Spatial Modeling and Analysis

Lecture 2

Introduction to Raster data Model and Analysis

Lecturer: Dr. Tingneyuc Sekac BS, Mphil, P.hd,
PNG University of Technology - 2022

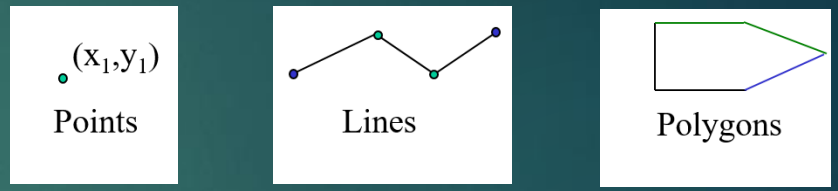
Lecture Outline

1. GIS Data Model Concepts
2. Raster Data Model
3. Raster Data Type
4. Cell/Pixel Dimension
5. Definition of a Grid and Cell Size
6. Discrete Data (integer) and floating data
7. Advantages and Disadvantages of Raster Data Model
8. Introduction to Types of spatial analysis (Raster)
9. References

1. GIS Data Model Concepts

- There are two main data models or conceptualizations used for spatial data:

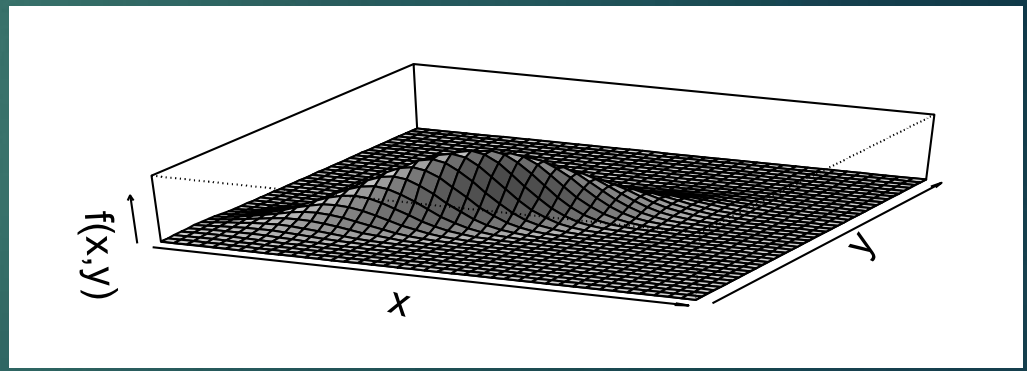
- Vector data model
- Raster data model



- "*raster is faster but vector is corrector*" Joseph Berry

- **Vector data model** use discrete objects such as point, lines and polygons to represent the geometry of the real-world entities, discrete entities

  E.g. a road, river, city and towns, lakes or wetlands, farm land, etc.

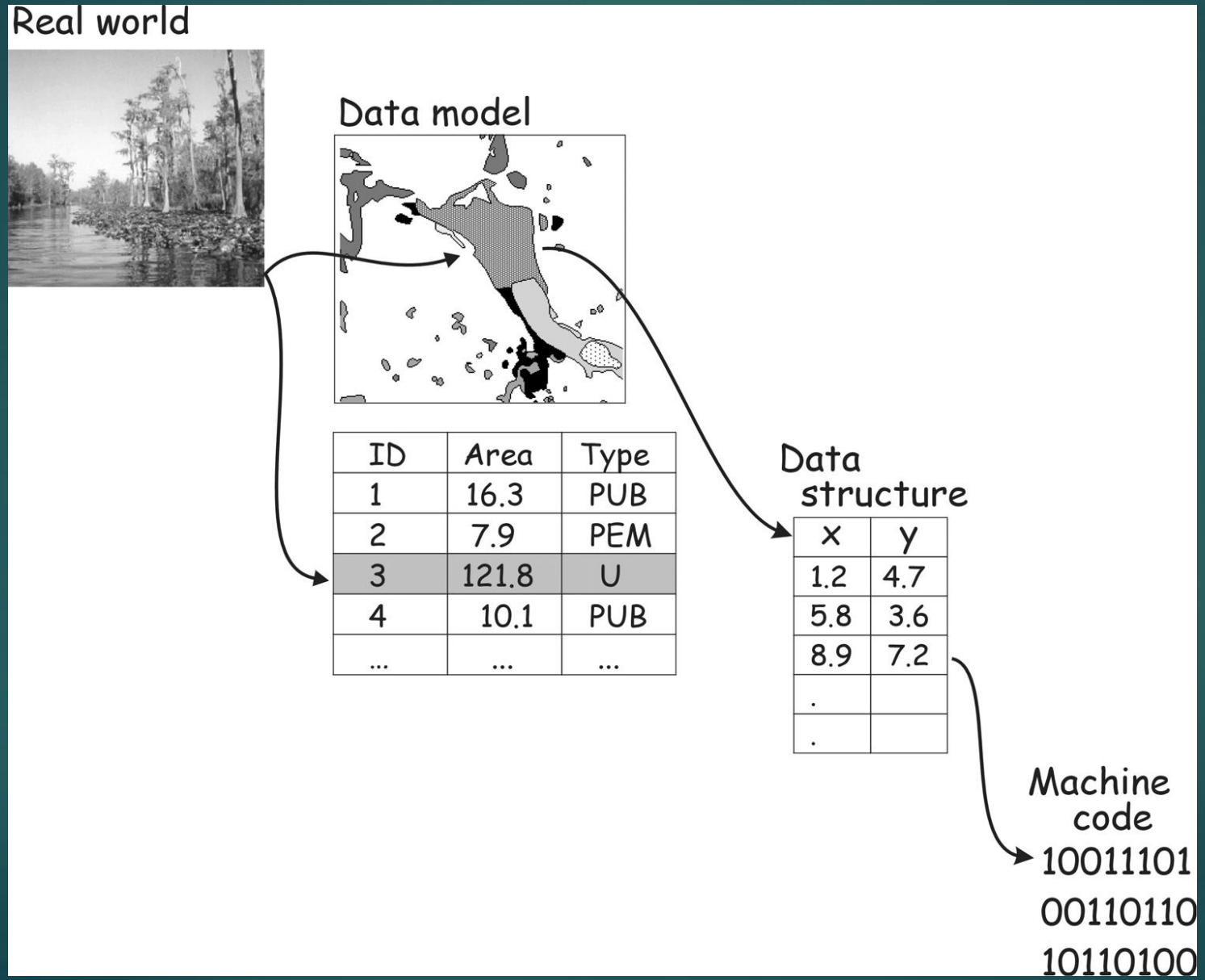


- **Raster data model** represents continuous phenomena that may change continuously across a region

  E.g. Elevation, rainfall, temperature, soil moisture, etc.

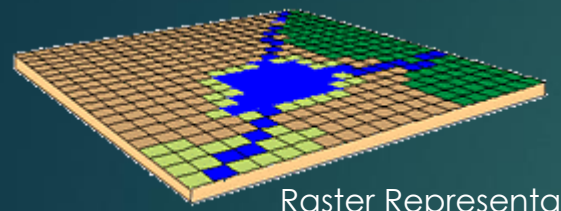
Source: Maputo, Mozambique, 2018- Geoinformation and Sectoral Statistics Section

1.2. Illustration



Source: Bolstad, 2016

1.3. Raster vs. Vector Data Model



Raster Representation

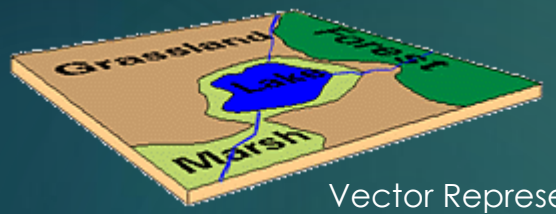
Raster Data Attributes

ROW	COL	LU_CODE
1	1	2
1	2	2
1	3	2
1	4	2
...		
1	19	101

LU_CODE	LAND_COVER	NAME	CELLS
1	forest	Sherwood Forest	100
2	grassland	Marshall Field	150
100	lake	Blue Lake	75
3	marsh	Okeefenokee Swamp	55
101	river	Suwanee River	20

In a attribute table, Each Cell has a coordinate representation that is indicated by 'LU_CODE'

LU_CODE associated with Geographical Entities



Vector Representation

Vector Data Attributes

Polygon Attribute

AREA	PERIMETER	LU_ID	LU_CODE	NAME
200	500	1	1	Sherwood Forest
1250	10000	2	2	Marshall Field
175	250	3	100	Blue Lake
100	295	4	3	Okeefenokee Swamp

Line Attribute

LENGTH	LU_ID	LU_CODE	NAME
45	4	101	Suwanee River
50	5	101	Suwanee River
35	6	101	Suwanee River

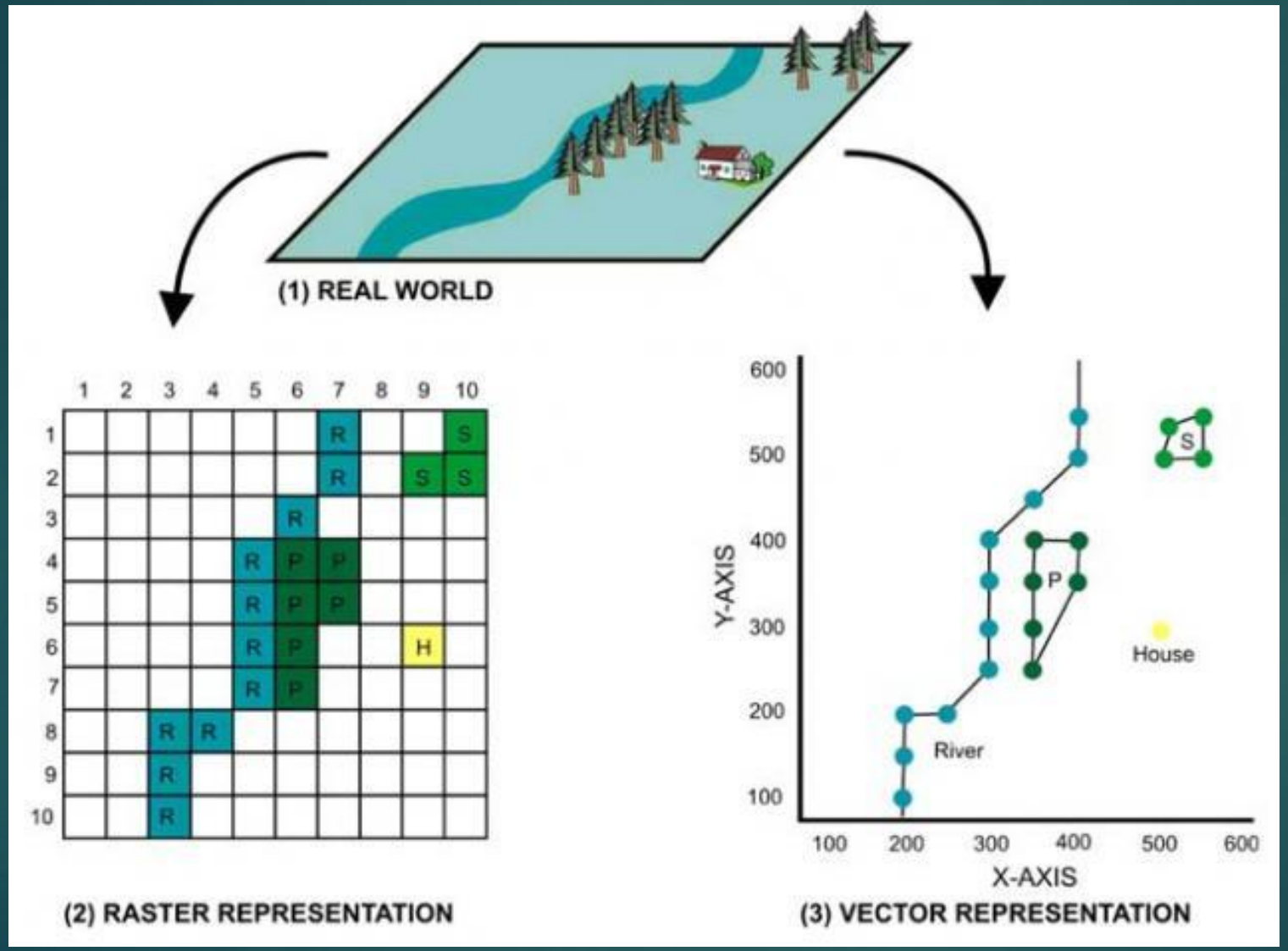
Real World

The Real World landscape consist of;

- Grassland
- Hill
- Lake
- Rivers
- Forested land
- marsh

Source: http://gis.washington.edu/phurvitz/professional/SSI/attrib.html
http://gis.washington.edu/phurvitz/professional/SSI/datatype.html

1.4. Raster vs Vector Illustration



Source: Maputo, Mozambique, 2018- Geoinformation and Sectoral Statistics Section

1.5. Summary of Data Model Concepts

* Vector data model and Raster data model can represent same phenomena;

  E.g. Elevation represented as surface (continuous field) using raster grid or as lines representing contours of equal elevation (discrete objects), or as points of height (Z values).

* Data can be converted from one conceptual view to another;

  E.g. raster data layer can be derived from contour lines, point cloud.

* Selection of raster or vector model depends on the application or type of operations to be performed;

  E.g. Elevation represented as surface (continuous field) in raster - to easily determine slope, or as discrete contours if printed maps of topography

2. Raster Data Model

- In a Raster data Model, the world is defined as a regular set of cells in a grid pattern.
- Whatever the geographical phenomena or entities on the planet earth are simple represented by cells/ grid cells/pixels
- The cells are square and evenly spaced in the x and y directions.
- The geographical phenomena or entities of interest are represented by attribute values associated with each cell location.
- Raster data models are the natural means to represent “continuous” spatial features or phenomena.
- Continuous” spatial features Example;

- Elevation,
- precipitation,
- temperature
- slope,
- pollutant concentration.
- Other

Source: Bolstad, 2016

2.1. Raster Model Representation

* Raster data models represent continuous phenomena or spatial features

  E.g. Elevation/DEM, bathymetry, precipitation, slope, etc.

* Raster data model may also be used to represent discrete data

  E.g. Land cover: forest, wetlands, urban areas

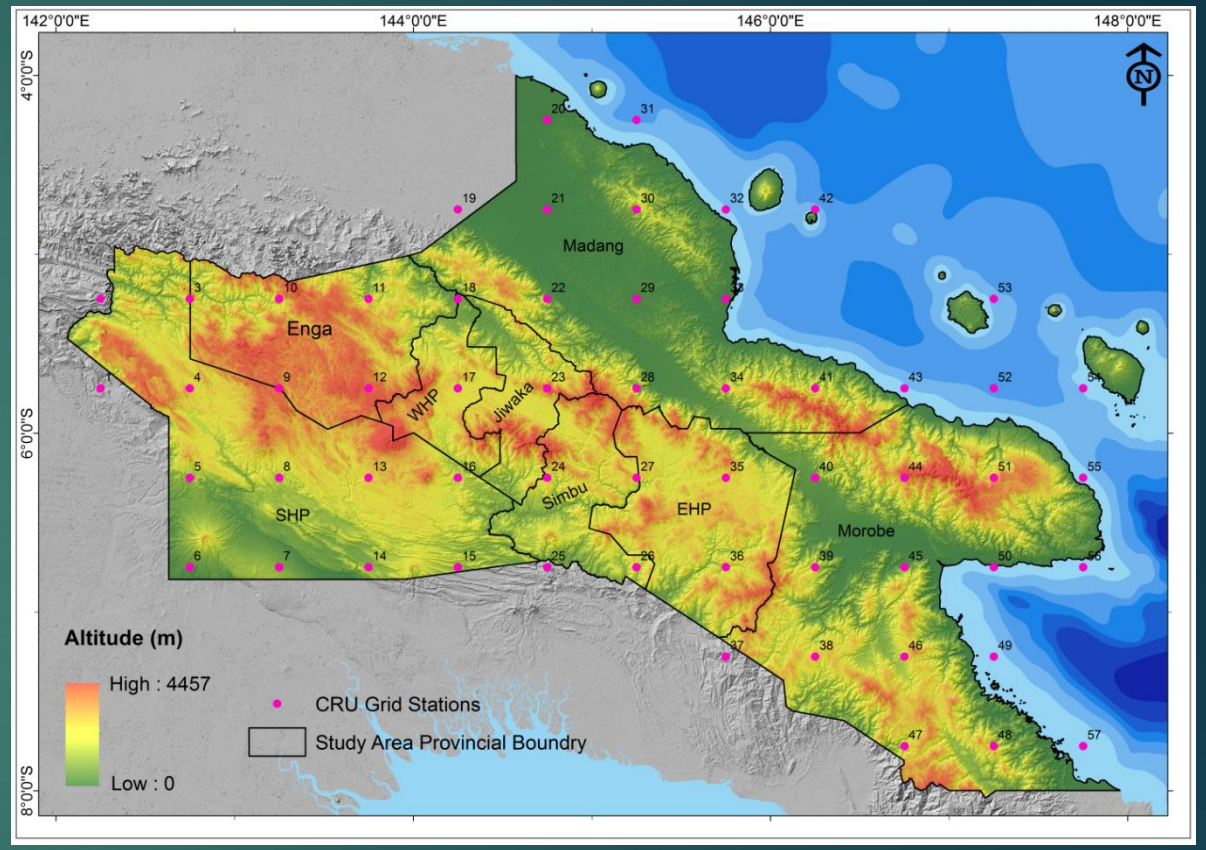


Image Source: T.Sekac (2022)

3. Raster Data type

* Raster data are classified into various :

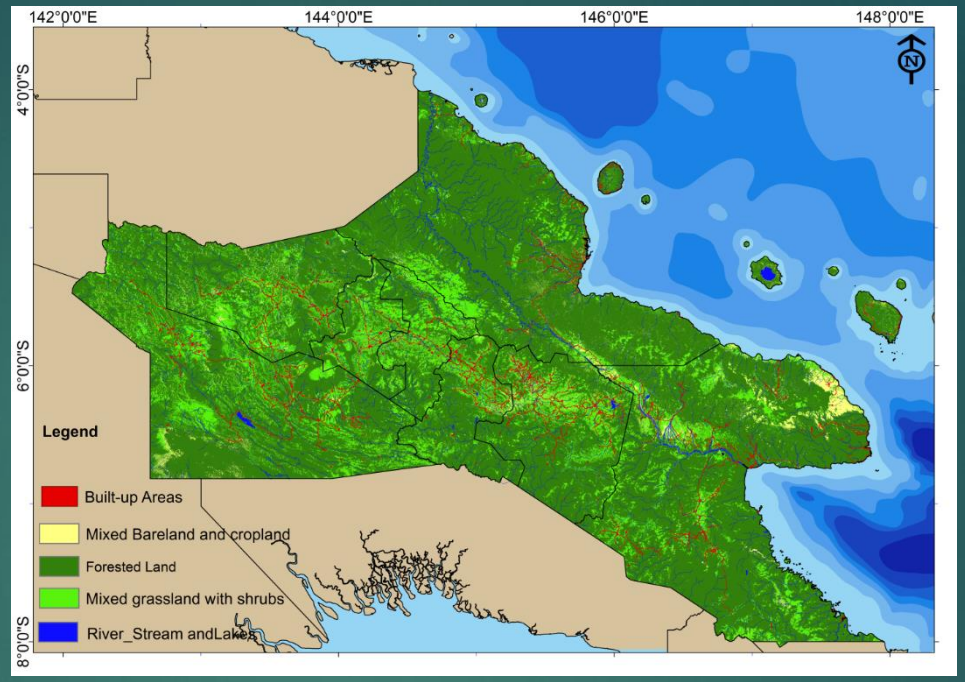
- Digital aerial photographs (from drone, Plane and Helicopter,
- Satellite Imagery (Lansat and SPOT, Quick bird, Geoeye)
- DEM
- digital pictures,
- scanned maps
- NETCDF Files (Satellite data)

3.1. What are Raster type

1. Thematic Raster

Represents map data:

- Rainfall
- Land use
- Soil map
- Lithology
- Temperature



2. Image Raster

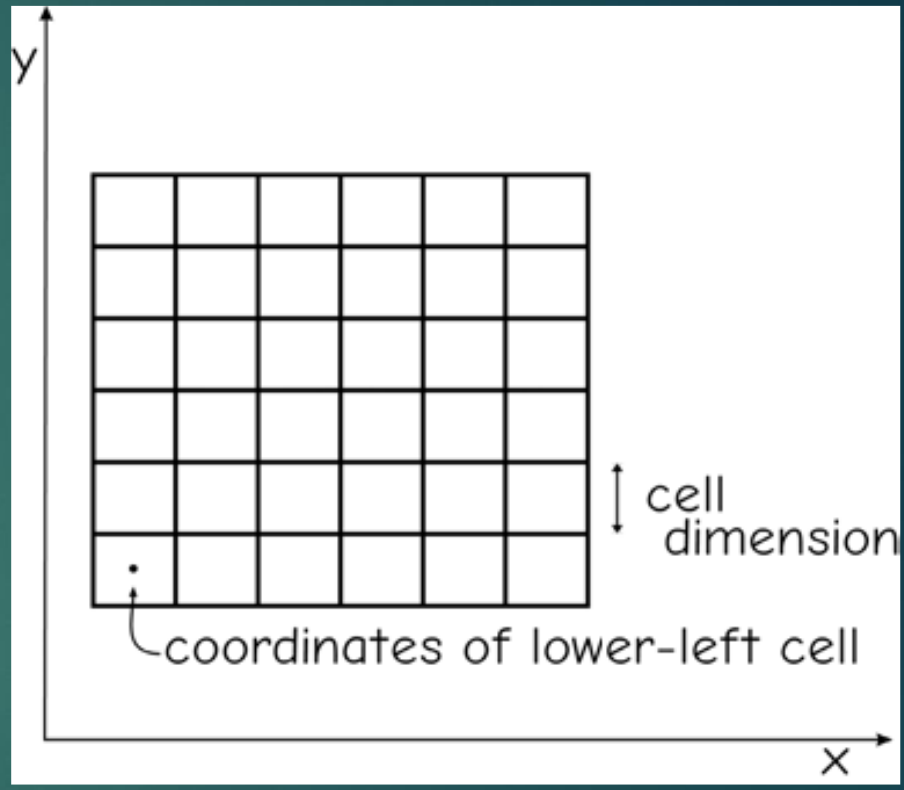
Image Raster Contains:

- Satellite Image
- Aerial Image

Image Source: T.Sekac (2022)

4. Cell/Pixel Dimension

- Raster data sets have a cell dimension, defining the edge length for each square cell.
- For example, the cell dimension may be specified as a square 10 meters on each side.
- The cells are usually oriented parallel to the x and y directions, and the coordinate are specified for each Conner location.
- Each cell contains a value representing information on each geographical entities.
- Raster data is geo-reference by real world coordinate systems.
- Raster can be Aerial photographs, satellite imagery, scanned maps, etc…



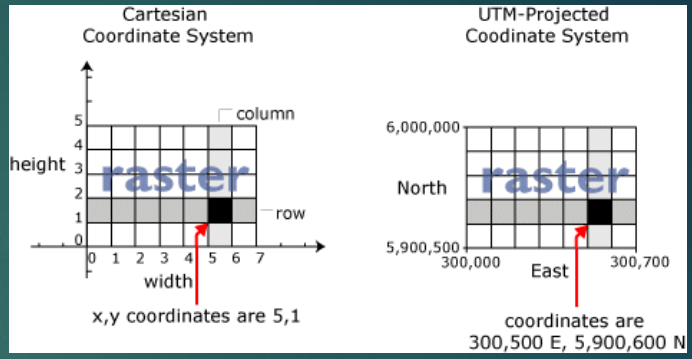
Source: Bolstad, 2016

4.1. Elements in Raster Data Model

- In a broader concept, a raster data model can be simply address as Grid.

Grid Consist of;

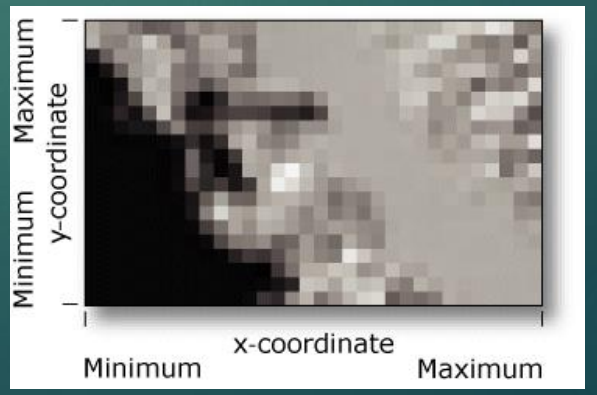
- Rows
- Columns
- Cells



- The rows and columns origin is the lower left Conner of the grid.
- The Rows Function as y – coordinates
- The Column functions as X – coordinates
- In line with rows and column in the raster, a cell is defined by its location.
- A raster cell stores a single value, however this can be extended by using raster bands.
- Raster Model divides the area into grid cells or pixel.
- The grid cells are measured with attribution information.
- Each cell can also represents Points, lines and polygon
- Resolution depends on grid cell size.

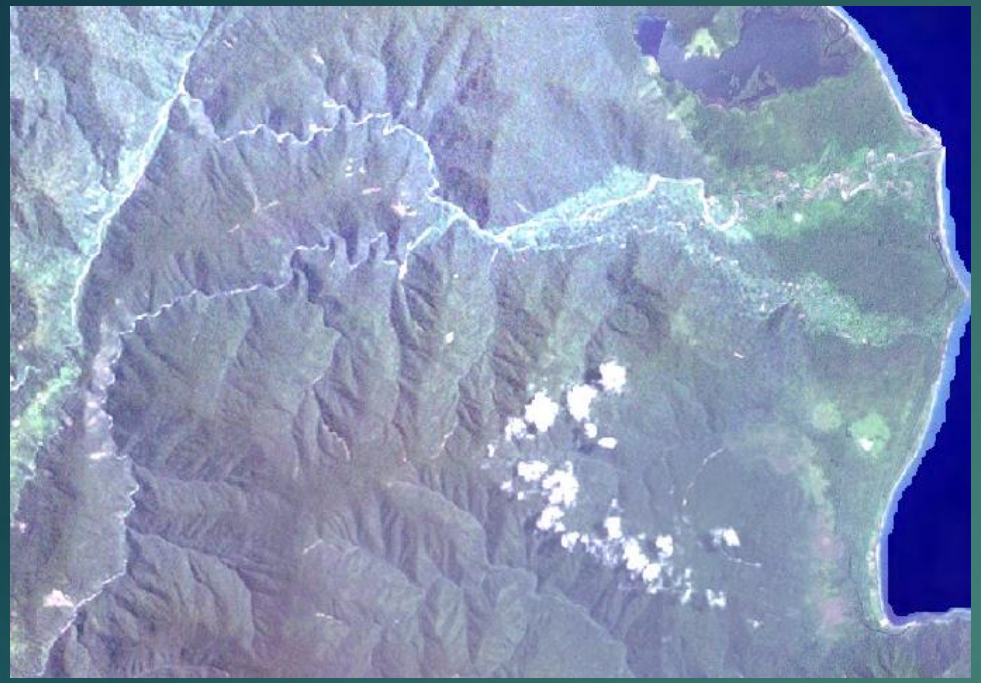
a	a	a	a	r	f	f	a	a	a	a	a
a	a	a	a	r	f	f	a	a	a	a	a
a	a	a	f	r	f	f	a	a	a	a	a
a	a	a	r	r	f	f	a	a	a	a	a
a	a	a	r	f	f	f	a	a	a	a	a
a	f	f	r	f	f	f	a	a	a	a	a
a	f	f	r	f	u	f	a	a	a	a	a
h	h	h	h	h	h	h	h	h	h	h	h
f	f	r	u	u	u	u	a	a	a	a	a
f	f	r	f	u	u	a	a	a	a	a	a
f	f	f	r	f	f	a	a	a	a	a	a
f	f	f	f	r	f	a	a	a	a	a	a

a = agriculture
f = forest
h = highways
u = developed
r = river



Source: Bolstad, 2016 and ESRI, 2016

4.2. Each Cell Contains Information



Lansat 8 Satellite Image

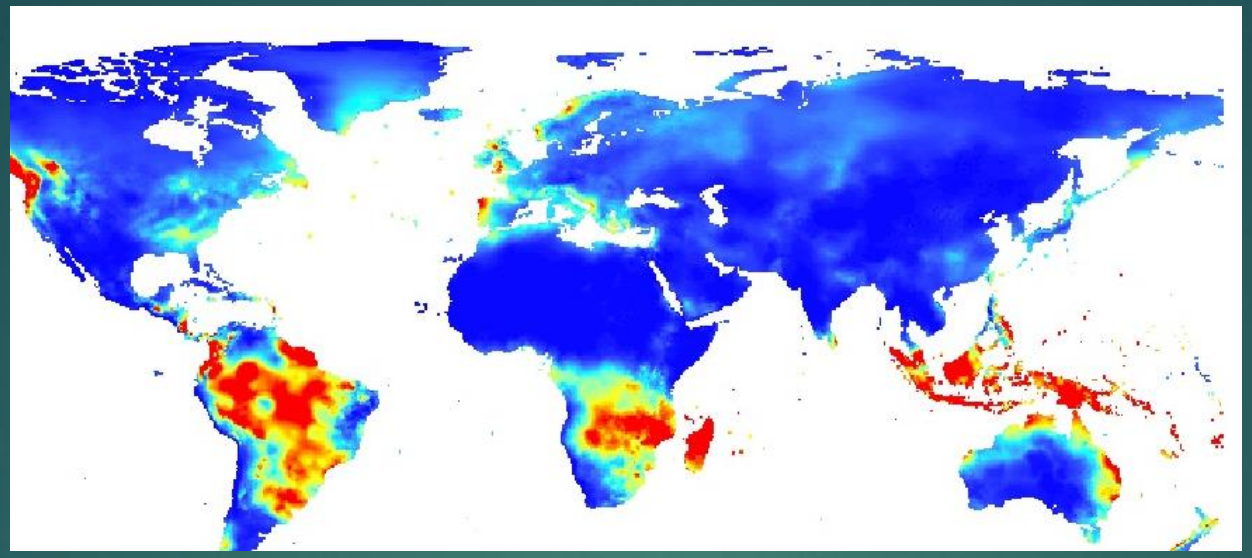
Each raster cell represents a given area
on the ground and is assigned a value that
Represents geographical entities on the ground.

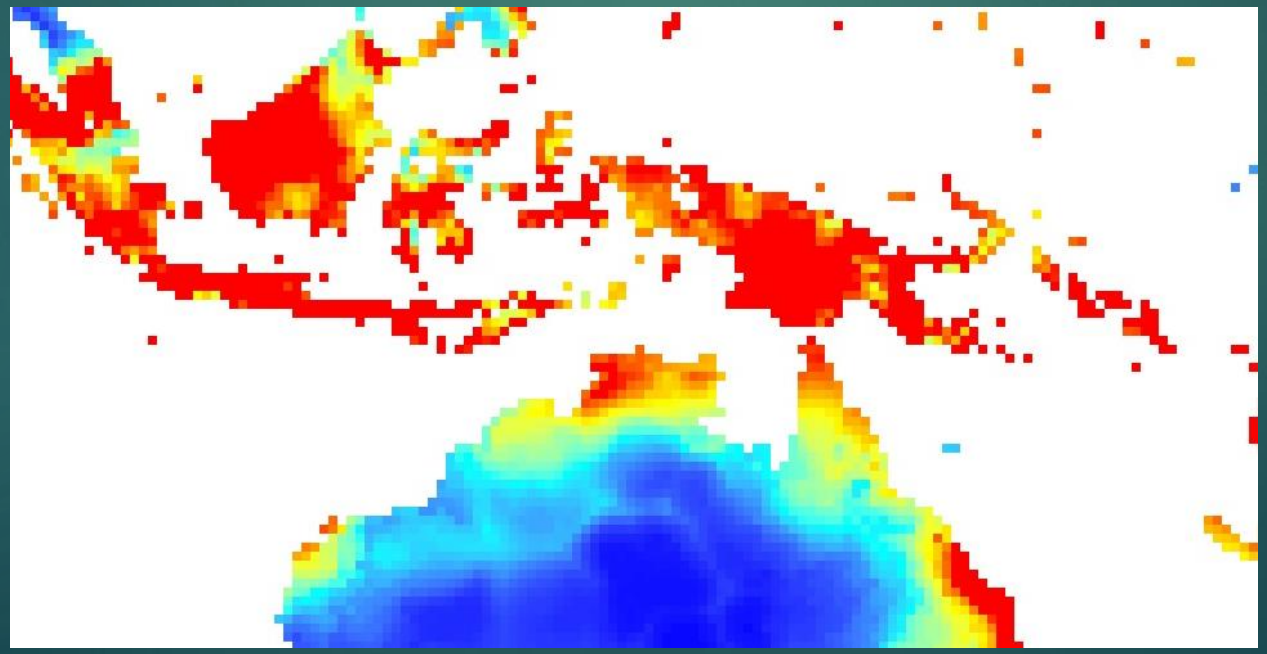
Image Source: T.Sekac (2022)

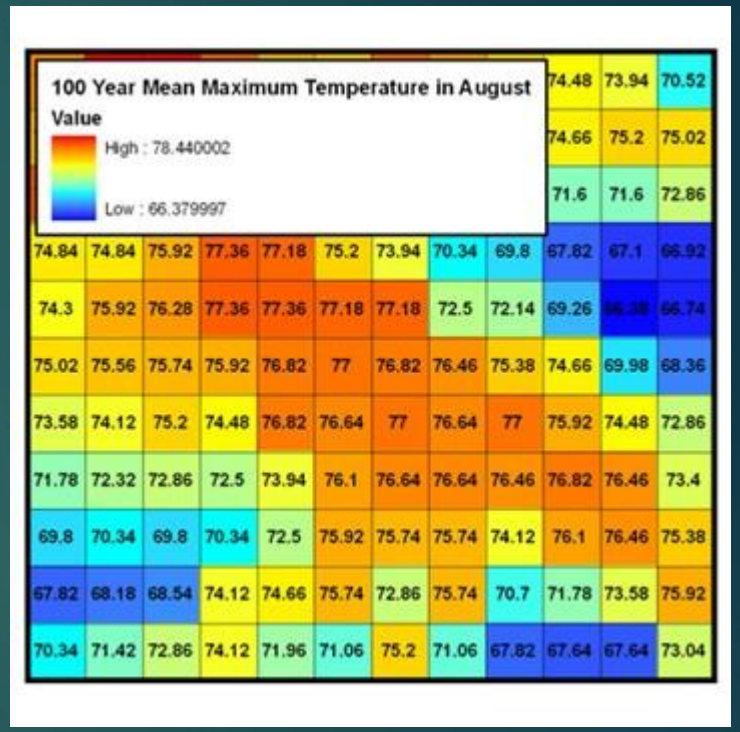
4.3. Different Raster Data - Representation

CRU – Temperature Data.

Each cell representing Temperature difference.







4.4. GRID Value/CELL Value

Raster model uses grid cells for representing continuous phenomena

Cell Value

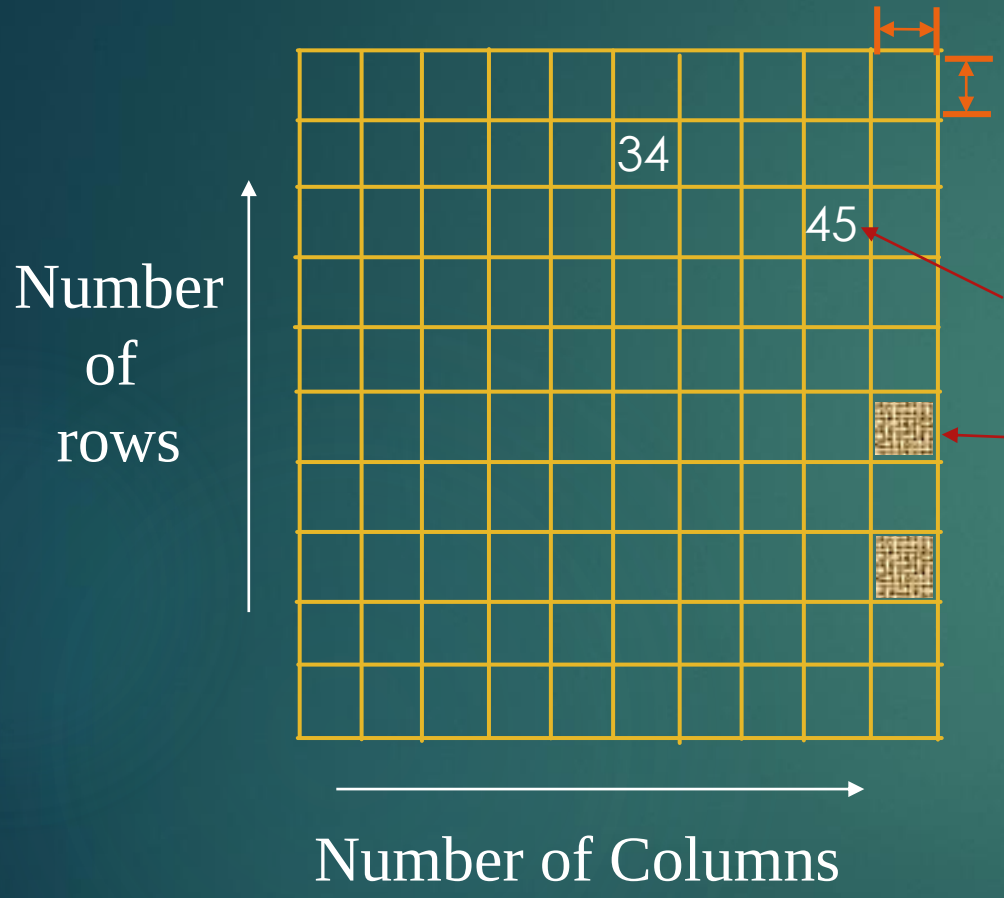
Positive or negative, integer Or floating point

Integer values are best used To represent categorical (discrete) data (Example Land Use Land Cover)

Floating-point values to represent Continuous surfaces (example Digital Elevation Model

Cells can also have a No Data Value to represent the Absence of data.

5. Definition of a Grid and Cell Size



Cell size Determines Resolution: example 30mx30m Spatial Resolution

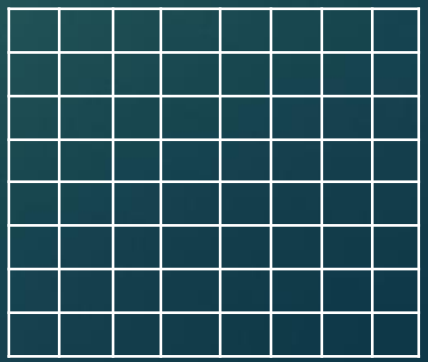
Cells that Have data – Grid Value

Cells that Have No data

80 meter Spatial Resolution (4 cells – Less Spatial Details)

30 meter Spatial Resolution (16 cells)

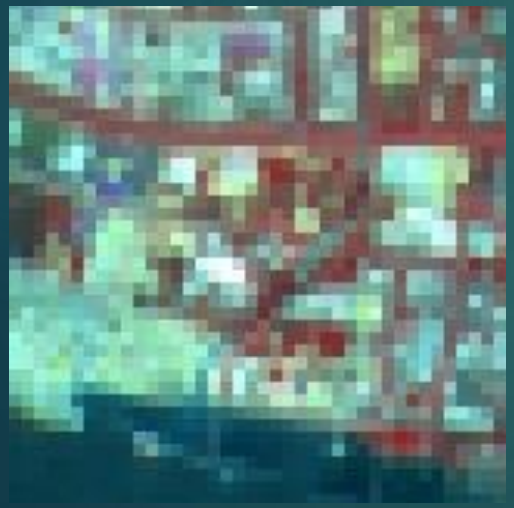
10 meter Spatial Resolution (64 cells – more spatial detaisl)



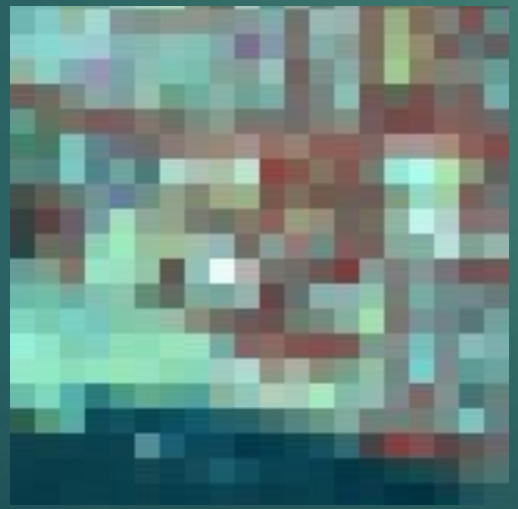
5.1. Same extent Raster Image at 4 different Resolution

Pixel size of 10 m
Width: 160 pixels
Height: 160 Pixels

Pixel size of 20 m
Width: 80 pixels
Height: 80 Pixels



Pixel size of 40 m
Width: 40 pixels
Height: 40 Pixels

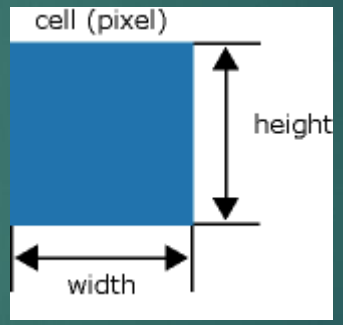


Pixel size of 80 m
Width: 20 pixels
Height: 20 Pixels

There is a common relation between spatial detail (resolution) and data volume in raster data sets.

Smaller cells provide greater spatial detail, but at the cost of larger data sets.

Bigger cells provide less spatial details with smaller data size and lesser cost.



High spatial resolution, the number of pixel increase

Low spatial resolution, the number of pixel decrease

Source: ESRI, 2016

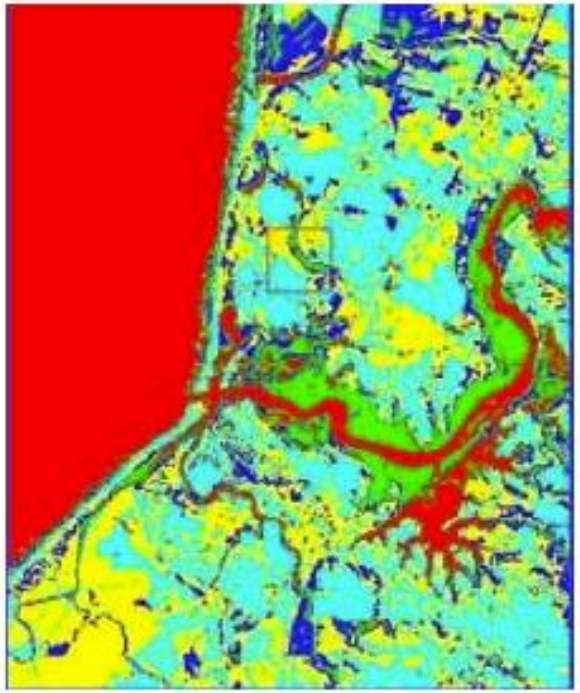
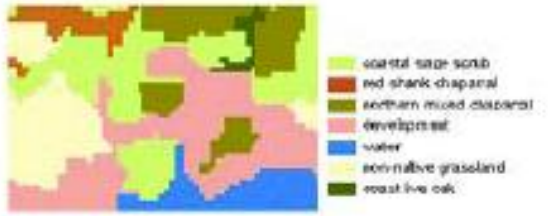
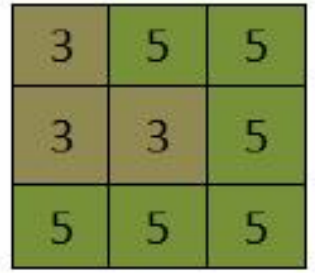
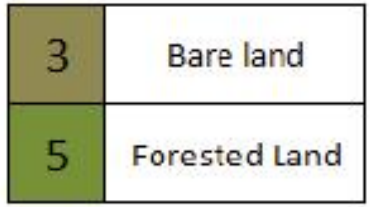
6. Discrete Data (integer)

Raster Data

Attributes of Raster Data – Discrete data

Value	Count	Land use type
1	7	Agriculture
3	10	River
5	8	Built-Up
7	8	Grassland

6.1. Representing Discrete Values

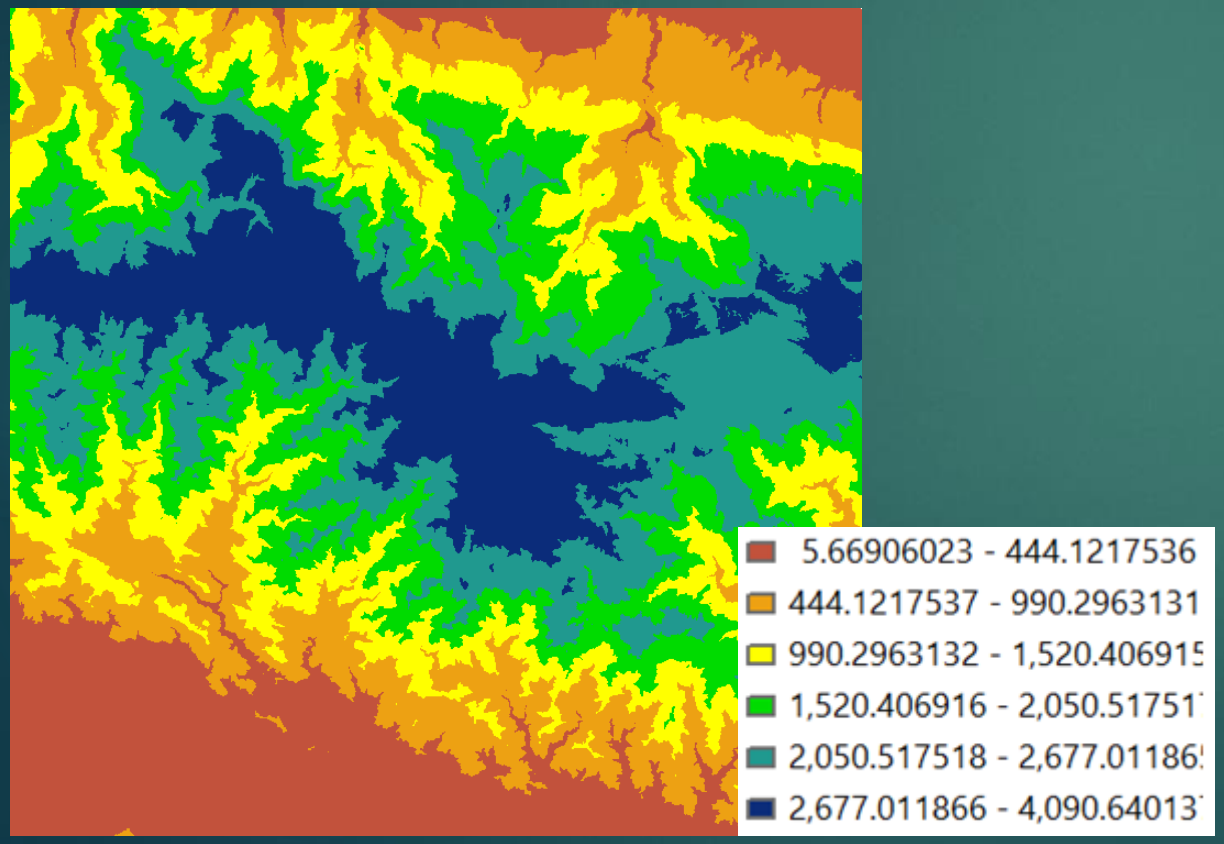
			 
Satellite Image of a Coastline	*Classification Result*	*Close-up of Classification Result*	*Raster Values Exposed*

Source: https://vector.geospatial.science/textbook/section-three-raster-data

6.2. GRID Value/CELL Value Representing Floating Values

A geographical space is define in a grid as a matric of square cells.

For a particular unit of space, each cell holds a numeric value that measures a geographic attribute.



23,90	768.90	4080.67	212.67	7.79
200.09	78.56	256.89	67.12	482.01
12.345	1006.65	87.09	19.67	3945.67
3000.04	310.78	78.78	1450.87	800.34
95.67	78.90	342.98	186.07	495.34

Image Source: T.Sekac (2022)

7. Advantages and Disadvantages of Raster Data Model

Advantages:

- The raster data can be very easy to produce
- The workflow and analysis are easy
- Represents Continuous features

- Disadvantages:

- Hard to represent objects less than cell size.
- Finer resolution generates huge data.
- Highly generalized representation of discrete features.
- Limited interactivity and more primitive analysis algorithm

Alam, 2012

8. Introduction to Types of spatial analysis (Raster):

- Mapping distance:
  - Cost Weighted Distance
  - Straight line/Euclidean distance
  - Shortest (or least-cost) path.
- Interpolating to raster:
  - Triangulated Irregular Network (TIN)
  - Thiessen polygon
  - IDW
  - Spline
  - Kriging
- Surface analysis:
  - Contour
  - Slop
  - Aspect
  - Hill shade
  - Viewshed
  - cut/fill
- Statistics:
  - Cell statistics
  - Neighborhood statistics
  - Zonal statistics
- Reclassification:
- Overlay:
- Raster Calculator (map algebra):
- Conversion of vector and raster:

8.1. Mapping Distance

Overview:

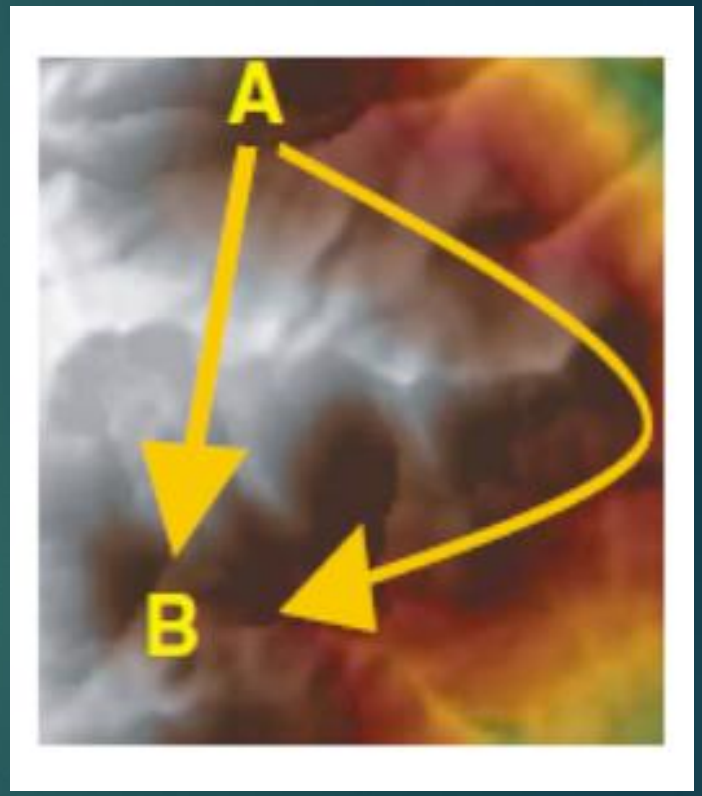
Mapping Distance allows you to identify which cells belong to which source based on straight line distance function or cost weighted distance function.

The Straight Line Distance function measures the straight line distance from each cell to the closest or nearest source.

The ***Cost Weighted Distance*** function modifies the Straight Line Distance by some other factor, which is a cost to travel through any given cell.

For example, it may be shorter to climb over the mountain to the destination, but it is faster to walk around it. Cost can be money, time, or preference.

The output from cost weighted distance function is further used to compute **least cost path** to come up with most cost effective path/route, for example to construct and connect roads between point A and point B.



Source: ESRI, 2001,2002

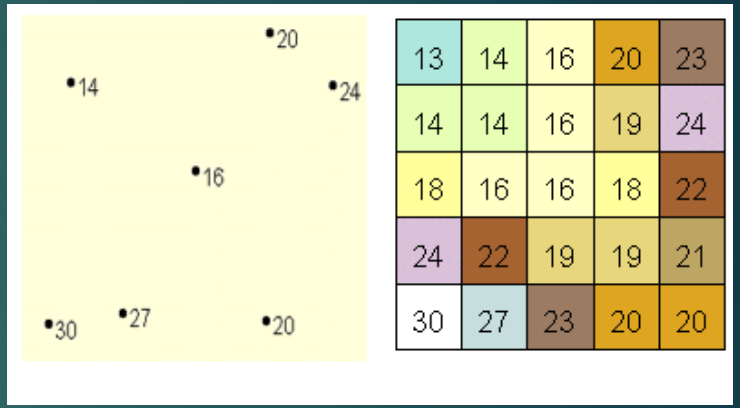
8.2. Interpolation to Raster

Overview:

- Interpolation predicts values for cells in a raster from a limited number of sample data points.

- It can be used to predict unknown values for any geographic point data:

- elevation,
- rainfall,
- Wi-Fi Strength
- chemical concentrations,
- noise levels, .



- The illustration on the left shows a point dataset of known rainfall-level values. The illustration on the right shows a raster interpolated from these points.

- Base on different application, multiple interpolation techniques are used:
  - Triangulated Irregular Network (TIN)
  - Thiessen polygon
  - IDW
  - Spline
  - Kriging

Source: ESRI, 2001,2002

8.3. Surface Analysis

Overview:

In GIS and spatial analysis;

- A **surface** is a vector or raster dataset that contains an attribute value for every locale throughout its extent.

- All raster datasets are surfaces, but not all vector datasets are surfaces.
- Surface analysis are mainly carried out in Raster data sets.
- Why vector because, vector are used to create surfaces also.
- Surfaces are commonly used in a geographic information system (GIS) to visualize phenomena such as

- elevation,
- temperature,
- slope,
- aspect,
- rainfall, and many more.

Source: Saylor Academy (2012), Chapter 8.3

8.4. Raster Statistics

Overview:

Raster Statistics means to perform statistical calculation between multiple raster.

For example;

- use the tools to analyze a certain phenomenon over time to determine the yearly mean or minimum rainfall for a 25-year period for allocating agricultural aid money or to
- calculate the range of temperatures between years to understand global warming.
- Calculate relationship between two raster (Temperature – NDVI relationship)
- Further to that, variability, majority, mean, minimum, maximum, etc… are also referred as raster statistical calculations.
- There are several geoprocessing tools that are use to calculate statistics (and the histogram) for raster datasets.
- Chain them together with other tools in models or create a Python script.

8.5. Reclassification

Overview:

One way to convert surface data into more usable information for an analysis is to reclassify the surface

Reclassification is the process of reassigning one or more values in a raster dataset to new output values.

The Reclassify tool is available in the Spatial Analyst extension in both ArcMap and ArcGIS Pro.

Reclassification can be seen as a method of:

1. Raster analysis
2. Maps preparation and generation.

8.6. Overlay Analysis

Overview:

- Overlay analysis can be performed in both Raster and vector data.
- Overlay analysis is operation in GIS for superimposing the multiple layer of datasets that representing different themes together for analyzing or identifying relationship of each layer.
- Overlay analysis represent the composite map by the combination of different attribute and geometry of datasets or entity.
- Overlay is the operations of comparing variables among multiple coverages.
- In the overlay analysis new spatial data sets are created by merging data from two or more input data layers.

Source: https://www.shivajicollege.ac.in/sPanel/uploads/econtent/7da3d1789fe2bb41d4875946dfcb0243.pdf

8.7. Raster Calculator

Overview:

- Raster calculator is one of the tool in GIS environment that are used to execute multiple raster related calculations.
- The new raster images are outputted from calculation.
- Complete raster equations and formulas can be involve for calculations to solve geographical problems
- The listed Raster data **Layers and variables** are use in the expression.
- The available Numerical values and mathematical operators are added to the expression during calculation of raster and variables.
- A list of commonly used conditional and mathematical tools is provided, allowing to easily add them to the expression.

8.8. Conversion of Vector and Raster

Overview:

- Conversion tool in ArcGIS can be use to perform such task.

- Any conversion has a special purpose to serve and special task to perform after conversion:

- Change pixel size
- Overlay
- Calculate area
- Modify attribute table
- Etc..

- In conversion the concept of:

- Rasterization
- Vectorization are perform.

9. References

Saylor Academy (2012). Surface Analysis: Spatial Interpolation. Retrieve from: https://saylordotorg.github.io/text_essentials-of-geographic-information-systems/s12-03-surface-analysis-spatial-inter.html

Alam, M., (2012). Modeling Wetland Connectivity And Vulnerability To Wetland corridor Loss: Retrieve from: https://www.researchgate.net/publication/315694261_MODELING_WETLAND_CONNECTIVITY_AND_VULNERABILITY_TO_WETLAND-CORRIDOR_LOSS

Maputo, Mozambique (2018), Geospatial Data Models. GIS for Spatial Planning
Training for Ministry of Transport Mozambique (Lecture Notes). *Geoinformation and Sectoral Statistics Section* - United nations Economic commissions for Africa.

Bolstad, P., (2016). GIS Fundamentals (Book). A First Text on Geographic Information Systems
Fifth Edition.

ESRI (2016). What is Raster Data?. Retrieve from: https://desktop.arcgis.com/en/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm

ESRI Book (2001 – 2002) Using ArcGIS Spatial Analist. Retrieve from: http://downloads2.esri.com/support/documentation/ao_/Using_ArcGIS_Spatial_Analyst.pdf

LearnGIS (2022). Section Three – Raster Data. Retrive from: https://vector.geospatial.science/textbook/section-three-raster-data