Geographical Information System

Unit 5: Spatial Analysis

5.1 Vector data analysis: geo-processing, overlay analysis, buffering, network analysis

5.2 Raster analysis: local operations, focal operations, zonal operations, re-sampling, mosaic and clip, distance measurement

5.2 Spatial interpolation techniques, geo-statistics, GIS modeling

5.3 GIS programming and customization: Opening and exploring Model Builder, Python script tools, Customizing QGIS with Python

Geographical Information System

Introduction

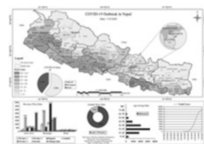
(Spatial analysis)

Geographical Information System

Spatial analysis: is a fundamental component of a GIS that allows for an in-depth study of the topological and geometric properties of a dataset or datasets.

All GIS provide functions for analyses of data chosen and for storing results of such analysis. Data may be analyzed at various levels:

1. Store attribute data in table: Are stored for presentation in reports or for use in other computer system.
2. Operation on geometric data: Are performed on geometric data, either in search mode or for computational purpose.
3. Operations on attribute tables: Like arithmetic, Boolean, statistical are performed in attribute tables.
4. Both Geometry and Attribute table: Are used jointly to:
   a. Compile new set of data based on original and derived attributes.
   b. Compile new sets of data based on geographic relationships.



Geographical Information System

Spatial analysis:

Geographic data are mainly categories in two types Vector and Raster data and accordingly the spatial analysis can be categories in two types.

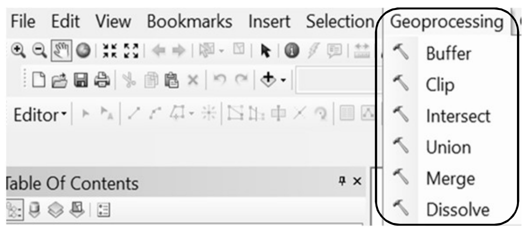
1. Geospatial Analysis: Vector operation
   a. Single Layer Analysis
   b. Multiple Layer Analysis
2. Geospatial Analysis : Raster Data
   a. Basic Geoprocessing with Raster
      i. Single Layer Analysis
      ii. Multiple Layer Analysis
   b. Scale of Analysis
      i. Local Operations
      ii. Neighborhood Operations
      iii. Zonal Operations
      iv. Global Operations
3. Surface Analysis:
   a. Spatial Interpolation-Creating Surfaces
   b. Terrain Mapping

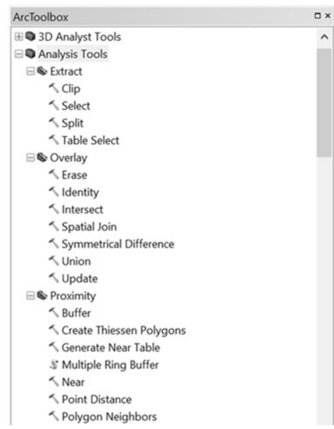
Note: Interpolation - calculated based on the numbers or items.

Geographical Information System

5.1 Vector data analysis:

- Geo-processing,
- Overlay Analysis,
- Buffering,
- Network Analysis





Geographical Information System

Geo-processing:

- Geoprocessing is a suite of tools provided by many geographic information system (GIS) software packages that allow the user to automate many of the routine tasks associated with manipulating GIS data.
- Geoprocessing usually involves the input of one or more feature datasets, followed by a spatially explicit analysis, and resulting in an output feature dataset.

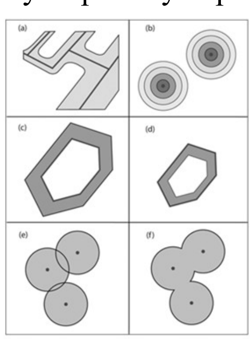
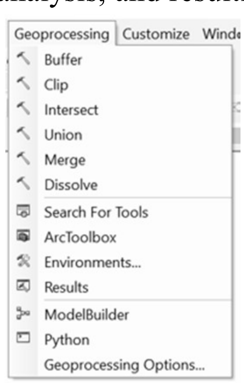
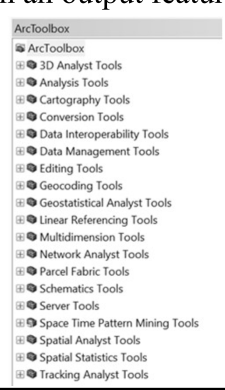


Figure "Additional Buffer Options around Red Features: (a) Variable Width Buffers, (b) Multiple Ring Buffers, (c) Doughnut Buffer, (d) Setback Buffer, (e) Nondissolved Buffer, (f) Dissolved Buffer"



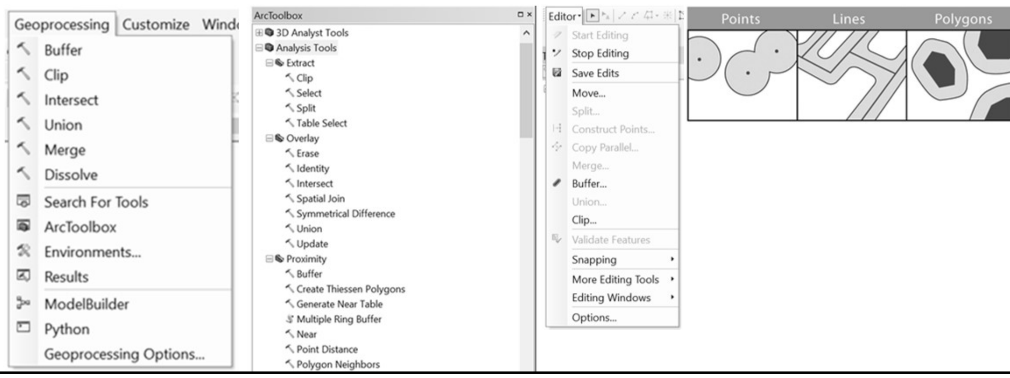


Geographical Information System

Buffering - Single Layer Analysis

Single layer analyses are those that are undertaken on an individual feature dataset.

Buffers are common vector analysis tools used to address questions of proximity in a GIS and can be used on points, lines, or polygons. Figure "Buffers around Red Point, Line, and Polygon Features”.



Geographical Information System

Buffering- Single Layer Analysis

- Buffers are frequently used to create zones of a specified width around points, lines, and polygons.
- Vector buffering options include (a) constant or variable widths, (b) multiple rings, (c) doughnuts, (d) setbacks, (e) Nondissolve and (f) dissolve.
- Common single layer geoprocessing operations on vector layers include dissolve, merge, append, and select.

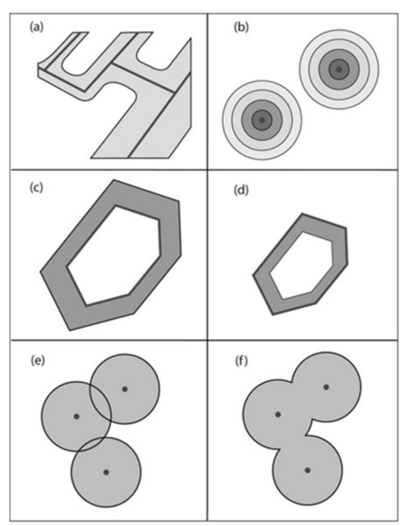


Figure "Additional Buffer Options around Red Features: (a) Variable Width Buffers, (b) Multiple Ring Buffers, (c) Doughnut Buffer, (d) Setback Buffer, (e) Nondissolved Buffer, (f) Dissolved Buffer"

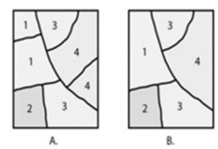
Geographical Information System

Spatial analysis - Single Layer Analysis

Dissolve operation:

The **dissolve** operation combines adjacent polygon features in a single feature dataset based on a single predetermined attribute. For example, part (a) of Figure "Single Layer Geoprocessing Functions" shows the boundaries of seven different parcels of land, owned by four different families (labeled 1 through 4). The dissolve tool automatically combines all adjacent features with the same attribute values. The result is an output layer with the same extent as the original but without all of the unnecessary, intervening line segments. The dissolved output layer is much easier to visually interpret when the map is classified according to the dissolved field.

(a) Dissolve



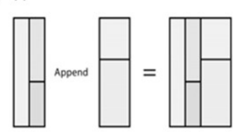
Geographical Information System

Spatial analysis - Single Layer Analysis

Append operation:

The append operation creates an output polygon layer by combining the spatial extent of two or more layers (part (b) of Figure "Single Layer Geoprocessing Functions"). For use with point, line, and polygon datasets, the output layer will be the same feature type as the input layers (which must each be the same feature type as well). Unlike the dissolve tool, append does not remove the boundary lines between appended layers (in the case of lines and polygons). Therefore, it is often useful to perform a dissolve after the use of the append tool to remove these potentially unnecessary dividing lines. Append is frequently used to mosaic (a pattern or picture made using many small pieces) data layers to create a single map for analysis and/or display.

(b) Append



Geographical Information System

Spatial analysis - Single Layer Analysis

Select and Merge operation:

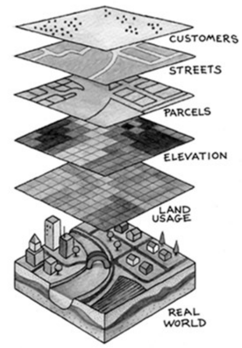
The select operation creates an output layer based on a user-defined query that selects particular features from the input layer. The output layer contains only those features that are selected during the query. For example, a city planner may choose to perform a select on all areas that are zoned "residential" so he or she can quickly assess which areas in town are suitable for a proposed housing development.

Finally, the merge operation combines features within a point, line, or polygon layer into a single feature with identical attribute information. Often, the original features will have different values for a given attribute. In this case, the first attribute encountered is carried over into the attribute table, and the remaining attributes are lost. This operation is particularly useful when polygons are found to be unintentionally overlapping. Merge will conveniently combine these features into a single entity.

Geographical Information System

Overlay:

- Overlay is a GIS operation that superimposes multiple data sets (representing different themes) together for the purpose of identifying relationships between them.. An overlay creates a composite map by combining the geometry and attributes of the input data sets.
- Overlay Operations involve combining spatial and attribute data from two or more spatial data layers. "Stacking data"
- Very powerful and popular operations



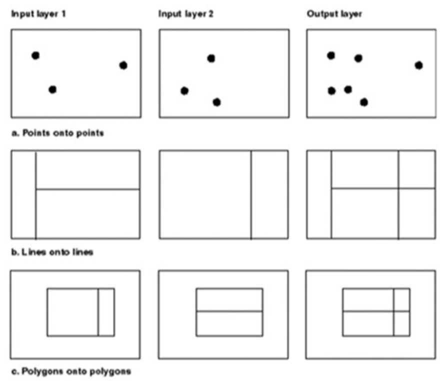
Examples of overlays?

Overlays require that data be in the same coordinate system

Geographical Information System

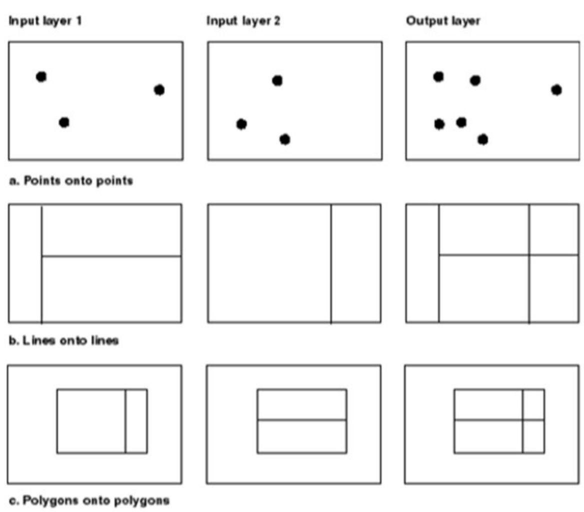
Overlay: The basics

- Two of more input layers
- Can be any combination of point, line or polygon
- Spatial data and attribute data are both combined to create an output layer
- Feature type of the output layer will depend on the combinations of the input layers



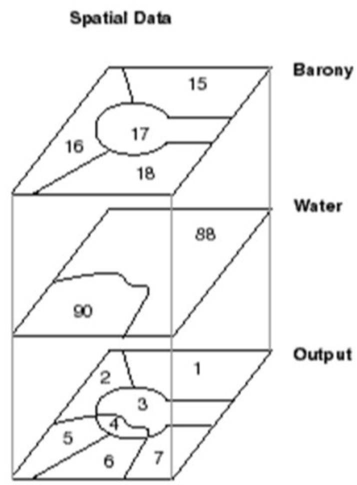
Geographical Information System

Overlay: The basics



Geographical Information System

Spatial and attribute data:



Attribute Data

Barony

BARONY-ID	NAME
15	Antrim
16	Glenarm
17	Belfast
18	Toome

Water

WATER-ID	WATER
88	LAND
90	WATER

Output

OUTPUT-ID	BARONY-ID	NAME	WATER-ID	WATER
1	15	Antrim	88	LAND
2	16	Glenarm	88	LAND
3	17	Belfast	88	LAND
4	17	Belfast	90	WATER
5	16	Glenarm	90	WATER
6	18	Toome	90	WATER
7	18	Toome	88	LAND

Geographical Information System

Overlay: Multiple Layer Analysis

Map overlay refers to a group of procedures and techniques used in combining information from different data layers. This is an important capability of most GIS environments. Map overlays involve at least two input layers and result in at least one new output layer. A basic set of overlay tools include clipping, intersecting and unioning.

Overlay is the core part of GIS analysis operation. It combines several spatial features to generate new spatial elements. In other word, overlay can be defined as a spatial operation, which combines different geographic layers to generate new information. Overlay is done using Arithmetic, Boolean, and Relational operators, and is performed in both vector and raster domain.

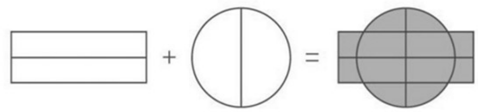
1. Vector overlay:

This overlay operation combines the geometries and attribute of two feature layers to create the output. The geometric of output represent the geometric intersection of features from the input layers.

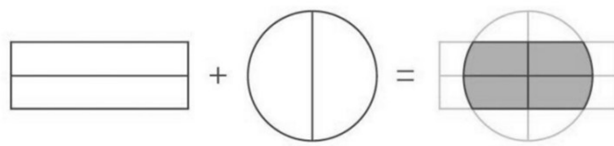
Geographical Information System

Overlay:

Union: Preserves all feature from the inputs. The area extent of output combines area extends of both input layers. It requires that both input layers be polygon layers.



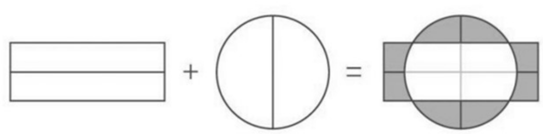
Intersection: It preserves only those features that fall within the area extent common to inputs. We preferred it because, any feature on its output has attribute data from both of its inputs.



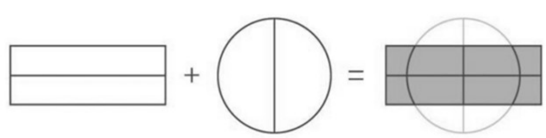
Geographical Information System

Overlay:

Symmetrical difference: Preserve features that fall within the area extent that is common to only one of the input. It is opposite to intersect in terms of outputs area extent. This method requires that both input layers be polygon layers.



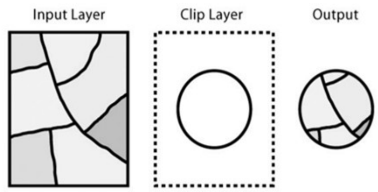
Identity: Identity preserve only features that fall within the area of layer defined as the input layer. The other layer is called identity layer. Input layer may contain point, lines, polygon and identity layer is polygon layer.



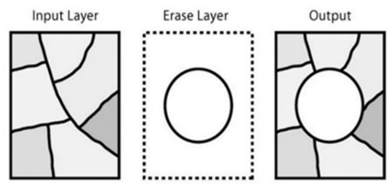
Geographical Information System

Overlay:

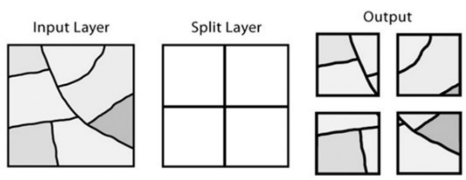
Clip:



Erase:

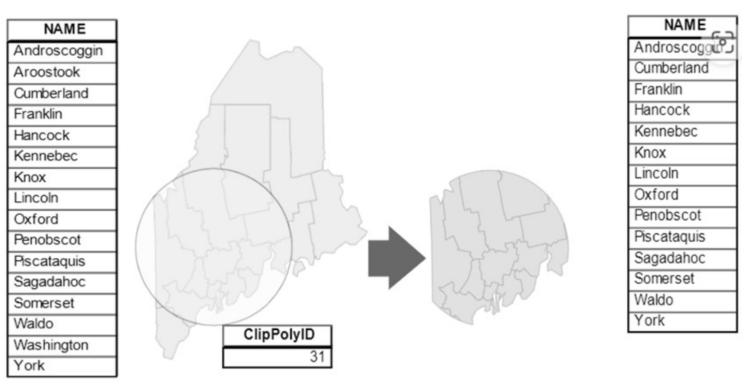


Split:



Geographical Information System

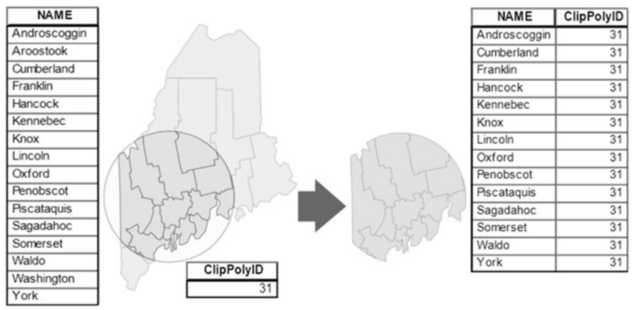
Clip: Clipping takes one GIS layer (the clip feature) and another GIS layer (the to-be-clipped input feature). The output is a clipped version of the original input layer. The output attributes table is a subset of the original attributes table where only records for the clipped polygons are preserved.



Note that the output layer is limited to the Name polygon geometry and its attributes (and does not include the clipping circle polygon).

Intersect:

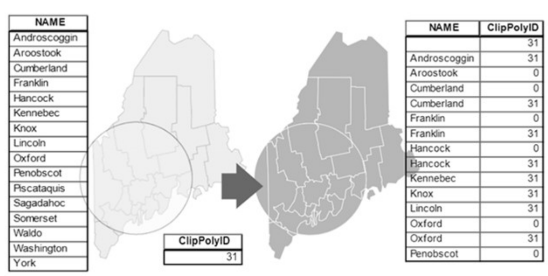
Intersecting takes both layers as inputs then outputs the features from both layers that share the same spatial extent. Note that the output attribute table inherits attributes from both input layers (this differs from clipping where attributes from just one layer are carried through).



The NAME polygon layer is intersected with the circle polygon. The output layer combines both intersecting geometries and attributes.

Union

Unioning overlays both input layers and outputs all features from the two layers. Features that overlap are intersected creating new polygons. This overlay usually produces more polygons than are present in both input layers combined. The output attributes table contains attribute values from both input features (note that only a subset of the output attributes table is shown in the following figure).



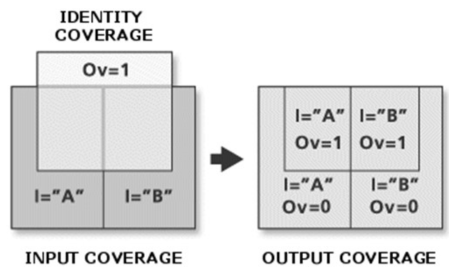
NAME polygon layer is unioned with the circle polygon. The output layer combines both (complete) geometries and attributes. Where spatial overlaps do not occur, most software will either assign a NULL value or a 0.

Geographical Information System

Identity tool

To use the Identity tool, the input coverage can be a point, line, or polygon coverage. The output coverage will be the same feature type as the input coverage.

All features of the input coverage will be preserved in the output coverage. This means that the input coverage acts like a cookie cutter on the identity coverage.



Geographical Information System

Identity tool

The three feature types are affected differently by the identity coverage:

- Polygons: Input coverage arcs are split at their intersections with polygons of the identity coverage.
- Lines: Identity coverage arcs are used to split input coverages where they overlap.
- Points: All input coverage points are saved in the output coverage, and the output coverage PAT file lists the identity coverage polygon within which each point falls.

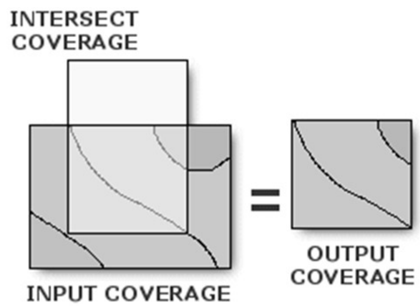
The Intersect and Union tools are similar to the Identity tool. The only difference is the features that remain in the output coverage.

Geographical Information System

Intersect tool

To use the Intersect tool, the input coverage can be a point, line, or polygon coverage. The output coverage will be the same feature type as the input coverage.

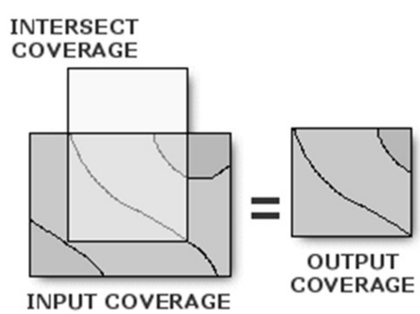
Only those features contained by polygons in the intersect coverage will be preserved in the output coverage.



Geographical Information System

Intersect tool

The Intersect tool is similar to the Clip tool; however, the Clip tool does not transfer any attributes from the clip coverage to the output. The Intersect tool is also similar to the Identity and Union tools. The only difference is the features that remain in the output coverage.

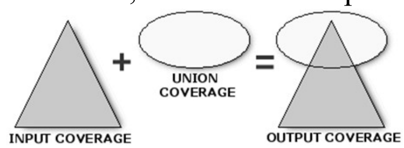


Geographical Information System

Union tool

Only polygon coverages can be used by the Union tool.

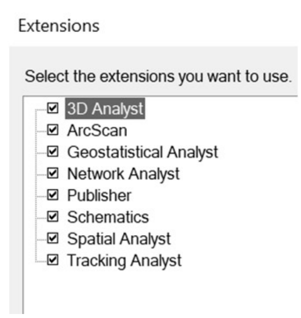
The output coverage contains the polygon features from both the input and union coverage. Unlike Identity and Intersect, Union never clips any data.

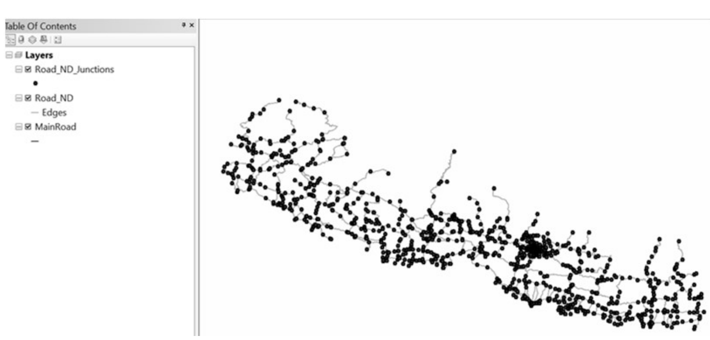


The Union tool is similar to the Intersect and Identity tools. The only difference is the features that remain in the output coverage.

Geographical Information System

Network Analysis





Geographical Information System

Network Analysis: Route Optimization and Shortest Path

Route Optimization and Shortest Path,

Route optimization in GIS (Geographic Information Systems) is a process of finding the most efficient or optimal route between two or more locations on a map, considering various constraints and criteria.

It is a crucial aspect of spatial analysis and is used in various real-world applications, such as logistics, transportation, emergency services, urban planning, and field workforce management.

The main goal of route optimization is to minimize travel time, distance, cost, or any other relevant criteria while ensuring that the constraints of the problem are met.

Geographical Information System

Route Optimization and Shortest Path:

Some common constraints in route optimization may include:

Time windows: Certain locations may have specific time frames within which a vehicle or a person can visit them. For example, a delivery truck may be allowed to deliver goods to a store only during specific hours.

Vehicle capacity: In logistics, vehicles may have limitations on the maximum quantity of goods they can carry. Route optimization considers this capacity constraint to ensure that the total demand for goods along the route does not exceed the vehicle's capacity.

Traffic conditions: Real-time traffic data or historical traffic patterns can be incorporated into route optimization algorithms to avoid congested areas and choose routes with less traffic, thereby minimizing travel time.

Road network restrictions: Some roads may have weight limits, height restrictions, or may be one-way, which needs to be taken into account during route planning.

Cost factors: Different routes may have varying costs associated with them, such as tolls, fuel costs, or vehicle maintenance expenses. Route optimization algorithms can consider these factors to find cost-effective routes.

Geographical Information System

Route Optimization and Shortest Path:

By leveraging GIS data and route optimization, businesses and organizations can enhance their operational efficiency, reduce transportation costs, improve response times, and make better-informed decisions in their spatial planning and resource allocation.

Example: Find the shortest Network Analysis in Arc GIS

Geographical Information System

2. Raster overlay:

- Raster overlay is simpler than vector overlay and can be carried out directly on cell values.
- It is more efficient than vector overlay, as extent of calculation is much less.
- Both raster layer must have identical geometry i.e. cell size must be same, there must be no relative rotation of transfer between grids.
- Attribute are representation of thematic layers.
- There is no need to distinguish between polygons, lines, points, because all raster data comprise cells.
- Arithmetic, logical, statistical operation may be performed directly during overlay process.
- Deviation is carried by transformation and resampling to the same cell size.
- New composite cells are composed from original cells and registered as a new thematic layer.
- There is no formation of smaller errorneous polygon.

Geographical Information System

5.2 Raster analysis

- Local operations,
- Focal operations,
- Zonal operations,
- Re-sampling,
- Mosaic and clip,
- Distance measurement

Spatial analysis:

Geographic data are mainly categories in two types Vector and Raster data and accordingly the spatial analysis can be categories in two types.

1. Geospatial Analysis: Vector operation
   a. Single Layer Analysis
   b. Multiple Layer Analysis
2. Geospatial Analysis : Raster Data
   a. Basic Geoprocessing with Raster
      i. Single Layer Analysis
      ii. Multiple Layer Analysis
   b. Scale of Analysis
      i. Local Operations
      ii. Neighborhood Operations
      iii. Zonal Operations
      iv. Global Operations
3. Surface Analysis:
   a. Spatial Interpolation-Creating Surfaces
   b. Terrain Mapping

Note: Interpolation - calculated based on the numbers or items.

Geographical Information System

Spatial analysis:

Geospatial Analysis : Raster Data

a. Basic Geoprocessing with Raster
   i. Single Layer Analysis
   ii. Multiple Layer Analysis

b. Scale of Analysis
   i. Local Operations
   ii. Neighborhood Operations
   iii. Zonal Operations
   iv. Global Operations

Surface Analysis:

a. Spatial Interpolation-Creating Surfaces
b. Terrain Mapping

Note: Interpolation - calculated based on the numbers or items.

Geographical Information System

Raster Data Analysis:

The raster data model uses a regular grid to cover the space and the value in each grid cell to represent the characteristic of a spatial phenomenon at the cell location. This simple data structure of a raster with fixed cell locations not only is computationally efficient, but also facilitates a large variety of data analysis operations. This is why raster data are typically used in geographic information system (GIS), involving heavy computation such as building environmental models.

In contrast with vector data analysis, which uses points, lines, and polygons, raster data analysis uses cells and rasters. Raster data analysis can be performed at the level of individual cells, or groups of cells, or cells within an entire raster. Some raster data operations use a single raster; others use two or more rasters. An important consideration in raster data analysis is the type of cell value. Statistics such as mean and standard deviation are designed for numeric values, whereas others such as majority (the most frequent cell value) are designed for both numeric and categorical values.

Geographical Information System

Raster Data Analysis: Four numbering systems in raster are the following:

Ratio: The ratio scale has a reference point, usually zero, and the numbers within the scale are comparable. For example, elevation values are ratio numbers, and an elevation of 50 meters is half as high as 100 meters.

Interval: The values in an interval scale are relative to one another; however, there is not a common reference point. For example, a pH scale is of type interval, where the higher the value is above the neutral value of 7, the more alkaline it is, and the lower the value is below 7, the more acidic it is. However, the values are not fully comparable. For example, a pH of 2 is not twice as acidic as a pH of 4.

Ordinal: An ordinal scale establishes order, such as who came in first, second, and third in a race. Order is established, but the assigned order values cannot be directly compared. For example, the person who came in first was not necessarily twice as fast as the person who came in second.

Nominal(categorical): There is no relationship between the assigned values in the nominal scale. For example, land-use values, which are nominal values, cannot be compared to one another. A land use of 8 is probably not twice as much as a land use of 4.

Geographical Information System

Overview of Raster Overlay:

Raster overlay, also known as raster combination or raster stacking, is a fundamental concept in GIS. It involves combining or overlaying multiple raster datasets to create a new composite raster layer.

Each input raster contains pixel values that represent specific attributes, such as elevation, land cover, temperature, or any other geographic phenomenon. The process of raster overlay allows you to integrate and analyze these attributes simultaneously.

Geographical Information System

Overview of Raster Overlay:

Raster overlay typically involves two or more input raster datasets, each with its own set of attributes. These attributes are represented by pixel values arranged in a grid pattern. Examples of raster datasets include satellite imagery, digital elevation models (DEMs), and thematic maps.

Cell Resolution and Alignment: Raster overlay requires that the input raster datasets have the same cell size (resolution) and align properly. If the resolutions are different, resampling may be necessary to adjust the cell sizes. Alignment ensures that corresponding cells in different layers represent the same geographic location.

Geographical Information System

Overview of Raster Overlay:

Data Compatibility: Before performing raster overlay, it's important to ensure that the data types, coordinate systems, and projections of the input raster layers are compatible. If they are not, preprocessing steps may be needed to standardize the data.

Cell Resolution and Alignment: Raster overlay requires that the input raster datasets have the same cell size (resolution) and align properly. If the resolutions are different, resampling may be necessary to adjust the cell sizes. Alignment ensures that corresponding cells in different layers represent the same geographic location.

Output Raster: The result of the raster overlay is a new composite raster layer that represents the combined attributes of the input layers. The pixel values in the output raster are determined by the specific overlay operation performed.

Geographical Information System

Overview of Raster Overlay:

Overlay Operations: There are several overlay operations that can be performed to combine raster datasets. Some common operations include:

- **Summation or Addition**: Adding the pixel values of corresponding cells in two or more raster layers.
- **Subtraction**: Subtracting the pixel values of one raster layer from another.
- **Multiplication**: Multiplying the pixel values of corresponding cells in two or more raster layers.
- **Division**: Dividing the pixel values of one raster layer by another.
- **Boolean Operations**: Combining raster layers using logical operations (AND, OR, NOT, XOR) to create binary outputs.
- **Zonal Statistics**: Calculating statistics (e.g., mean, sum) for values within specified zones or areas of interest.

Geographical Information System

MAP ALGEBRA:

Map Algebra refers to a set of mathematical and logical operations performed on spatial data layers (maps) to generate new layers or to analyze and manipulate existing layers. It is a fundamental concept in GIS that allows you to combine, overlay, and analyze different spatial datasets to derive meaningful information and insights.

Map Algebra operations involve applying mathematical operators and functions to raster datasets (gridded representations of geographic information). These operations can include simple arithmetic operations (addition, subtraction, multiplication, division), as well as more complex operations like convolution, focal statistics, and reclassification.

Geographical Information System

MAP ALGEBRA:

Map Algebra refers to a set of mathematical and logical operations performed on spatial data layers (maps) to generate new layers or to analyze and manipulate existing layers. It is a fundamental concept in GIS that allows you to combine, overlay, and analyze different spatial datasets to derive meaningful information and insights.

Map Algebra operations involve applying mathematical operators and functions to raster datasets (gridded representations of geographic information). These operations can include simple arithmetic operations (addition, subtraction, multiplication, division), as well as more complex operations like convolution, focal statistics, and reclassification.

Geographical Information System

MAP ALGEBRA: Some common operations in Map Algebra include:

- **Overlay Operations:** These involve combining multiple layers to create a new layer. Examples include union, intersection, difference, and xor operations.
- **Local Operations:** These operations are performed on a cell-by-cell basis and involve applying mathematical functions or operators to the values of corresponding cells in multiple input layers. Examples include addition, subtraction, multiplication, division, and more.
- **Zonal Operations:** These operations involve aggregating or summarizing data within defined zones or regions. For example, calculating the mean elevation within different watersheds.
- **Focal Operations:** These operations involve applying a function to a neighborhood of cells around each cell in a raster. Common examples include calculating the mean, maximum, minimum, or standard deviation within a specified radius.

Geographical Information System

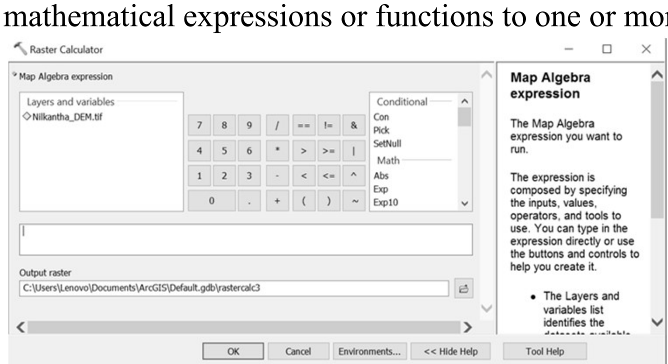
MAP ALGEBRA: Some common operations in Map Algebra include:

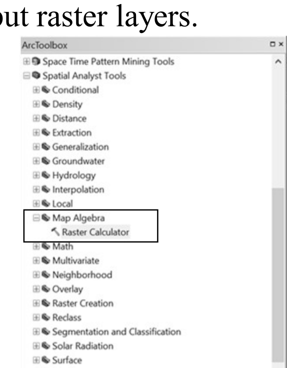
- **Reclassification:** This operation involves reassigning values within a raster based on defined criteria. It is often used to simplify or recategorize data.
- **Boolean Operations:** These operations involve logical comparisons between raster layers, resulting in binary outputs. Examples include AND, OR, and NOT operations.
- **Distance and Cost Surfaces:** These operations involve calculating distance or cost values from specific features, such as calculating travel time or cost from a set of points.

Geographical Information System

Raster Calculator:

Raster Calculator is a tool commonly used in GIS software to perform mathematical operations on raster datasets. Raster datasets consist of cells or pixels arranged in a grid, where each cell contains a value representing a specific attribute, such as elevation, temperature, land cover type, etc. The Raster Calculator allows you to create new raster layers by applying various mathematical expressions or functions to one or more input raster layers.





Geographical Information System

Raster Calculator: Basic overview of how Raster Calculator works and how you might use it:

- **Input Raster Layers:** You start by specifying the raster layers (bands) you want to use in your calculations. These layers should have the same spatial extent (same geographic area covered) and cell resolution.
- **Mathematical Expressions:** You create mathematical expressions that define how you want to manipulate the values in the input raster layers. These expressions can involve basic arithmetic operations (addition, subtraction, multiplication, division), as well as more complex operations like trigonometric functions, conditional statements, and statistical functions.
- **Output Raster:** The Raster Calculator generates a new raster layer as output, where each cell's value is determined by evaluating the mathematical expression using the corresponding cell values from the input raster layers.

Geographical Information System

Raster Data Analysis:

Various types of data are stored in raster format. Raster data analysis, however operates only on raster data supported by a GIS package. Therefore, for some raster data, we must process them first before analysis. We will discuss the basic tools for raster data analysis.

- Data Analysis Environment - The analysis environment including the area for analysis and the output cell size.
- Four common types of raster data analysis - Local Operations, Neighborhood Operations, Zonal Operations, and Physical Distance Measures.
- Raster operations that do not fit into the common classification of raster data analysis.
- Map Algebra - which allows complex raster data operations.
- Overlay and Buffering - to compare vector and raster based operations.

Geographical Information System

Raster Data Analysis: Data Analysis Environment

A raster operation may involve two or more rasters, it is necessary to define the data analysis environment by specifying its area extent and output cell size.

The area extent for analysis may correspond to a specific raster, or an area defined by its minimum and maximum x-, y-coordinates, or a combination of rasters.

Given a combination of rasters with different area extents, the area extent for analysis can be based on the union or intersect of the rasters. The union option uses an area extent that encompasses all input rasters, whereas the intersect option uses an area extent that is common to all input rasters.

Geographical Information System

Raster Data Analysis: Data Analysis Environment

An analysis mask, either a feature layer or a raster, can also determine the area extent for analysis. An analysis mask limits analysis to its area coverage. For example, to limit soil erosion analysis to only private lands, we can prepare a mask of either a feature layer showing private lands or a raster separating private lands (e.g., with a cell value of 1) from others (e.g., with a cell value of no data).

We can define the output cell size at any scale deemed suitable. Typically, the output cell size is set to be equal to, or larger than, the largest cell size among the input rasters. This follows the rationale that the resolution of the output should correspond to that of the lowest-resolution input raster. For instance, if the input cell sizes range from 10 to 30 meters, the output cell size should be 30 meters or larger.

Geographical Information System

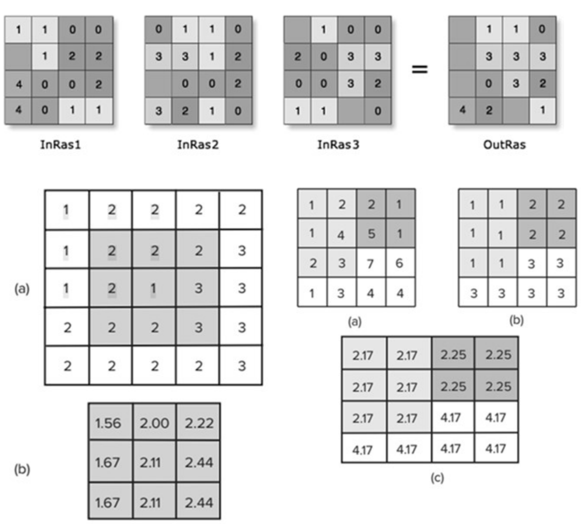
Data Analysis Environment

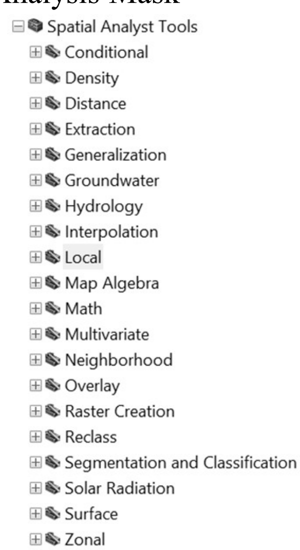
Local
Neighborhood
Zonal Operation

Geographical Information System

Raster Data Analysis: Data Analysis Environment

Area extent: Local, Neighborhood, Zonal Operation, Analysis Mask





Geographical Information System

Raster Data Analysis: Local Operations

Local operations are cell-by-cell operations. A local operation can create a new raster from either a single input raster or multiple input rasters. The cell values of the new raster are computed by a function relating the input to the output or are assigned by a classification table.

Local Operations with a Single Raster - A local operation computes each cell value in the output raster as a function of the cell value in the input raster at the same location. The function may involve a GIS tool, a mathematical operator, and/or a constant. A large number of mathematical operators are available in a GIS package.

For example, converting a floating-point raster to an integer raster is a simple local operation that uses the Integer operator to truncate the cell value at the decimal point on a cell-by-cell basis. Converting a slope raster measured in percent to one measured in degrees is also a local operation but requires a more complex mathematical expression.

Geographical Information System

Raster Data Analysis: Local Operations

Reclassification - A local operation, reclassification creates a new raster by classification. Reclassification is also referred to as recoding, or transforming, through lookup tables. Two reclassification methods may be used.

- The first method is a one-to-one change, meaning that a cell value in the input raster is assigned a new value in the output raster. For example, irrigated cropland in a land-use raster is assigned a value of 1 in the output raster.
- The second method assigns a new value to a range of cell values in the input raster. For example, cells with population densities between 0 and 25 persons per square mile in a population density raster are assigned a value of 1 in the output raster and so on.
- An integer raster can be reclassified by either method, but a floating-point raster can only be reclassified by the second method.

Geographical Information System

Raster Data Analysis: Local Operations

Reclassification - Reclassification serves three main purposes.

First, reclassification can create a simplified raster. For example, instead of having continuous slope values, a raster can have 1 for slopes of 0 to 10 percent, 2 for 10 to 20 percent, and so on.

Second, reclassification can create a new raster that contains a unique category or value such as slopes of 10 to 20 percent.

Third, reclassification can create a new raster that shows the ranking of cell values in the input raster. For example, a reclassified raster can show the ranking of 1 to 5, with 1 being least suitable and 5 being most suitable.

Geographical Information System

Local geoprocessing functions:

Some local operations do not involve statistics or computation. A local operation called Combine assigns a unique output value to each unique combination of input values.

Suppose a slope raster has three cell values (0 to 20 percent, 20 to 40 percent, and greater than 40 percent slope), and an aspect raster has four cell values (north, east, south, and west aspects). The Combine operation creates an output raster with a value for each unique combination of slope and aspect, such as 1 for greater than 40 percent slope and the south aspect, 2 for 20 to 40 percent slope and the south aspect, and so on.

Input Raster

456	416	364	326	243
448	364	315	276	218
359	325	268	234	164
306	296	201	133	44
274	231	184	65	5

↓

Output Raster (x10)

4560	4160	3640	3260	2430
4480	3640	3150	2760	2180
3590	3250	2680	2340	1640
3060	2960	2010	1330	440
2740	2310	1840	650	50

Figure: Local Operation on a Raster Dataset

Geographical Information System

Local geoprocessing functions:

- Local operations can be performed on single or multiple rasters. When used on a single raster, a local operation usually takes the form of applying some mathematical transformation to each individual cell in the grid. For example, a researcher may obtain a digital elevation model (DEM) with each cell value representing elevation in feet. If it is preferred to represent those elevations in meters, a simple, arithmetic transformation (original elevation in feet * 0.3048 = new elevation in meters) of each cell value can be performed locally to accomplish this task.

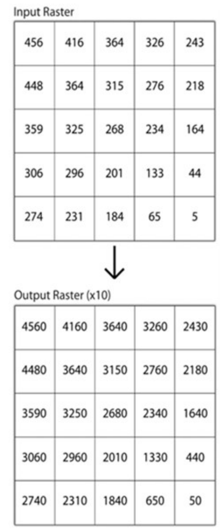


Figure: Local Operation on a Raster Dataset

Geographical Information System

Raster Data Analysis: Local Operations with Multiple Rasters

Local operations with multiple rasters are also referred to as compositing, overlaying, or superimposing maps. Because local operations can work with multiple rasters, they are the equivalent of vector-based overlay operations. A greater variety of local operations have multiple input rasters than have a single input raster. Besides mathematical operators that can be used on individual rasters, other measures that are based on the cell values or their frequencies in the input rasters can also be derived and stored in the output raster. Some of these measures are, however, limited to rasters with numeric data.

Geographical Information System

Raster Data Analysis: Local Operations with Multiple Rasters

Summary statistics, including maximum, minimum, range, sum, mean, median, and standard deviation, are measures that apply to rasters with numeric data. Figure 12.3, for example, shows a local operation that calculates the mean from three input rasters. If a cell contains no data in one of the input rasters, the cell also carries no data in the output raster by default.

Other measures that are suitable for rasters with numeric or categorical data are statistics such as majority, minority, and number of unique values. For each cell, a majority output raster tabulates the most frequent cell value among the input rasters, a minority raster tabulates the least frequent cell value, and a variety raster tabulates the number of different cell values. Figure 12.4, for example, shows the output with the majority statistics from three input rasters.

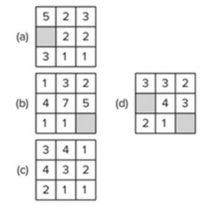


Figure 12.3
The cell value in (*d*) is the mean calculated from three input rasters (*a*, *b*, and *c*) in a local operation. The shaded cells have no data.

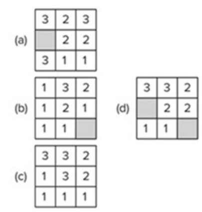


Figure 12.4
The cell value in (*d*) is the majority statistic derived from three input rasters (*a*, *b*, and *c*) in a local operation. The shaded cells have no data.

Geographical Information System

Raster Data Analysis: Local Operations with Multiple Rasters

Some local operations do not involve statistics or computation. A local operation called Combine assigns a unique output value to each unique combination of input values. Suppose a slope raster has three cell values (0 to 20 percent, 20 to 40 percent, and greater than 40 percent slope), and an aspect raster has four cell values (north, east, south, and west aspects). The Combine operation creates an output raster with a value for each unique combination of slope and aspect, such as 1 for greater than 40 percent slope and the south aspect, 2 for 20 to 40 percent slope and the south aspect, and so on (Figure 12.5).

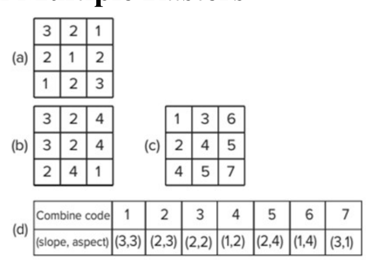
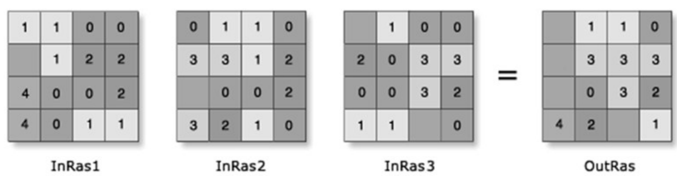


Figure 12.5
Each cell value in (*c*) represents a unique combination of cell values in (*a*) and (*b*). The combination codes and their representations are shown in (*d*).

Geographical Information System

Local geoprocessing functions:

- Determines the largest value of the inputs on a cell-by-cell basis. If all the inputs are integer, the output is integer. If any of the inputs are floating point, the output is floating point. (Note: Empty cell store 99)



Geographical Information System

Local geoprocessing functions:

- When applied to multiple rasters, it becomes possible to perform such analyses as changes over time. Given two rasters containing information on groundwater depth on a parcel of land at Year 2000 and Year 2010, it is simple to subtract these values and place the difference in an output raster that will note the change in groundwater between those two times (Figure "Local Operation on a Raster Dataset"). These local analyses can become somewhat more complicated however, as the number of input rasters increase. For example, the Universal Soil Loss Equation (USLE) applies a local mathematical formula to several overlying rasters including rainfall intensity, erodibility of the soil, slope, cultivation type, and vegetation type to determine the average soil loss (in tons) in a grid cell.

Geographical Information System

Local geoprocessing functions:

- The Local geoprocessing functions are those where the value at each cell location on the output raster is a function of the values from all the inputs at that location. With these geoprocessing functions, you can combine the input rasters, calculate a statistic on them, or evaluate a criterion for each cell on the output raster based on the values of each cell from multiple input rasters.
- To perform the calculation, a local geoprocessing function only needs, for each of the input rasters, the value at that location, as well as (in some cases) a comparison value. Once the result is generated, the calculation is made for the next cell location, and the process is iterated until all cells have been processed.

Geographical Information System

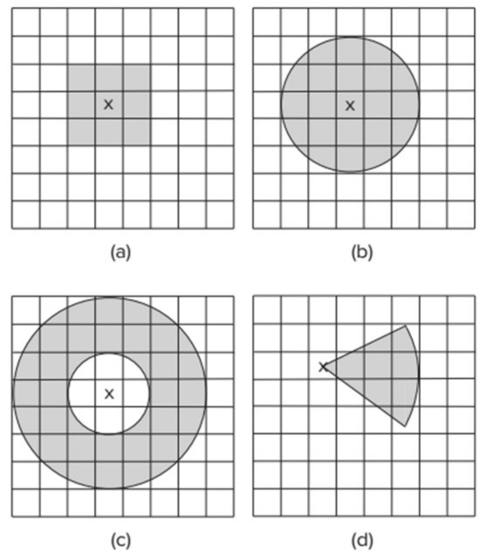
Raster Data Analysis: Neighborhood Operations

A neighborhood operation, also called a focal operation, involves a focal cell and a set of its surrounding cells. The surrounding cells are chosen for their distance and/or directional relationship to the focal cell. A required parameter for neighborhood operations is the type of neighborhood. Common neighborhoods include rectangles, circles, annuluses, and wedges (Figure 12.6). A rectangle is defined by its width and height in cells, such as a 3-by-3 area centered at the focal cell. A circle extends from the focal cell with a specified radius. An annulus or doughnut-shaped neighborhood consists of the ring area between a smaller circle and a larger circle centered at the focal cell. And a wedge consists of a piece of a circle centered at the focal cell. As shown in Figure 12.6, some cells are only partially covered in the defined neighborhood. The general rule is to include a cell if the center of the cell falls within the neighborhood. Although irregular neighborhoods such as symmetric and discontinuous neighborhoods have been proposed in the literature (e.g., Guan and Clarke 2010), they are not available as options in GIS packages.

Geographical Information System

Raster Data Analysis: Neighborhood Operations

Figure 12.6: Four common neighborhood types: rectangle (a), circle (b), annulus (c), and wedge (d). The cell marked with an x is the focal cell.



Geographical Information System

Raster Data Analysis: Neighborhood Operations

Figure 12.7: The cell values in (b) are the neighborhood means of the shaded cells in (a) using a 3-by-3 neighborhood. For example, 1.56 in the output raster is calculated from (1 + 2 + 2 + 1 + 2 + 2 + 1 + 2 + 1)/9.

(a)

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

(b)

1.56	2.00	2.22
1.67	2.11	2.44
1.67	2.11	2.44

Geographical Information System

Raster Data Analysis: Neighborhood Operations

Neighborhood Statistics: A neighborhood operation typically uses the cell values within the neighborhood in computation, and then assigns the computed value to the focal cell. To complete a neighborhood operation on a raster, the focal cell is moved from one cell to another until all cells are visited. Different rules devised by GIS software developers are applied to focal cells on the margin of a raster, where a neighborhood such as a 3-by-3 rectangle cannot be used. A simple rule is to use only cell values available within the neighborhood (e.g., 6 instead of 9) for computation. Although a neighborhood operation works on a single raster, its process is similar to that of a local operation with multiple rasters. Instead of using cell values from different input rasters, a neighborhood operation uses the cell values from a defined neighborhood.

The output from a neighborhood operation can show summary statistics including maximum, minimum, range, sum, mean, median, and standard deviation, as well as tabulation of measures such as majority, minority, and variety. These statistics and measures are the same as those from local operations with multiple rasters

Geographical Information System

Raster Data Analysis: Neighborhood Operations

Neighborhood Statistics:

A block operation is a neighborhood operation that uses a rectangle (block) and assigns the calculated value to all block cells in the output raster. Therefore, a block operation differs from a regular neighborhood operation because it does not move from cell to cell but from block to block.

Geographical Information System

Raster Data Analysis: Neighborhood Operations

Figure 12.8: The cell values in (b) are the neighborhood range statistics of the shaded cells in (a) using a 3-by-3 neighborhood. For example, the upper-left cell in the output raster has a cell value of 100, which is calculated from (200 − 100).

(a)

200	200	110	210	210
200	200	110	210	210
150	150	100	170	170
140	140	130	160	160
140	140	130	160	160

(b)

100	110	110
100	110	110
50	70	70

Figure 12.9: The cell values in (b) are the neighborhood majority statistics of the shaded cells in (a) using a 3-by-3 neighborhood. For example, the upper-left cell in the output raster has a cell value of 2 because there are five 2's and four 1's in its neighborhood.

(a)

1	2	2	2	2
1	2	2	2	3
1	2	1	3	3
2	2	2	3	3
2	2	2	2	3

(b)

2	2	2
2	2	3
2	2	3

Geographical Information System

Raster Data Analysis: Zonal Operations

A zonal operation works with groups of cells of same values or like features. These groups are called zones. Zones may be contiguous or noncontiguous. A contiguous zone includes cells that are spatially connected, whereas a noncontiguous zone includes separate regions of cells. A watershed raster is an example of a contiguous zone, in which cells that belong to the same watershed are spatially connected. A land use raster is an example of a noncontiguous zone, in which one type of land use may appear in different parts of the raster.

Geographical Information System

Raster Data Analysis: Zonal Operations -Zonal Statistics:

A zonal operation may work with a single raster or two rasters. Given a single input raster, zonal operations measure the geometry of each zone in the raster, such as area, perimeter, thickness, and centroid (Figure 12.10). The area is the sum of the cells that fall within the zone times the cell size.

The perimeter of a contiguous zone is the length of its boundary, and the perimeter of a noncontiguous zone is the sum of the length of each part. The thickness calculates the radius (in cells) of the largest circle that can be drawn within each zone. And the centroid is the geometric center of a zone located at the intersection of the major axis and the minor axis of an ellipse that best approximates the zone. Given two rasters in a zonal operation, one input raster and one zonal raster, a zonal operation produces an output raster, which summarizes the cell values in the input raster for each zone in the zonal raster. The summary statistics and measures include area, minimum, maximum, sum, range, mean, standard deviation, median, majority, minority, and variety. (The last four measures are not available if the input raster is a floating-point raster.) Figure 12.11 shows a zonal operation of computing the mean by zone. Figure 12.11b is the zonal raster with three zones, Figure 12.11a is the input raster, and Figure 12.11c is the output raster.

Geographical Information System

Raster Data Analysis: Zonal Operations -Zonal Statistics:

Figure 12.10: Thickness and centroid for two large watersheds (zones). Area is measured in square kilometers, and perimeter and thickness are measured in kilometers. The centroid of each zone is marked with an x

Zone	Area	Perimeter	Thickness
1	36,224	1708	77.6
2	48,268	1464	77.4

1	2	2	1
1	4	5	1
2	3	7	6
1	3	4	4

(a)

1	1	2	2
1	1	2	2
1	1	3	3
3	3	3	3

(b)

2.17	2.17	2.25	2.25
2.17	2.17	2.25	2.25
2.17	2.17	4.17	4.17
4.17	4.17	4.17	4.17

(c)

Figure 12.11: The cell values in (c) are the zonal means derived from an input raster (a) and a zonal raster (b). For example, 2.17 is the mean of {1, 1, 2, 2, 4, 3} for zone 1.

Geographical Information System

Raster Data Analysis: Applications of Zonal Operations

Measures of zonal geometry such as area, perimeter, thickness, and centroid are particularly useful for studies of landscape ecology.

Many other geometric measures can be derived from area and perimeter. For example, the perimeter-area ratio (i.e., zonalperimeter/zonalarea) is a simple measure of shape complexity used in landscape ecology.

Zonal operations with two rasters can generate useful descriptive statistics for comparison purposes. For example, to compare topographic characteristics of different soil textures, we can use a soil raster that contains the categories of sand, loam, and clay as the zonal raster and slope, aspect, and elevation as the input rasters. By running a series of zonal operations, we can summarize the slope, aspect, and elevation characteristics associated with the three soil textures.

Geographical Information System

Raster Data Analysis: Physical distance measure perations

In a GIS, distances may be expressed as physical distances or cost distances.
The physical distance measures the straight-line or Euclidean distance, whereas the cost distance measures the cost for traversing the physical distance. The distinction between the two types of distance measures is important in real-world applications.
A truck driver, for example, is more interested in the time or the fuel cost for covering a route than in its physical distance. The cost distance in this case is based on not only the physical distance but also the speed limit and road condition.

Geographical Information System

Raster Data Analysis: Physical distance measure operations

Physical distance measure operations calculate straight-line distances away from cells designated as the source cells. For example, to get the distance between cells (1, 1) and (3, 3) in Figure 12.12, we can use the following formula:

$$\text{cell size} \times \sqrt{(3-1)^2 + (3-1)^2}$$

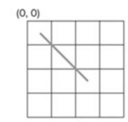


Figure 12.12
A straight-line distance is measured from a cell center to another cell center. This illustration shows the straight-line distance between cell (1,1) and cell (3,3).

or cell size × 2.828. If the cell size were 30 meters, the distance would be 84.84 meters. operations.

Geographical Information System

Raster Data Analysis: Physical distance measure operations

A physical distance measure operation essentially buffers the source cells with wavelike continuous distances over the entire raster (Figure 12.13) or to a specified maximum distance. This is why physical distance measure operations are also called extended neighborhood operations or global (i.e., the entire raster).

Figure 12.13
Continuous distance measures from a stream network.

Geographical Information System

Introduction to Interpolation:

GIS data modeling

- What is a model? A model is a simplified representation of a phenomenon or a system.
- Several types of models have already been discussed.

  A map is a model. So are the vector and raster data models for representing spatial features and the relational model for representing a database system.
- A model helps us better understand a phenomenon or a system by retaining the significant features and relationships of reality.
- First, we studied with models using geospatial data.
- Second, the emphasis is on the use of GIS in modeling rather than the models.

Geographical Information System

Introduction to Interpolation:

GIS data modeling

- What is a model? A model is a simplified representation of a phenomenon or a system.
- Several types of models have already been discussed.

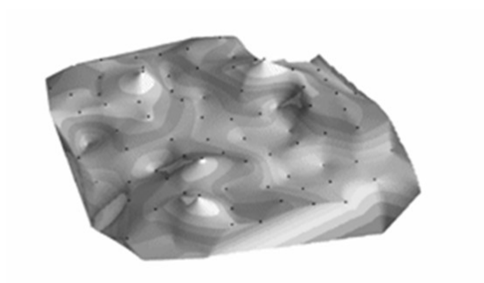
  A map is a model. So are the vector and raster data models for representing spatial features and the relational model for representing a database system.
- A model helps us better understand a phenomenon or a system by retaining the significant features and relationships of reality.
- First, we studied with models using geospatial data.
- Second, the emphasis is on the use of GIS in modeling rather than the models.

Geographic Information System

Surface modeling

- Surface modeling - A representation of a geographic feature or phenomenon that can be measured continuously across some part of the earth's surface (for example, elevation). A surface model is an approximation of a surface, generalized from sample data. Surface models are stored and displayed as rasters, TINs, or terrains.

GIS supports display of raster and TIN surface models and provides analysis tools in the Spatial Analyst, 3D Analyst, and Geostatistical Analyst extensions to create, analyze, and extract information from surfaces.



Geographic Information System

Surface modeling

Surface modeling - A representation of a geographic feature or phenomenon that can be measured continuously across some part of the earth's surface (for example, elevation). A surface model is an approximation of a surface, generalized from sample data. Surface models are stored and displayed as rasters, TINs, or terrains.

- What are surfaces?
- Creating surfaces
- Analyzing surfaces
- Extracting information from surfaces

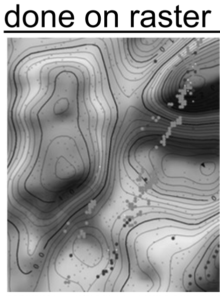
Surface modeling - What are surfaces?

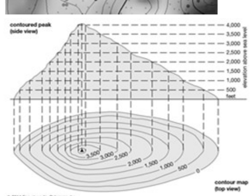
Surfaces represent phenomena that have values at every point across their extent.

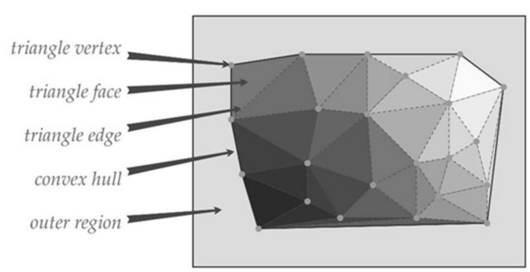
- The values at the infinite number of points across the surface are derived from a limited set of sample values. These may be based on direct measurement, such as height values for an elevation surface, or temperature values for a temperature surface; between these measured locations, values are assigned to the surface by interpolation.
- Surfaces can also be mathematically derived from other data, such as slope and aspect surfaces derived from an elevation surface, a surface of distance from bus stops in a city, or surfaces showing concentration of criminal activity or probability of lightning strikes.

5.1 Surface modeling - What are surfaces?

Surfaces - can be represented using contour lines or isolines, arrays of points, TINs, and rasters; <u>however, most surface analysis in GIS is done on raster or TIN data</u>.







Geographic Information System

Surface modeling

The Interpolation geoprocessing functions create a continuous (or prediction) surface from sampled point values that represents some measure, such as the height, concentration, or magnitude (for example, elevation, acidity, or noise level). Surface interpolation geoprocessing functions make predictions from sample measurements for all locations in an output raster dataset, whether or not a measurement has been taken at the location.

Geographic Information System

Surface modeling - Creating surfaces?

There are tools to create surfaces from vector features or from other surfaces. There are several ways to create surfaces, including

- interpolating values stored at measured point locations,
- interpolating a surface of the density of a given phenomenon or feature type from the numbers of features in an area,
- deriving surfaces of distance (or direction) from a feature or features, or
- deriving a surface from another surface (slope raster from elevation).

Geographic Information System

Introduction to Interpolation - Creating surfaces

Interpolation in GIS is a technique used to estimate values at unmeasured locations within a set of known data points. It is especially useful when working with spatial data that has a continuous variation, such as elevation, temperature, or population density. Interpolation helps create a smooth representation of these continuous phenomena, making it easier to visualize and analyze the data.

There are several methods of interpolation commonly used in GIS:

Geographic Information System

Introduction to Interpolation - Creating surfaces

There are several methods of interpolation commonly used in GIS:

- **Inverse Distance Weighting (IDW)**: IDW is a simple method that assigns values to unmeasured locations based on the weighted average of nearby known values. The weight decreases as the distance from the unknown point increases. The parameter that controls the rate of weight decrease is called the "power" or "exponent."
- **Kriging**: Kriging is a more sophisticated technique that takes into account both the spatial correlation and the variability of the data. It estimates values by considering the semivariogram, which describes the spatial autocorrelation of the data. Kriging can provide estimates along with measures of uncertainty.

Geographic Information System

Introduction to Interpolation - Creating surfaces

There are several methods of interpolation commonly used in GIS:

- **Natural Neighbor Interpolation**: This method calculates the value at an unmeasured location based on the values of its nearest neighbors. The values are weighted according to the proportion of the area that belongs to each neighbor. This approach can create smooth surfaces and works well for irregularly spaced data points.
- **Triangulated Irregular Network (TIN) Interpolation**: TIN interpolation divides the data points into triangles and constructs a surface by connecting the vertices of these triangles. It creates a natural representation of the terrain and is often used for elevation data.

Geographic Information System

Spatial Autocorrelation:

Spatial autocorrelation is a concept commonly used in GIS and spatial analysis to understand the degree of similarity or dissimilarity between spatially related data points. It refers to the tendency of similar values to cluster together on a map, indicating that nearby locations tend to have similar attributes or values.

In simpler terms, if spatial autocorrelation is high, it means that similar values are often found near each other on the map. If spatial autocorrelation is low, there's no clear pattern of similarity or dissimilarity among neighboring locations.

Geographic Information System

DEM:

DEM stands for "Digital Elevation Model," and it is a fundamental component in Geographic Information Systems (GIS) and remote sensing. A DEM is a representation of the Earth's surface, usually in a gridded format, where each cell in the grid contains elevation information. In other words, a DEM provides a digital representation of the topography of a landscape.

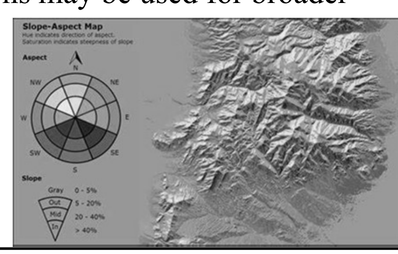
DEM data is obtained through various remote sensing techniques, such as LiDAR (Light Detection and Ranging) and photogrammetry. These techniques involve collecting elevation measurements from aircraft, satellites, or ground-based instruments. Once collected, the elevation data is processed and interpolated to create a continuous elevation model.

Geographic Information System

DEM:

Some key points about DEMs in GIS:

- **Applications:** DEMs have numerous applications in GIS and geospatial analysis. They are used for terrain analysis, hydrological modeling, floodplain mapping, viewshed analysis, slope and aspect calculations, and more.
- **Resolution:** DEMs can vary in resolution, with finer resolutions providing more detailed elevation information. Higher-resolution DEMs are suitable for localized analyses, while coarser resolutions may be used for broader regional assessments.



Geographic Information System

DEM:

Some key points about DEMs in GIS:

- **Derived Data:** From a DEM, various derived datasets can be generated. For example:
  - **Slope and Aspect:** Slope represents the steepness of the terrain at a given point, while aspect indicates the direction the slope faces.
  - **Hillshades:** These are used to visualize the terrain using light and shadow effects, simulating how sunlight interacts with the landscape.
  - **Contours:** Lines connecting points of equal elevation on the landscape are known as contour lines. These lines help visualize elevation changes.

Geographic Information System

Spatial Statistics:

Spatial statistics in Geographic Information Systems (GIS) involve the analysis of spatial patterns and relationships within geospatial data. It helps to uncover hidden insights, identify trends, and make informed decisions based on the spatial distribution of data. Spatial statistics are widely used in various fields, including urban planning, environmental science, epidemiology, natural resource management, and more.

Some key concepts and techniques in spatial statistics within GIS:

Point Patterns Analysis: This involves studying the spatial arrangement of individual points, such as the distribution of trees in a forest or the locations of disease cases.

Spatial Autocorrelation: Spatial autocorrelation assesses whether similar values are clustered together in space. It helps to identify areas with high or low similarity among neighboring locations.

Geographic Information System

Spatial Statistics:

Some key concepts and techniques in spatial statistics within GIS:

Spatial Interpolation: Interpolation methods estimate values at unobserved locations based on data from observed points. Common techniques include Inverse Distance Weighting (IDW), kriging, and spline interpolation.

Spatial Clustering: Spatial clustering identifies areas where similar values are concentrated.

Spatial Regression: Spatial regression analyzes the relationships between spatially distributed variables, considering spatial dependencies. It accounts for spatial autocorrelation in the regression model.

Hotspot Analysis: Hotspot analysis identifies statistically significant clusters of high or low values in a dataset. It's commonly used in crime analysis, disease mapping, and environmental monitoring.

Spatial Join: Spatial join combines attributes from one dataset with another

Geographic Information System

Spatial Statistics:

Some key concepts and techniques in spatial statistics within GIS:

Spatial Join: Spatial join combines attributes from one dataset with another based on their spatial relationships. This is useful for overlaying different types of data to perform analysis or visualization.

Network Analysis: Network analysis focuses on studying the connectivity and accessibility of geographic features through networks, such as road networks or utility networks. It's often used for route optimization and facility location analysis.

Buffer Analysis: Buffer analysis creates zones around points, lines, or polygons to study the influence or interaction of features within a specified distance.

Geographic Information System

Spatial Statistics:

Some key concepts and techniques in spatial statistics within GIS:

- Spatial Statistics Tools
  - Analyzing Patterns
    - Average Nearest Neighbor
    - High/Low Clustering (Getis-Ord General G)
    - Incremental Spatial Autocorrelation
    - Multi-Distance Spatial Cluster Analysis (Ripleys K Function)
    - Spatial Autocorrelation (Morans I)
  - Mapping Clusters
    - Cluster and Outlier Analysis (Anselin Local Morans I)
    - Grouping Analysis
    - Hot Spot Analysis (Getis-Ord Gi*)
    - Optimized Hot Spot Analysis
    - Similarity Search
  - Measuring Geographic Distributions
  - Modeling Spatial Relationships
    - Exploratory Regression
    - Generate Network Spatial Weights
    - Generate Spatial Weights Matrix
    - Geographically Weighted Regression
    - Ordinary Least Squares
  - Rendering
  - Utilities

Geographical Information System

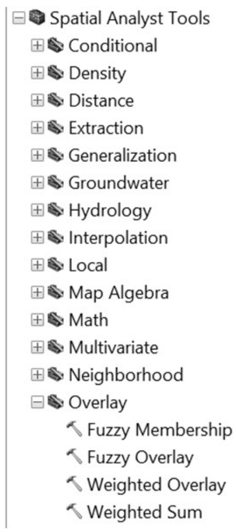
Raster Weighted Overlay:

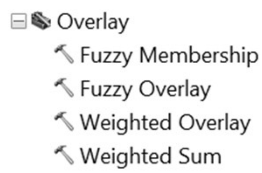
There are various approaches to weight and transform input criteria when performing overlay analysis.
The three main approaches available to weight and add the transformed input are Weighted Overlay, Weighted Sum, and Fuzzy Overlay.
In Weighted Overlay analysis, a series of tools can complement the Weighted Overlay tool to follow the general overlay analysis steps. The Weighted Overlay tool scales the input data on a defined scale (the default being 1 to 9), weights the input rasters, and adds them together. The more favorable locations for each input criterion will be re-classed to the higher values such as 9. In the Weighted Overlay tool, the weights assigned to the input rasters must equal 100 percent. The layers are multiplied by the appropriate multiplier, and for each cell, the resulting values are added together. Weighted Overlay assumes that more favorable factors result in the higher values in the output raster, therefore identifying these locations as being the best.

Geographical Information System

Raster Weighted Overlay:

In Weighted Overlay analysis,





Geographical Information System

Raster Weighted Overlay: Example

Weighted overlay analysis is a technique commonly used in Geographic Information Systems (GIS) to combine multiple raster layers into a single composite raster layer. This composite layer is created by assigning weights to each input raster layer based on their relative importance in a decision-making process. Here's a simple example to illustrate the concept:

Let's say we're interested in finding the best location for a new residential development, and we have three factors to consider: proximity to schools, proximity to parks, and distance from industrial areas. Each factor is represented as a raster layer with values ranging from 0 to 1, where 1 indicates the most desirable conditions and 0 indicates the least desirable conditions.

Geographical Information System

Raster Weighted Overlay: Example

Proximity to Schools: Raster layer where cells closer to schools have higher values (e.g., 1 for cells right next to schools, decreasing to 0 as distance increases).
Proximity to Parks: Similar raster layer representing proximity to parks.
Distance from Industrial Areas: Inverted raster layer where cells closer to industrial areas have lower values (e.g., 0 for cells right next to industrial areas, increasing to 1 as distance increases).
Next, we assign weights to each factor based on their importance. Let's say we decide that proximity to schools is most important (weight = 0.4), followed by proximity to parks (weight = 0.3), and distance from industrial areas (weight = 0.3).
Now, we perform the weighted overlay: Multiply each raster layer by its corresponding weight.

Weighted proximity to schools = Proximity to Schools * 0.4
Weighted proximity to parks = Proximity to Parks * 0.3
Weighted distance from industrial areas = Distance from Industrial Areas * 0.3

Add the weighted raster layers together to get the composite suitability raster layer:
Composite Suitability = Weighted proximity to schools + Weighted proximity to parks + Weighted distance from industrial areas
The resulting composite suitability raster layer will have values ranging from 0 to 1, where higher values indicate more suitable locations for the new residential development based on the weighted combination of the three factors.

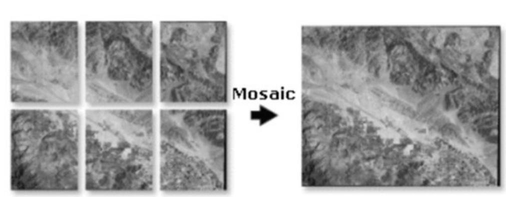
Geographical Information System

Mosaic and Aggregate tools:

In GIS (Geographic Information Systems), mosaic and aggregate are two important tools used for data manipulation and visualization. Let's explore each tool in more detail:

Mosaic:

Mosaic is a process of combining multiple raster datasets into a single seamless mosaic dataset. This tool is commonly used when you have several individual raster files that represent different parts of the same area or have the same extent but different attributes. Mosaicking allows you to create a continuous and consistent representation of the entire study area.



Geographical Information System

Mosaic and Aggregate tools:

Key features and uses of the Mosaic tool:

Combining multiple raster datasets: Mosaic tool can merge overlapping or adjacent raster datasets to create a single composite image.

Handling overlapping pixels: When raster datasets have overlapping areas, the Mosaic tool provides methods to determine how to blend the overlapping pixels, such as using the maximum or minimum value, average, or blending based on transparency.

Pyramid generation: Mosaicking often involves creating pyramid layers (pyramiding) for efficient visualization and rendering at different scales.

Geographical Information System

Mosaic and Aggregate tools:

Aggregate:

The Aggregate tool is used to resample raster data to a lower resolution, effectively reducing the amount of data and creating a new raster dataset at a coarser scale. This can be useful for managing large datasets, speeding up processing, and creating generalized representations of geographic information.

Key features and uses of the Aggregate tool:

Resampling to coarser resolution: By aggregating data, you reduce the spatial detail of the raster, which can be helpful for visualization, data compression, and faster analysis.

Maintaining statistics: When aggregating raster data, statistics such as minimum, maximum, mean, or majority are often computed for the new pixel values based on the original raster cells that fall within the aggregated cells.

Mosaic and Aggregate tools:

Mosaic and Aggregate:

Both Mosaic and Aggregate tools are commonly found in GIS software packages and are essential for preprocessing and data preparation before analysis or visualization. They help improve the efficiency and effectiveness of handling large raster datasets, making it easier to work with vast amounts of geographic information.