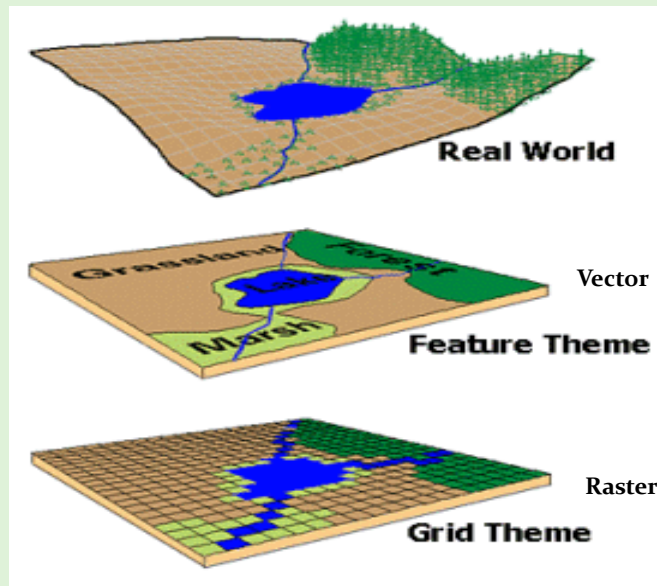


# RASTER AND VECTOR DATA

Dr. Dipanwita Dutta

1

## Methods of representing geographic space



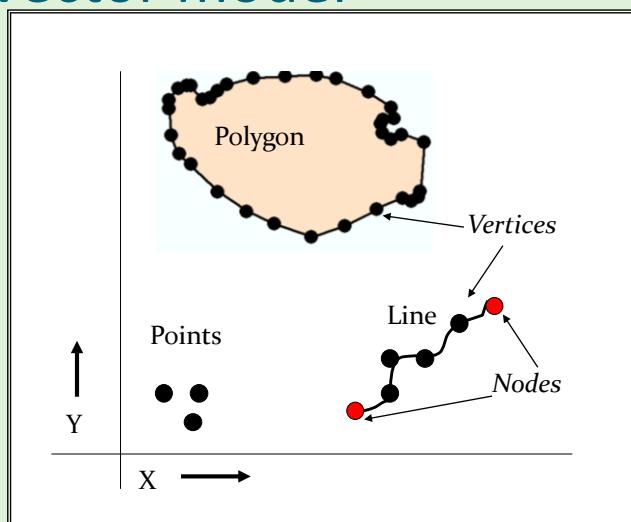
2

## What is a Database?

- A database is a set of computer files that stores information in an organized, structured format
- The information is organized in records and fields
- Information in a database is related so questions can be asked such as:
- List all of the courses that are 4000 level or higher
- List the name and address for all people whose last names begin with "T"

3

## Vector model



Features are stored as a series of x-y coordinates in a rectangular coordinate system.

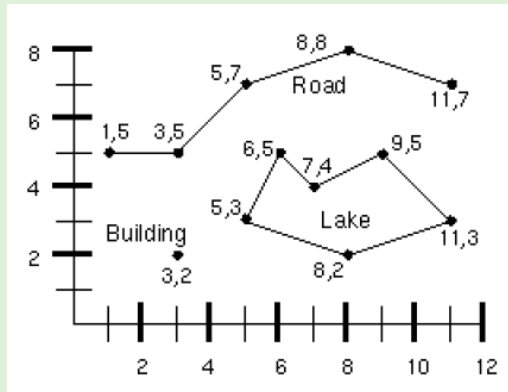
Many different coordinate systems may be used.

Copyright © 2006 by Manbeth  
H. Price

1-4

4

## Vector model



The coordinate pairs 1,5 3,5 5,7 8,8 and 11,7 represent a line (road)

The coordinate pairs 6,5 7,4 9,5 11,3 8,2 5,3 and 6,5 represent a polygon (lake).

The first and last coordinates of the polygon are the same; a polygon always closes.

5

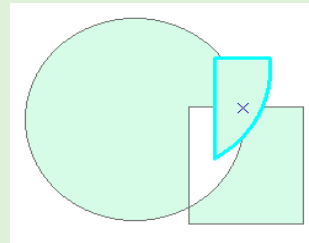
## Vector Model Advantages

- High resolution
- Better suited for map output
- Contains in-depth attribute data (discrete objects)
- Good for networking – pipelines (Topology)
- Smaller file sizes
- More accurate and defined boundaries than Raster
  - A “smoother” look (more aesthetically pleasing)

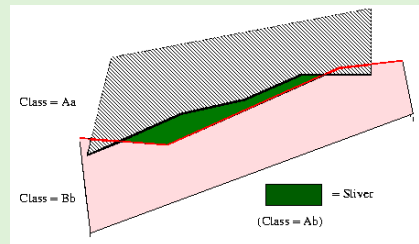
6

## Vector Model Disadvantages

- Complex data structures
  - “Hard” for CPU to process
- Inputting data in digital format is time-intensive, *expensive*
- many quality control problems (E.g., “overlap” & “slivers”)
- Overlay operations are more complex



Overlap



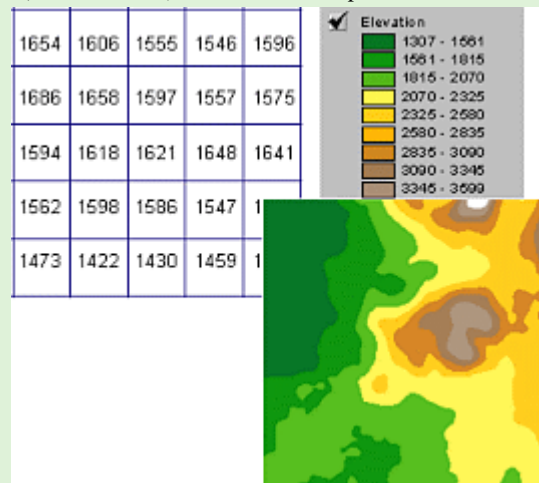
Slivers

© Showker & Krimmer, 2007

7

## Raster Model

A **grid** defines geographic space as a matrix of identically-sized square cells. Each cell holds a numeric value that measures a geographic attribute (like elevation) for that unit of space.



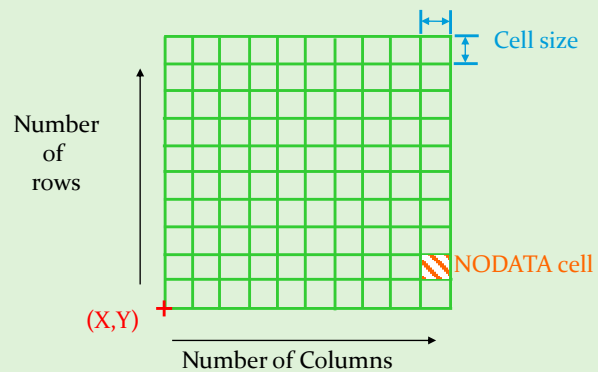
8

## The grid data structure

- Grid size is defined by **extent**, **spacing** and **no data value** information
  - Number of rows, number of column
  - Cell sizes (X and Y)
  - Top, left , bottom and right coordinates
- Grid values
  - Real (floating decimal point)
  - Integer (may have associated attribute table)

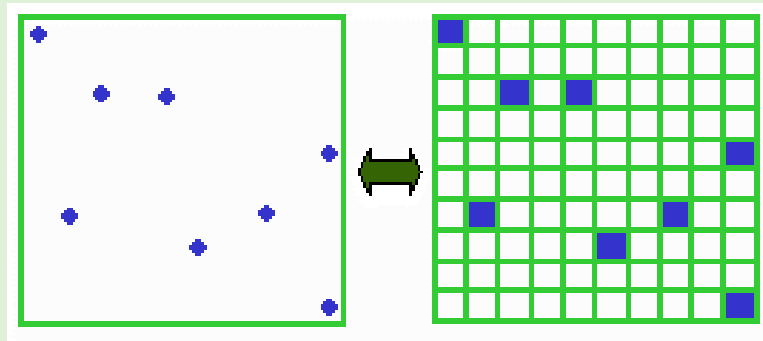
9

## Definition of a Grid



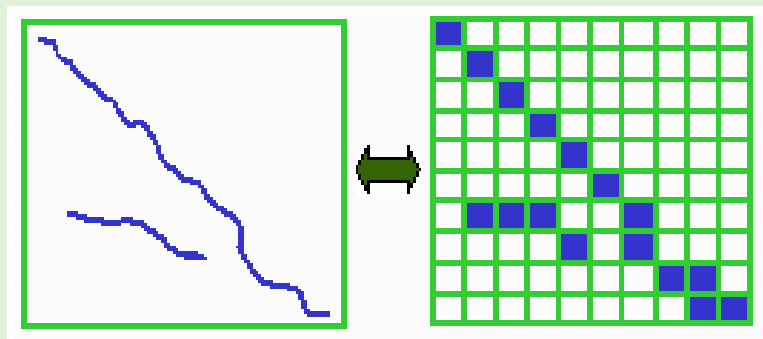
10

## Points as Cells



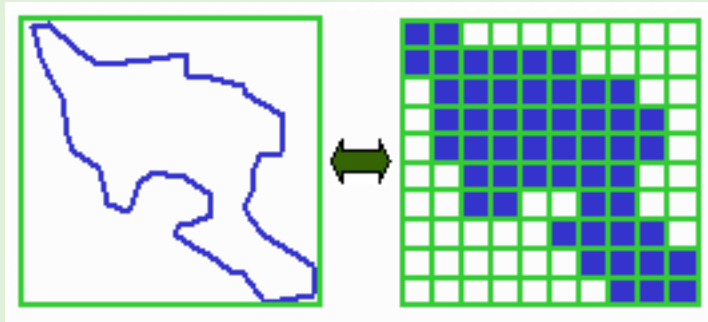
11

## Line as a Sequence of Cells



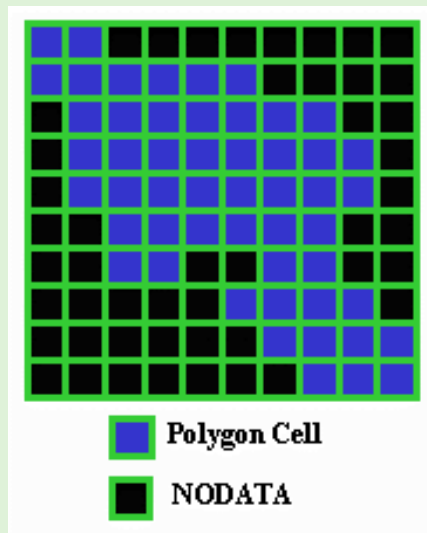
12

## Polygon as a Zone of Cells



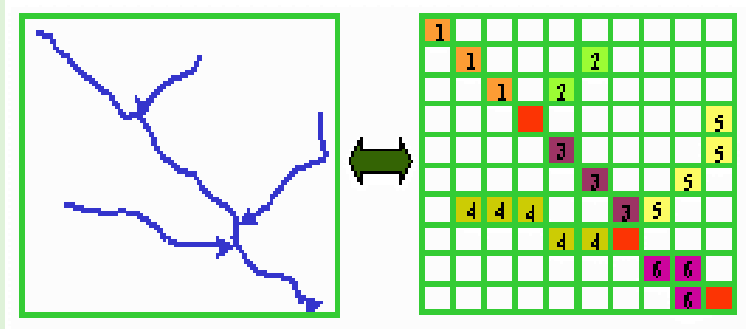
13

## NODATA Cells



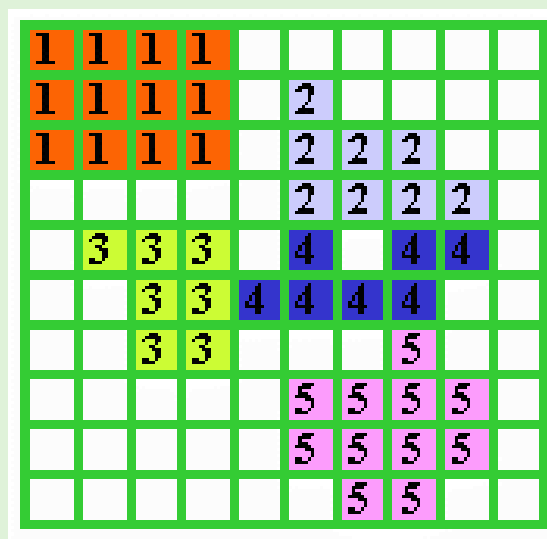
14

## Cell Networks



15

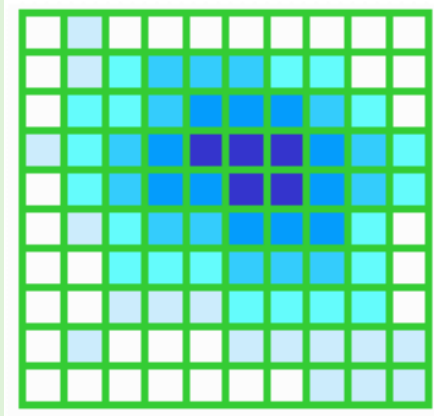
## Grid Zones



16



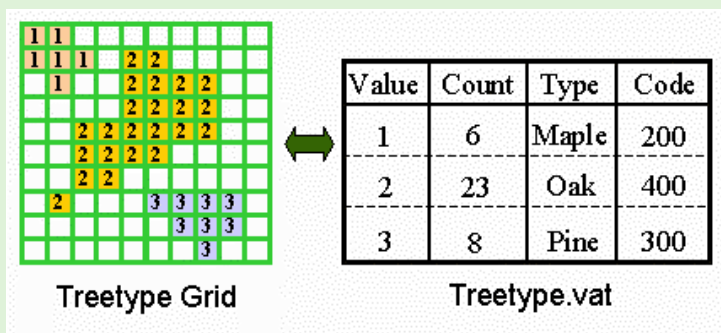
## Floating Point Grids



Continuous data surfaces using floating point or decimal numbers

17

## Value attribute table for categorical (integer) grid data



Attributes of grid zones

18

## Raster Model Advantages

- Simple data structure
  - Numerical attributes (bit maps) are “easy” for CPU to process
- Spatial analysis functions are often simpler, faster
  - Overlay data sets for quick(er) visual analysis
- Better for representing continuous surfaces
  - E.g., terrain, elevation

“Raster is faster”

19

19

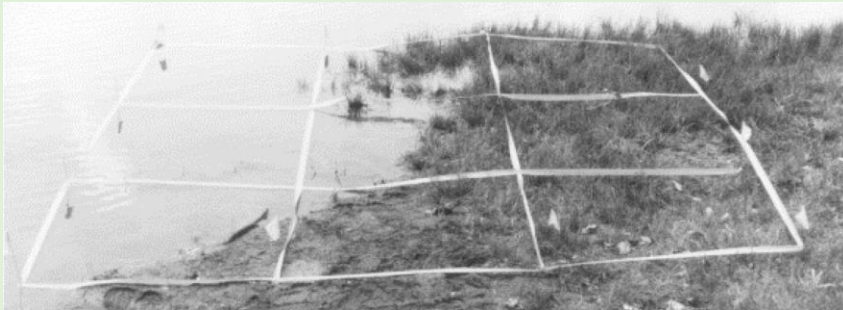
## Raster Model Disadvantages

- Requires lots of computer “storage” space
- Resolution (“coarseness”) of data can be a factor
  - Size of the grid cells not appropriate, or less useful
- Networks are hard(er) to model because boundaries are not (as) discrete
- Attributes being displayed can be difficult to understand
  - May take lots of visual interpretation
- The “mixed pixel” problem
  - Only one attribute per grid cell

20

20

## Raster Model Mixed Pixel Problem



**Water dominates**

W	W	G
W	W	G
W	W	G

**Winner takes all**

W	G	G
W	W	G
W	G	G

**Edges separate**

W	E	G
W	E	G
E	E	G

21

21

### Raster

1. Points, line & polygons everything in the form of Pixels.
2. Large file size.
3. Networks are not so well represented.
4. Only one pixel value represents each grid cell.
5. Generalization of features (like boundaries) hence accuracy may decrease.
6. Simulations and modeling is easier (spatial analysis, terrain modeling etc.).
7. Maintaining is easier.
8. Excellent for representing data containing continuous values (like land use, elevation etc.)
9. Coordinate-system transformations take more time and consume a lot of memory.
10. Grid cells or pixel makes simpler data structure.

### Vector

1. Represented by point, line and polygon.
2. Relatively small file size (small data volume)
3. Excellent representation of networks.
4. A large no. of attributes can be attached, hence more information intensive and a number of thematic maps can be prepared from a single layer.
5. Features are more detailed & accurate.
6. Creating, cleaning and updating data is more time and labour consuming.
7. Topology-based analysis & operations are easier to perform (like network analysis etc.).
8. Can not represent continuous values like land use, elevation etc very well.
9. Assigning projection and transformations are less time taking and consumes less memory of the computer system.
10. Topology makes data structure complex.

22

# Acquiring Data for Models

## Vector

- Digitize data
- Digital Line Graphs
- GPS/Mobile GIS
- Download data via internet or GIS server

Vector and raster data models complement ~~not complete~~ with ~~each other~~

## Raster

- Satellites- LANDSAT, IKONOS, SPOT, etc.
- Aerial Photographs
- Radar/Lidar
- Scanned Images
- Digital Raster Graphs (Scanned Topographical Maps)

23