of nearby sample points.

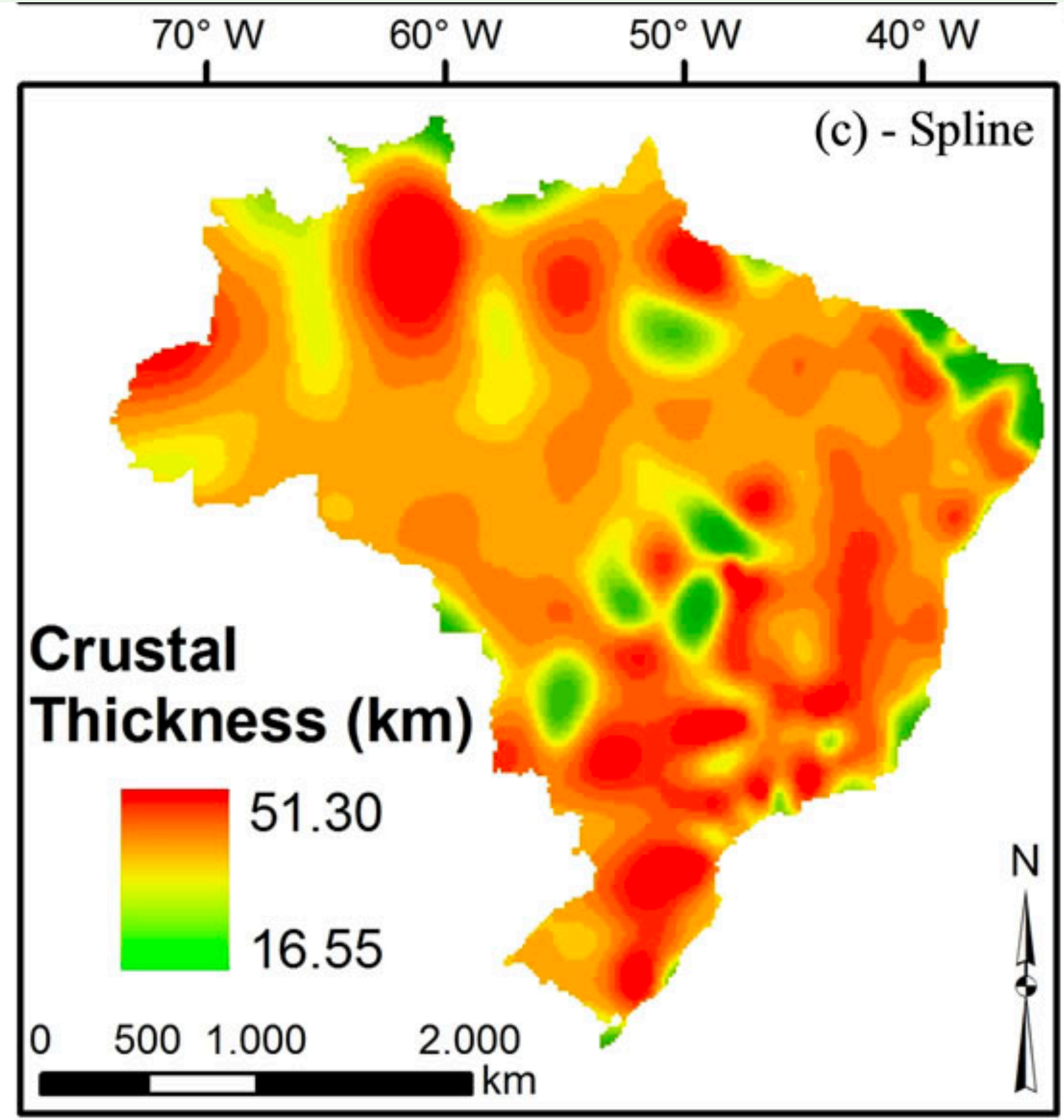


- Conceptually, it is like bending a sheet of rubber to pass through the points, while minimizing the total curvature of the surface.



- This method is best for gently varying surfaces (water table heights or pollution concentrations.


80° W
70° W
60° W
50° W
40° W
10° N
0°
10° S
20° S
30° S
40° S
N
S
W
E
1
2
3
4
5
6
7
8
Legend
RF
Gravimetry
Seismic
1 - Borborema
2 - Tocantins
3 - Paraná
4 - Mantiqueira
5 - Amazonas Norte
6 - Parnaíba
7 - Amazonas Sul
8 - São Francisco
0 250 500 1.000
km



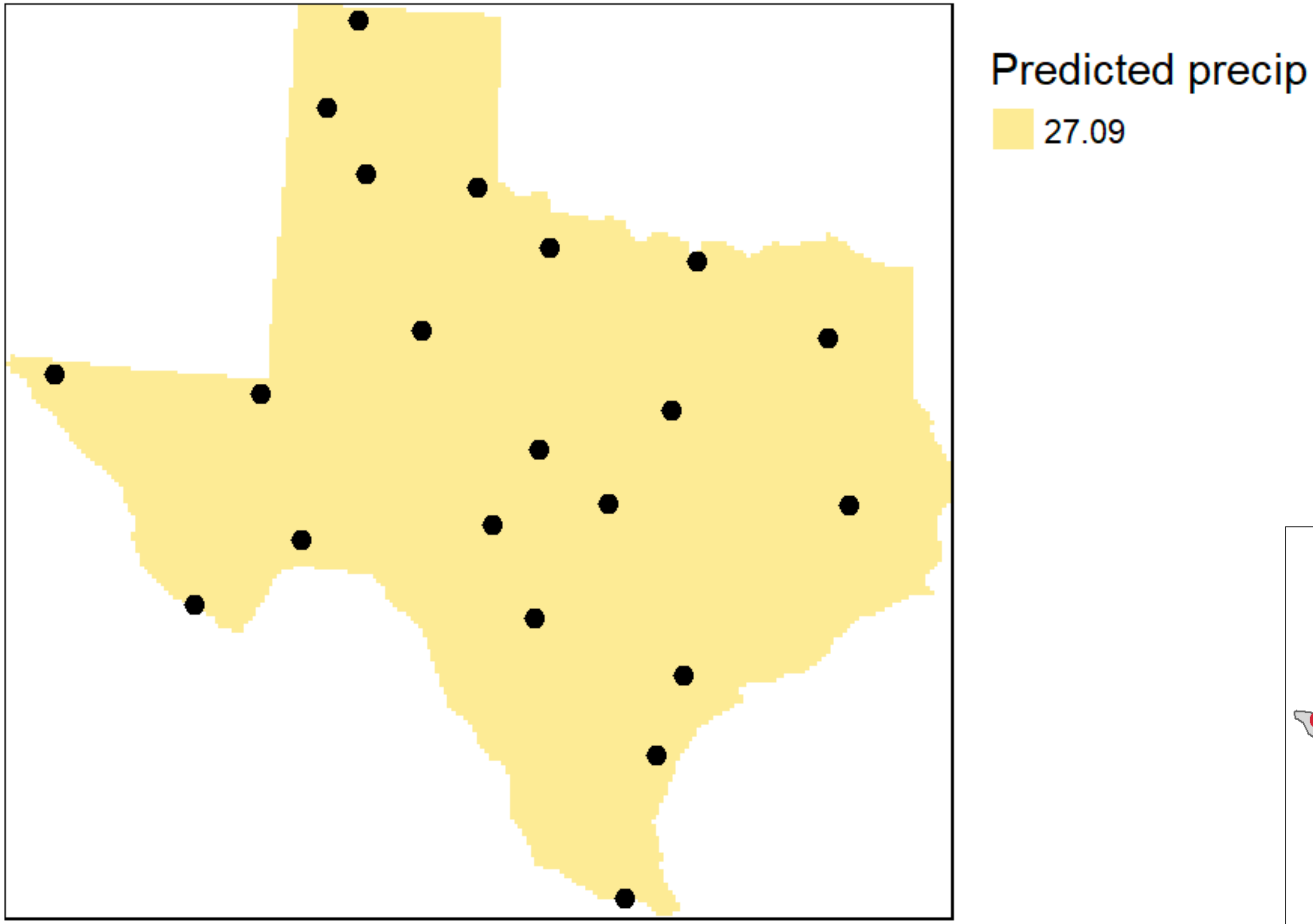
70° W
60° W
50° W
40° W
(c) - Spline
Crustal Thickness (km)
51.30
16.55
N
0
500
1.000
2.000
km

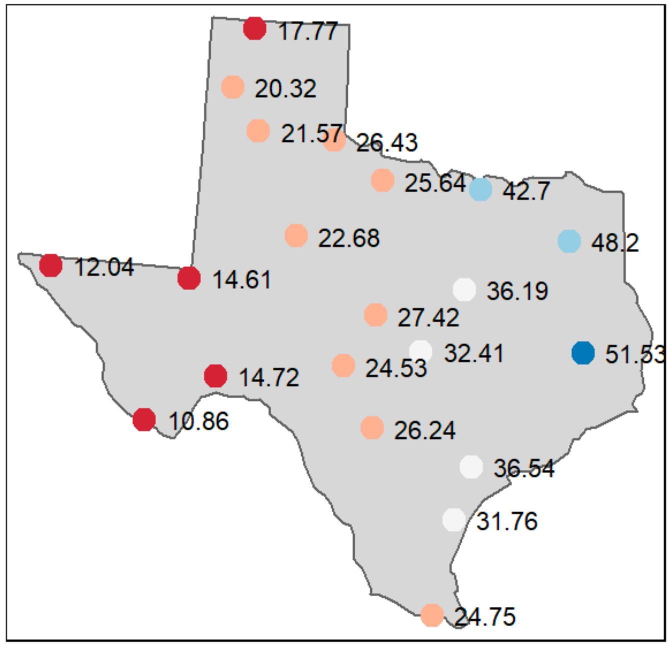
Statistical Methods

Global Interpolation Methods

- Global interpolation
  - use all the samples to estimate at an unsampled location
- Trend surface (polynomial equations)
  - A first-order trend surface fits a 3D plane to all the samples
    - $Z = b_0 + b_1x + b_2y$
  - Higher order polynomial functions could fit a more complex surface to the samples
    - $Z = b_0 + b_1x + b_2y + b_3xy + b_4x^2 + b_5y^2$
  - Regression with coordinates or their functions

Zero-order polynomial

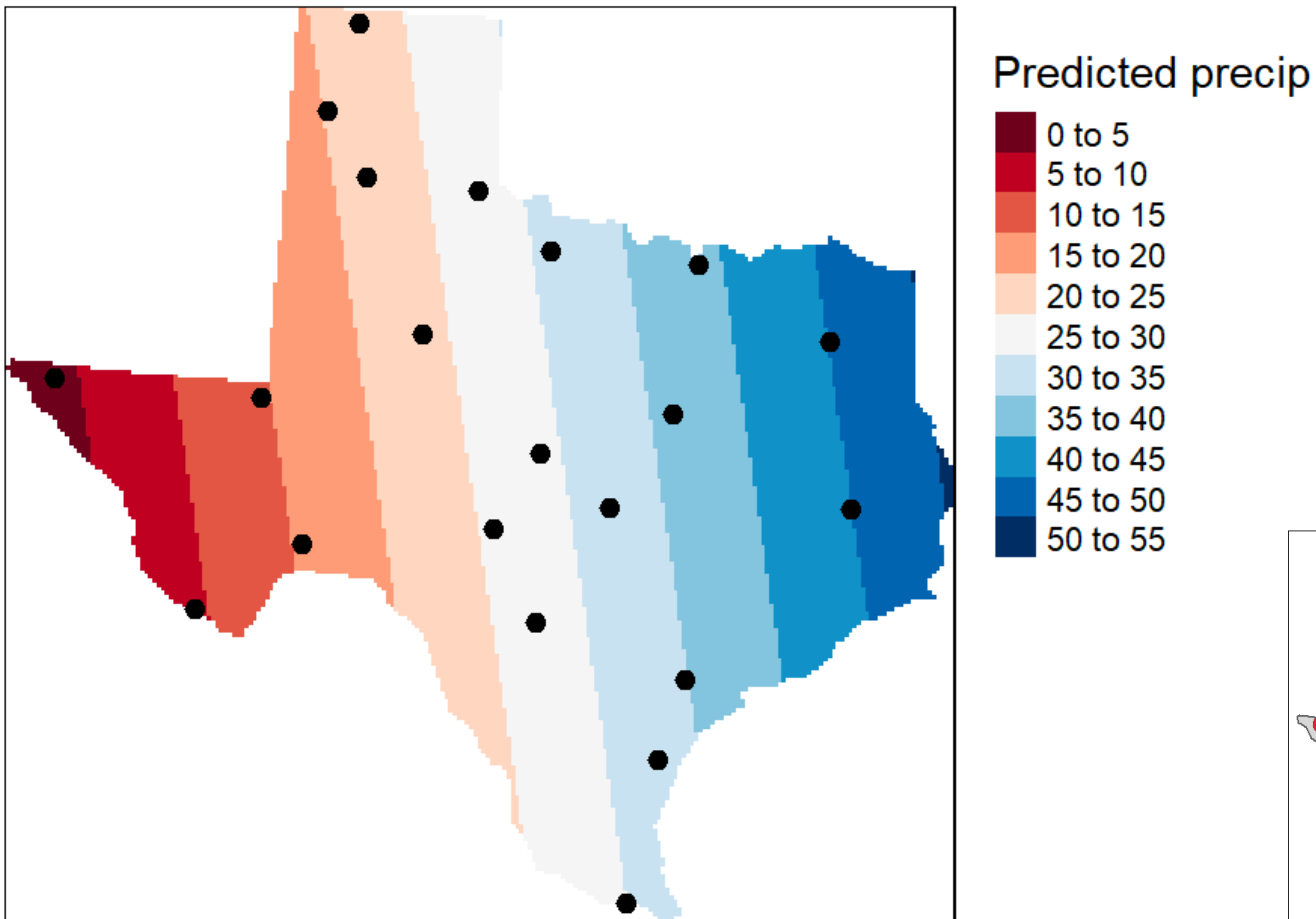


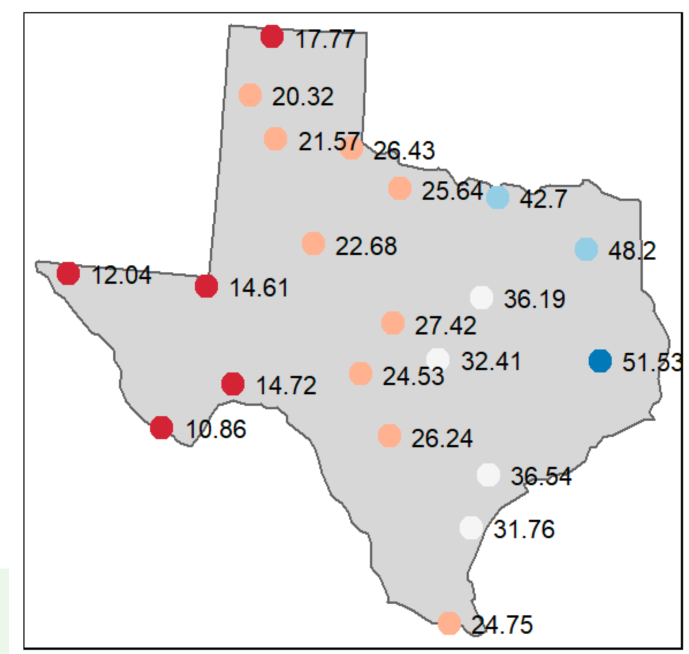


$$z = b_0 + b_1 x^0 + b_2 y^0$$

$$z = b_0$$

First-order polynomial

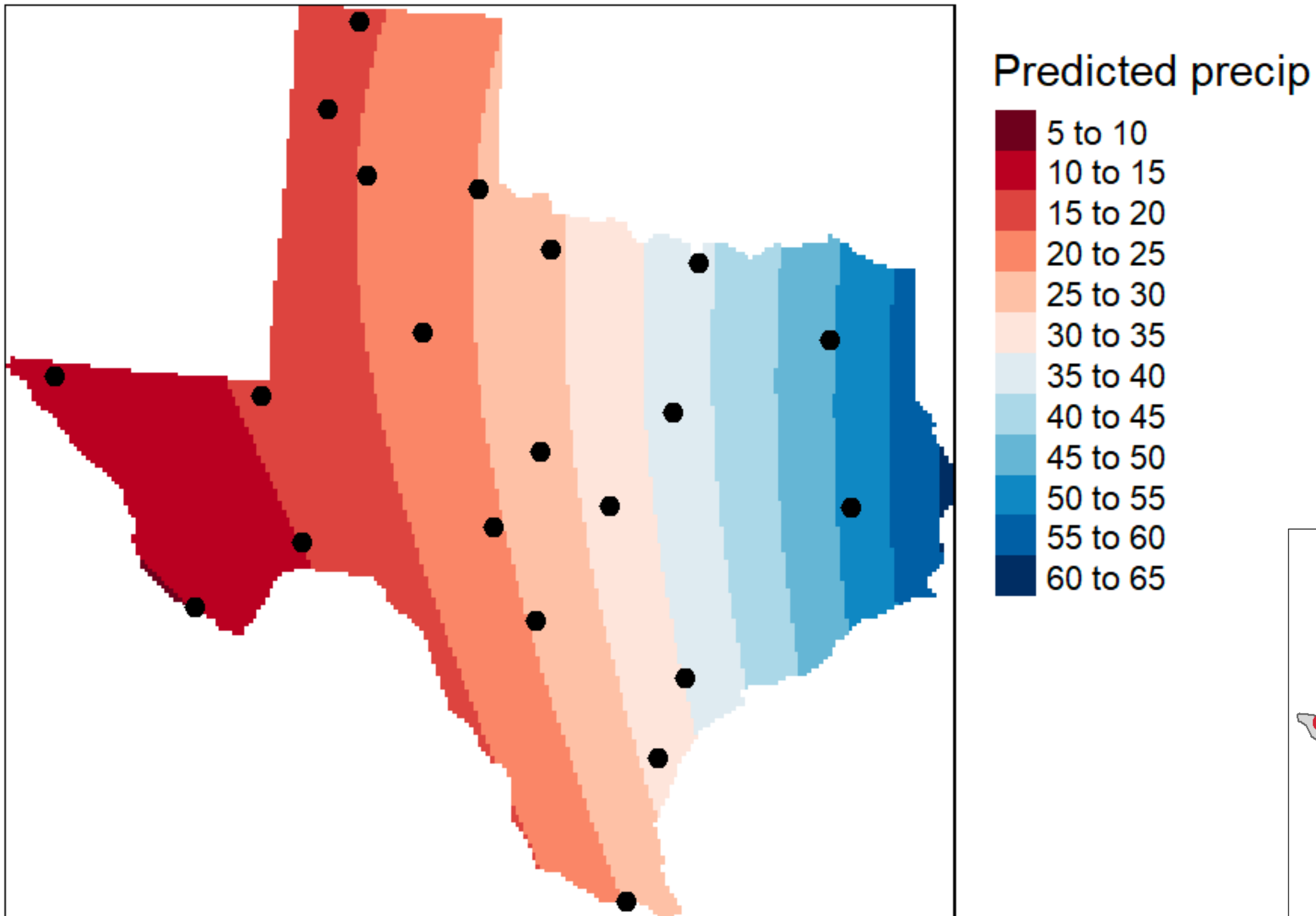


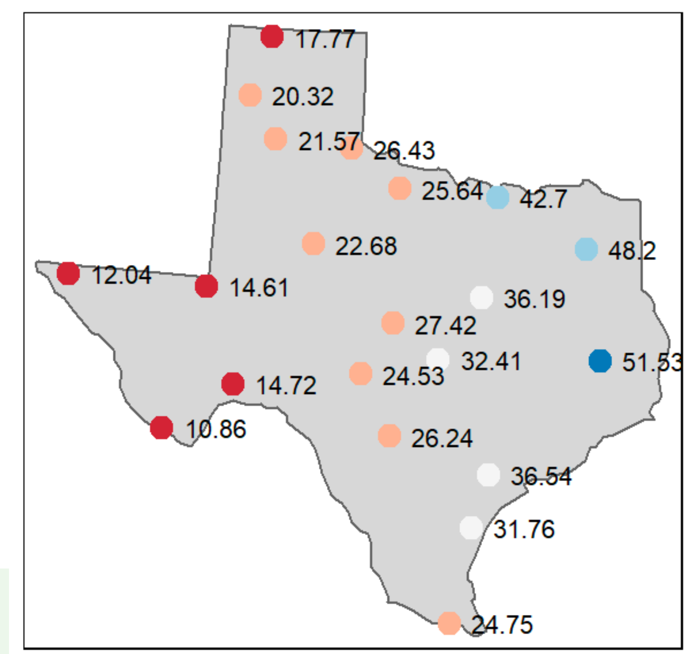


$$z = b_0 + b_1 x^1 + b_2 y^1$$

$$z = b_0 + b_1 x + b_2 y$$

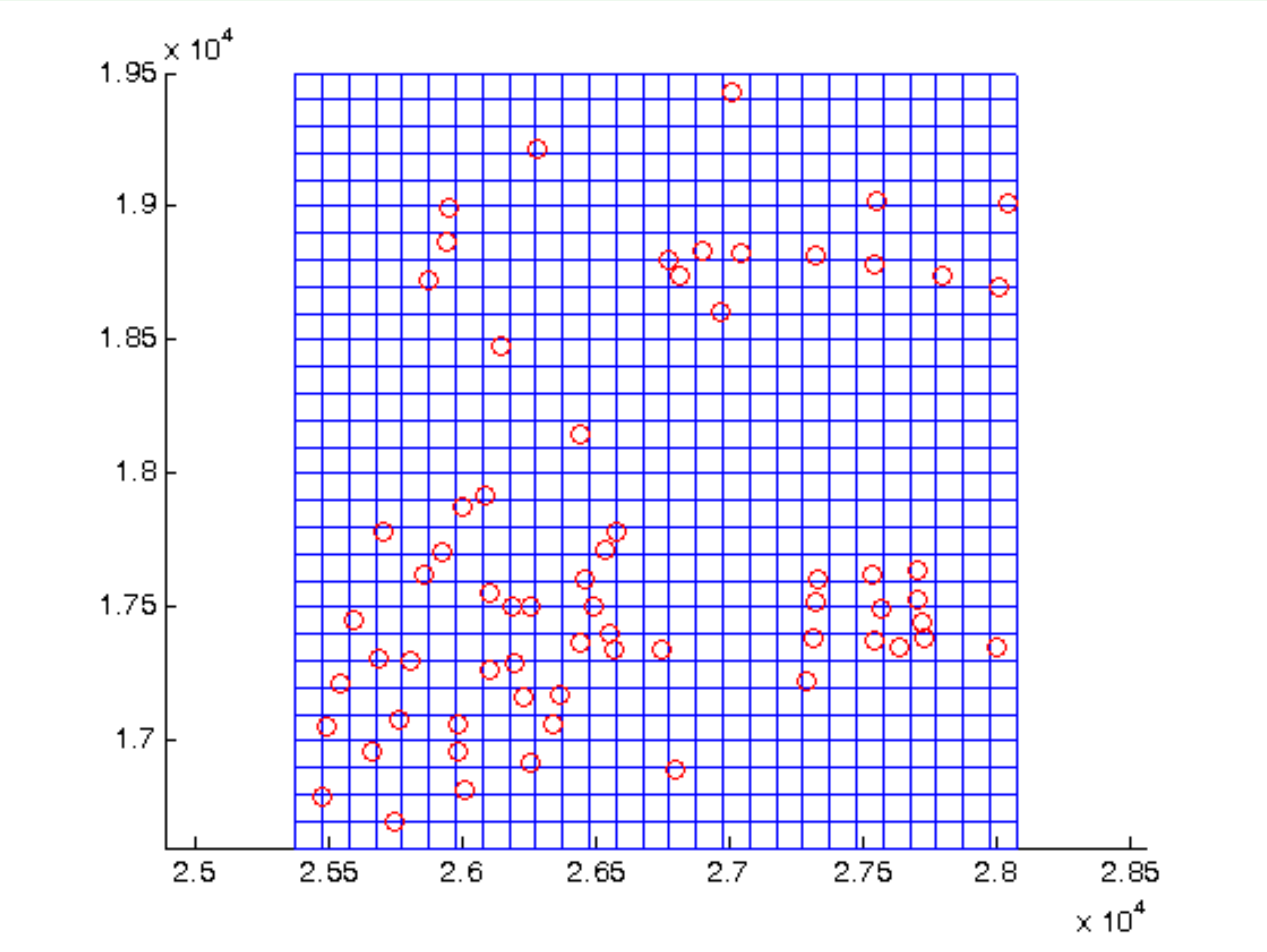
Second-order polynomial



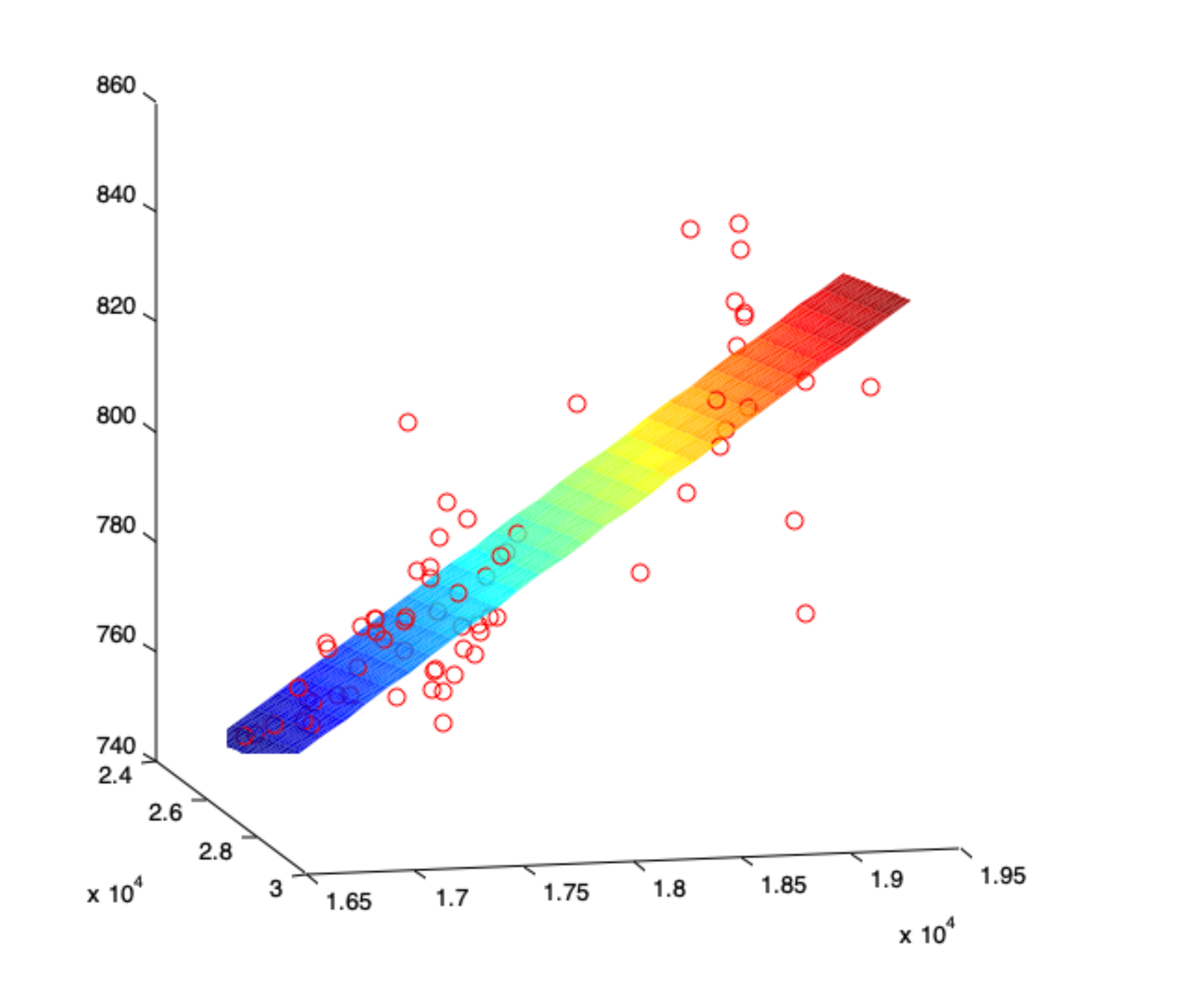


$$z = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy$$

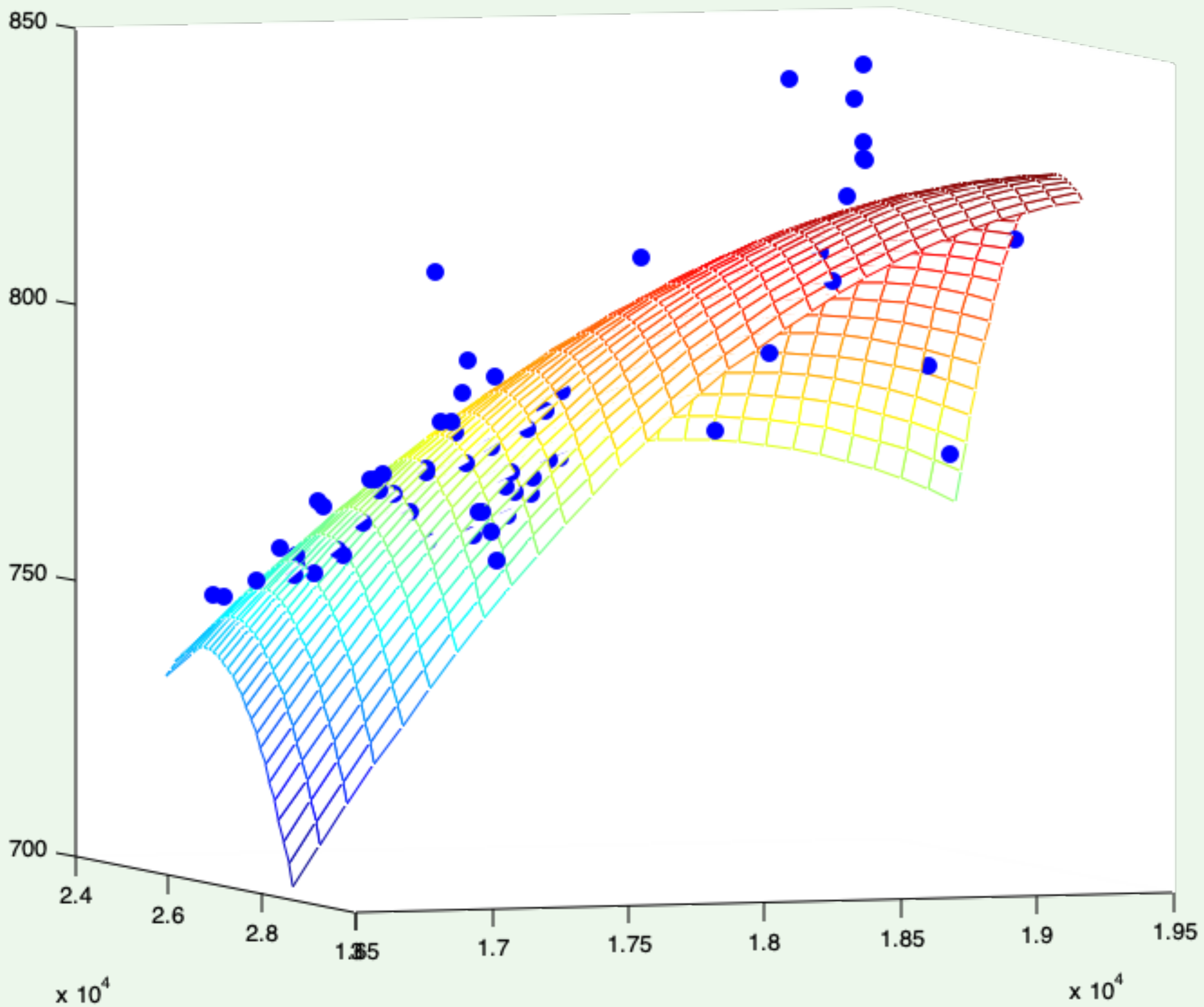
Interpolate Ground Water Table Elevation

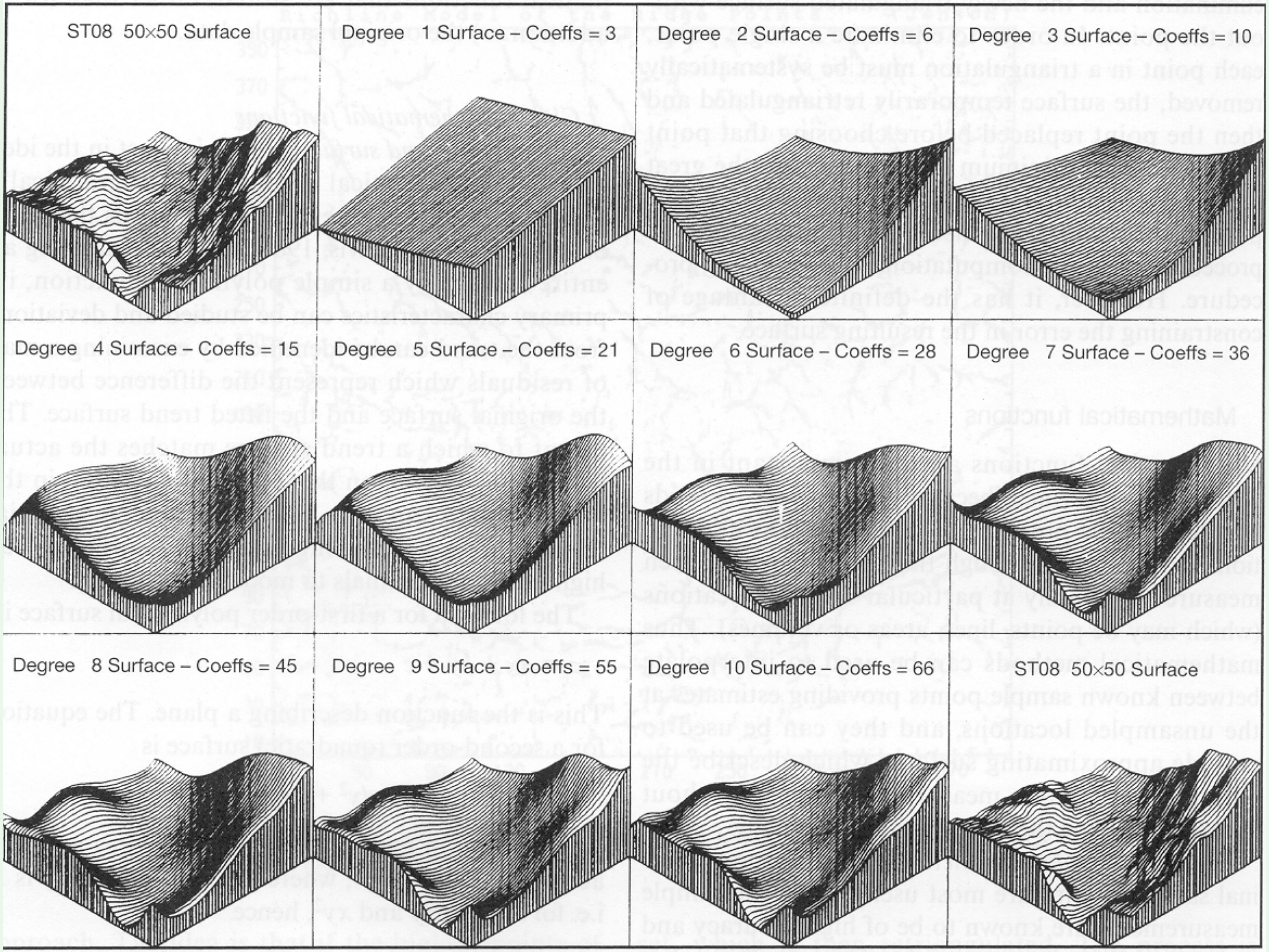


First-order polynomial



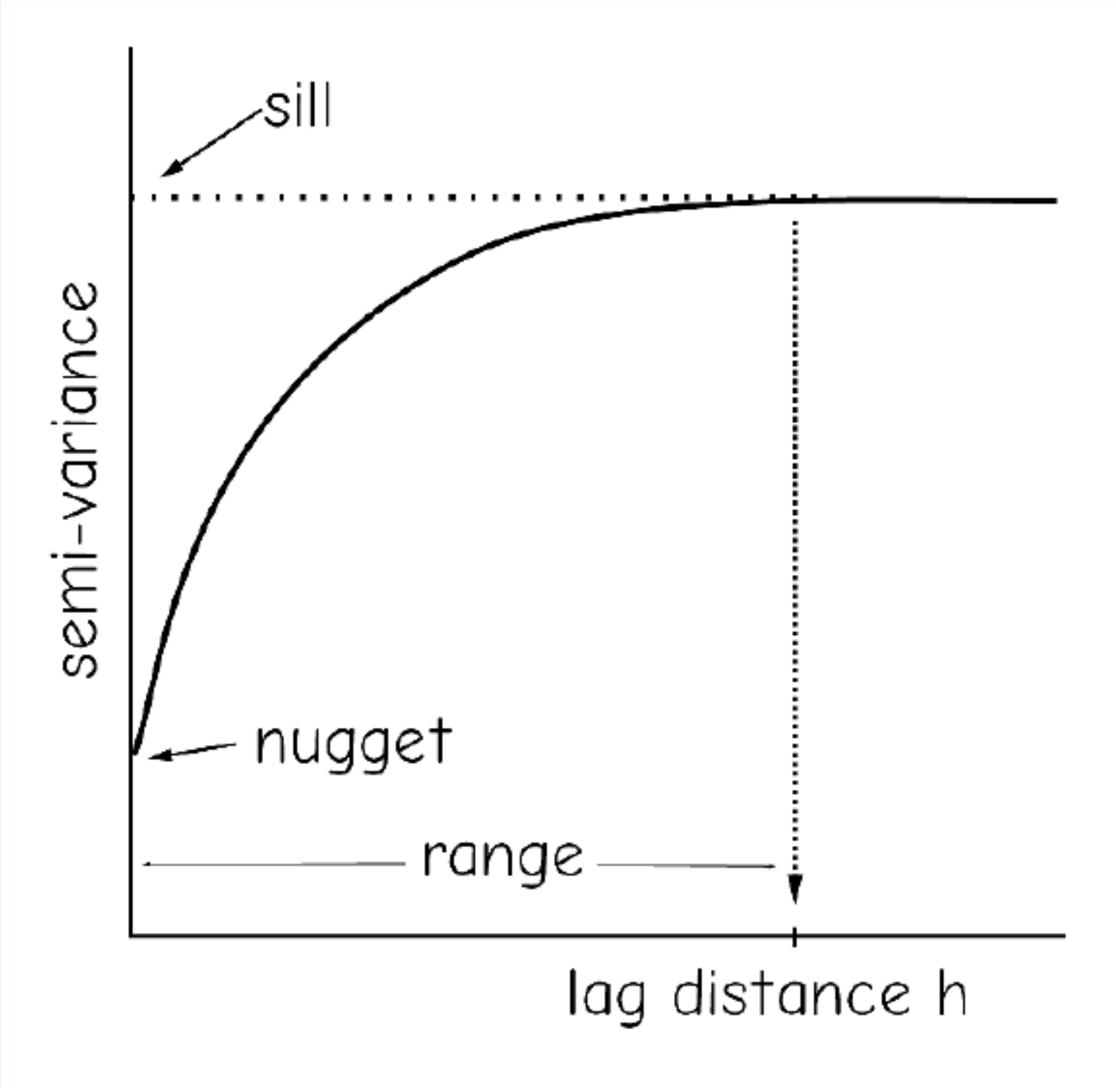
Second-order polynomial

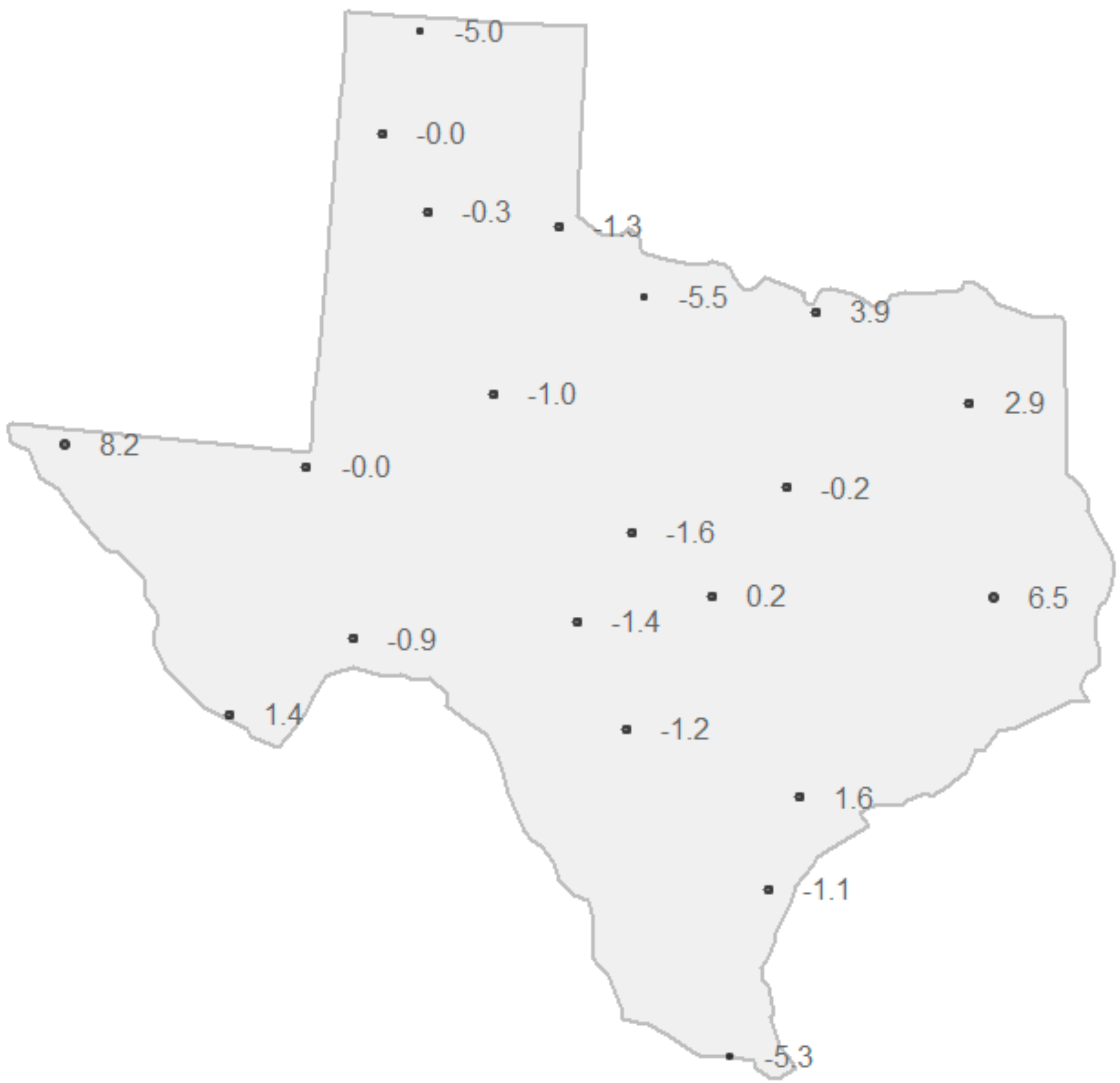



ST08 50×50 Surface
Degree 1 Surface – Coeffs = 3
Degree 2 Surface – Coeffs = 6
Degree 3 Surface – Coeffs = 10
Degree 4 Surface – Coeffs = 15
Degree 5 Surface – Coeffs = 21
Degree 6 Surface – Coeffs = 28
Degree 7 Surface – Coeffs = 36
Degree 8 Surface – Coeffs = 45
Degree 9 Surface – Coeffs = 55
Degree 10 Surface – Coeffs = 66
ST08 50×50 Surface

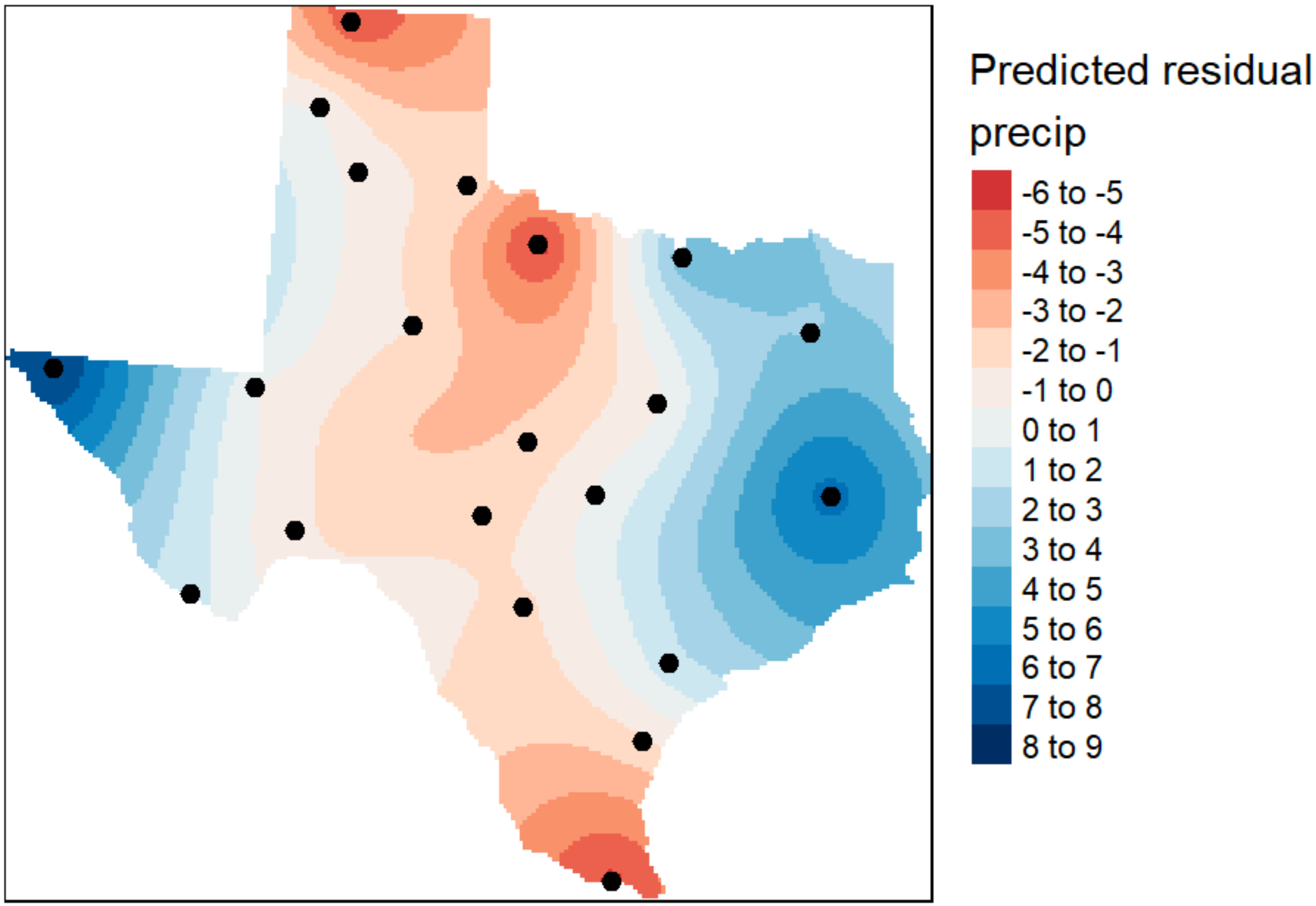
Kriging

- Weights determined based on spatial autocorrelation
- Search radius can be estimated based variogram
- Uses both global and local data
  - Variogram uses all the samples
  - Weights use local samples

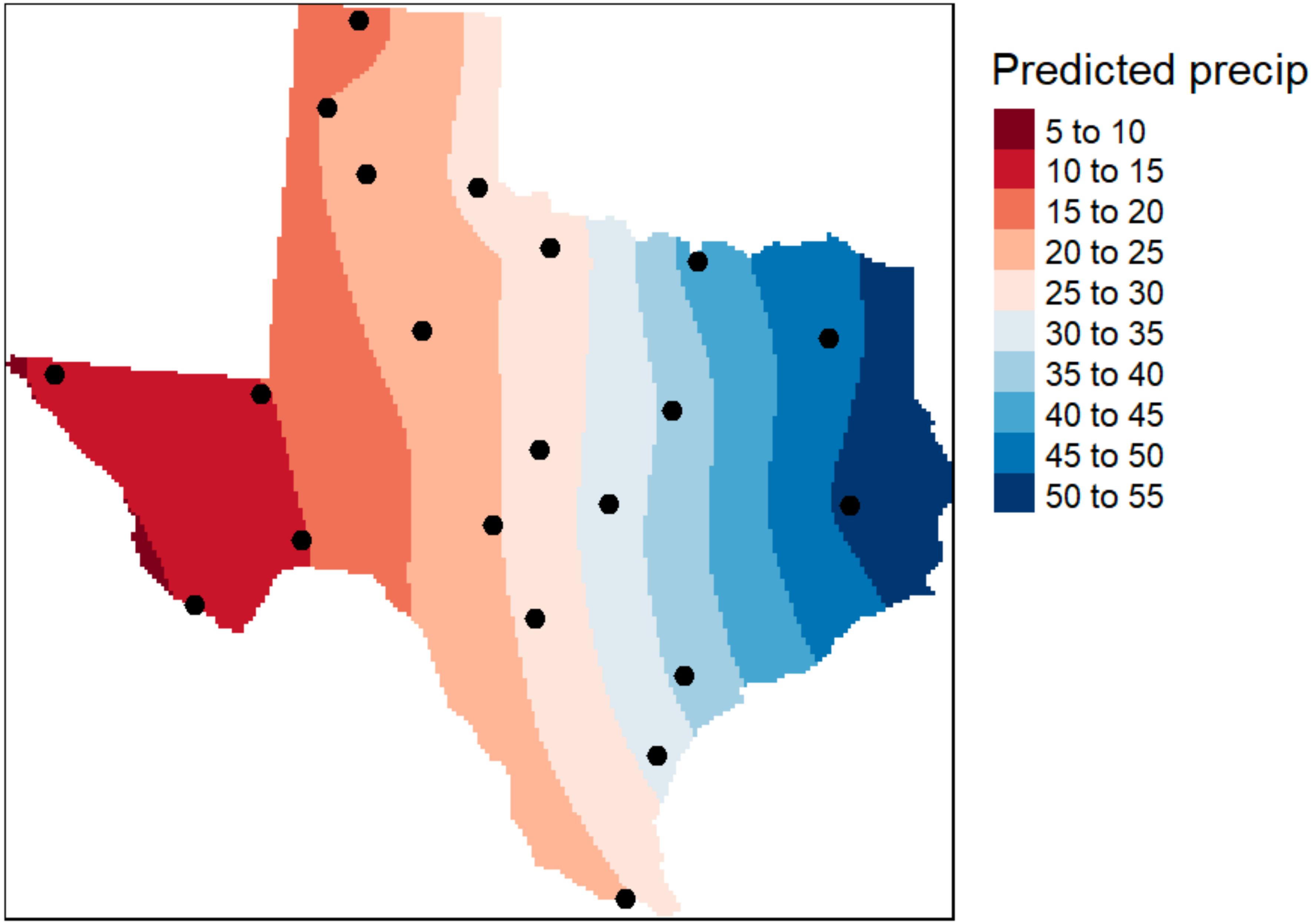




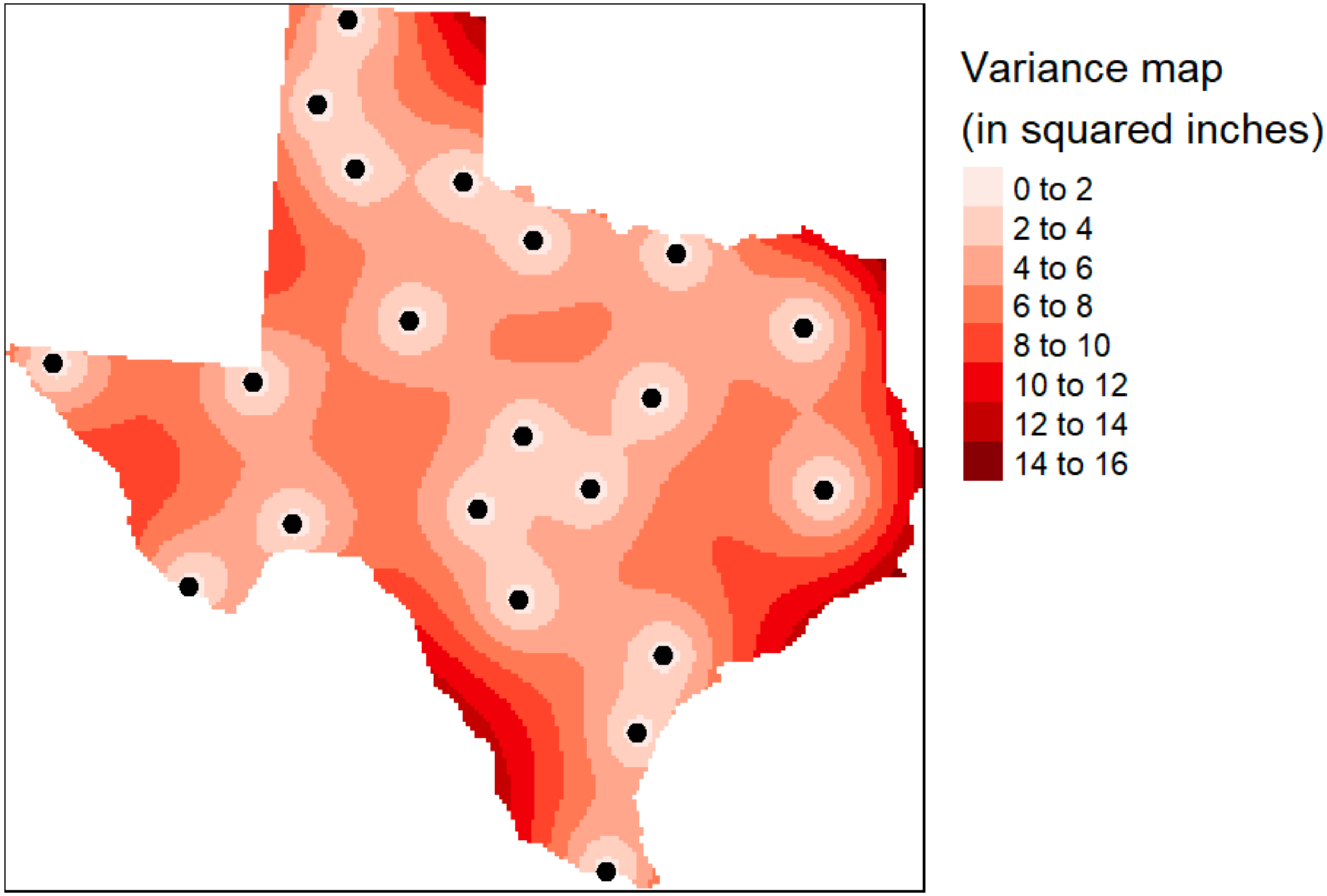
Detrended precipitation



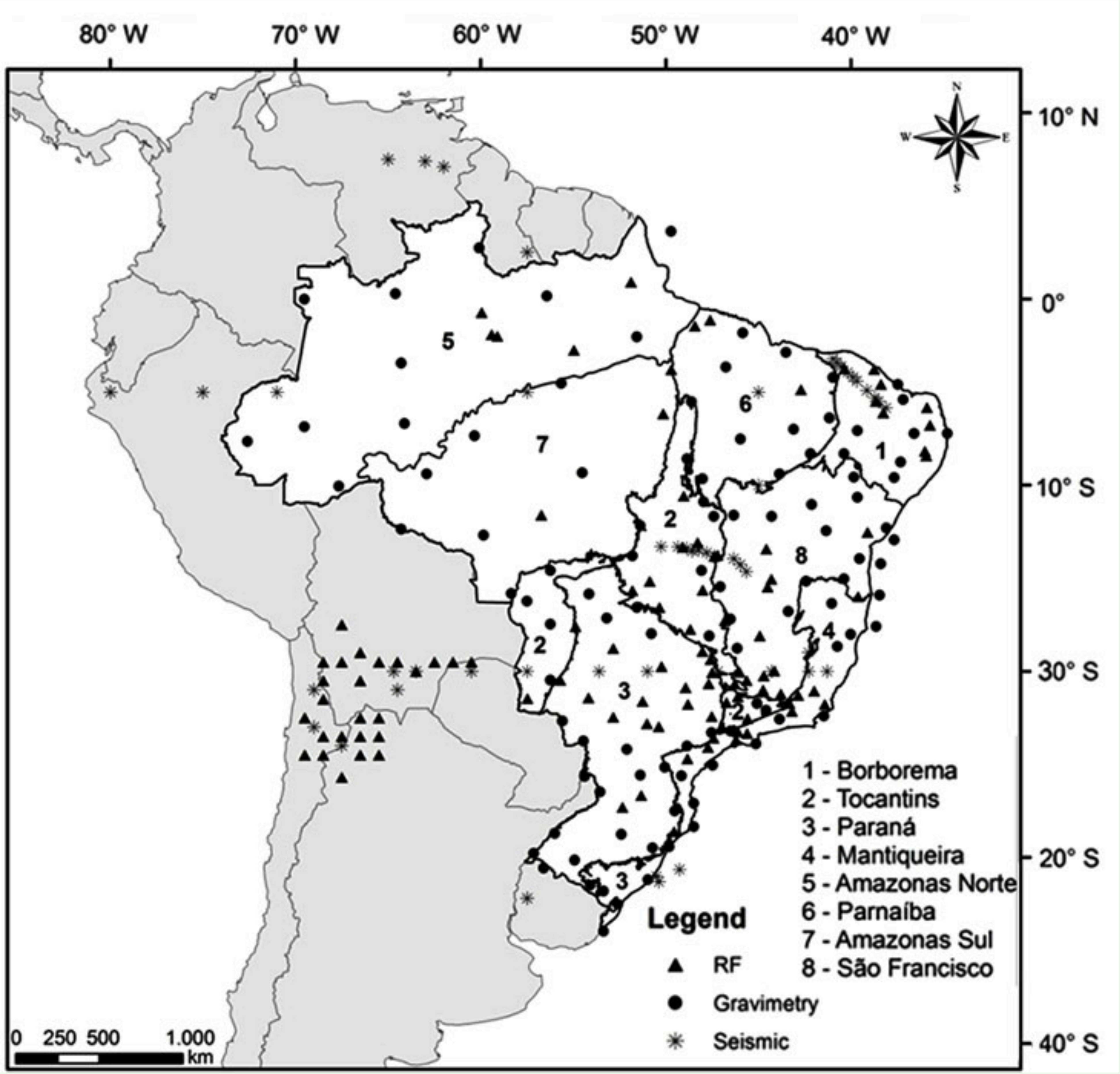
Predicted precipitation from local factors

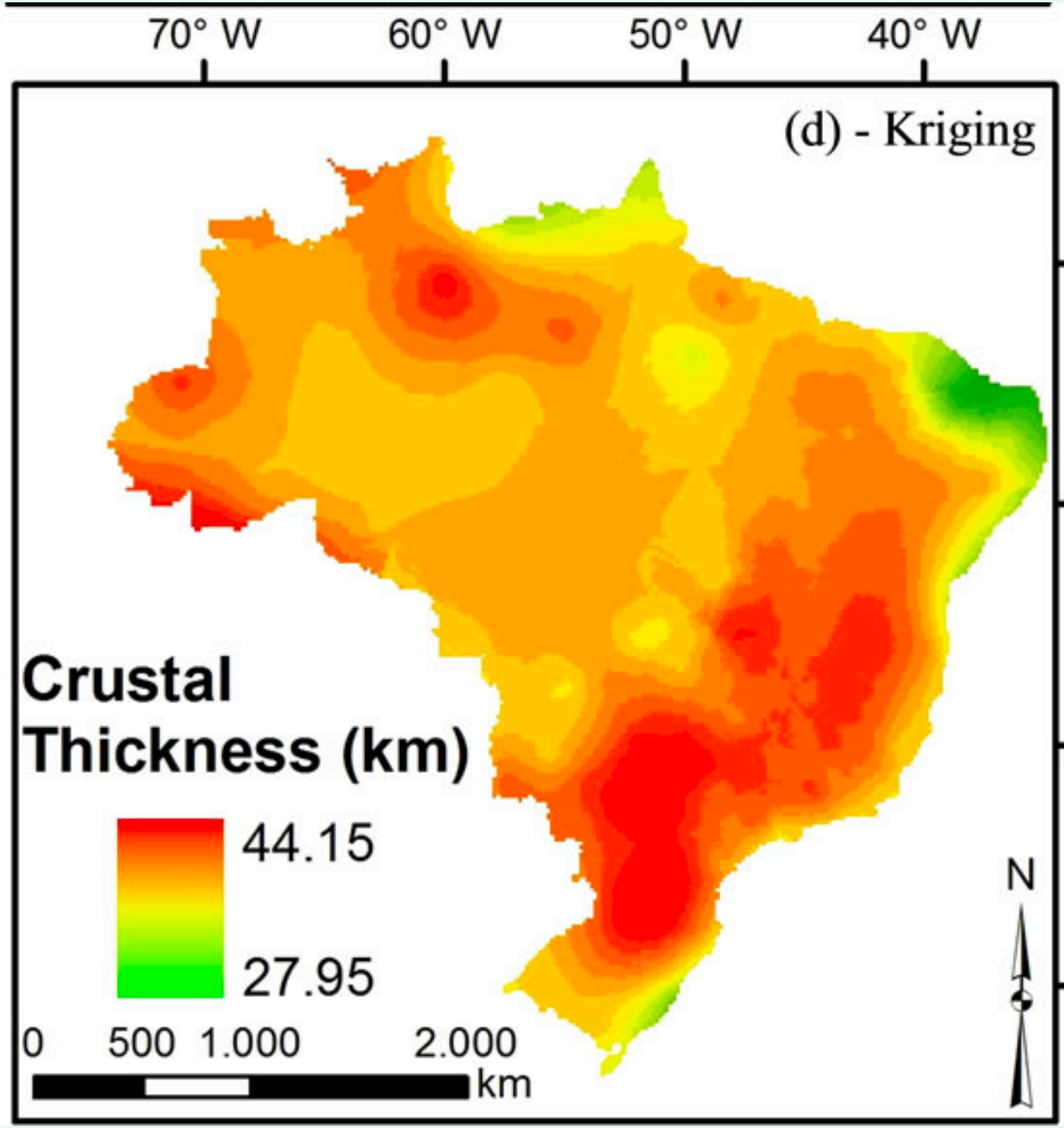


Predicted precipitation from local factors combined with trend



Uncertainty of interpolated values


80° W
70° W
60° W
50° W
40° W
10° N
0°
10° S
20° S
30° S
40° S
5
6
7
1
2
8
4
3
2
3
1 - Borborema
2 - Tocantins
3 - Paraná
4 - Mantiqueira
5 - Amazonas Norte
6 - Parnaíba
7 - Amazonas Sul
8 - São Francisco
Legend
RF
Gravimetry
Seismic
0 250 500 1.000
km

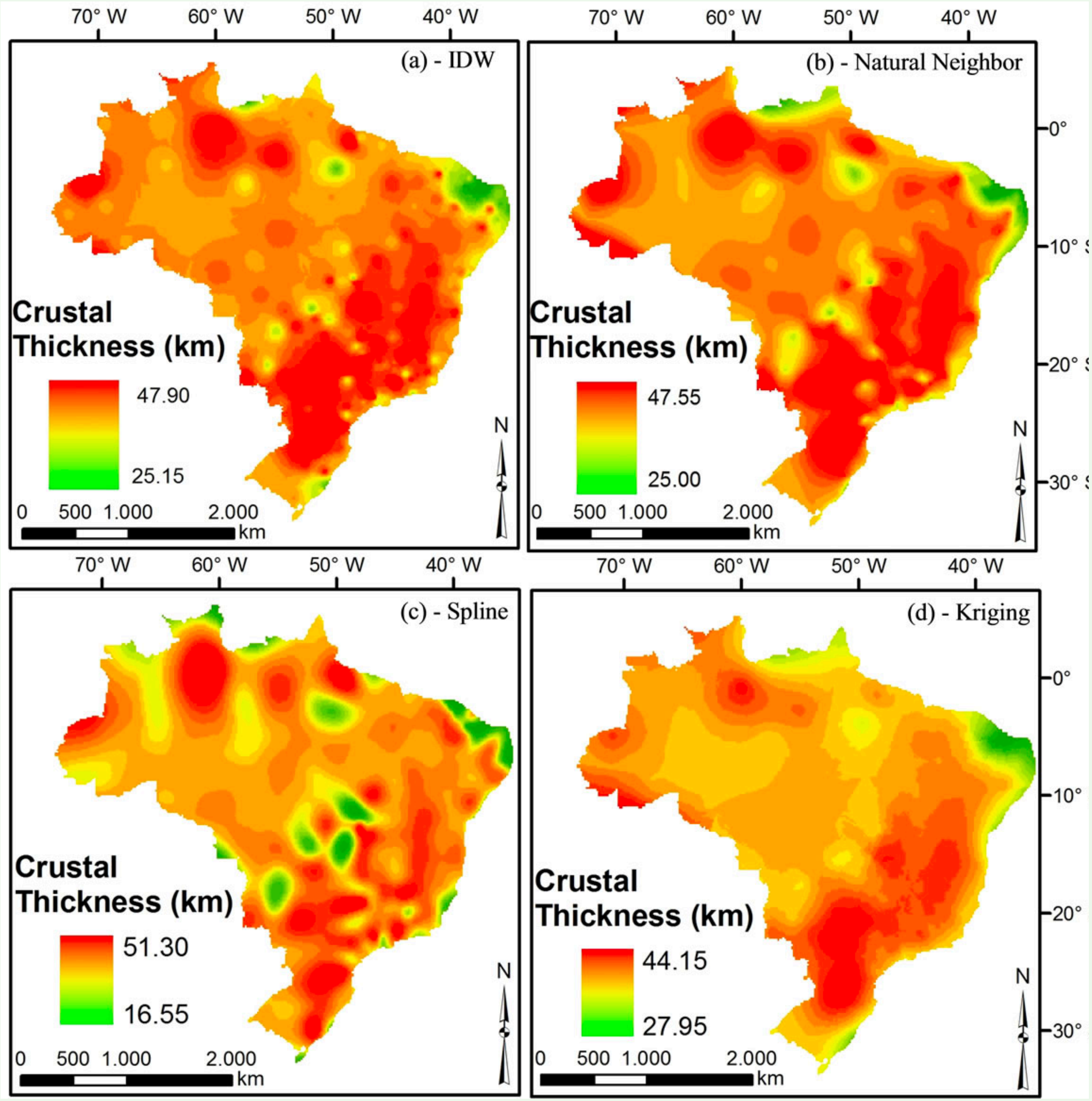

70° W
60° W
50° W
40° W
(d) - Kriging
Crustal
Thickness (km)
44.15
27.95
0
500
1.000
2.000
km
N

Accuracy Evaluation

- Accuracy is typically judged by withholding some sample points
- Divide sample points into two groups
  - One for interpolation and another for validation
  - Waste precious sample data
- Cross validation
  - Leave one sample point for validation
  - Use all the sample points left
  - Computational intensive

Interpolation Method Summary

- Methods differ in weighting methods and the number of samples used
- Each produces different results (with the same data)
- No single method has been shown to be more accurate in every application



(a) - IDW
(b) - Natural Neighbor
(c) - Spline
(d) - Kriging
Crustal Thickness (km)
47.90
25.15
47.55
25.00
51.30
16.55
44.15
27.95
70° W
60° W
50° W
40° W
0°
10° S
20° S
30° S
0 500 1.000 2.000 km
N