

INDIAN INSTITUTE OF REMOTE SENSING, DEHRADUN

Spatial Data Analysis- Functionality and Tools

Kapil Oberai

Scientist/Engineer SF, Geoinformatics Department

kapil@iirs.gov.in



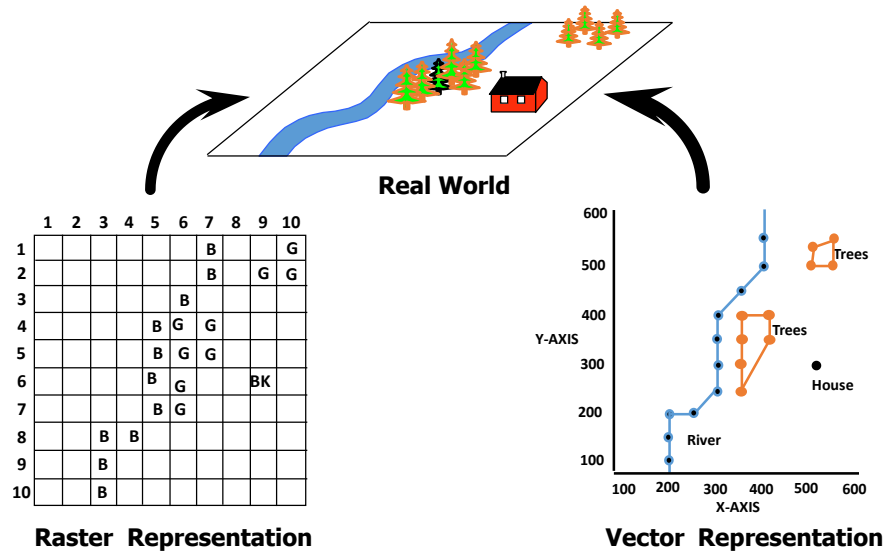
INDIAN INSTITUTE OF REMOTE SENSING, DEHRADUN



Lecture Outline

- Review GIS Data Models
- What is Spatial Data Analysis?
- Broad Classification of analytical GIS capabilities
- Spatial Data Analysis: Vector Based Operations
- Spatial Data Analysis: Raster Based Operations
- Comparison of Vector and Raster-based Spatial Data Analysis
- Example- Site Suitability

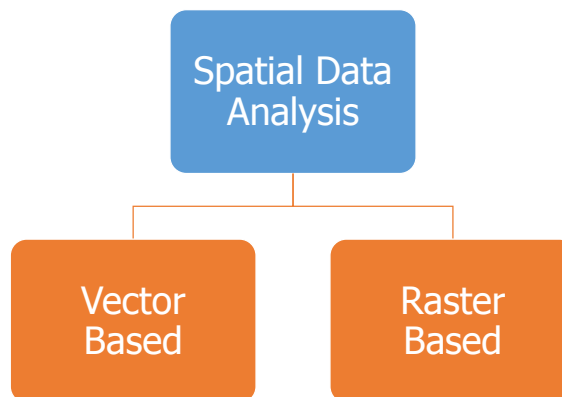
GIS Data Model



3

Spatial Data Analysis

- It can be defined as **computing new information that provides new insight from the existing stored data.**



4

History- Spatial Analysis is not new

London cholera epidemic 1854



Soho

- + Cholera death
- Water pump



Source: <http://www.ph.ucla.edu/epi/snow/snowcricketarticle.html>

5

Broad Classification of analytical GIS capabilities

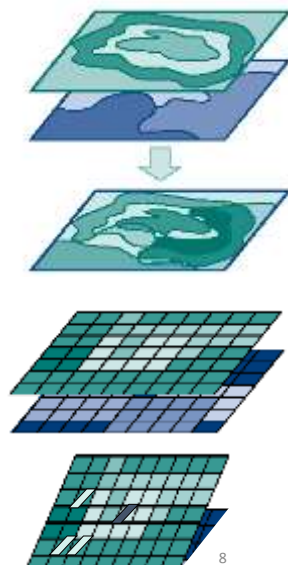
- Classification, retrieval and measurements
 - All functions in this category are performed on a single (vector or raster) data layer, often using the associated attribute data.
- Overlay functions
 - Combination of two or more spatial data layers.
- Neighborhood functions
 - Evaluate the characteristics of an area *surrounding the features location*

7



Overlay Operation

- Standard overlay operators
 - take **two input** data layers;
 - assume they are **georeferenced** in the same system;
 - **overlap** in study area.
- If either **condition is not met**, the use of an overlay operator is **senseless**.
- The **principle** is to:
 - **compare the characteristics** of the same location in both data **layers**, and
 - to **produce a new output** value for each location.

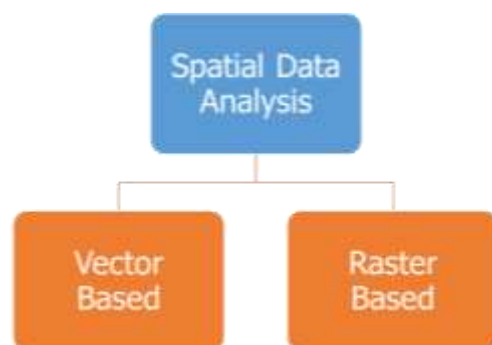


8



Overlay Operations

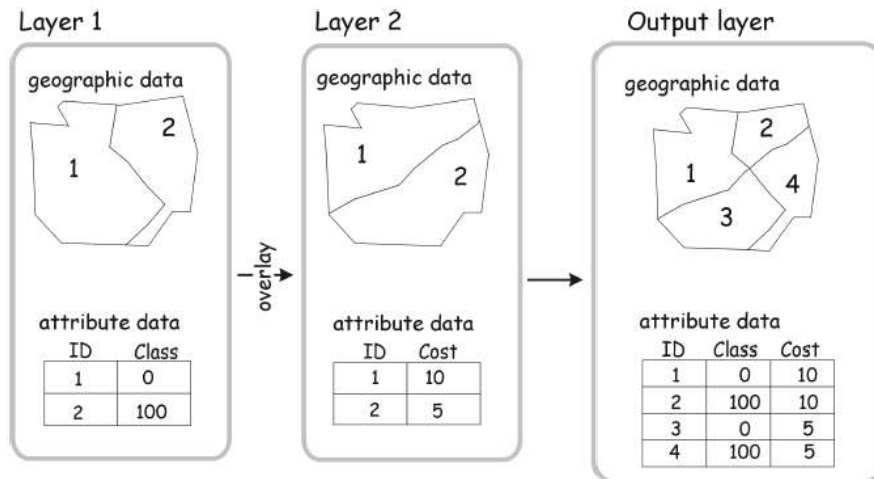
- **Vector** (point, lines, polygons)
 - Intersection
 - Clip
 - Union
- **Raster**
 - Arithmetic operators
 - Comparison and logical operators
 - Conditional



9

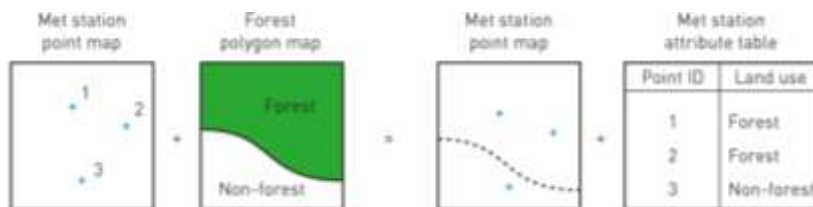
Vector based analysis- Overlay

General Principles of vector overlay

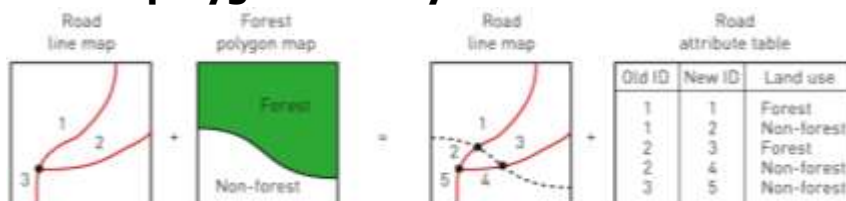


Vector based analysis Contd...

Point-in-polygon overlay

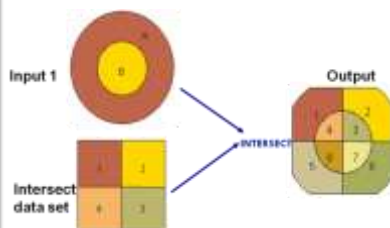


Line-in-polygon overlay



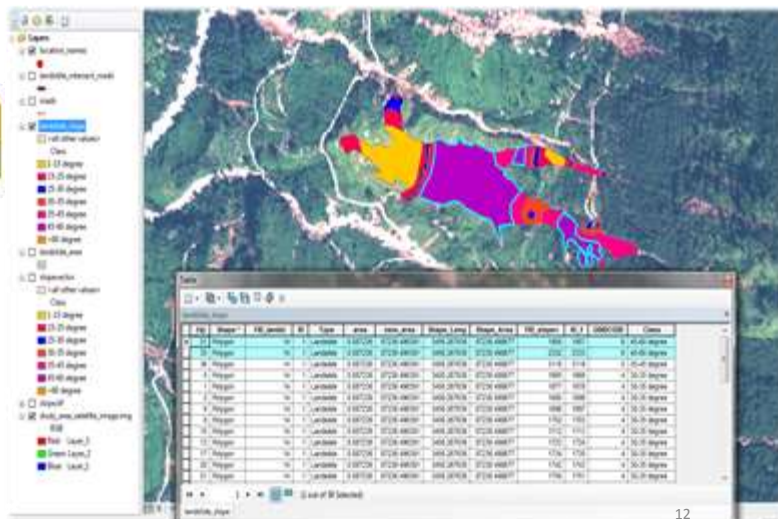
Vector based analysis Contd...

■ Intersect



OUTPUT	INPUT 1	INPUT 2
1	A	1
2	A	2
3	B	2
4	B	1
5	A	4
6	B	4
7	B	3
8	A	3

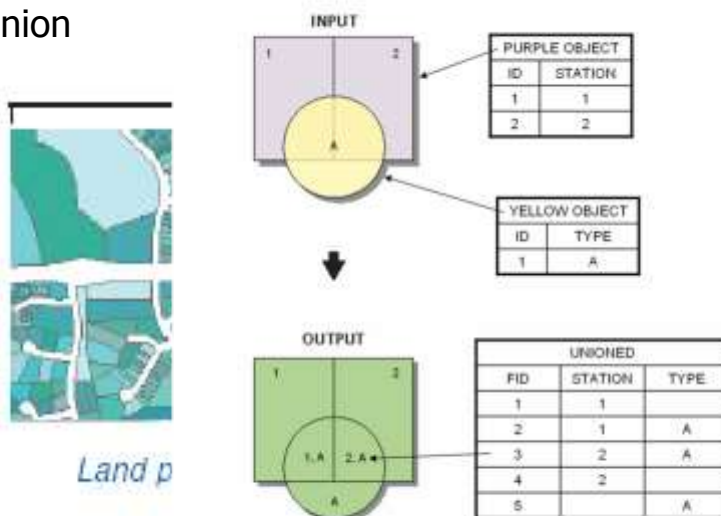
Source: ArcGIS Desktop Documentation (ESRI)



12

Vector based analysis Contd...

■ Union



Input from Union



parcels and soil

Source: <https://www.researchgate.net/figure/The-GIS-Union-tool-The-tool-creates-a-new-polygon-shape-file-where-intersecting-polygons-are-unioned> Source: ArcGIS Desktop Documentation (ESRI)

13



Vector based analysis Contd...

- Clip (Cookie Cutter)

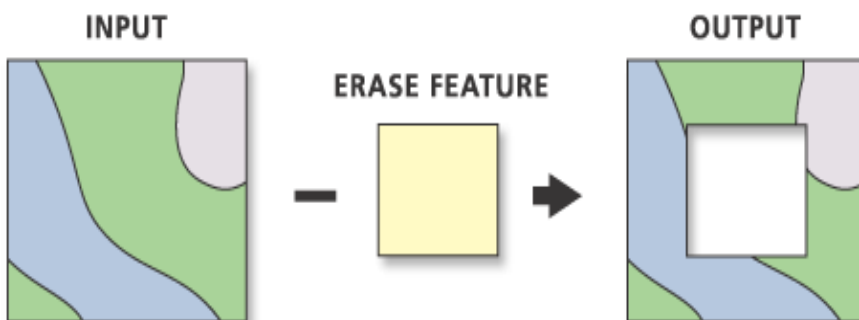


14



Vector based analysis Contd...

- Erase



Source: ArcGIS Desktop Documentation (ESRI)

15

CLIP VS ERASE

- **Clip** extracts features **inside** the boundary
- **Erase** keeps features **outside** the boundary

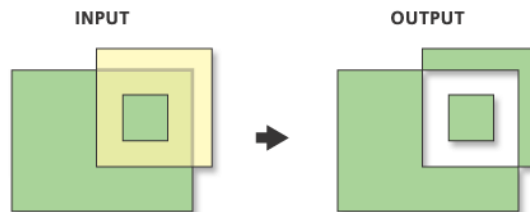


Source: <https://learnqis.org/textbook/section-two-overlay-analysis>

16

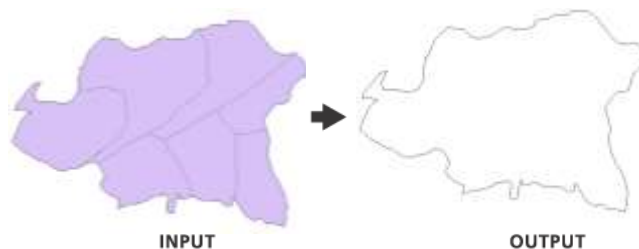
Vector based analysis Contd...

- Symmetrical Difference



Source: ArcGIS Desktop Documentation (ESRI)

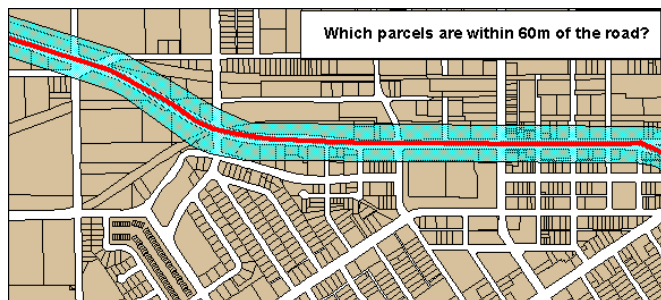
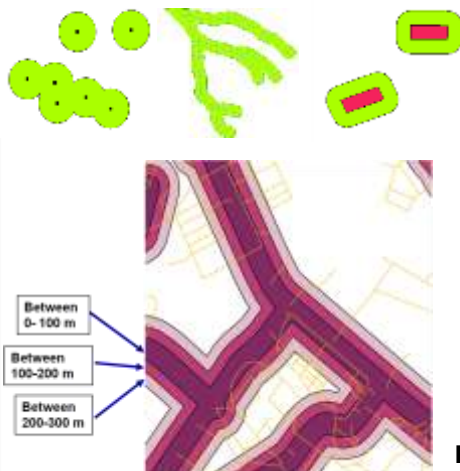
- Dissolve



17

Vector based analysis

- **Buffering:** also called proximity analysis is based on the distance derived from certain selected features.



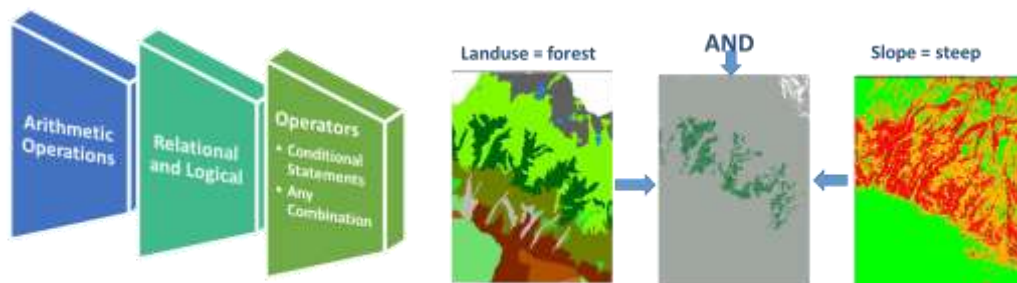
Multiple Ring Buffer

19

Raster based analysis

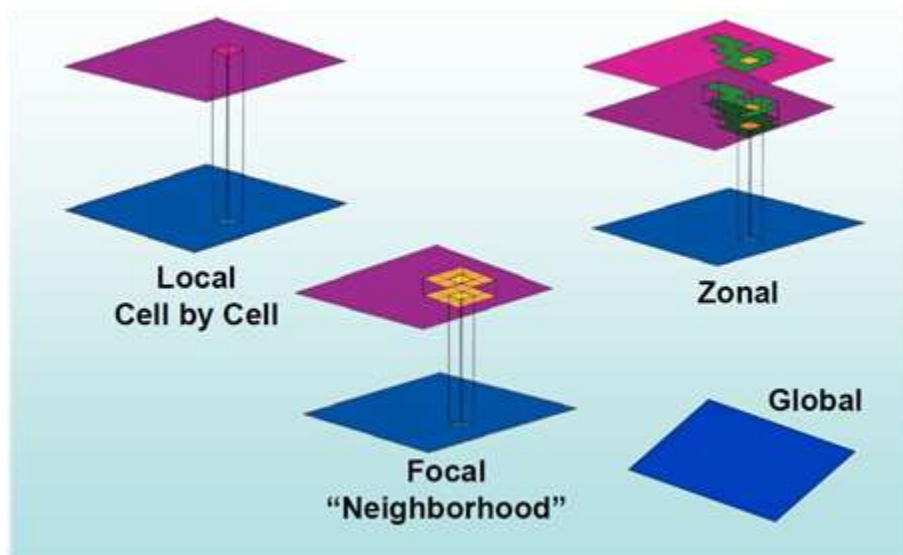
- Overlay operations – RASTER
 - New cell values are calculated using **map algebra**
 - Performed on cell-by-cell basis
 - No geometric calculation

Output_raster_name = Raster_expression



20

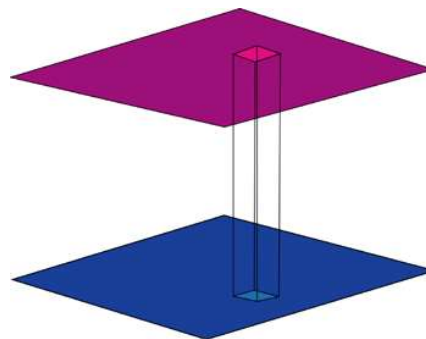
Raster Operations



Source: ArcGIS Desktop Documentation (ESRI)

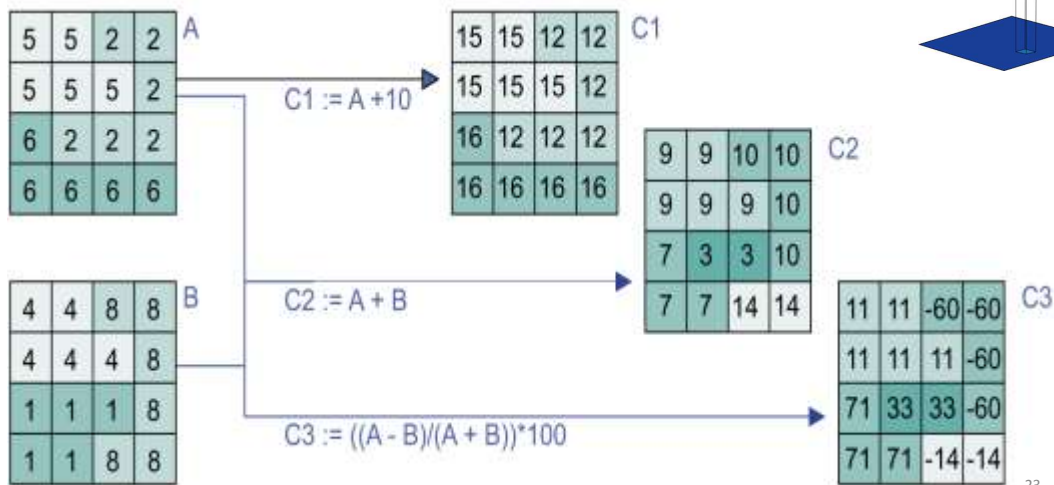
Local: Cell by Cell Operations

- Performed on Cell by cell basis
- Computes output cell values as a function of the input cell values
- Can be done using single or multiple raster's
- "No data" cells not included in calculations
- Common examples: overlays and reclassification



Local: Cell by Cell Operations

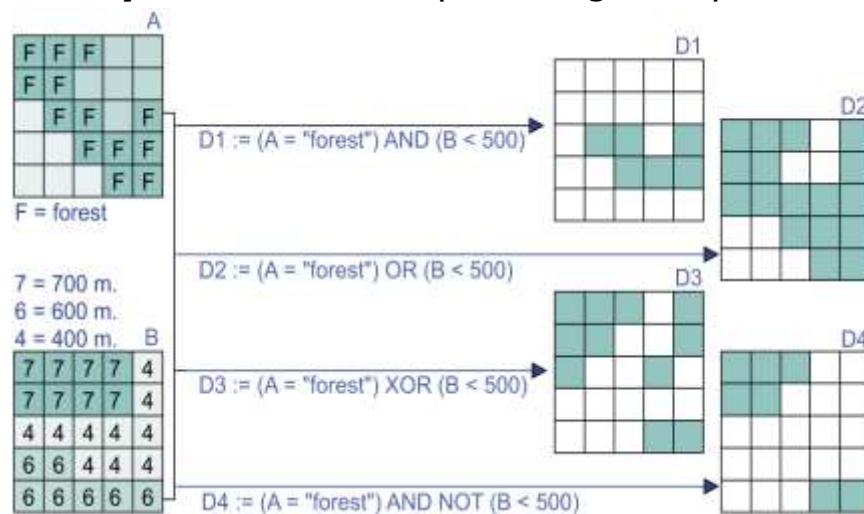
Overlay RASTER – Examples of arithmetic operator



23

Local: Cell by Cell Operations

Overlay RASTER – Examples of logical expressions

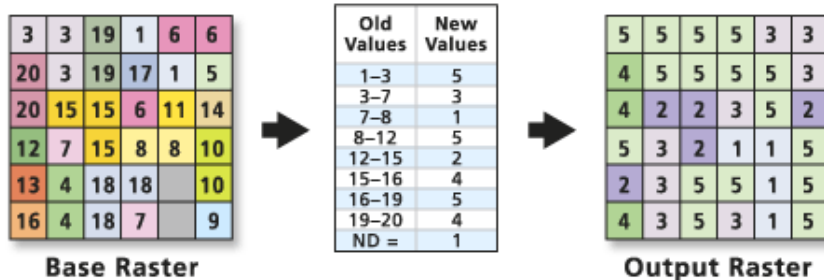


24

Raster Operations Contd...

Local

- **Reclassification:** is the process of reassigning a value, a range of values, or a list of values in a raster to new output values



Value = NoData

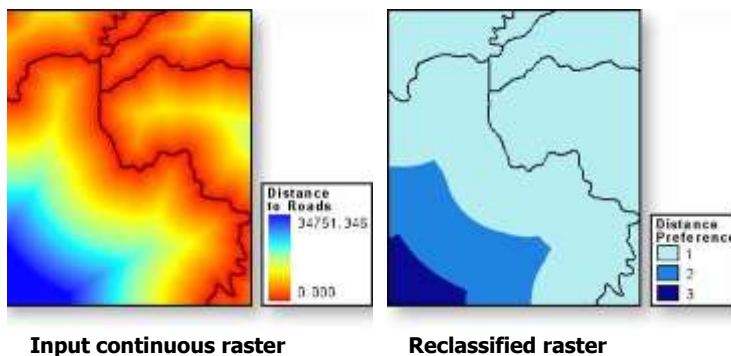
25

Source: ArcGIS Desktop Documentation (ESRI)

Raster Operations Contd...

Local

- **Reclassification**



- Reclassification of continuous data involves replacing a range of values with a new values.
- For example, a raster depicting distance from roads can be reclassified into three distance zones.

26

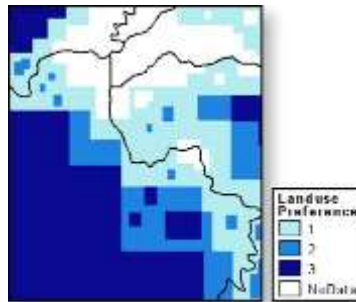
Source: <http://www.geography.hunter.cuny.edu/~jochen/GTECH361/lectures/lecture11/concepts/Reclassifying%20raster%20data.htm>

Raster Operations Contd...

- Local
 - Reclassification



Input discrete raster



Reclassified raster

- Reclassification of categorical data involves replacing individual values with new values.
- For example, land use values can be reclassified into preference values of low (1), medium (2), and high (3).
- Land use values not desired in the analysis are given values of NoData.

Source: <http://www.geography.hunter.cuny.edu/~jochen/GTECH361/lectures/lecture11/concepts/Reclassifying%20raster%20data.htm>

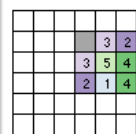
27

Focal (Neighbourhood) Operation

- Value of the output cell determined by the cells in a specified neighbourhood around each input cell



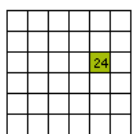
Input processing raster



Input processing raster



Output raster



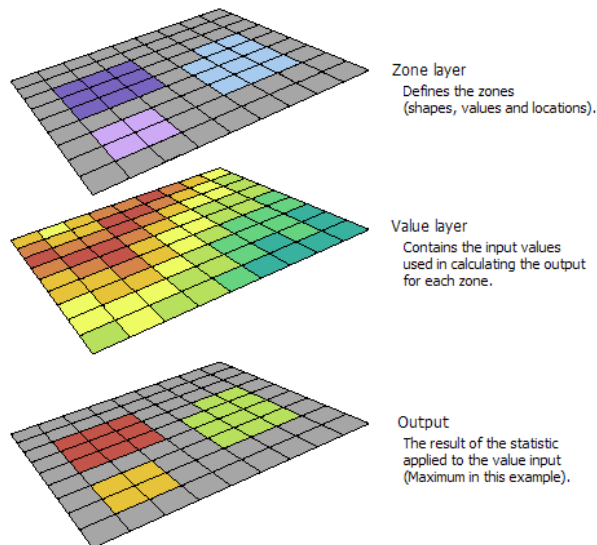
Output raster

- Focal Statistics:** calculating a Sum statistic, consider the processing cell with a value of 5 in the diagram.
- A rectangular 3 by 3 cell neighbourhood shape is specified.
- The sum of the values of the neighbouring cells ($3 + 2 + 3 + 4 + 2 + 1 + 4 = 19$) plus the value of the processing cell (5) equals 24 ($19 + 5 = 24$).
- So a value of 24 is given to the cell in the output raster

Source: <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/how-focal-statistics-works.htm>

Raster Operations Contd...

■ Zonal



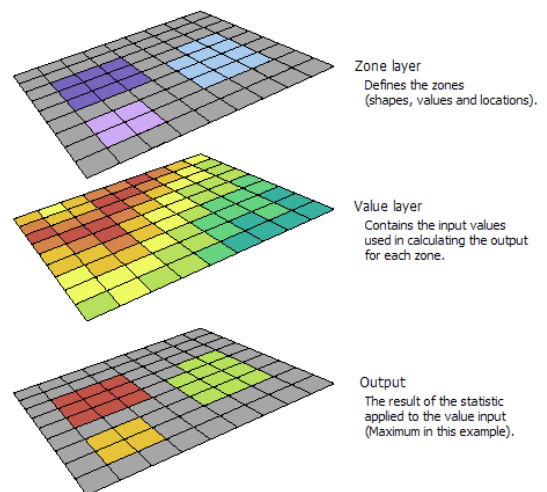
Example inputs and output from Zonal Statistics

29

Source: ArcGIS Desktop Documentation (ESRI)

Outputs of Zonal Operations

- **Raster layer**
 - All the cells within a zone have the same value on the output raster layer
- **Table**
 - Each row in the table contains the statistics for a zone.
 - The first column is the value (or ID) of each zone.
 - The table can be joined back to the zone layer.



Example inputs and output from Zonal Statistics

Source: ArcGIS Desktop Documentation (ESRI)

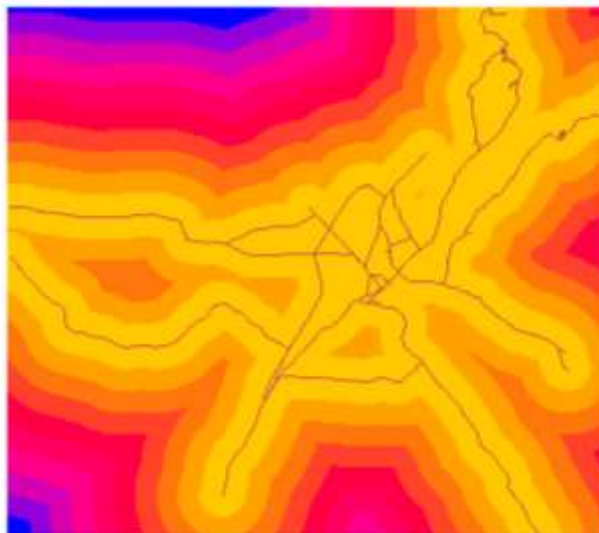
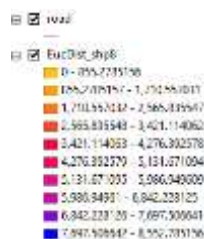


Raster Operations Contd...

Global

Euclidean distance:

Operations assign to each cell in the output raster dataset its distance from the closest source cell.



31



Quiz

1) Dissolve operations requires two input layers

a) True

b) False

2) Re-classification is an example of _____ operation in GIS

a) Local

b) Focal

c) Zonal

d) Global

3) Computing NDVI is an example of _____ operation in GIS.

a) Local

b) Focal

c) Zonal

d) Global

4) Euclidean distance is an example of _____ operation in GIS

a) Local

b) Focal

c) Zonal

d) Global

5) Erase Operation keeps features outside the boundary

a) True

b) False

32

Quiz

- 1) Dissolve operations requires two input layers
a) True
b) False
- 2) Re-classification is an example of _____ operation in GIS
a) Local
b) Focal
c) Zonal
d) Global
- 3) Computing NDVI is an example of _____ operation in GIS.
a) Local
b) Focal
c) Zonal
- d) Global
- 4) Euclidean distance is an example of _____ operation in GIS
a) Local
b) Focal
c) Zonal
d) Global
- 5) Erase Operation keeps features outside the boundary
a) True
b) False

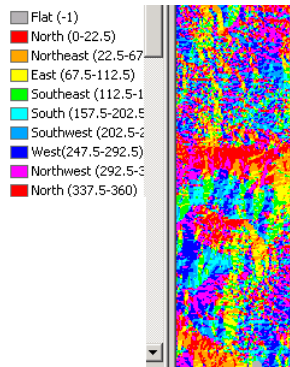
33

Terrain Analysis

Slope:



Aspect:

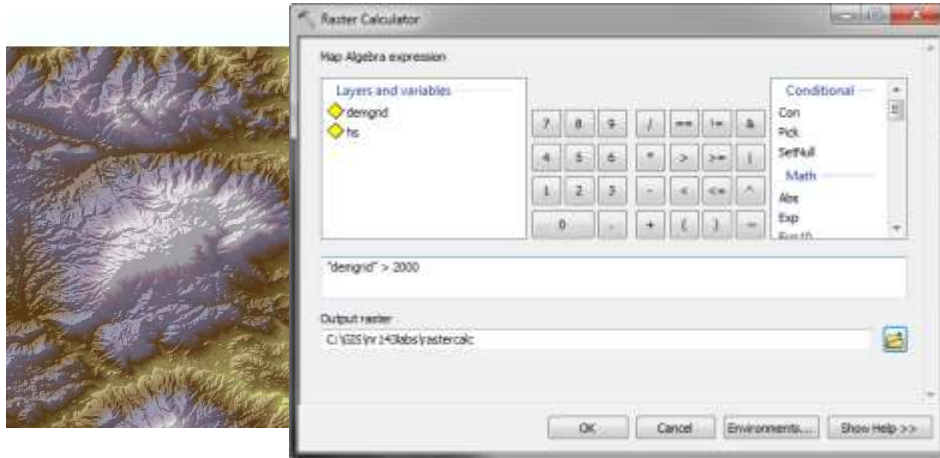


Contours:



Raster Operations Contd...

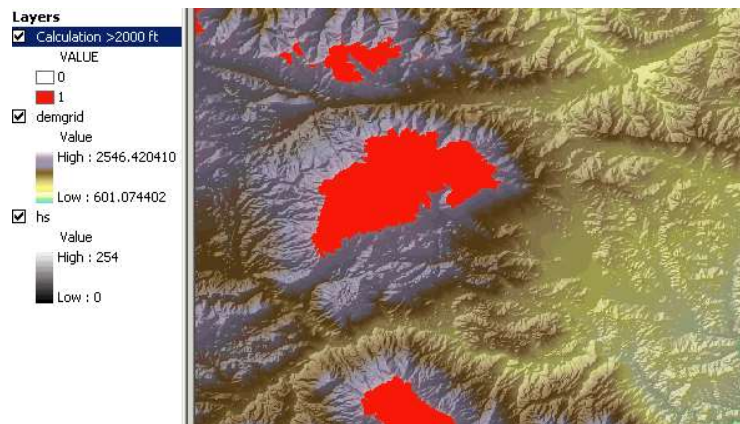
- Map Query Examples
 - Single layer numeric example: elevation > 2000 ft.



35

Raster Operations Contd...

- Map Query Examples
 - Single layer numeric example: elevation > 2000 ft.
 - Results in a binary True/False layer



36



COMPARISON OF VECTOR AND RASTER-BASED DATA ANALYSIS

- **Vector and raster data analysis** are the **two basic types of GIS analysis**.
- **Treated separately** because
 - **GIS software cannot run** them together in the same operation.
- GIS software provide **tools to convert vector to raster data and vice-versa**.
- GIS software **allow use of vector data in some raster data operations** (e.g., extraction operations)
 - the data are converted into raster data before the operation starts.
- **Each GIS project is different in terms of data sources and objectives**.
- Therefore choose the type of data analysis that is efficient and appropriate.
 - depends on the data availability and its format, GIS Software, data complexity and efficiency

37



COMPARISON OF VECTOR AND RASTER-BASED DATA ANALYSIS

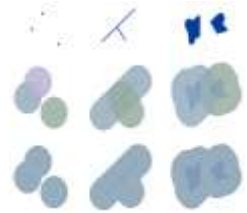
- **Lets consider difference using two popular operations: overlay and buffering**.
- **Overlay: local operation** with **multiple raster's** is often **compared to vector overlay** operation.
- Differences exist between them:
 - **vector-based overlay** must **compute intersections between features and insert points at the intersections**.
 - **not necessary** for **raster-based** local operation; input have **same cell size and extent**.
 - **Computation is less complicated than calculating line intersections**; even if resampled
 - **raster-based local operation** has access to **various tools and operators** to create the **output** whereas a **vector-based overlay** only **combines attributes** from input layers.
 - **Any computations** with the attributes **must follow the overlay operation**.
 - So, **raster-based overlay is often preferred** for projects that involve **large number of layers** and a **considerable amount of computation**.
 - **However an vector overlay operation has advantages** that it can **combine multiple attributes from each input layer**.
 - Once combined, **all attributes can be queried and analyzed individually or in combination** unlike raster operation- **one raster for each attribute**

38



COMPARISON OF VECTOR AND RASTER-BASED DATA ANALYSIS

- **Buffering:** Vector-based buffering and raster physical distance measure operation are similar; both measure distances from select features
- But they differ in following aspects:
 - **buffering uses (X,Y) coordinates** in measuring distances;
 - **raster-based operation uses cells in measuring physical distances.**
 - **buffering can create more accurate buffer zones** than a raster-based operation.
 - **accuracy difference can be important**, e.g. implementing riparian zone management programs.
 - Second, **buffering operation is more flexible**- create **multiple rings** (buffer zones), whereas a **raster-based operation creates continuous distance measures.**
 - Additional **data processing (Reclassification)** is required to **define buffer zones** from continuous distance measures.
 - **buffering operation** has option of creating **separate buffer zones for each select feature or a dissolved buffer zone** for all select features.
 - **difficult to create and manipulate separate distance measures** using a **raster-based operation.**



39



Site Selection/Suitability

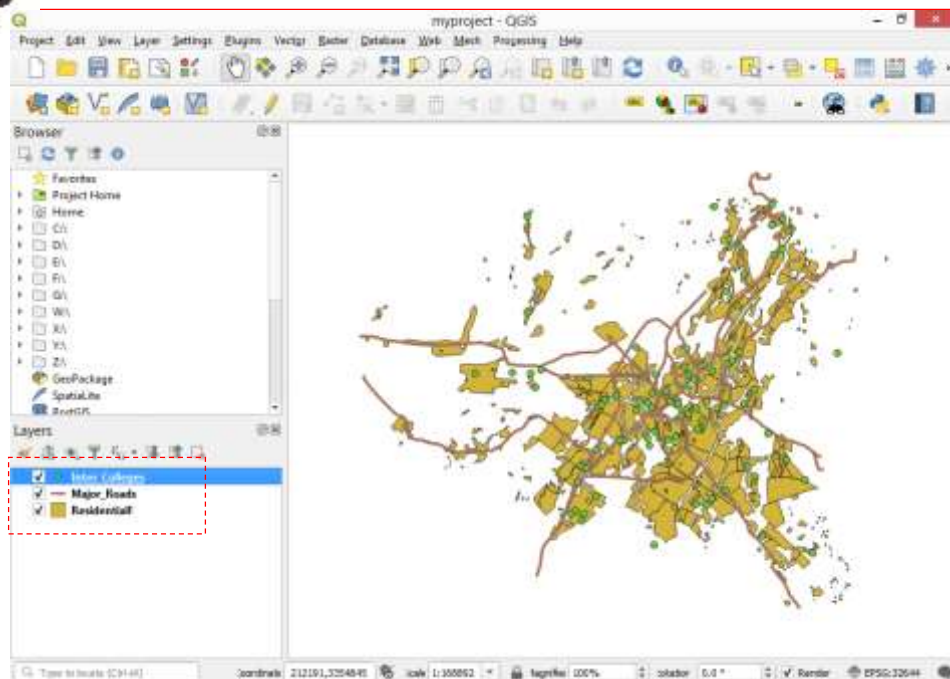
40

Site suitability analysis for construction of school

Problem: Select **suitable sites** for construction of a **new school (Intermediate School)** considering following **criteria**:

1. The school should be 1km away from existing Inter colleges.
2. School should be at least 1 km away from major roads to avoid accidents however for convenience of transportation, it should be within 2 km range.
3. To gather a good number of students and to avoid construction of school in congested areas following preference is given to residential areas:
 - a. Medium density residential area: 1st preference
 - b. Minimum and High density areas: 2nd preference
 - c. Highest density residential area: 3rd preference.

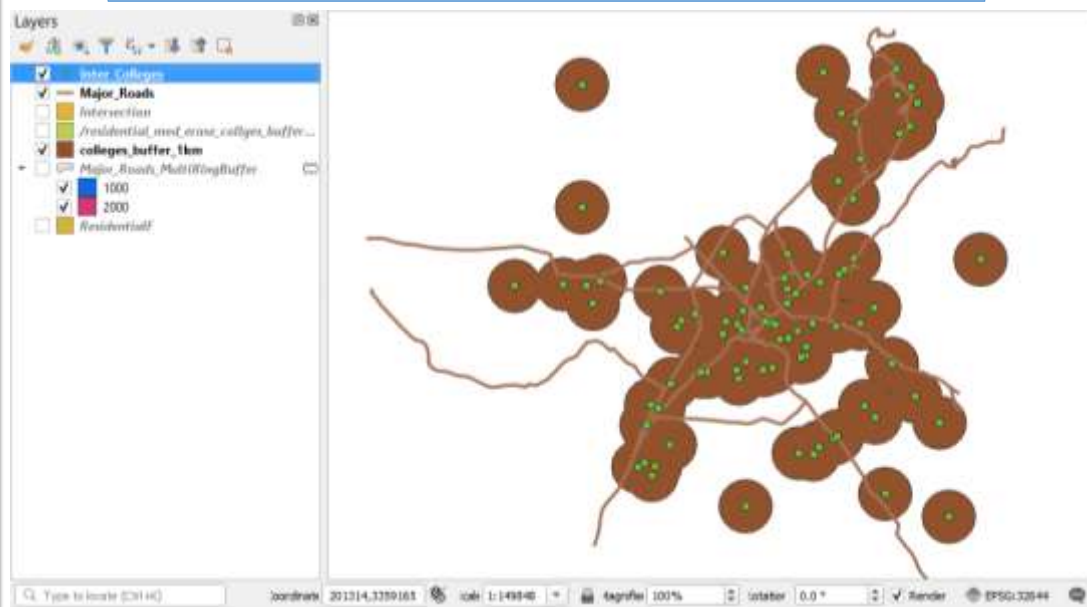
41



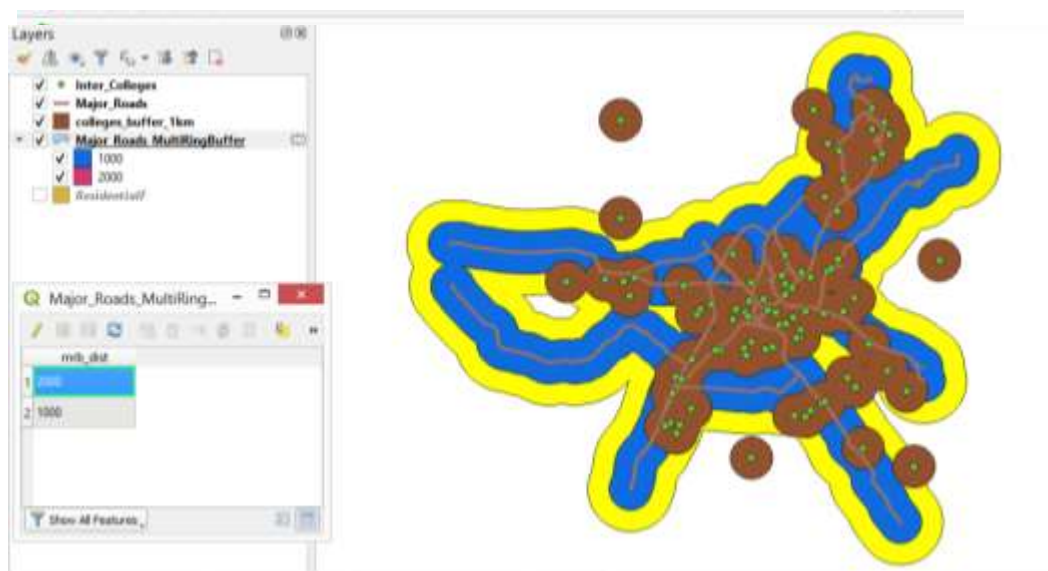
42



1. The school should be 1000 meters away from existing Inter colleges.

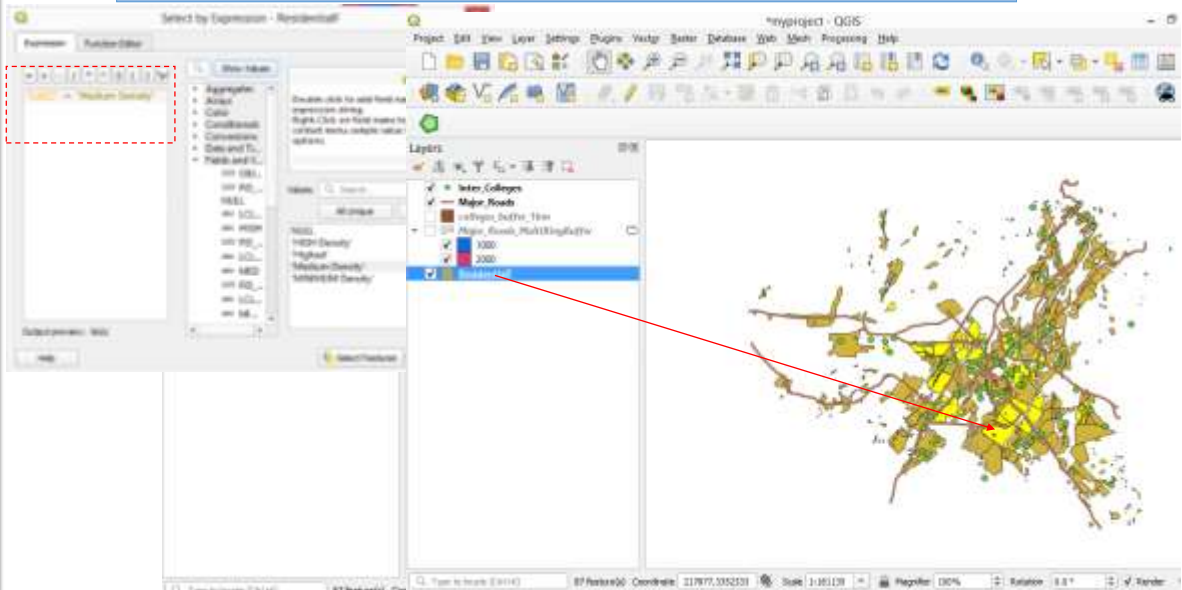


2. School should be at least 1 km away from major roads to avoid accidents however for convenience of transportation, it should be within 2 km range.





3. To find out the area which satisfies third criteria of residential density, we need to select only medium density areas which are highly suitable



Site suitability analysis for construction of school

To select suitable sites for construction of a new school (Intermediate School) considering following criteria:

1. The school should be 1000 meters away from existing Inter colleges.

2. School should be at least 1 km away from major roads to avoid accidents however for convenience of transportation, it should be within 2 km range.

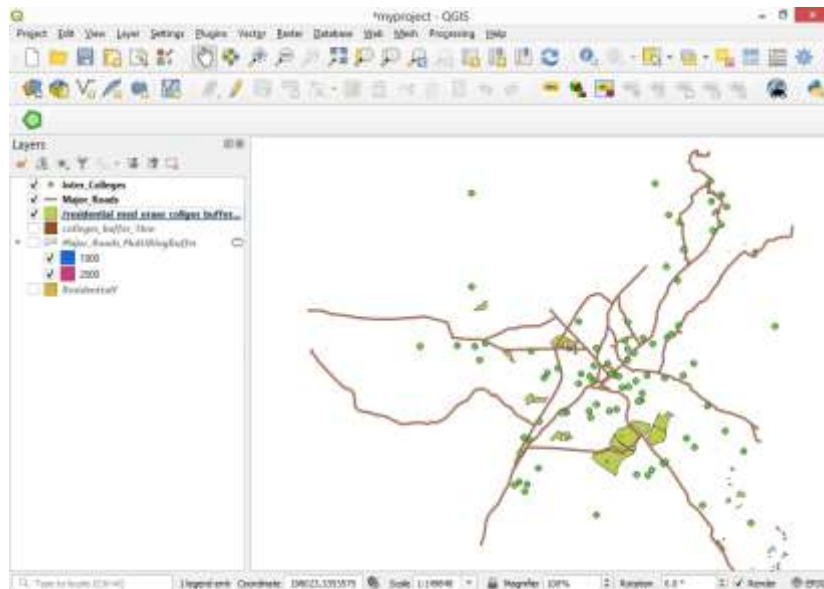
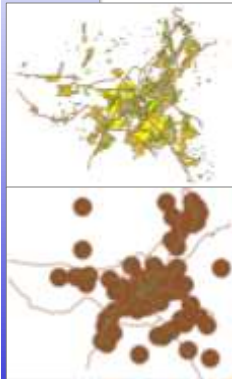
3. To gather a good number of students and to avoid construction of school in congested areas following preference is given to residential areas:

a. Medium density residential area: 1st preference

b. Minimum and High density areas: 2nd preference

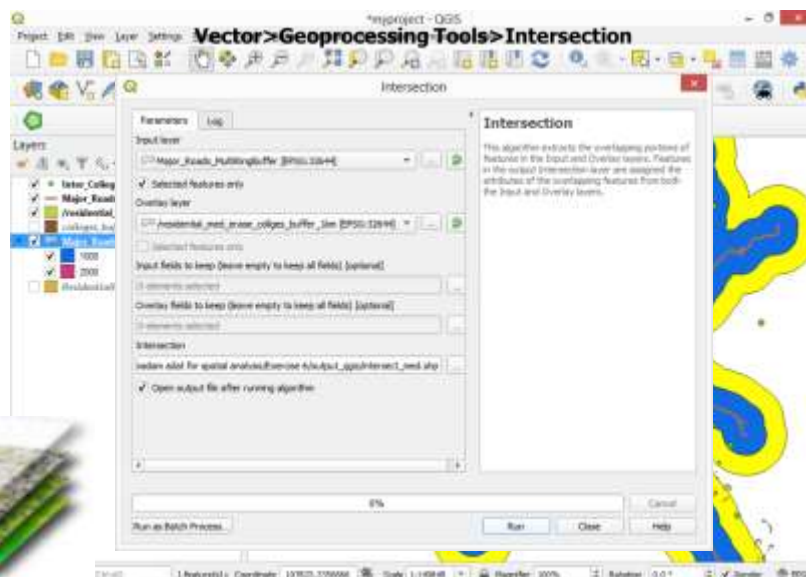
c. Highest density residential area: 3rd preference.

To select the sites which satisfy 1st and 3rd criteria, the Inter College buffer area need to be removed from selected residential area (medium density)



47

Finally to select the areas which also satisfy road (second) criteria, overlay operation is used

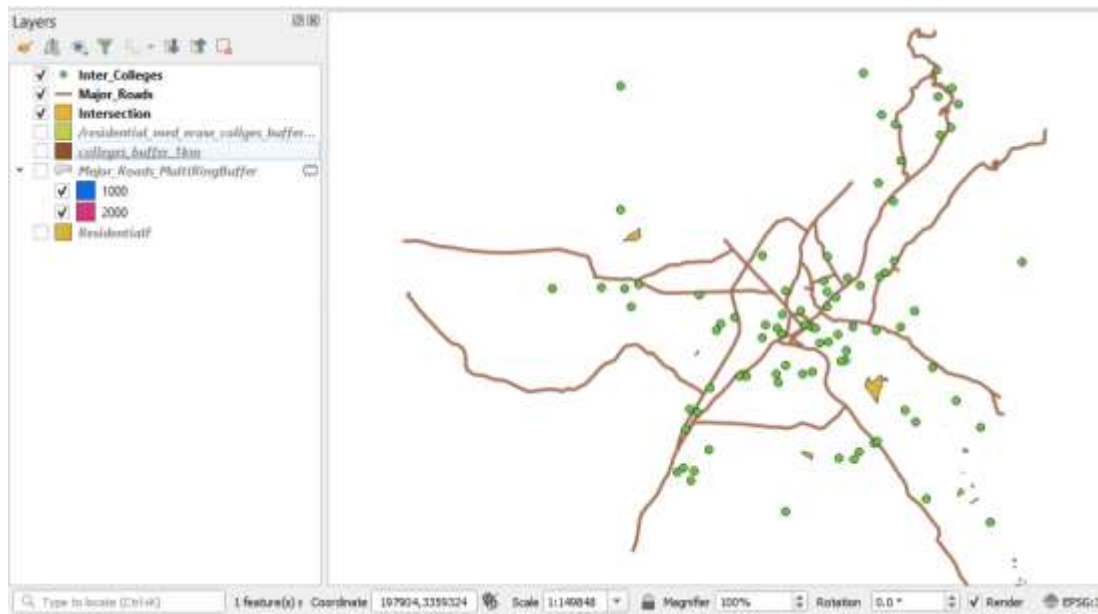


48





Final Result



Thank You

Contact Details of the Faculty:

Email- Kapil@iirs.gov.in
Tel-01352524128