

Describing and Analyzing Fields

Outline

- Quick Reminder
- Representing Fields
- Spatial Interpolation

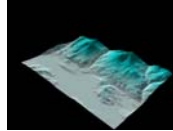
Quick Reminder:

- **Field view:** the world consists of attributes that are continuously variable and measurable across space (e.g. elevation, population density)
- **Object view:** the world consists point, line, and area objects each with a bundle of properties (attributes).

Fields

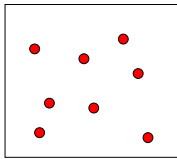
- **Assumptions**
 - **Continuity**: for every location s_i , there is a measurable z_i at the same place
 - **Single-valued**: for each location, there is only one value of z

$$z_i = f(s_i) = f(x_i, y_i)$$



Representing Fields

- **Step1: Sampling the Real Surface**



Sampling points

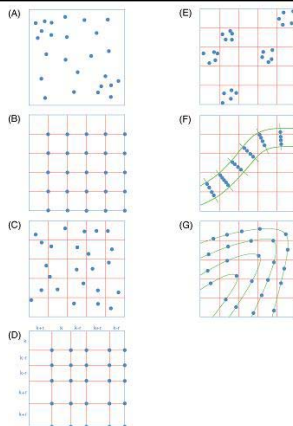
Sampling: The data constitute a **sample** of the underlying, continuous field.

For example:

Weather station (e.g. temperature)
Pollution monitor station (e.g. CO₂ concentration)

- **Types of samples**
 - Random samples
 - Systematic samples
 - Clustered samples

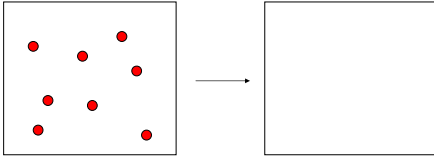
- Random
- Systematic
- Systematic with local random allocation
- Systematic with random variation in grid space
- Clustered
- Transect
- Contour



Representing Fields

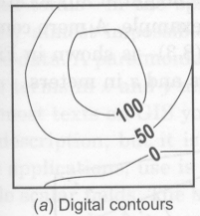
• Step2: Continuous Surface Description

Interpolation: **reconstruct** the underlying continuous field of data from the limited evidence of the control points (samples)



Field Data Models

• Digital contours



- ✓ The most common and the oldest method for terrain mapping
- ✓ Contour lines connect points of **equal value** (e.g. elevation)
- ✓ Contour **interval** represents the vertical distance between contour lines
- ✓ The arrangement and pattern of contour lines reflect the topography

(a) Digital contours

Field Data Models

• Digital contours

Advantages: easy to acquire and display

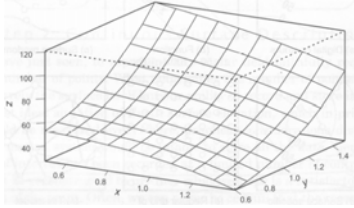
Disadvantages

- 1) The accuracy attainable depends on the scale of the original map, together with the spatial and vertical accuracy of the source map contours
- 2) All information on surface detail falling between contours is lost
- 3) The method over-samples steep slopes with many contours
- 4) Many processing operations are difficult to automate

Field Data Models

• Mathematical Functions

$$z_i = f(\mathbf{s}_i) = f(x_i, y_i) = -12x_i^3 + 10x_i^2y_i - 14x_iy_i^2 + 25y_i^3 + 50$$



Field Data Models

• Mathematical Functions

Advantages:

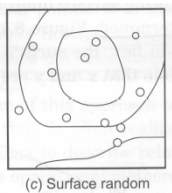
- 1) Compact (with only several parameters)
- 2) Some processing operations (e.g. slope and curvature calculation), are easily carried out

Disadvantages:

- 1) Arbitrary chosen functions are often used
- 2) Many functions give values which are very unlikely or even impossible

Field Data Models

• Point systems

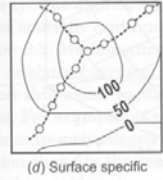


1) **Surface random samples**: the control point locations are chosen without reference to the shape of the surface being sampled.

✓ The result is an irregular scatter of control points that may or may not capture significant features of the surface relief.

Field Data Models

• Point systems

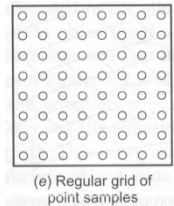


2) **Surface specific sampling:** points are located at places judged to be important in defining the surface detail (e.g. peaks, saddle points)

✓ The surface specific points provide more information about the structural properties of the surface.

Field Data Models

• Point systems



3) **Grid sampling:** field heights across a regular grid of (x,y) coordinates are recorded

Advantages:

- 1) A uniform density of point data, easy to be integrated with raster data.
- 2) Easy to calculate and map surface properties, such as slope and aspect

Disadvantages

- 1) Large data storage and difficulties associated with sampling

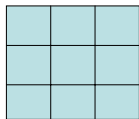
Field Data Models

• Digital Elevation Model (DEM)

- represents a regular array of elevation points
- Can be obtained from U.S.G.S.
- Alternative sources for DEMs come from satellite images, radar data, and LIDAR (light detection and ranging) data



Point-based

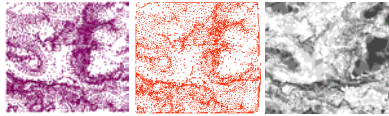
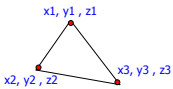


Raster-based

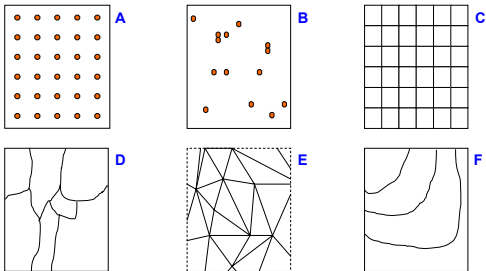
Field Data Models

• Triangulated Irregular Networks (TIN)

- approximates the land surface with a series of **non-overlapping** triangles
- z values along with x, y -coordinates are stored at nodes that make up triangles.
- TIN is based on an irregular distribution of elevation points

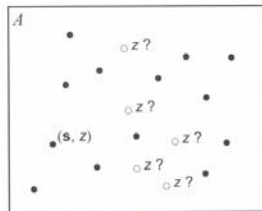


Summary I



Spatial Interpolation

- Spatial interpolation is the **prediction** of exact values of attributes at un-sampled locations from measurements made at control (**sample**) points within the same area
- Spatial associations (SA)
 - **Deterministic**: assume SA is known
 - **Stochastic**: estimate SA from sample points



Spatial Interpolation = **Educated** guesswork

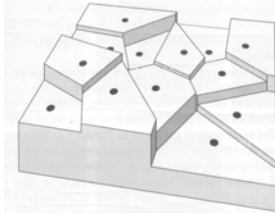
Near vs. Distant

- Known spatial associations (SA) → Spatial autocorrelation (positive): 'near' locations are more relevant
- In different interpolation methods, nearness concept appears in various form
 - As a way to distinguish different methods
 - As a way to think about how useful different methods are likely to be

Interpolation Methods

• Proximity Polygons: 'Near' means nearest

1. Construct proximity polygons for the sample locations.
2. Assuming each polygon has a uniform z value equal to the value at the control point for that polygon



Advantages: simple, follows the first law of geography
Disadvantages: it does not produce a continuous field of estimates
→ blocky

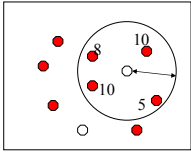
Exercise 16

• Proximity Polygon interpolation

Interpolation Methods

Local Spatial Average (LSA)

– utilizes the points within a fixed distance (d) of the location (nearness) which value to be determined.



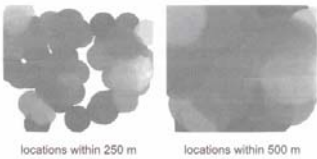
The value equals the average value of the points who are within the fixed distance

$$\hat{z}_j = \frac{1}{m} \sum_{i=1}^m z_i$$

- Point with known value
- Point with unknown value

Interpolation Methods

Local Spatial Average



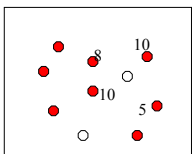
Problems:

- 1) Some locations are not within the chosen distance of any samples, it is not possible to derive a surface
- 2) The result is not properly continuous, dependent on the number of points within the distance
- 3) Each point has the same weight
- 4) Average effect



Interpolation Methods

Local Spatial Average



An alternative way is to use a specified number (k) of nearest neighbors (nearness) to estimate the value of a point.

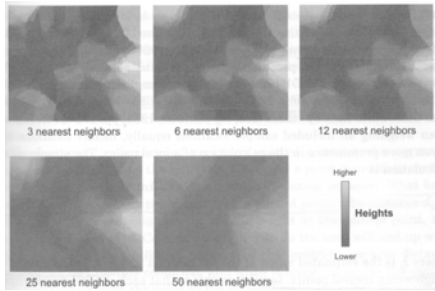
- 3 nearest neighbor
- 6 nearest neighbor
- 12
- ...

$$\hat{z}_j = \frac{1}{m} \sum_{i=1}^m z_i$$

- Point with known value
- Point with unknown value

Interpolation Methods

- Local Spatial Average



Question to Ponder?

- Why does surface become smoother when d or k increases?
- Central Limit Theorem:
 - $- d$ or k increases \rightarrow Samples # increases \rightarrow variation in the estimates decreases \rightarrow surface smoother

Exercise 17

- LSA interpolation

Interpolation Methods

• Inverse-Distance-Weighted (IDW) Spatial Average

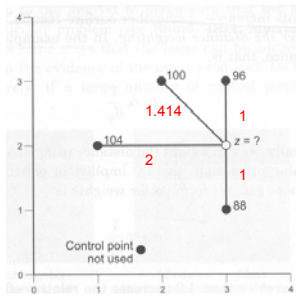
Local spatial average: $\hat{z}_j = \frac{1}{m} \sum_{i=1}^m z_i$ •Same weights

IDW:
$$\hat{z}_j = \frac{\sum_{i=1}^m w_{ij} z_i}{\sum_{i=1}^m 1/d_{ij}}$$

- Different weights: decline with distance
- The estimate is a weighted average

Interpolation Methods

• IDW



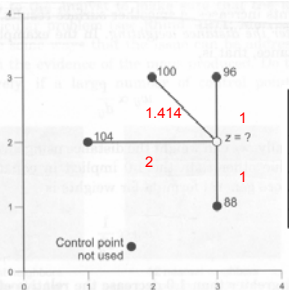
$$\hat{z}_j = \frac{\sum_{i=1}^m w_{ij} z_i}{\sum_{i=1}^m 1/d_{ij}}$$

$$w_{ij} = \frac{1/d_{ij}}{\sum_{i=1}^m 1/d_{ij}}$$

What is the value of z
With IDW method?
With Local spatial average?

Interpolation Methods

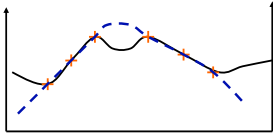
• IDW



Height, z_i	Distance, d_i	Inverse distance, $1/d_i$	Weight, w_i	$w_i z_i$
104	2.000	0.500	0.156	16.2
100	1.414	0.707	0.221	22.0
96	1.000	1.000	0.312	29.9
88	1.000	1.000	0.312	27.4
Totals:		3.207	1.000	95.6

Issues with LSA & IDW

- The range of interpolated values cannot exceed the range of observed values
 - it is important to position sample points to include the extremes of the field
 - this can be very difficult



This set of six data points clearly suggests a hill profile (dashed line). But in areas where there is little or no data the interpolator will move towards the overall mean (solid line).

Interpolation Methods

• IDW

Parameters change results of IDW interpolation

- 1) The **grid size** (finer or coarser)
- 2) The choices of **neighboring control points** (how many neighbors, or circle radius)
 - Same as **Local Spatial Average interpolation**
- 3) The **distance weighting** (distance $(1/d_{ij})$ or distance squares $(1/d_{ij}^2)$), or a form of distance functions (e.g. inverse negative exponential $(\exp(-kd_{ij}))$)

Choice of parameter is often somewhat arbitrary!

Interpolation Methods

• IDW

Tips for parameter choice

- 1) If you have a relatively large # of control points, use some for interpolation, and the remaining for error checking
- 2) Repeat many parameter setting and choose those with least error

Exercise 18

- IDW interpolation

Interpolation Methods

- **Spline**
 - finds contours that are the smoothest possible curves that can be fitted and still honor all the data
- **Kriging**: Geostatistical method
 1. Estimation of spatial association (Variogram)
 2. Estimation of the point value (z) using spatial association information

One Last Thing

- **Spatial interpolation (SI)** is different from **density estimation (DE)**
 - **SI**: known \rightarrow unknown; samples; **must** associate with a z value at the sample locations
 - **DE**: density of events; entire point set (not samples); **may** associate with a value (population)

Summary II

- Spatial interpolation methods
 - Proximity polygon
 - Local spatial average
 - IDW
 - Spline
 - Kriging
