UNIT 15

INTRODUCTION TO RASTER DATA ANALYSIS

Structure

15.1 INTRODUCTION

You have been introduced to vector data analysis in the previous unit. You have also learnt about fundamental concepts related to raster data and its sources in Unit 5 and Unit 6 of MGY-101. Now, in this unit, you will be introduced to raster data analysis. The raster data analysis uses a regular grid to cover the space. As you have read in the previous units, the value in each grid represents the characteristic of a spatial phenomenon at the cell location. We can perform raster data analysis at the level of individual cells, or groups of cells, or cells within a raster in a geographic information system. In raster data analysis, the type of cell value is an important consideration. The simple raster data structure with fixed cell locations is computationally efficient and facilitates a large variety of data analysis operations. Some of its operations use a single raster and others use two or more rasters. On the contrary, vector data analysis uses points, lines, and polygons.

In this unit, we will learn about various ways of raster data analysis such as, spatial measurements, neighbourhood functions and spatial interpolation.

Expected Learning Outcomes

After studying this unit, you should be able to:

- get familiarised with spatial measurements such as distance and area in raster data analysis;
- list the operations involved in raster data analysis;
- describe the local, zonal and buffer operations, and reclassification;
- know about neighbourhood functions and overlay; and
- get familiarised with spatial interpolation and digital elevation data.

15.2 OVERVIEW OF RASTER DATA ANALYSIS

Spatial analysis is often referred to as the 'heart' of the analysis can be conducted on both raster and vector data. You have read about vector data analysis in the previous unit and now you will learn about raster data analysis. GIS provides a powerful basis for raster analysis. Some of the raster operations described here will be familiar to people who regularly use software for processing digital photographs or scanned documents, or for processing the images captured by remote sensing satellites. Vector datasets are similar in form to those used in computer-assisted design software (CAD), but it is unusual to find comparable methods of analysis in these environments. In raster data GIS world is represented as a series of layers (Fig. 15.1). Each layer divides the same project area into the same set of rectangular or square cells or pixels**.** Each layer contains a unique set of information.

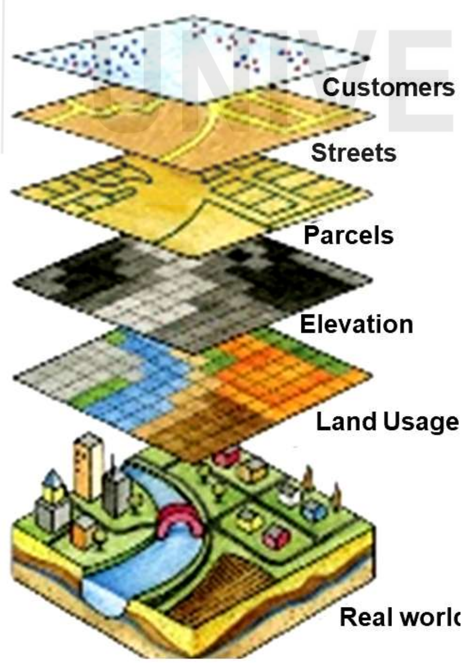
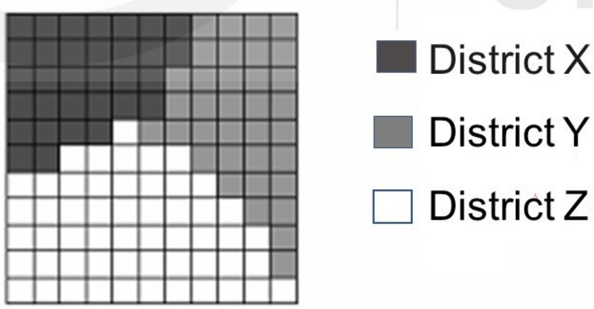


Fig. 15.1: Real-world and raster layers providing a different set of information.
(Source: https://osdi.orsac.gov.in/osdi/images/services/spatial%20analysis.jpg)

It is important to set the data analysis environment as a first step while performing raster data analysis. Data analysis environment can be defined by specifying its areal extent and output cell size because a raster operation may involve two or more rasters. The areal 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. The areal extent for raster analysis is based on the intersection or union of the rasters. The intersection or the union option use an area extent which is common to all input rasters. An operation **masking** either a feature layer or a raster can also determine the areal extent of analysis. Masking a raster is an operation required to represent and/or analyse only a subset of pixels included in a specific area or region.

Let us discuss an example- to limit soil erosion analysis for lands occupied by private parties, we can prepare a mask of either a feature layer showing private lands or a raster separating private lands. So, we can define the output cell size at any scale which can be equal to, or larger than, the largest cell size among the input rasters.

This makes it very easy to compare layers, since the pixels in each layer exactly coincide. But it also means that pixel size limits the ability to represent, since no feature on the Earth's surface is smaller than a single pixel that cannot be represented. We can represent smaller features by reducing the pixel size, but only at the cost of rapidly increasing data volume. Each layer records the values of one variable or attribute, such as land use or county name, for each pixel. The recorded values might be the average value over the pixel, the value at the exact center of the pixel, or the commonest value found in the pixel. If there are many attributes to record (for example, the census reports hundreds of attributes for each county in India), then we must create separate layers for each attribute. However, some raster GIS allow more efficient solutions to this problem, in which a single layer is used to record the district ID for each pixel. Fig. 15.2 gives the many attributes corresponding to that ID.



Districts	Population	Tax rate
X	2019	6%
Y	4511	7.5%
Z	7633	8.8%

Fig. 15.2: By using raster cells to store the IDs of polygons with multiple attributes in a related table it is possible to store many attributes in a single raster layer.

Raster data analysis includes following basic types of operations. We will briefly discuss about them in following sections.

- Spatial Measurements
- Local Operations
- Reclassification
- Neighbourhood Operations
- Zonal Operations
- Buffer Operations
- Overlay
- Spatial Interpolation
- Digital Elevation Data

15.3 SPATIAL MEASUREMENTS

Calculating lengths, perimeters and areas is a very common application of GIS. All measurements from spatial data are an approximation. You have learnt about the spatial measurements in vector data in the previous unit. The measurement of the length of a river or a road from a digital map is a relatively simple task. We can obtain different measurements depending on the type of spatial data used, i.e. raster or vector and the measurement methods employed. The vector data comprise straight-line segments (lines appearing ascurves on the screen are stored as a collection of short straight-line segments). In the raster data, all entities are estimated using a grid cell representation. With raster data, there is more than one answer to the question:

‘What is the distance between X and Y?’ -where X and Y are two ends of a straight line. The answer depends on the measurement method used. Normally the shortest path, or Euclidean distance, is calculated by drawing a straight line between the end points of a line-X and Y and creating a right-angled triangle so that Pythagorean geometry can be used (Fig.15.3a). Thus, the distance XY is calculated using the formula:

$XY_2 = XZ_2 + ZY_2$

Therefore:

$XY = XZ_2 + ZY_2$

‘Manhattan’ distance’ can be calculated. This is the distance along raster cell sides from one point to the other. This name comes from the way in which you would get across a city, like Manhattan, consisting of dense ‘blocks’ of buildings.

Third method of calculating distance in a raster GIS uses ‘proximity’. In this method, concentric equidistant zones are established around the start point X (Fig.15.3c). The resulting image shows the shortest straight-line distance from every point on the map (including end point Y) to the location of interest (X). Thus, the distance from A to B can be ascertained.

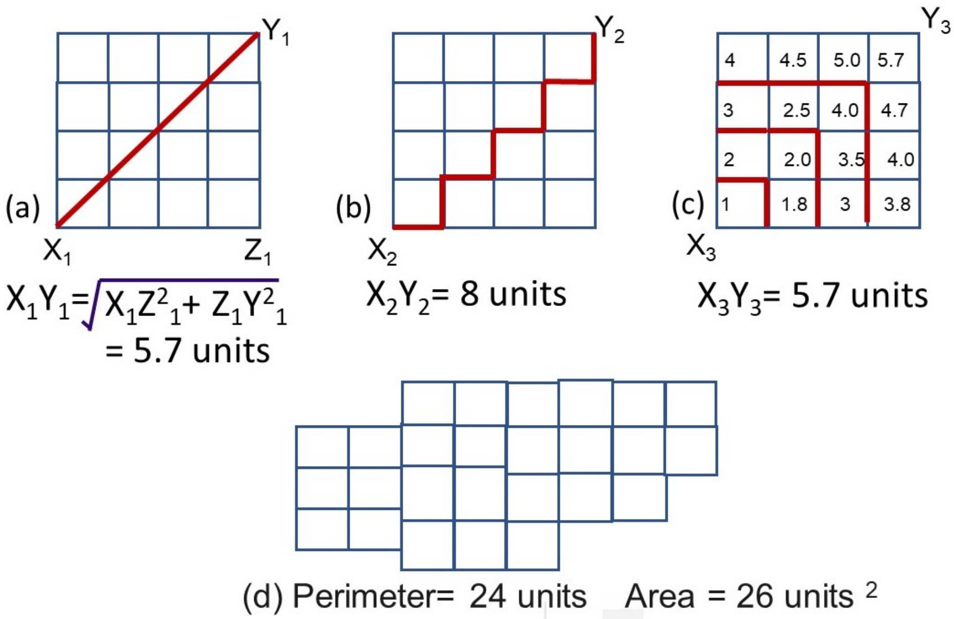


Fig. 15.3: Raster GIS measurements: a) Pythagorean distance; b) Manhattan distance; c) proximity distances; and d) perimeter and area.

The perimeter measurement in a raster GIS, the number of cell sides make up the boundary of a feature is multiplied by the known resolution of the raster grid. For area calculations, the number of cells a feature occupies is multiplied by the known area of an individual grid cell (Fig. 15.3d). Area and perimeter calculations in raster data can be affected by the cell size, origin and orientation of the raster grid.

15.4 LOCAL OPERATIONS

The local operations constituting the core of raster data analysis, are based on point-by-point or cell-by-cell operations. A new raster from either a single input raster or multiple input rasters are created by local operations. The cell values of the new raster can be computed by a function relating the input to the output or are assigned by a classification table. The most important among local operations *i.e.* overlay analysis. In the raster-based overaly analysis, either the logical or arithmetic operators are used. The logical methods use operators AND, OR, and XOR (exclusive OR). Mathematically AND multiplies the individual cells. Whereas logical OR and XOR add individual values of corresponding cells. The appropriate coding of the features in the input layers is the most important consideration in a raster overlay. The simplest kind of local operations occur when a single raster layer is processed. This is often done to apply a simple reclassification; such as “All areas of Soil Classes 1, 3, and 5 are suitable building sites for residential development; all other areas are not.” This could be operationalized by reclassifying the soil class layer, assigning 1 to all pixels that currently have Soil Class 1, 3, or 5, and 0 to all other pixels. GIS software with well-developed raster capabilities, allows the user to access a wide range of operations of this type that creates a new layer through a local operation on an existing layer. Local operations with multiple rasters are referred to as overlaying, compositing, or superimposing maps. Summary statistics, including maximum, minimum, range, sum, mean, median, and standard deviation, are measures that apply to rasters with numeric data (Fig. 15.4).

Arithmetic	+, -, /, *, absolute, integer, floating-point
Logarithmic	Exponentials, logarithms
Trignometric	sin, cos, tan, arcsin, arccos, arctan
Power	Square, square root, power

Fig. 15.4: Arithmetic, logarithmic, trigonometric and power functions for local operations.

15.5 RECLASSIFICATION

We have briefly understood recoding/reclassification in Unit 17 Post-Classification Methods in MGY-102. Reclassification is also referred to as recoding or transforming, and can create a new raster by classification in a local operation. There are two reclassification methods that may be used:

- First method includes one-to-one change which means that a cell value in the input raster is allotted a new value in the output raster;
- Second method designates a new value to a range of cell values in the input raster.

Let us understand by an example - cells with population densities between 0 and 50 persons per square kilometre 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. Floating-point cell values in raster data are used to store continuous data. This type of data may also be referred to as non-discrete data, surface data, or field data. The numbers are stored in a format that allows many significant digits and great range. The second method assigns a new value to a range of cell values in the input raster.

Reclassification serves three main purposes:

1. It can create a simplified raster, e.g. instead of having continuous slope values, a raster can have 1 for slopes of 0 to 20 per cent, 2 for 20 to 40 per cent, etc;
2. It can create a new raster that contains a unique category or value such as slopes of 20 to 40 per cent; and
3. It can create a new raster that shows the ranking of cell values in the input raster, e.g. a reclassified raster can show the ranking of 1 to 5, minimum value being least suitable and maximum being most suitable.

Reclassification is an important variation of query ideas in GIS. It can be used in place of a query in a raster GIS. Suppose in a raster land use image we have a query 'Where are all the forested areas?' This could be obtained using a query or by reclassifying the image. Reclassification may result in a new image.

Let us consider an example- if cells representing the forested area in an original image had a value of 10, the resulting reclassification can be:

Cells with values = forested area (value 10) take the new value of 1

Cells with values ≠ of forested artichokes the new value of 0

Now the reclassification would result in a new image with all areas of forestry coded with value 1. All the areas that are not forested are coded with the value 0. This is a Boolean image. Thus, the reclassification produces a two-code image from a complex original image. It is a very useful technique as it allows the resulting image (containing only values 1and 0) to be used in further analysis.

Boolean expressions used in GIS are AND, OR, NOT and XOR. These are explained best with the help of Venn diagrams, where each circle in the diagram represents the set of data meeting a specific criterion (Fig.15.5). In the diagrams, X is the set of schools that are in the 'residential school' category, and Y is the set of Schools that have more than 30 classrooms.

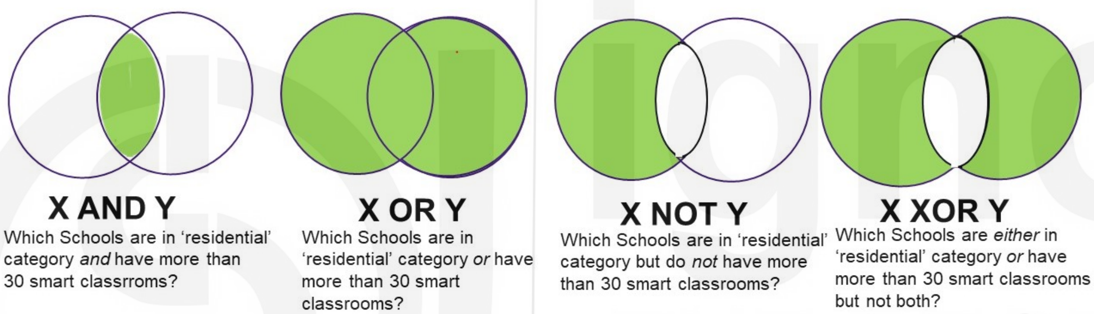


Fig. 15.5: Boolean operators: Venn diagrams.

In an example of case study, reclassification was used to allocate a new value to different land use classes based on their ecological importance (Table 15.1).

Table 15.1: Reclassification values for the land use data layer.

Land use	Original value	New value after reclassification: Boolean example	New value after reclassification: Weighing example
Forested area	10	1	5
Water	11	0	2
Settlement	12	0	1
Agricultural land	13	0	4

In that case the forested and agricultural land was taken to have a relatively high conservation value. Whereas, the water and settlement was considered having lower conservation value. After implementation of a set of rules, a new image was produced after:

- Cells with original value = 10 take thenew value of 5
- Cells with original value = 11 takes thenew value of 2
- Cells with original value = 12 takes thenew value of 1

- Cells with original value = 13 should take the new value of 4

The resulting map helped the planner to identify areas of high conservation value. However, the new classes (1, 2, 4 and 5) are still simply labels (ordinal values) and care needs to be taken to ensure appropriate further analysis of the image.

In the previous sections, we have read about spatial measurements, reclassification and local operations. Before reading further, let us spend five minutes to check our progress.

SAQ I

a) List the operations involved in raster data analysis.

b) Mention three methods of measurement of raster data.

c) Local operations are based on ________________analysis.

d) What are the two reclassification methods?

15.6 NEIGHBOURHOOD OPERATIONS

Neighbourhood operations are also called as focal operations. They involve a focal cell and a set of its surrounding cells. We choose these cells for their distance and/or directional relationship to the focal cell. The common neighbourhoods include circles, rectangles, wedges, and annuluses (Fig. 15.6).

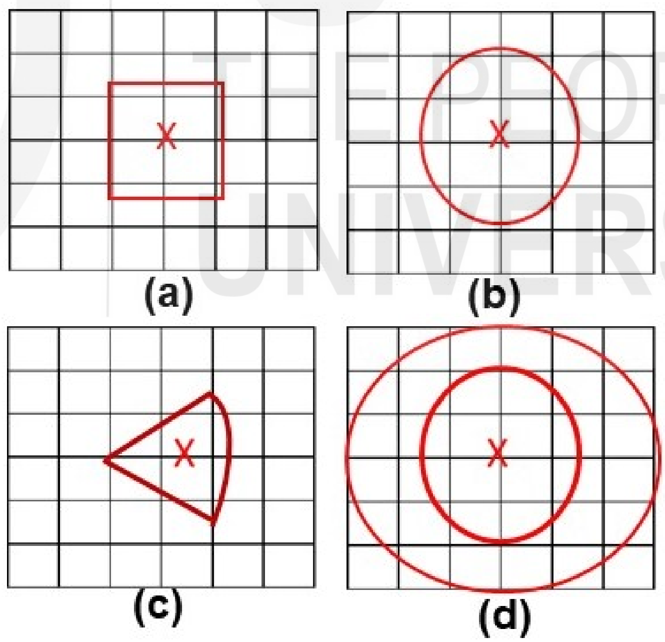


Fig.15.6: Four common neighbourhood types: a) rectangle; b) circle; c) annulus; and d) wedge. The cell marked with an x is the focal cell.

The neighbourhood or focal operations produce results by analyzing each pixel in relation to its immediate neighbours. Some of these operations work on all eight immediate neighbours, but others focus only on the four neighbours that share a common edge and ignore the four diagonal neighbours. Sometimes we distinguish these options by referring to moves in the game of chess - the eight-neighbour case is called the queen's case, and the four-neighbour case is called the rook's case (Fig. 15.7).

Fig. 15.7: Two definitions of a raster cell's neighbourhood. In the rook's case only the four cells that share an edge are neighbours, but in the queen's case the neighbourhood includes the four diagonal neighbours.

Convolutions are the most useful among the focal operations because the output is like an averaging over the immediate neighbourhood. For example, we might produce an additional layer in which each pixel's value is the average over the values of the pixel and its 8 queen's case neighbours. Fig. 15.8 shows a simple instance of this. The result of this operation produces an output layer that is smoother than the input, by reducing gradients, lowering peaks and filling valleys. The layer has been filtered to remove some of the variation between pixels and to expose more general trends. Repeated application of the convolution eventually smoothens the data.

10	9	7	4	5
9	6	5	3	4
6	3	2	2	3

8.5	7.5	5.7	4.7	4.0
7.2	6.3	4.6	3.9	3.5
6.0	5.2	3.5	3.2	3.0

Fig. 15.8: Application of a convolution filter. Each pixel's value in the new layer is the average of its queen's case neighbourhood in the input layer. The result is smoother, picking up more of the general trend. At the edge, where part of the neighbourhood is missing, averages are over fewer than nine pixels.

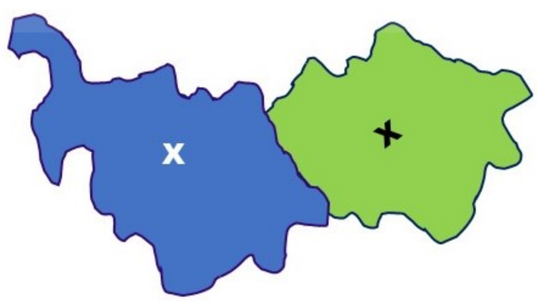
Convolutions are very useful in remote sensing, where they are used to removing noise from images. Of course, averaging can only be used if the values in each pixel are numeric and measured on continuous scales. Thus, this operation will make no sense if the values represent classes of land, such as soil classes. But in this case, it is possible to filter by selecting the commonest class in the neighbourhood, rather than by averaging (Fig. 15.8). Note also that special rules have to be adopted at the edge, where cells have fewer than the full complement of neighbours.

Typically, a neighbourhood operation uses the cell values and then designates the computed value to the focal cell. To complete a neighbourhood operation on a raster, the focal cell is moved from one cell to another until all cells are visited. Different rules devised by software developers apply to focal cells on the margin of a raster, where a neighbourhood such as a 3-by-3 rectangle cannot be used. A simple rule is to use only cell values available within the neighbourhood (e.g., 6 instead of 9) for computation. Although a neighbourhood operation works on a single raster, its process is like that of a local operation with multiple rasters. A neighbourhood operation uses the cell values from a defined neighbourhood instead of using cell values from different input rasters. The output obtained from a neighbourhood operation can show summary statistics including maximum, minimum, range, sum, median, mean, and standard deviation, as well as a tabulation of measures such as majority, minority, and variety.

15.7 ZONAL OPERATIONS

Raster data is represented in pixels; therefore, line objects and area objects are displayed as collections of pixel values, with no clear linkage between them. If we simply count the pixels assigned to Class 2 on a land usage layer, for example, we get the total area of Class 2, not the separate areas of the individual patches of Class 2. Zonal operations attempt to address this, by focusing on operations that examine a layer as a collection of zones, each zone being defined by contiguous pixels of the same class. Thus, you can know that the zonal operation works with groups of cells of the same values or similar features and these groups are called zones.

Zones may be contiguous or non-contiguous. A contiguous zone comprises cells that are spatially connected, while a non-contiguous zone includes separate regions of cells. An example of a contiguous zone is watershed raster in which the cells belong to the same zone and are spatially connected. An example of a non-contiguous zone is a land use raster, in which one specific type of land use may **appear in different parts of the raster.** Zonal operations can measure the areas of patches or their perimeter lengths, and the results are returned as new layers in which each pixel is given the measure, and evaluated over the zone of which the pixel is a part (Fig. 15.9).



Zone	Area	Perimeter	Thickness
1	37,504	1604	75.1
2	42,654	1456	75.3

Fig. 15.9: Thickness and centroid for two cultivated areas. Area is measured in square kilometres, and perimeter and thickness are measured in kilometres. The centroid of each zone is marked with an X.

15.8 BUFFER OPERATIONS

Buffering, as already discussed in the previous unit, is used to identify a zone of interest around an entity or set of entities. Buffering lines and areas create new areas. It is another powerful form of local operation which forms the raster equivalent of the buffer operation and also supports a range of other operations concerned with finding routes across surfaces. To determine a buffer on a raster, it is necessary only to determine which pixels lie within the buffer distance of the object. Figure15.10 shows how this works in the cases of a point and a line. But suppose we make the problem a little more complex, by asking for pixels that are within a certain travel time of a point and allowing travel speed to be determined by an additional layer. Figure15.11 shows such a layer, and the result of determining how far it is possible to travel in given numbers of minutes. The operation is also known as **spreading**. Of course, this assumes that travel is possible in all directions, but this would be appropriate with an aircraft looking for the best route across the Atlantic, or a ship in the open ocean.

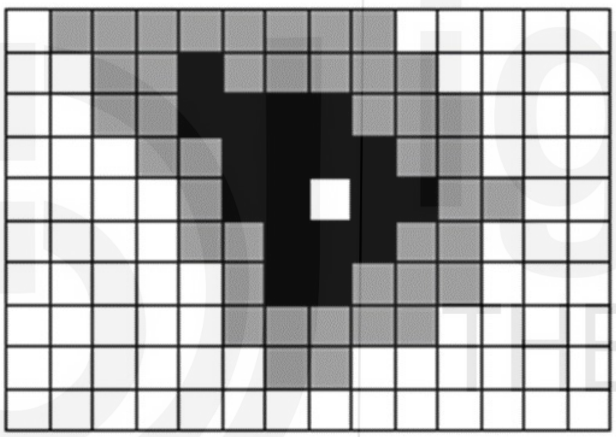


Fig. 15.10: The buffer operation in its raster form, for a point and a line (indicated by the white cells).

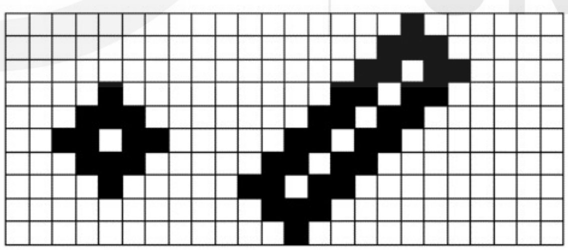


Fig. 15.11: Illustration of a spreading operation over a variable friction layer. From the central white cell, all cells in the black area can be reached in 10 minutes, and all cells in the grey area can be reached in 20 minutes.

The creation of a zone of interest around an entity is the most common example is buffering. Other neighbourhood functions include data filtering, which involves the recalculation of cells in a raster image based on the characteristics of neighbours. This method is often used to find the best routes for power transmission lines, pipelines, highways, and military vehicles. In all the cases, the assumption of travel in all directions is reasonably accurate. Origin and destination points for the route are defined. A raster GIS is used to create a layer of travel speed or friction. Because this is not uniform, the best route between the origin and destination is not necessarily a straight line. Instead, the best route is modelled as a series of moves in the raster, from each cell to the most appropriate to its eight neighbours, until a complete route of the least total time or least total cost is found (Fig. 15.12). There have been many applications of this GIS method over the past three decades by highway departments, power utilities, and pipeline companies.

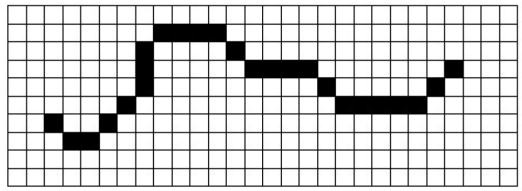


Fig. 15.12: Finding the least-cost path across a variable-cost layer from a origin to a destination, using queen's case moves between raster cells. This form of analysis is often applied to find routes for power lines, pipelines, or new highways.

15.9 OVERLAY

We had discussed the overlay operation in the previous unit. Now we will learn about overlay operation in raster data analysis. We have learnt that everything is represented by cells and grids in the raster data structure. For example, a point is represented by a single cell and a line is represented by a string of cells and an area by a group of cells. However, the methods of performing overlays in raster data differ from those in vector GIS. Overlay of raster datasets combines the pixel-based calculations or map algebra. This is an informal language with syntax like algebra. It can facilitate manipulation and analysis of raster data. Map algebra uses an expression to link the input and output. Apart from this, the expression can be composed of mathematical, GIS tools, operators, and constants. GIS tools include the basic tools of local, focal, zonal, and distance measurement operations. Input data layers may be added, subtracted, multiplied, or divided to produce output data using map algebra. To produce an output value, we perform the mathematical operations on individual cell values fromtwo or more input layers. Therefore, the most important consideration in raster overlay is the appropriate coding of spatial elements, viz. point, line and area features in the input data layers.

The overlay is a fast, straight forward and efficient data sets for operations. As you have read, it is known as a cell-by-cell combination or operation and is computationally less demanding. In raster datasets, overlay includes two or more different data that derive from a common grid. Each separate set of data possess usually possesses specified numerical values (Fig. 15.13). In the raster overlay these values from each set are mathematically merged to create a new set of values for a single output layer. The raster-based overlay is done to create sustainability assessments, risk surfaces, value assessments, and other procedures. For example, a raster-based overlay divides the habitat of an endangered species into a grid, and then, after getting data for multiple factors that influence the habitat, creates a risk surface to illustrate sections of the habitat needing protection. A new layer of values is produced from each pair of coincident cells. Values of these cells can be added, subtracted, divided or multiplied, we can extract the maximum value, mean value calculated, a logical expression computed, and so on. The output cell simply takes on a value equal to the result of the calculation.

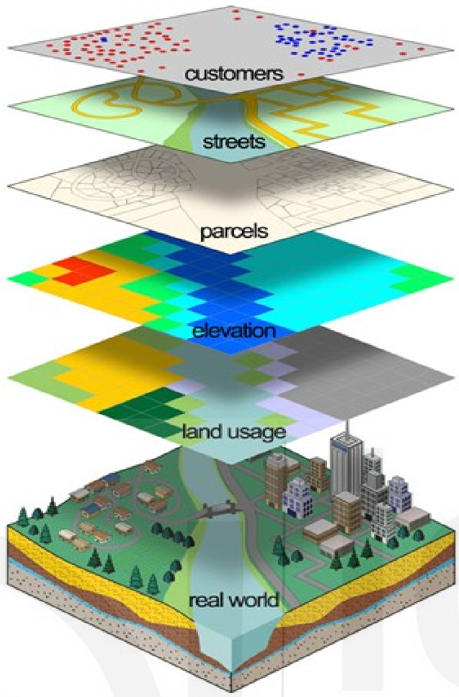


Fig. 15.13: Overlay of GIS layers. (Source: http://wiki.gis.com/wiki/images/1/11/Gis_layers.png)

A local operation with multiple rasters is frequently compared to a vector-based overlay operation. The two operations are similar in that they both use multiple data sets as inputs. However, important differences exist between them. This type of computation done in a vector-based overlay operation involving intersections between features and insert points is unnecessary for a raster-based local operation. This is because the input rasters have the same cell size and area extent. If the input rasters have to be first resampled to the same cell size, the computation is less complicated than calculating line intersections. The raster-based local operation has access to various tools and operations from the input layers to create the output as compared to a vector-based overlay operation. The raster-based overlay is often preferred for projects that involve numerous layers anda significant amount of computation.

Thus, a raster-based local operation is computationally more efficient than a vector-based overlay operation. However, the vector-based overlay operation has its advantages. This overlay operation can combine multiple attributes from each input layer. Once combined into a layer, all attributes can be queried and analysed individually or in combination. For example, a vegetation layer can have the attributes of height, crown closure, stratum, and crown diameter. Similarly, soil layers can have attributes such as depth, texture, organic matter, and pH value. An overlay operation combines the attributes of both layers into a single layer and allows all attributes to be queried and analysed. To query or analyse the same vegetation and soil attributes as mentioned above, a raster-based local operation would require one raster for each attribute. Whereas a vector-based overlay operation will be more efficient than a raster-based local operation if the data sets are to be analysed have many attributes. The raster overlay is affected by the resolution (cell size) and scale of measurement. The scale of measurement of both the input and analysis layers should be compatible. Basic arithmetic operators in raster overlay operations are addition, subtraction, multiplication, and division.

15.10 SPATIAL INTERPOLATION

Spatial interpolation may be defined as using points with known values at unknown points. It involves the calculation of the value for a query point or a raster cell with an unknown value from a set of known sample points values that are distributed across an area. A spatial data set would provide an observed value at every spatial location, in an ideal situation. Spatial interpolation is characteristically applied to a raster with estimates made for all cells in GIS applications. Spatial interpolation is therefore a means of creating surface data from sample points. Spatial interpolation is therefore a means of converting point data to surface data so that this data can be used with other surfaces for analysis and modelling. Satellite or aerial photography operates the same way in providing such type of data. More often, data are stratified, consisting of regularly spaced observations. The role of interpolation is to fill in the gaps between the observed data points. Thus, spatial interpolation is using points with known values to estimate values at other points. Spatial interpolation requires two basic inputs. They are known **points** (Fig. 15.14) and an **interpolation method**.

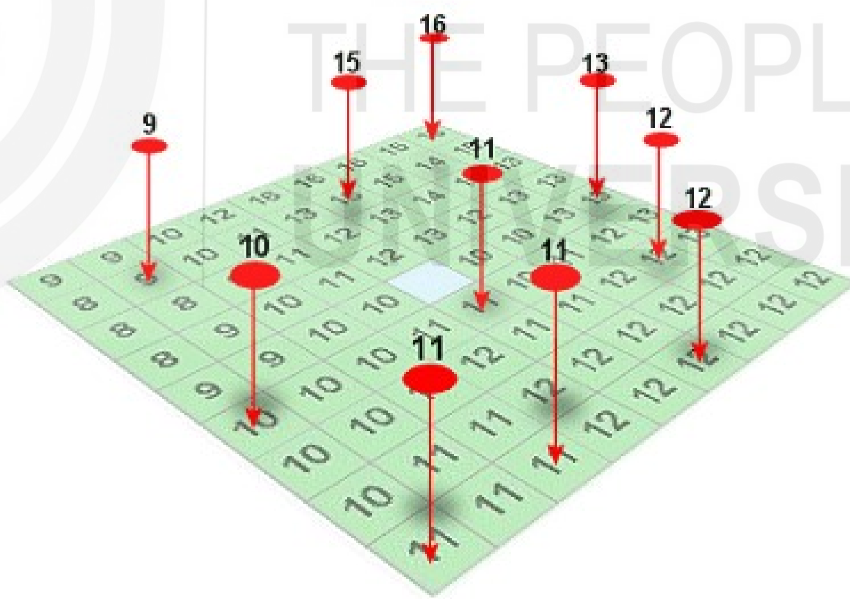


Fig. 15.14: Interpolating the sample points values creates a surface. The unknown value of the light-blue cell in the center will be estimated based on values of the surrounding sample points. (Source: www.geography.hunter.cuny.edu/~jochen/gtech361/lectures/lecture11/concepts/What%20is%20interpolation.htm)

In most of the cases, known points are actual points such as weather stations and archaeological sites. Control Points are points with known values. They are also called known points, sample points, or observations. The control points can influence the accuracy of spatial interpolation. They provide the data necessary for developing an interpolator for spatial interpolation. The value to be estimated at a point is more influenced by nearby known points than those farther away. Therefore, for effective estimation, control points should be well distributed within the study area.

Spatial interpolation methods can be categorized into:

1) First, we can group them into global and local methods. A global interpolation method uses each known point available to estimate an unknown value. A local interpolation method uses a sample of known points to estimate an unknown value.

2) Second, we can group spatial interpolation methods into exact and inexact interpolation (Fig.15.15). Exact interpolation predicts a value at the point location that is the same as it's known value. Exact interpolation generates a surface that passes through the control points. In exact interpolation predicts a value at the point location that differs from its known value.

3) Third, spatial interpolation methods may be deterministic or stochastic. **Deterministic interpolation** method provides no assessment of errors with predicted values. Whilst **stochastic interpolation** method considers some randomness as its variable and offers an assessment of prediction errors.

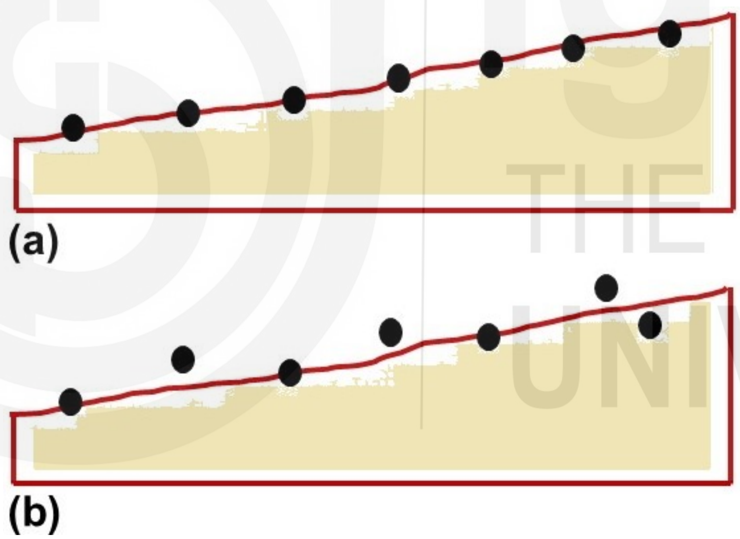


Fig. 15.15: a) Exact interpolation; and b) Inexact interpolation.

In the previous sections, we have learnt about neighbourhood operations, zonal operations, buffer operation, overlay and spatial operations in raster data analysis. Before reading further, let us spend five minutes to check our progress.

SAQ II

a) Neighbourhood operations are also called as ___________operations.

b) Draw the diagram of the queen's case.

c) Distinguish between contiguous or non-contiguous zonal operations.

d) Define spatial interpolation.

15.11 DIGITAL ELEVATION DATA

Digital Elevation data is also referred to as digital elevation models (DEMs), or digital terrain data. They are digital representations of the shape of the Earth's surface. Typically, digital elevation data comprises arrays of values that represent topographic elevations measured at equal intervals on the Earth's surface (Fig. 15.16). Such data can be analogous to digital remote sensing images, except that each pixel represents an elevation measurement rather than a brightness value. Digital elevation data can be manipulated, classified and analysed in many of the same ways that digital remote sensing data can be displayed as pictorial images.

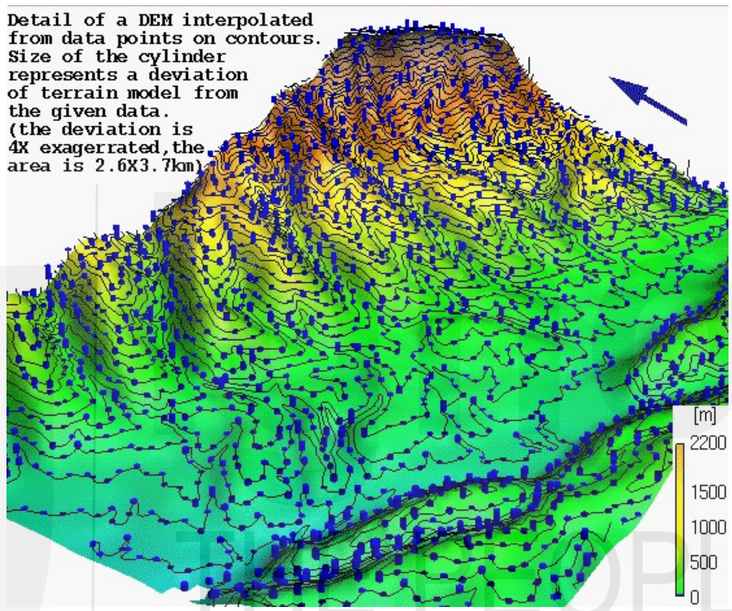


Fig. 15.16: DEM interpolated from data points on contours. (Source: http://fatra.cnr.ncsu.edu/~hmitaso/gmslab/lcgfin/img/elevpts.gif)

Slope and aspect can also estimate the pattern of surface water flow over a DEM (digital elevation model). It is a very useful operation in determining watershed boundaries and other aspects of surface hydrology. Each pixel's elevation is compared to those of its eight neighbours. If at least one neighbour is lower, then the GIS infers that water will flow to the lowest neighbour. If no neighbour is lower, the pixel is inferred to be a pit, in which a shallow lake will form. If this rule applies to every pixel in a DEM, the result is a tree-like network of flow directions (Fig. 15.17), with associated watersheds. There are many more advanced versions of this simple algorithm, along with a range of sophisticated methods for studying water flow (hydrology) using GIS.

If the input layer is a digital elevation model, with pixel values equal to terrain elevation, local operation can calculate slope and aspect. We normally do this by comparing each pixel's elevation with those of its eight neighbours and applying simple formulae (Burrough and McDonnell, 1998). The result is not the actual slope and aspect at each pixel's central point, but an average over the neighbourhood. Because of this, slope and aspect estimates always depend on the pixel size (often called the distance between adjacent postings) and will change if the pixel size changes. For this reason, it is always best to quote the pixel size when dealing with slope or aspect in a GIS.

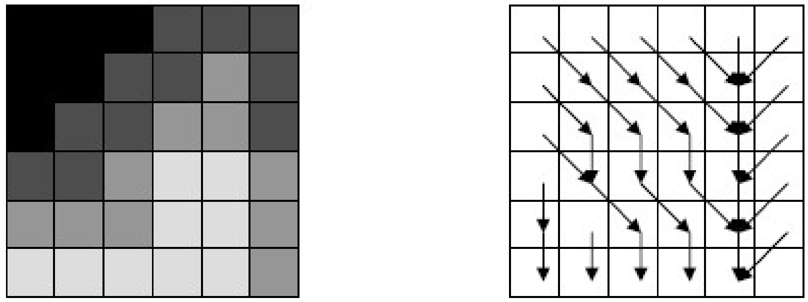


Fig. 15.17: A digital elevation model, and the result of inferring drainage directions using the simple rule "If at least one neighbour is lower, the flow goes to the lowest neighbour".

Viewshed Operation also works on a DEM and is used to compute the area that can be seen from a specified point, by a person at that point. When over one point is specified, the GIS could determine the area visible from at least one point, or in some military applications, it might be useful to determine the area invisible from all points. We often use the viewshed operation in planning. For example, if a forest management agency proposes a new clearcut it might be useful to determine whether the scar will be visible from certain key locations, such as public roads. If a newspaper mill is proposed, it might be useful to determine how visible it will be from surrounding areas. One complication of the viewshed operation as implemented on a DEM is that it is difficult to incorporate the effects of tree cover on visibility since a DEM records the elevation of the ground surface, not the top of the tree canopy.

15.12 SUMMARY

The vast number of options is available as methods of spatial analysis and modelling implemented in GIS. Any good GIS software provides many of these and more are available as add-ons from other vendors, agencies, or individuals. Let us sum up about what we have learnt in this unit:

- Spatial measurements, local operations, reclassification, neighbourhood operations, zonal operations, buffer operations, overlay, spatial interpolation and digital elevation data are the operations used in the raster data analysis.
- The local operations are based on point-by-point or cell-by-cell analysis. In the raster-based analysis, either the logical or arithmetic operators are used.
- Reclassification is also referred as recoding or transforming can create a new raster by classification in a local operation.
- Neighbourhood operations are also called as focal operations. We chose the focal cell for their distance and/or directional relationship to the focal cell.
- Zonal operation works with groups of cells of the samevalues or similar features and these groups are called zones. Zones may be contiguous or non-contiguous.
- Buffering, as already discussed in the previous unit, is used to identify a zone of interest around an entity or set of entities. If a point is buffered, a circular zone is created.

- Buffering lines and areas create new areas. It is a powerful form of local operation which forms the raster equivalent of the buffer operation and also supports a range of other operations concerned with finding routes across surfaces.
- Overlay of raster datasets combines Pixel-based calculations or Map Algebra. It can facilitate manipulation and analysis of raster data. Map algebra uses an expression to link the input and output.
- Spatial interpolation may be defined as approximating the values of properties at unsampled sites within the area covered by existing observations.
- Digital Elevation data, also referred to as digital elevation models (DEMs), or digital terrain data, are digital representations of the shape of the Earth's surface.

15.13 TERMINAL QUESTIONS

1. Discuss in brief spatial measurements used in calculating lengths, perimeters and areas in raster data.
2. Describe methods ofreclassification.
3. Give an account of neighbourhood operations employed in raster data analysis.
4. With the help of neat and well-labelled diagrams, explain buffer operations used in raster data analysis.

15.14 REFERENCES

- Burrough, P. A. and McDonnell, R. A. (1998) Principles of geographical information systems. New York: Oxford University Press. 332p.
- Chang, Kang-tsung, (2010) Introduction to Geographic Information Systems. 4th Ed., Tata McGraw Hill Education Pvt. Ltd. 450p.
- DeMers, M. N. (2000) Fundamentals of Geographic Information Systems. 2nd Ed., John Wiley & Sons, INC., 498p.
- Repeated application of the convolution eventually smoothens the data

(Data from the website was retrieved from 20th till 30th September 2022)

15.15 FURTHER/ SUGGESTED READINGS

- Chang, Kang-tsung, (2010) Introduction to Geographic Information Systems. 4th Ed., Tata McGraw Hill Education Pvt. Ltd. 450p.
- DeMers, Michael N., (2000) Fundamentals of Geographic Information Systems. 2nd Ed., John Wiley & Sons, INC., 498p.

15.16 ANSWERS

SAQ I

a) Euclidean distance, Manhattan' distance, proximity.

b) Spatial Measurements; Local Operations; Reclassification; Neighbourhood Operations; Zonal Operations; Buffer Operations; Overlay; Spatial Interpolation; Digital Elevation Data.

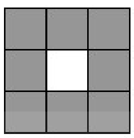
c) Cell-by-cell.

d) First method includes a one-to-one change, which means that a cell value in the input raster is allotted a new value in the output raster. The second method designates a new value to a range of cell values in the input raster.

SAQ II

a) Focal.

b) 2.



c) A contiguous zone comprises cells that are spatially connected, while a noncontiguous zone includes separate regions of cells.

d) Spatial interpolation may be defined as the process of approximating the values of properties at unsampled sites within the area covered by existing observations.

Terminal Questions

1. Please refer to section 15.3.
2. Please refer to section 15.5.
3. Please refer to section 15.6.
4. Please refer to section 15.8.