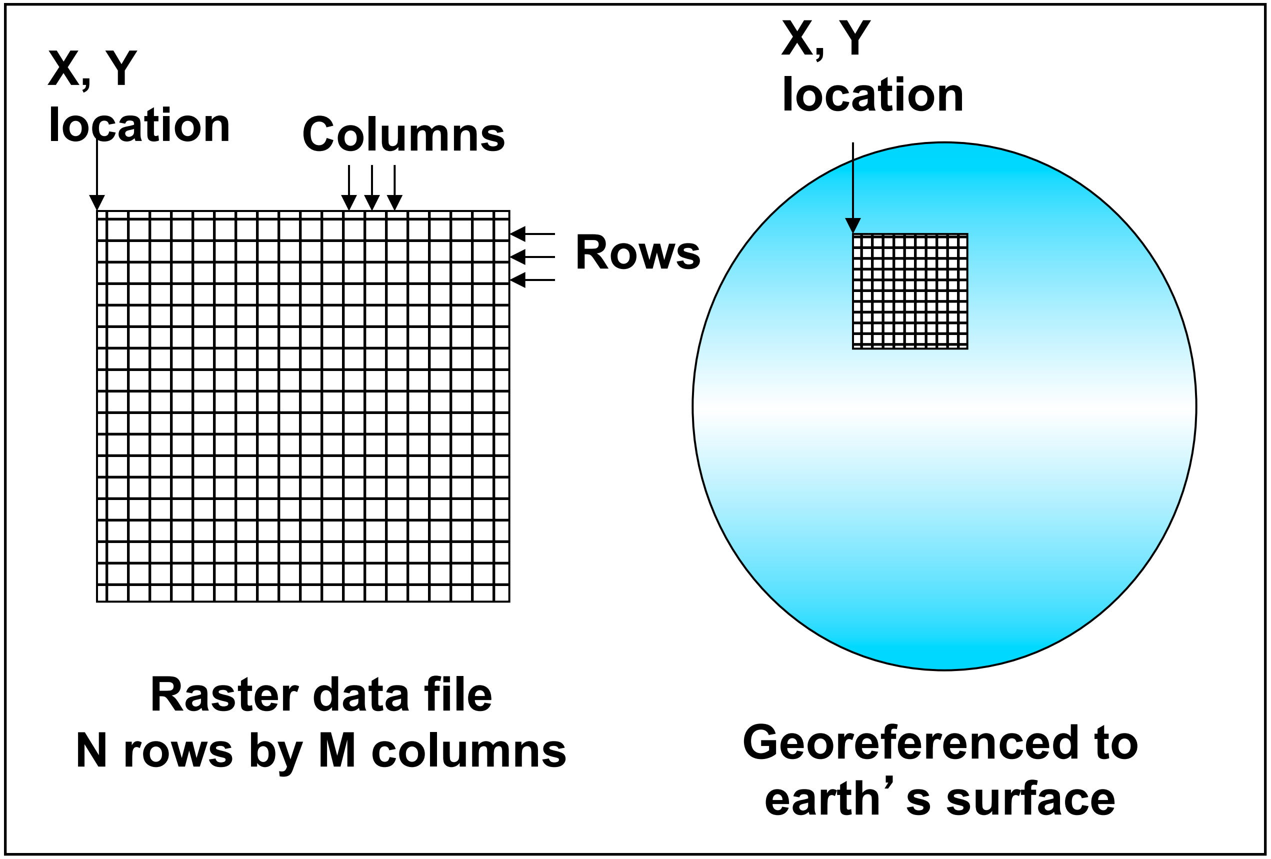
Raster Analysis—Part 1

Raster data

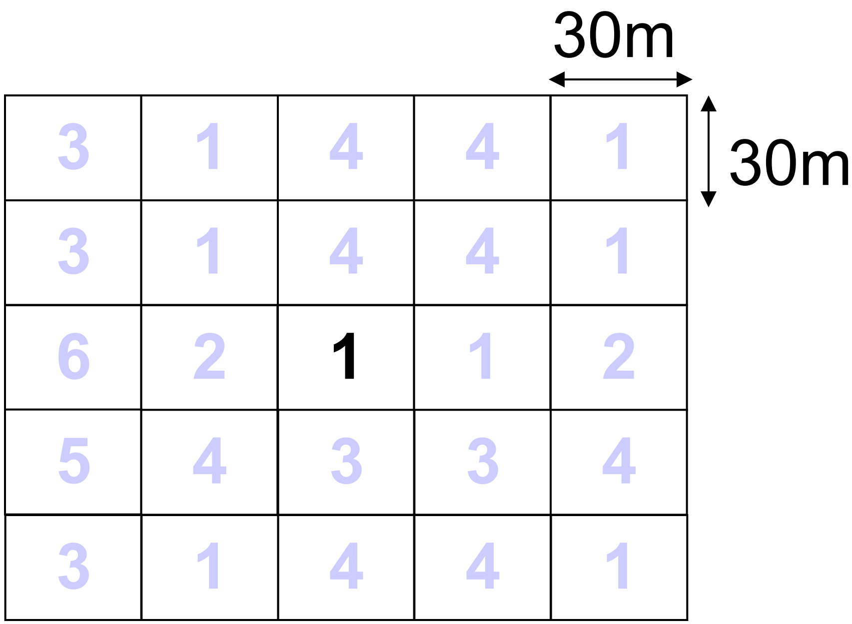
Desired Learning Objectives

- Compare and contrast the raster data model with the vector data model
- Define the following terms: pixel, cell, band, pixel depth, resampling, cell resolution
- Discuss situations in which spatial analyses is better done with raster data

The raster data model



Pixels or Cells



- Dimension of a pixel varies (resolution)
- Each pixel contains one numeric value
- Value represents some property of that pixel area, e.g. elevation or rainfall
- Values may be integers or floating point numbers

Unlike a polygon, each cell has only ONE attribute: its value. Storing multiple values means storing multiple rasters.

Binary data

0
1
10
11
100
101
110
111
1000
1001
1010
1011

- Most raster formats use binary storage
- Numbers are stored as a series of 0' s and 1' s representing numbers in base 2
- Binary values are grouped by eight

1 bit

10011101

one byte

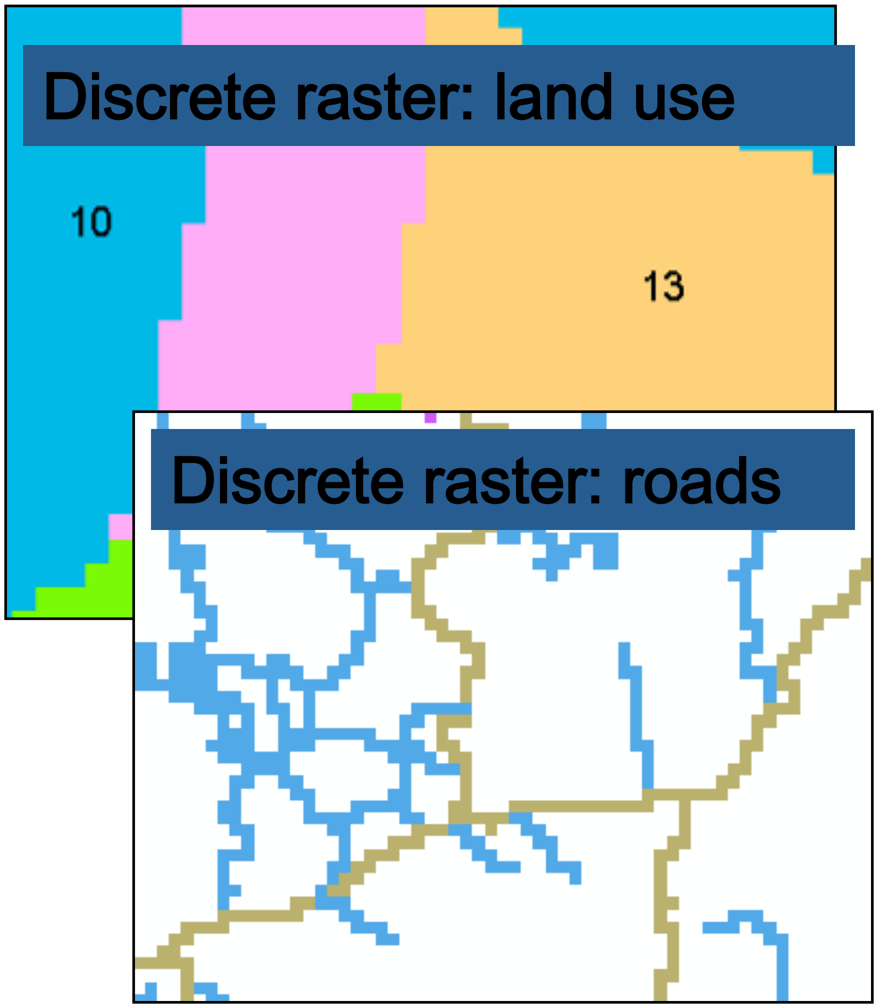
In base 2:
00000000 = 0
11111111 = 255
$2^8 = 256$

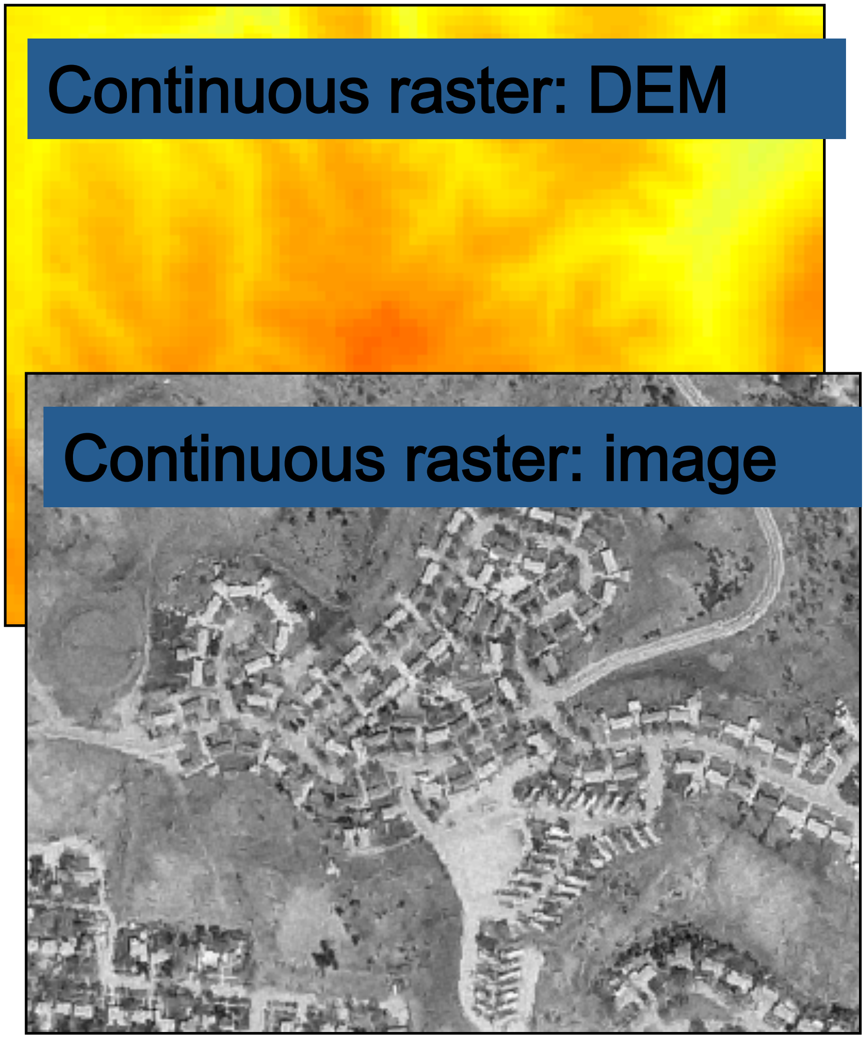
1111111111111111 = 65,565
$2^{16} = 65,566$
two bytes

Pixel depth

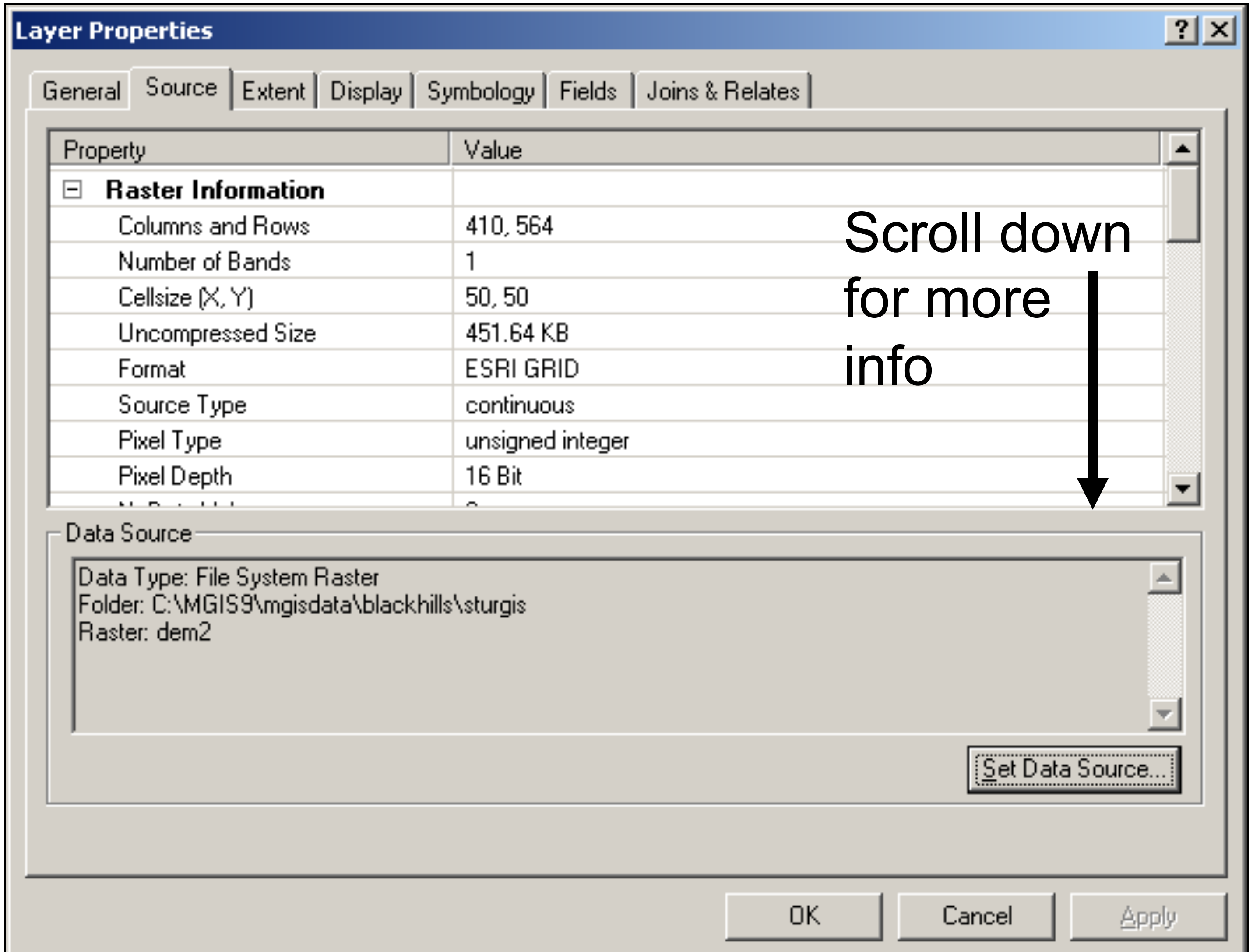
- The number of bytes used for each pixel
  - More bytes = larger numbers = more space
- Integer values
  - 8-bit pixel (one byte) stores 0 - 255
  - 16-bit pixel (two bytes) stores 0 - 65,565
  - 24-bit pixel (three bytes) stores 0 -16.7 million
- Floating point values
  - Required for decimal number storage
  - 32-bit pixel (four bytes)

Types of raster data





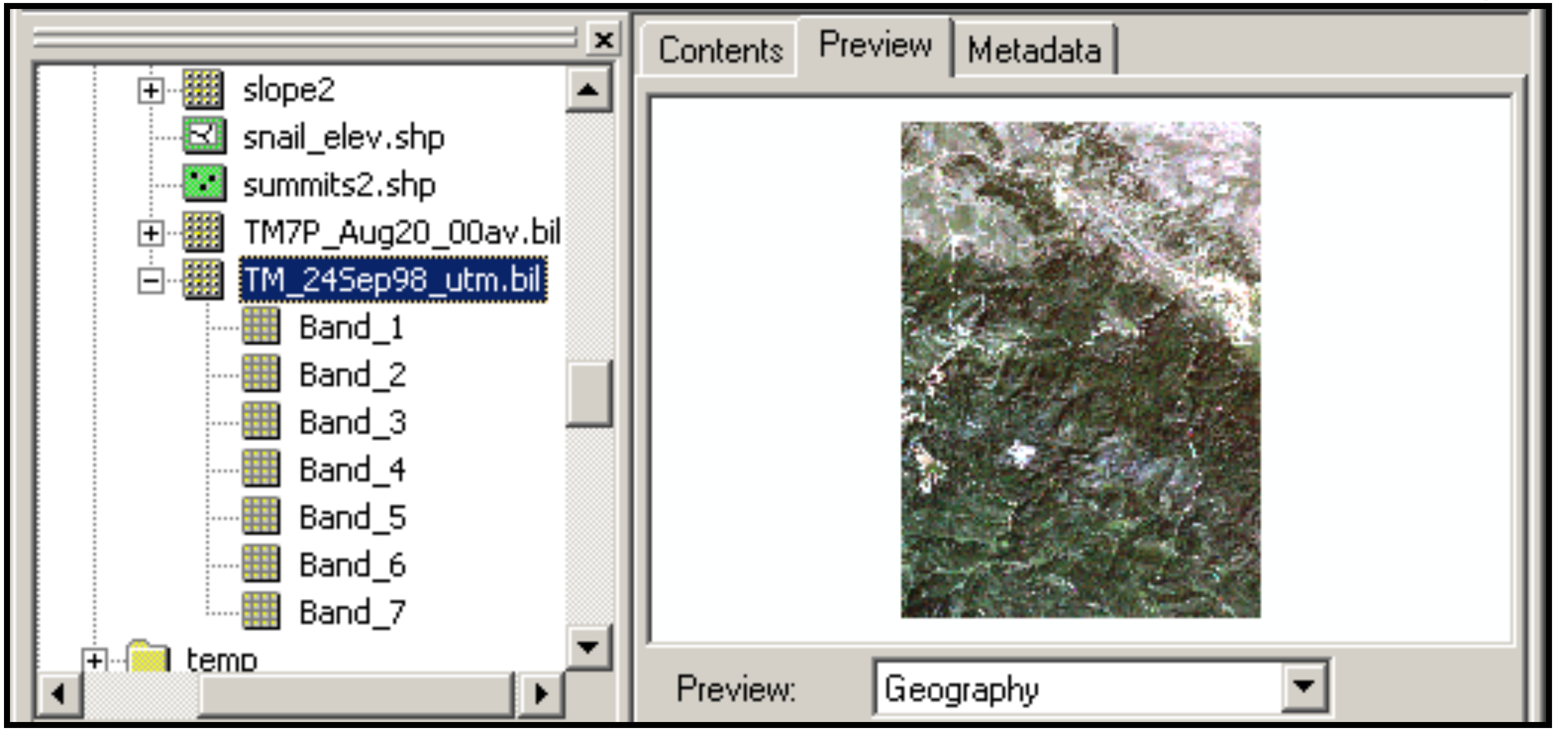
Raster Properties



Bands

- A single raster may include multiple arrays
- Most often used to store color images and satellite images

7-band Landsat satellite image

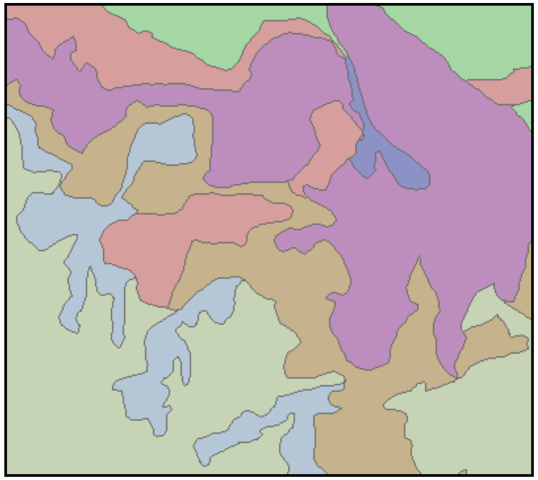


Why use rasters?

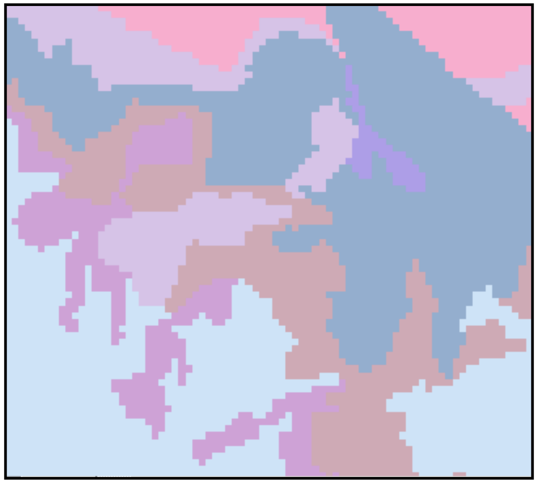
- Better at storing certain kinds of data
- Better at analyzing certain kinds of data
- Often faster analysis than vector
- Imagery desirable for certain maps
- BUT…
  - Coordinate precision generally lower
  - High precision has high storage costs
  - Cannot store multiple attributes in a single raster

Raster resolution

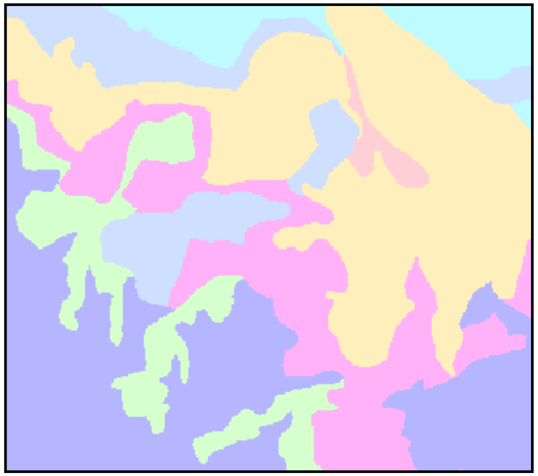
- Measured by cell size dimensions
- Storage space increases dramatically with increased resolution



Vector format



200 m raster

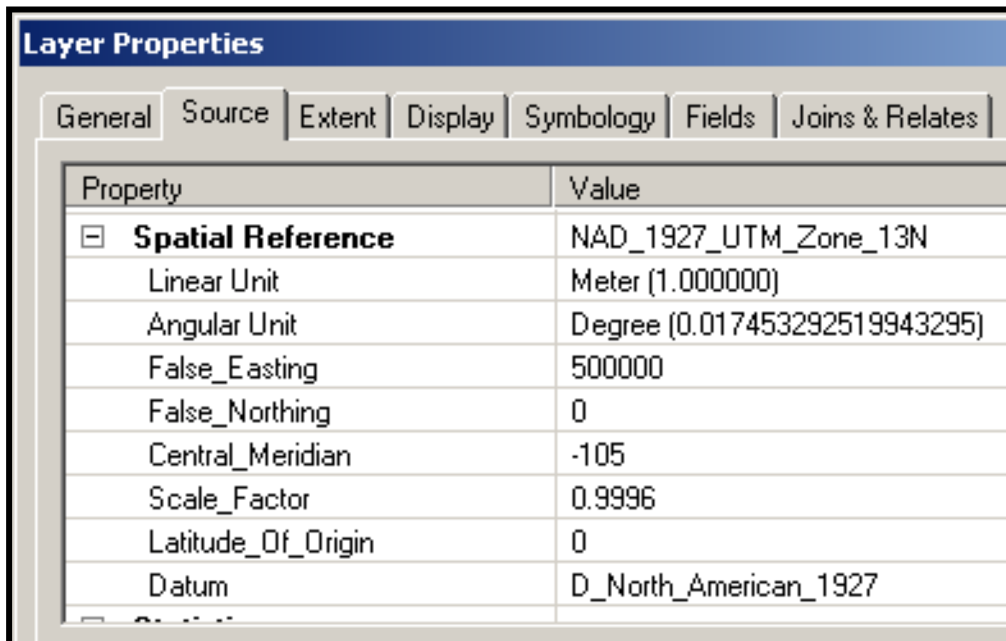


50 m raster

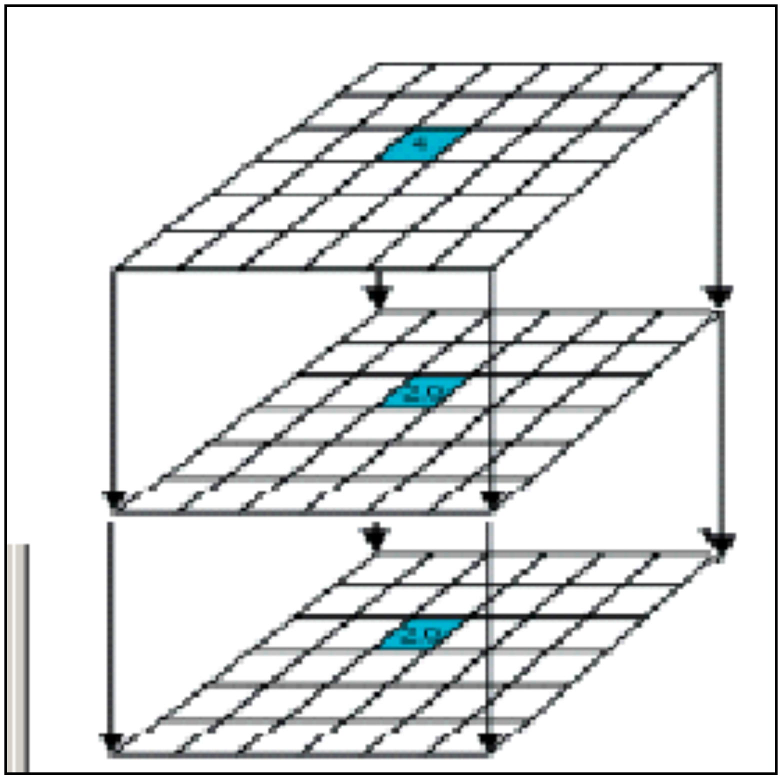
Cell size units

- Cell x-y resolution units are based on the raster's coordinate system definition
  - Decimal degrees*
  - Meters
  - Feet

* Because distances and areas are fundamental bases for raster analysis, use projected coordinate systems for rasters…

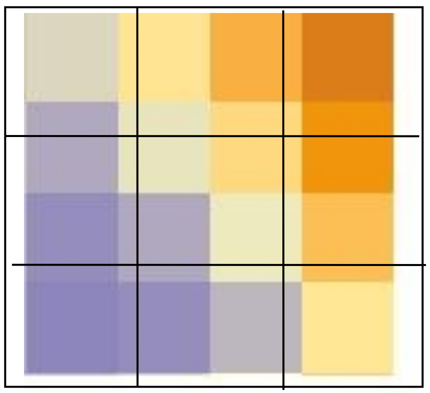


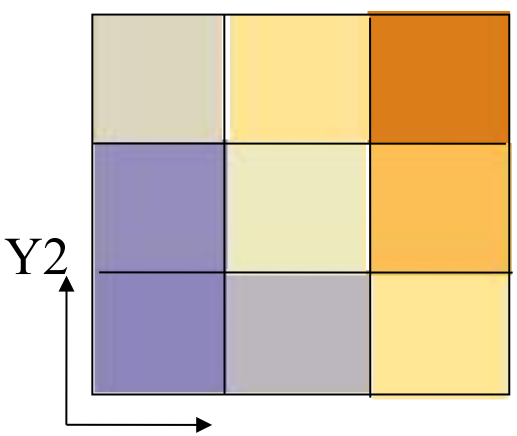
Raster analysis



- Raster analysis uses cell-by-cell functions on one or more input grids
- Cells must be the same size and line up spatially

Resampling





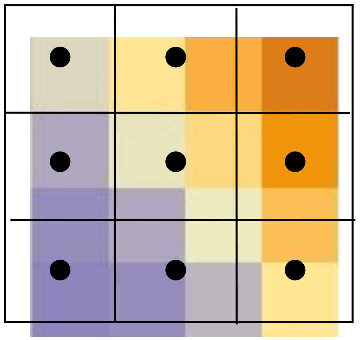
If input grids do not match, then one must be resampled to match the other

Resampling can degrade the accuracy of a raster even if the difference in cell size and location is small

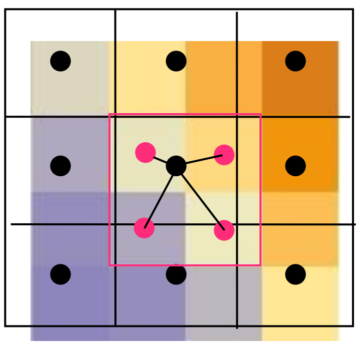
The new cell grid is determined, and the old cell values must be fit into the new structure somehow

Several methods are used for resampling

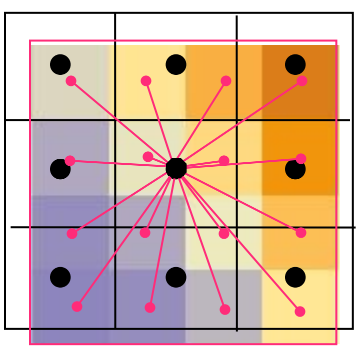
Resampling methods



Nearest neighbor resampling grabs the value from the old cell that falls at the center of the new cell. It preserves the original value and should always be used with categorical data, or when the original data values need to be preserved. It is the fastest method.



Bilinear resampling calculates a new value from the four cells that fall closest to the center of the new cell. It uses a distance-weighted algorithm based on the old cell centers. It is best used with continuous data such as elevation.



Cubic convolution resampling calculates a new value from the sixteen cells that fall closest to the center of the new cell. It uses a distance-weighted algorithm based on the old cell centers. It is best used with continuous data such as elevation. It is the most time-consuming method.

Raster Analysis—End Part 1