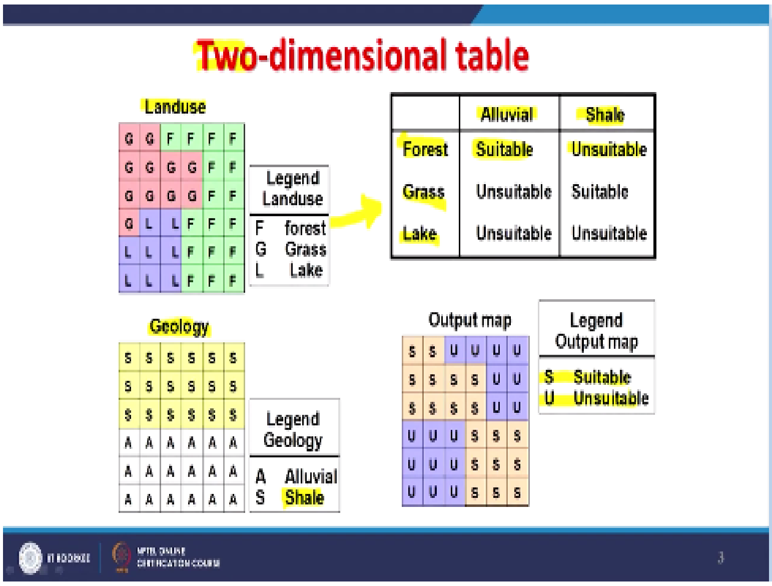
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology - Roorkee

Lecture - 22
GIS Analysis - 03

Hello everyone! and welcome to our new discussion which is on GIS analysis part 3. In the previous discussion, we touched about the overlay operations and also went to some extent in details. Today, we will also continue little bit on that and then we will take different analytical operation. So, these overlaying operations as discussed that this is based on set theory or Boolean logic.

(Refer Slide Time: 00:58)



And it has got you know various applications. And this is what in GIS is that you are having data in different layers and you can analyse using either just simple 2 maps or several maps together. So, we have yesterday discussed you know cross table. Similarly, it is also called in some other way, we can also think of 2-dimensional table which is quite close to the cross table.

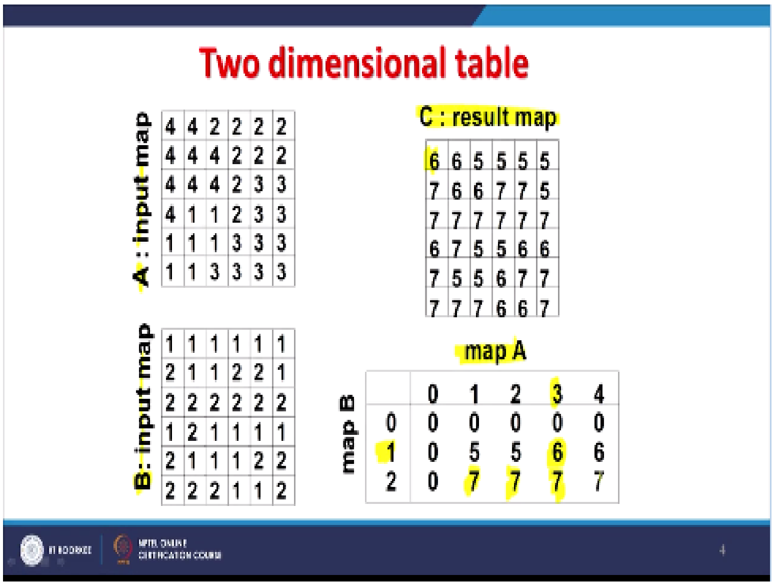
But here, these 2-dimensional tables can also be used for change detection studies. And this is all based on what overlaying operations. So, if we take example here that we are having one land use map, another one is geology map. And if I want to create a 2-dimensional table and that means I am overlaying these 2 and here the syntax part is not here. But how to create a 2-dimensional table? In cross table, we were doing operation little differently.

So, like against alluvial, now we are having here these land uses; forest, grass and Lake. So, this part has gone as rows and geology has gone as columns. And then we can decide that for certain purposes which type of forest or which type of land use and which type of geology is suitable. For example, if we are looking some areas where we require alluvial deposits as a you know soft sediments and also, we are looking forest then if it is getting satisfied then we get a suitable condition.

But if we are having say different methodology like here, the shale and though it is forest but it is unsuitable. So likewise, we can create a table and we can call as a suitability table where it is satisfying our conditions and also the same thing, we can display in the map. So, in a cross table, it's a different scenario whereas in 2-dimensional table, though it is all coming under the overlaying operations.

Some few things might be similar but here then what the output which we are getting is not binary because here, the conditions are very different that we have fixed the conditions. So, only what it is giving the output; the suitable or not suitable. So, in that way, we can create 2-dimensional tables. And these tables and of course the output maps can be very-2 useful to find out the suitability.

(Refer Slide Time: 04:07)

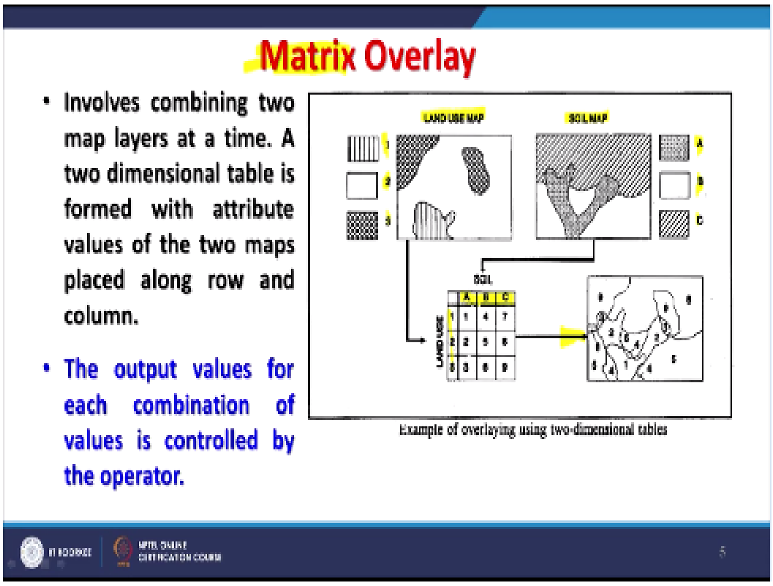


Now here, further if more variations exist between 2 input maps like this is input map A and this is input map B and what variations we are having from 1 to 4. Here just having 1 to 2.

The result map C would be something like this that it is just adding the things here and also you know, creating the suitability or a 2-dimensional table something like that.

So, if it is satisfying the conditions then it assigned the value likewise. So, like here, we are getting the value 6. So, category of the first map is the 3 and for the second map or B map, it is 1. So, when these 2 conditions are satisfying then you know weightage kind of thing has been assigned and likewise, we can assign different weightages here in our output map also. So, there are some applications of 2 dimensional tables as well through overlaying operations.

(Refer Slide Time: 05:22)

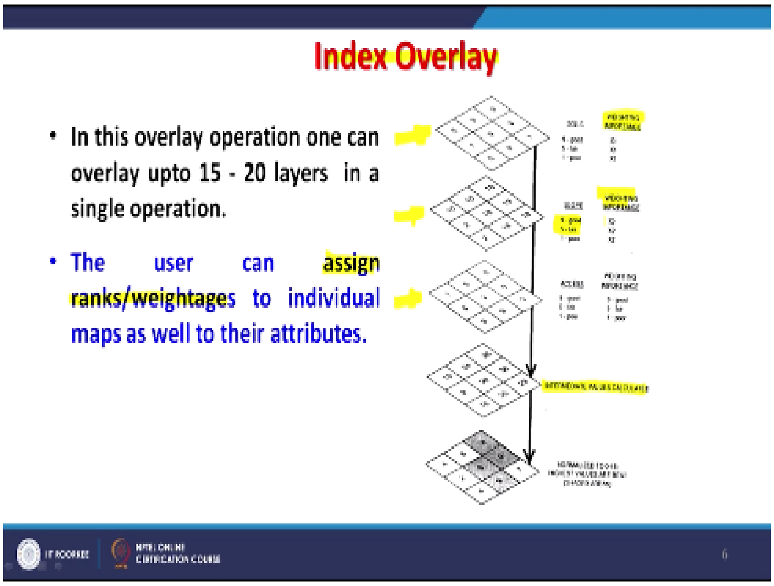


As you know that this raster data is nothing but a 2-dimensional matrix. So, the same way, it is handled a 2-dimensional matrix. So that means when we go for this matrix overlay, basically what we are talking about the raster data overlay which involves 2 maps or 2 raster layers at a time and 2-dimensional table is formed as we have seen some examples with attribute values of 2 maps placed along row and column.

So, one map attribute value becomes rows and another one becomes columns and likewise we can do it. The same way, the output values for each combination of values is controlled by the operator. Now, we can decide that how the output is going to be there. Here again, a land use map, soil map; 3 categories in the land use, soil map also having 3 categories. When these 2 are subjected to 2-dimensional table.

So, soil attributes are coming as columns and land use attributes are coming as rows. We develop a 2-dimensional table and same time, a map also. As you can see that you know number of polygons in the output map has definitely increased but these are not really vector boundaries but only in this schematic, it is shown as vector boundaries. Otherwise, what we are discussing is the matrix or raster one. So, one has to understand accordingly.

(Refer Slide Time: 07:03)



Now, there is another way of doing overlay is based on the index. So, these overlay operations can overlay up to 15 to 20 layers in a single operation. Because in 2-dimensional table; once we have said 2-dimensional table then only 2 maps can be overlaid. But here in index map, we can overlay as many as we can theoretically but practically, even overlaying 10 maps and the output which it would create, we have to really understand and utilise that one.

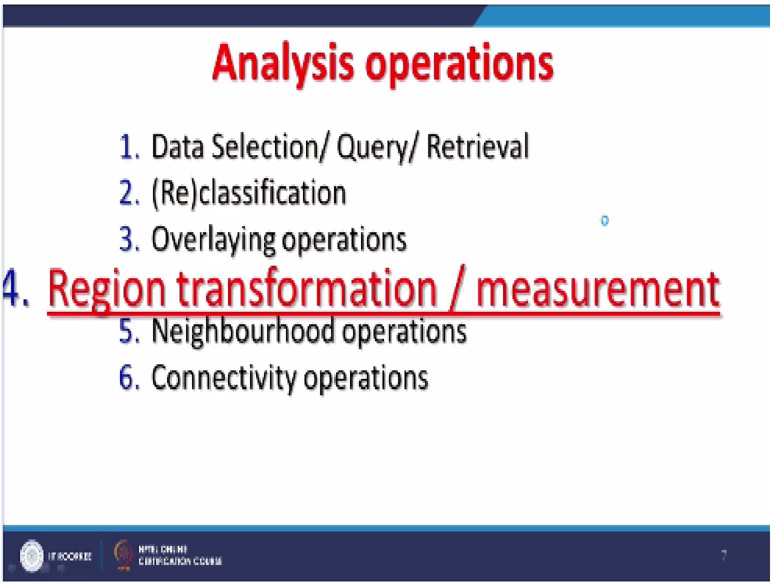
So, this user can assign the ranks or weightage, whatever we want to individual maps as well as for their attributes. And these index analyses, when we assign weightage to individual maps attributes and then the output maps can be used for various purposes, maybe for say groundwater exploration. Maybe for disasters related things like landslide hazard zonation or maybe for you know like erosional studies.

So, erosional studies; there also for different things, we can assign weightage like lithology will have some weightage. So, different attributes within the lithologic map will have different kind of weightage. Same with the solid type, same with the slope and likewise, when we combined these maps through this index overlay, we can create some predicted map which might so like soil erosion or maybe landslide hazard zonation and other thing.

As depicted here that for each layer; here the 5-layer scenario is presented. So, as we start from here and then we can have the intermediate calculation. So, input basically are 3 but then this 4th one becomes further as one layer for further analysis. So, we can assign different weightages here like weighting here for slope, for saying good, fair, poor and something like that.

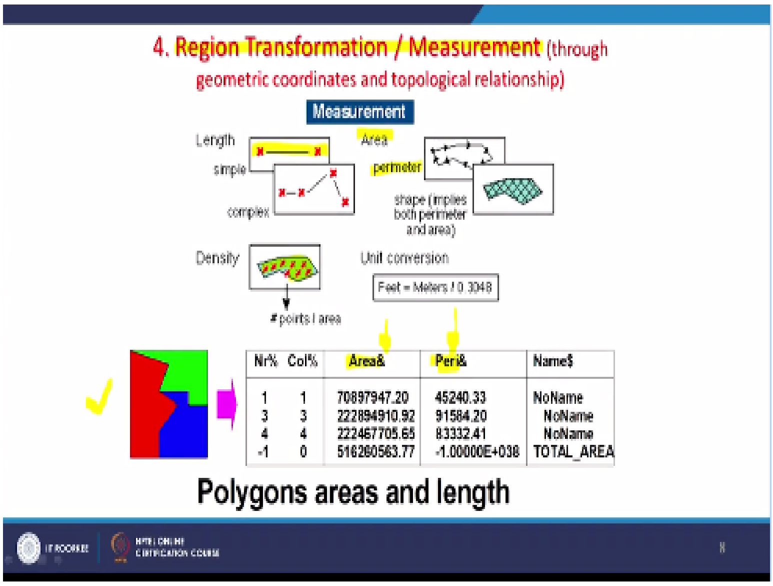
So, different attributes of different layers can be assigned weightages and then analysis can be performed through this concept of index overlay.

(Refer Slide Time: 09:40)



Now after these overlaying operations, we come to this Region transformation or measurement. Though we have discussed earlier in the primary operations or the simple measurement operations but here, the measurement operations are not simple.

(Refer Slide Time: 10:03)

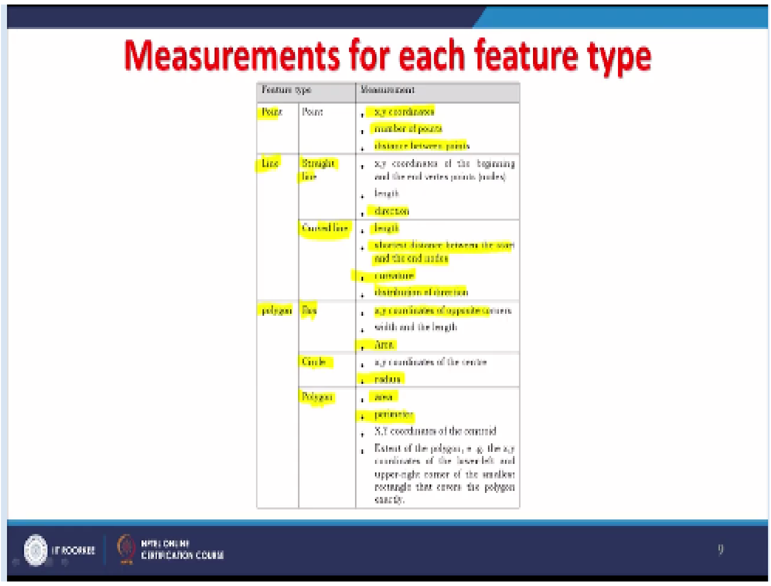


So, we will see how these things can be done that when in under this region transformation and this measurement like length can be measured, you know area can be measured then we go for the density measurement. How dense the points are located within an area? So that is why it is a little different. In the simple or primary operations, you measure either length or distance but here, we are going for even density measurements. Of course, area and perimeter are also measured at different stages in GIS operations.

Similarly, if one would like to have through this measurement, what one can do? Auto populate these areas. If we asked that I want to calculate area and perimeter or centroid for each polygon which is present in my map, it can be done quite easily on a GIS platform. See, this example is having just 3 polygons in a map but in real scenarios, a map may have 1000s of polygons and manually, it is not possible to calculate area or perimeter and that too accurately.

But if we involve GIS then definitely these measurements can be done very easily for even 1000s of polygons with in a theme.

(Refer Slide Time: 11:32)



Now as we know that measurement for different types of features vector entities, there will be different like for points, you would be measuring x, y coordinates number of points and distance between points also. So, basically in case of points, one can do this thing. Also, you can measure the distribution, assess the distribution, whether points are distributed systematically.

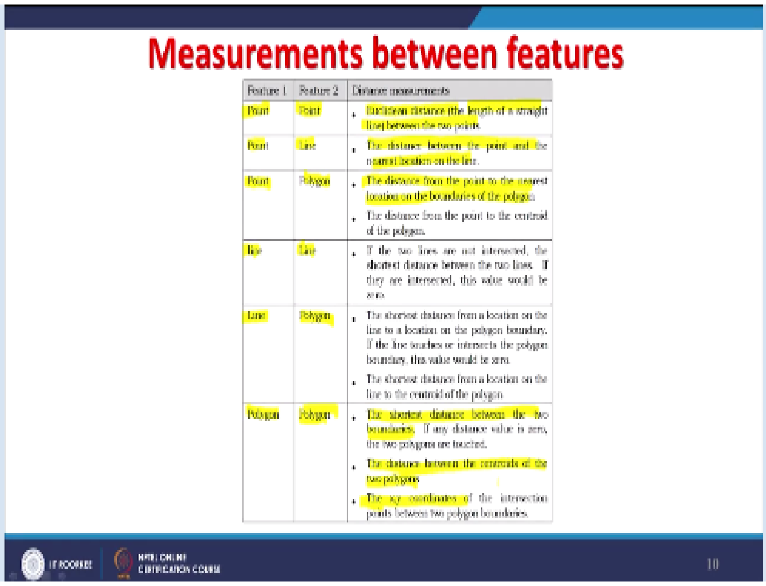
Points are distributed uniformly; points are clustered or randomly. Those kinds of measurements can also be done and comparisons with some other themes can also be done in case of point. Those things we can also put under this measurement category but if we are having line or polyline then we can measure the distance of a straight line very easily. Also, we can measure the length of a curved line without any problem plus also, we can measure the curvature.

Now, this curvature information for a polyline theme may be very-2 useful in some kind of projects. If I give the example from civil engineering like for curvature on the road or highways which is very-2 important for the safety of the vehicles. And similarly, curvature and slope, aspect especially, the curvature is also very important for rail track or rail route alignment. There, we cannot have a very sharp curvature or sharp bent for rail route.

So, all these things can be measured very easily. And if some you know criteria has been fixed that curvature has to be less than this, then automatically also GIS can be employed to find those areas where curvature is very sharp. So not only the length but its shortest distance between start and end points; end nodes, distribution of directions, all kinds of measurements can be done with polyline data. Then of course, you come the polygon.

So, polygon; obviously you can do the area, you can do the perimeter, you can do the radius, if it is circular or not, then also you can get the centroid. So, in case of circular then you get the radius. In case of polygon, you get the perimeter and area. In case of box means rectangular then you get the x,y coordinates and so on, so forth. So, likewise you do measurements depending on which type. The example which we have taken here is about the vector.

(Refer Slide Time: 14:17)

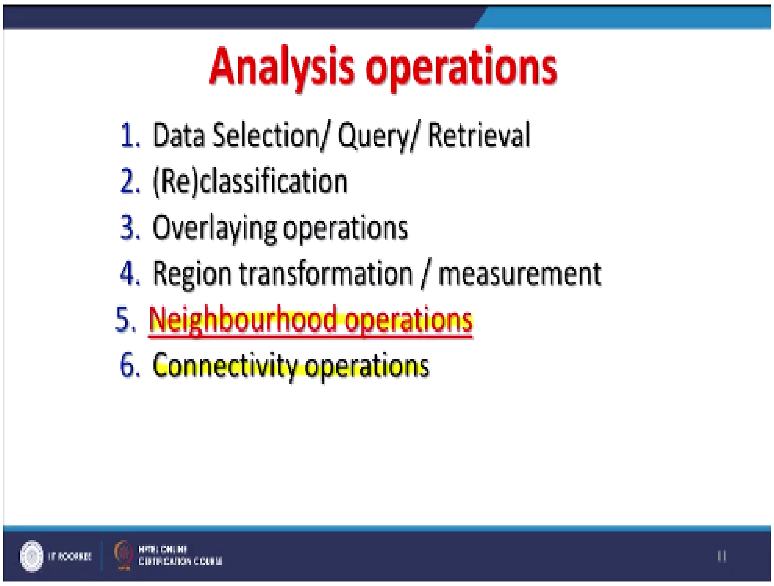


Now further in this list, again say between one point feature to another so 2. Now, we are talking between 2 different layers having point features. So, this Euclidean distance; the length is straight between lines of 2 points can be measured. And if we are having one theme point, another theme is line then the distance between this can also be measured. We may be having a scenario point and polygon then distance from point to the nearest location of the boundary of another layer can be done.

Then line versus line, line and polygon, polygon and polygon. So, all permutation, combinations which are possible, the measurement tools can be used. If we go polygon to polygon, again the shortest distance between 2 boundaries can be measured. The distance between you knows centroids of 2 polygons and x, y coordinates because every vector feature is always having either one pair or multiple pairs of x and y's.

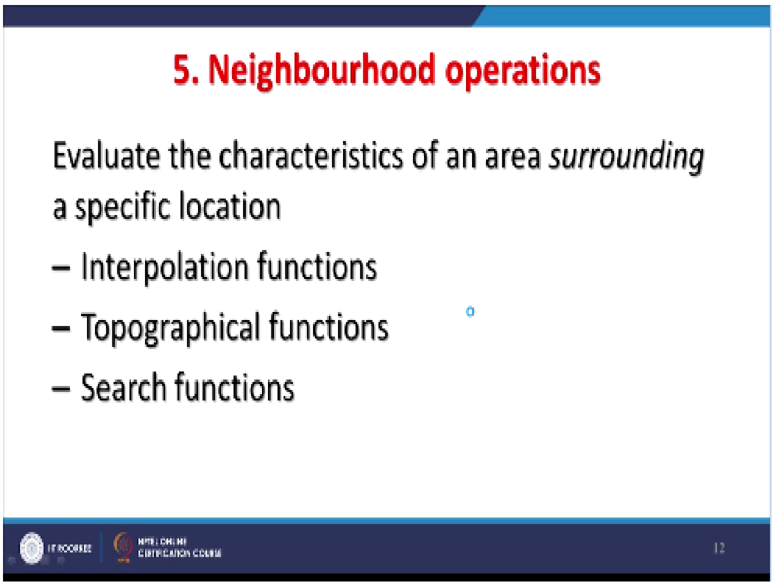
So, anytime you want to retrieve that information without any problem, that can be done quite easily. So, that basically completes this measurement part. And now we come to lengthier discussion about these next 2 topics; one is the neighbourhood operations and another one is the connectivity operations.

(Refer Slide Time: 15:45)



Connectivity's which involves the network analysis and for which we will have you know 2 lectures on this about the network analysis. Very-2 useful because as you know that network analysis can be done for various purposes. For example, it can be for road network, for rail network, even for fibre network or even for electric supply, you know power supply network. They are also electrical engineers have started using GIS for that purpose. So, we will be discussing little later but for time being, let us discuss about the neighbourhood operations.

(Refer Slide Time: 16:29)



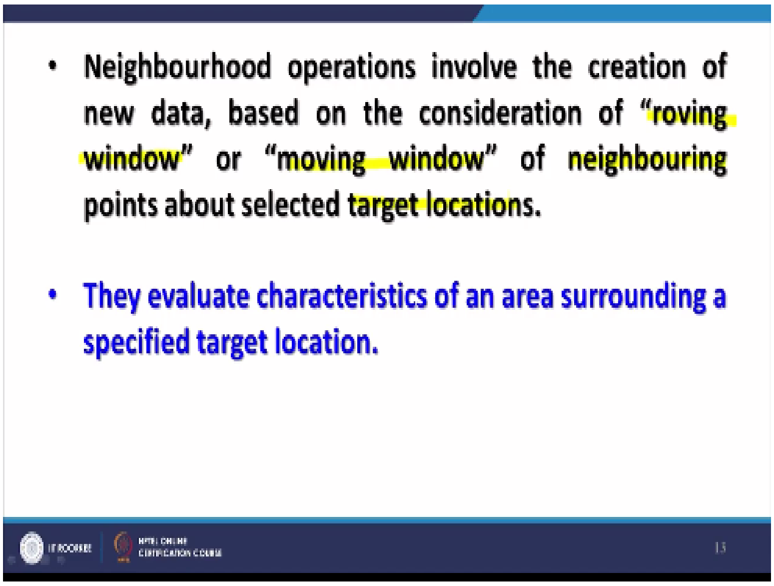
As the name implies that neighbourhood means looking in the surroundings around a specific location. So, what it does? Basically, evaluate the characteristics of an area surrounding a specific location. Location; we have to provide that this is my location. I want to evaluate the area or whatever the characteristics which are present in surrounding of this area.

For example, interpolation functions; they also fall under this neighbourhood operations. We have already discussed through 2 lectures about different interpolation techniques and their advantages and disadvantages also. And basic types like linear and non-linear interpolations and variants within these 2 and variants of individual interpolation techniques. Those things we have already discussed.

So, that part we will not but these comes under neighbourhood operations. So, somebody gets a new software or a literature and he is not getting interpolation topic separately. So, he should look interpolation menus or topic within this neighbourhood operations. Another type of neighbourhood operations which are topographic operations that involves lot of parameters or characteristics driving of a terrain for example, slope, aspect and watershed boundary and many other things.

So, we will have a discussion on that later about the topographic functions. And much more detailed discussion on surface hydrologic modelling and other things which again may fall broadly in this category of topographic functions. Then search functions that is also, we will be discussing that how as per user defined specifications, different layers can be searched within a GIS database to find out some suitable areas for certain purposes.

(Refer Slide Time: 18:37)



So, in neighbourhood operations as discussed briefly that it involves the creation of new data and that is based on the consideration of roving window or moving window. And this concept has come from digital image processing and that is called spatial filtering; SPATIAL, spatial filtering where we are having a window maybe 3*3 or a small matrix of 3*3 or maybe 5*5. It has to be odd number and for the centre cell, the value is calculated using whatever the value is one provided for the surrounding or in neighbourhood.

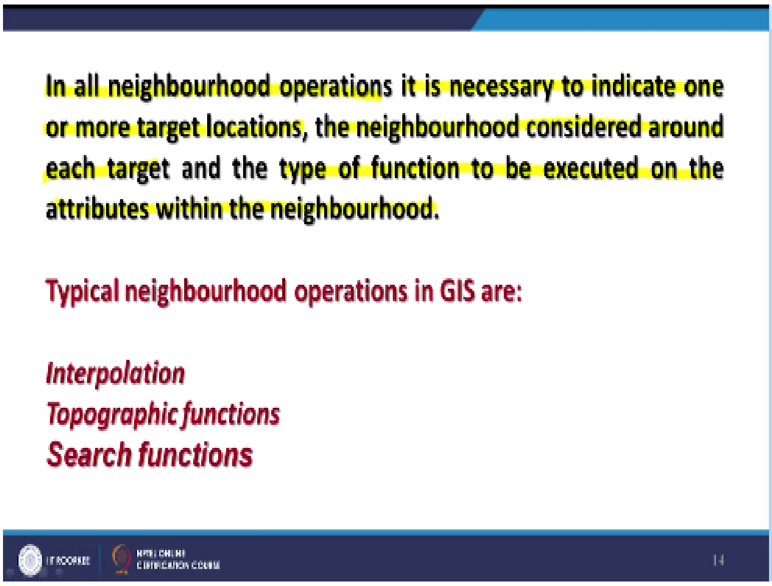
So likewise, we can have a roving window or moving window which goes through the entire raster data set and you know search the characteristics based on whatever the weightage, you have assigned to individual cells of your roving window and that is basically of course, looking the neighbourhood. And then for the selected target location; that is in case of roving or moving window, it is the centre cell.

Now here, what these windows are doing? They are evaluating characteristics of an area surrounding a specific target location. Target location as I have just said is the centre cell. Of course, these are the attributes and in case of raster, there will be only single attribute. So, that can be searched. Suppose I am having a digital elevation model and now, I want to calculate say slope map.

So, what it will do? It will search in the surrounding and whichever the cell is having less elevation value compared to the centre cell then you get the information that okay! the slope is in this direction. And if it continues then after that roving window or moving window, we can find out that what is the length of the slope and what is you know the degree of slope; how much the sloping surface it is?

And of course, the orientation that means with reference to north; that is the aspect can also be calculated through this neighbourhood operations. More specifically overall, they are neighbourhood operation but we put them as a topographic function.

(Refer Slide Time: 21:16)

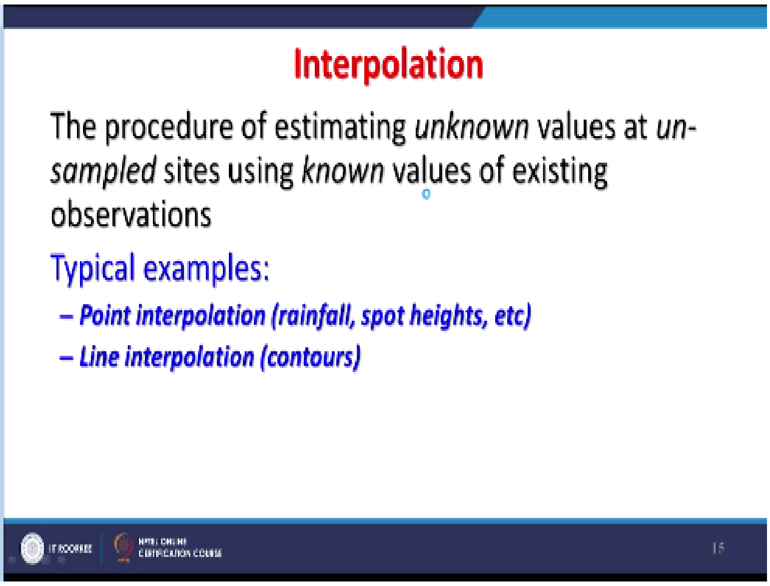


So, in all neighbourhood operations, it is necessary to indicate one or more target locations. In roving or moving windows, the target location always is single; the centre cell. And the neighbourhood considered around each target. Now, how much it will consider? It will depend on the size of this roving or moving window. As I have mentioned that it has to be the odd number in terms of rows and column.

So, it can be minimum of 3*3 and large, you can go for higher and higher but like 11*11 or 9*9 roving window. But larger than this roving window, more the time it would take and more smoothen product, it might create. So, sometimes that in case of terrain and slope, it may be misleading. So, it is always better either choose 3*3 roving window or maximum 5*5, that would be sufficient for most of the cases.

And the type of functions to be executed on attributes within the neighbourhood. Type of functions; how the weightage has been assigned in the neighbourhood that means in the surrounding of the centre cell of this roving window and that will decide that how this will be operated. Now, the typical neighbourhood operations in GIS are for example, interpolation I have already mentioned and topographic functions and search functions.

(Refer Slide Time: 22:57)

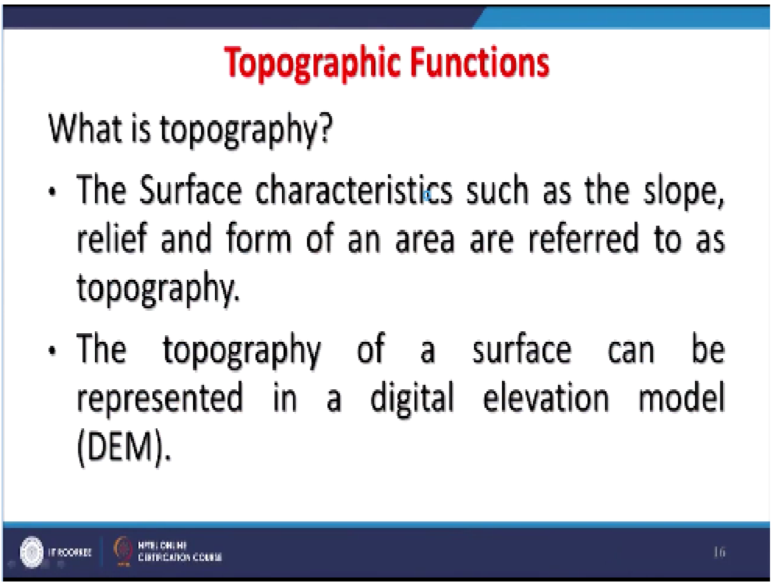


Interpolations; we have already discussed but very briefly, I will just touch for completeness related with this neighbourhood operation that this procedure of interpolation basically estimates values at unknown location or unknown sample sites using whatever known information is available or existing observations information is available. And interpolations, we use for various purposes, maybe for rainfall, maybe for you know terrain.

So, we can do the point interpolation. We can also do line interpolation that means using contours but the system will first convert your contours to the point and then from point, it will do the surface generation or interpolation. Now, second is among these functions of neighbourhood after interpolation, is topographic function.

So, before that, I would like to briefly discuss what is exactly topography. Those who are may not be from earth science or civil engineering domain. For them, it is necessary to very briefly understand what is basically topographic is?

(Refer Slide Time: 24:08)



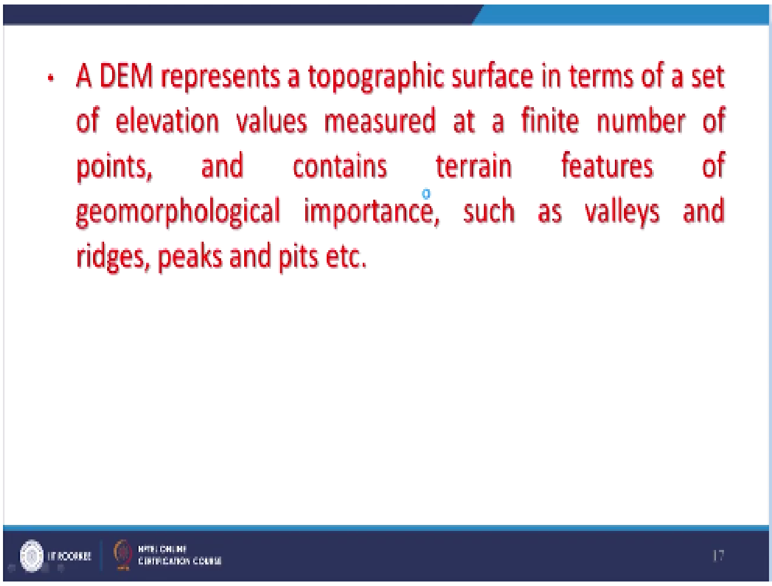
So, this is about the surface characteristics of a terrain. For example, what are these characters, can be slope, relief or form of an area are referred as topography. Sometimes we say, this is hilly terrain. So, on what basis we are saying hilly terrain because they are undulations. There are peaks, there are ridges, there are valleys so then we say, it is a hilly terrain.

If we go like in Himalayan terrain which is highly rugged so with in you know this horizontal distance, you may get a very high peak and very low valley. That means there might be even 1-kilometre or 2-kilometre difference between the peak and the valley in a very small horizontal distance. So that way, we cannot say it's a gentle slope. Then we say the terrain is having very high slopes.

So, these are basically surface characteristics of a terrain, not inside of that terrain but only about the surface. And most popular one is also the slope, aspect, relief form and other things. Now, as you know that the topography of a surface, that can be very nicely represented through digital elevation models. And we will have extensive discussion through several lectures on digital elevation model.

What exactly digital elevation model? How they are generated and how we can utilise or drive various characteristics of topography of a terrain using these digital elevation models. And we will be also discussing that from where I can get these digital elevation models? If I do not want to create by myself then how I can get and how to assess the quality of a digital elevation model? So that, we will be doing much later in this course.

(Refer Slide Time: 26:18)

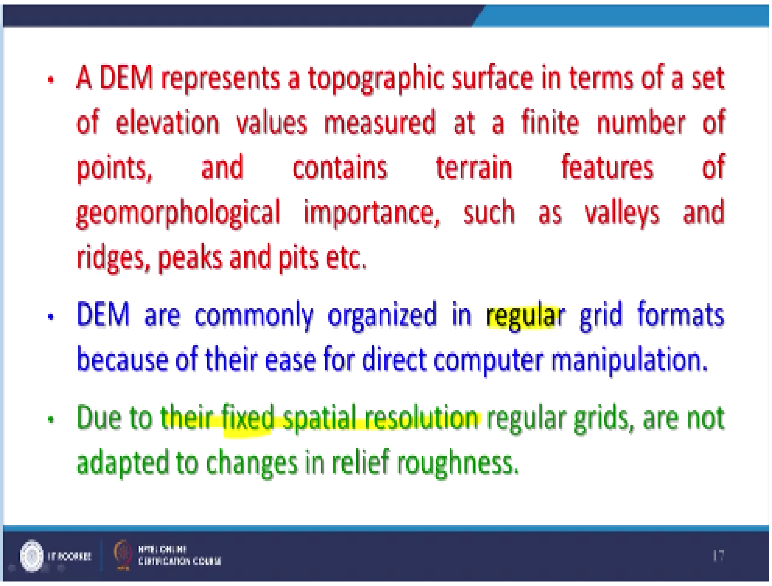


But for timing being because while discussing slope and aspect, we might be using this term DEM several times. So, let me very briefly also introduce the DEM which basically represents a topographic surface or terrain in terms of a set of elevation values which are measured at a finite number of points. Not at every point but finite number of points and then through interpolation techniques, a surface is created.

Now, this DEM also contains the terrain features of geomorphological importance which is very useful in geology, Earth sciences and in civil engineering or also in environmental engineering such as valleys, ridges, peaks, pits etcetera. These are all characteristics of topographic surface; these valleys ridges and other things. Now, DEM or digital elevation models are commonly organised in regular grid formats.

That means we are talking about the raster format. Not vector but of course, this topographic surface representation can also be done using TIN; triangulated irregular network.

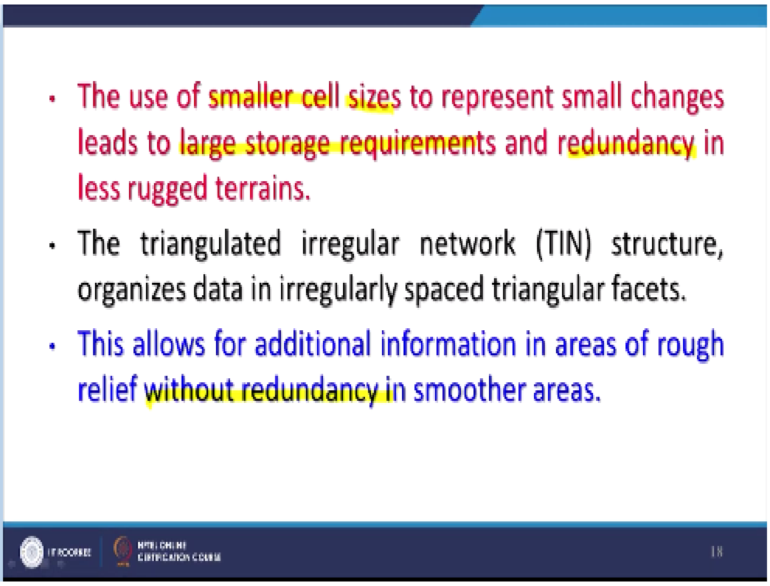
(Refer Slide Time: 27:41)



Though the model is completely different than what raster represents. Raster is having a regular grid whereas TIN is irregular by its name, it is the word irregular is there. And as you know that in case of DEM, spatial resolution is fixed within a regular grid. Though, 2 different digital elevation models may have different spatial resolution.

But within a one grid file or within one raster file, all cells will represent the same spatial resolution, that is why it is the fixed spatial resolution. And because of this fixed spatial resolution, these are not adaptable to changes in relief roughness as in case of TIN. So, these we have already discussed in details about raster and TIN advantages and disadvantages.

(Refer Slide Time: 28:44)

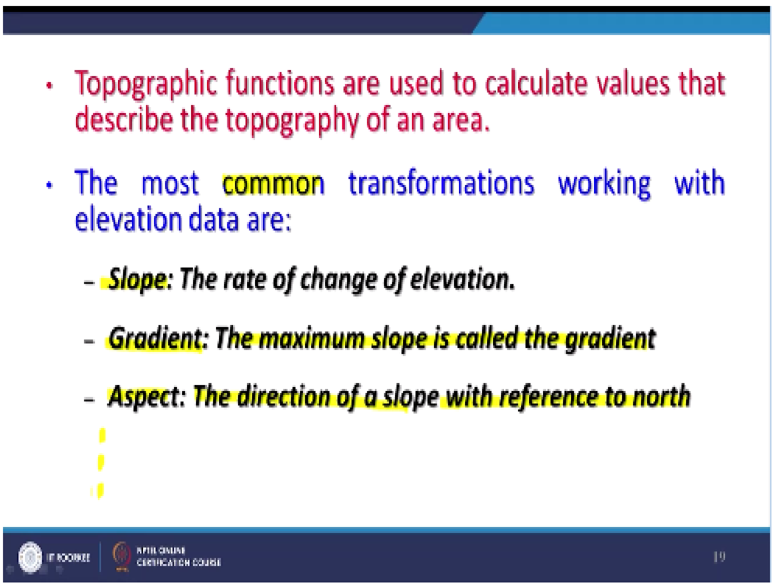


Now if we go for higher spatial resolution, that means we are going to represent a small ground area through smaller cell sizes of through this digital elevation model. If we go for that, then there might be large storage requirements. And also, we will be introducing redundancy especially in case of less rugged terrain. For example, if I am developing a very high spatial resolution digital elevation model for a part of Indo-Gangetic plain or near flat area then I am introducing lot of redundancy.

But this will not be true in case of small cell size or high spatial resolution digital elevation model for Himalayan terrain because their elevation values are changing very frequently and within very less physical distance, values might change. So, redundancy will not be there. As I have already mentioned that triangulated irregular network structures organised data in irregularly spaced triangle facets.

And this allows the additional information and the areas of roughness because it is adaptive to relief roughness and without any redundancy in the smoother area. In a smoother area, we will have larger triangles; in a highly rugged terrain, we will have smaller time. So, this TIN model is adaptable to the relief roughness.

(Refer Slide Time: 30:17)

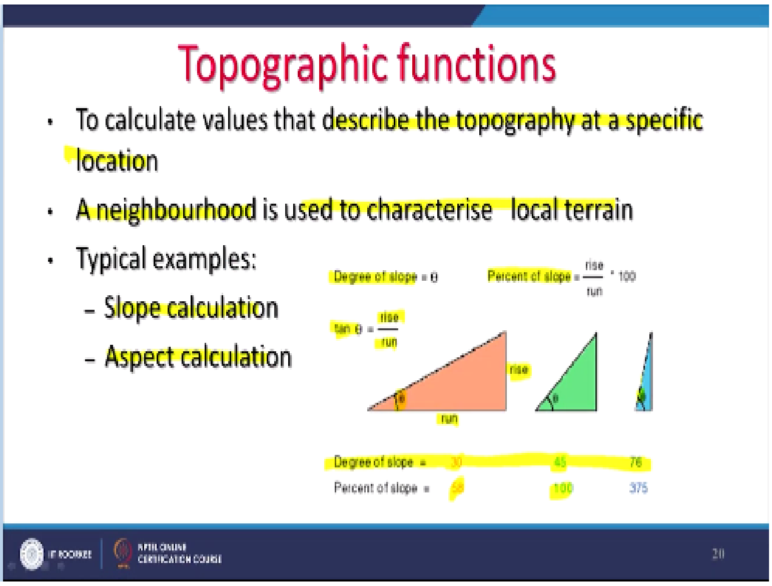


Now, what are the other derivatives which we can drive through these topographic functions of a terrain or using a digital elevation model. So, the most common transformation or derivatives which we can do using digital elevation model is of course slope which is the rate of change of elevation. And just simple trigonometry, we will discuss just after 1 or 2 slides. Then gradient; that the maximum slope is called the gradient.

And then aspect which is the orientation of sloping surface with reference to the north; that is the direction of a slope with reference to the north. These are the common characteristics or common derivatives of a digital elevation model through these topographic function's operations. But if I say there theoretically n number of derivatives or transformations are possible.

You should not be surprised because lot of new derivatives are coming. Even for Earth scientists people or for civil engineer, we can even use digital elevation models to find out the landforms also.

(Refer Slide Time: 31:37)



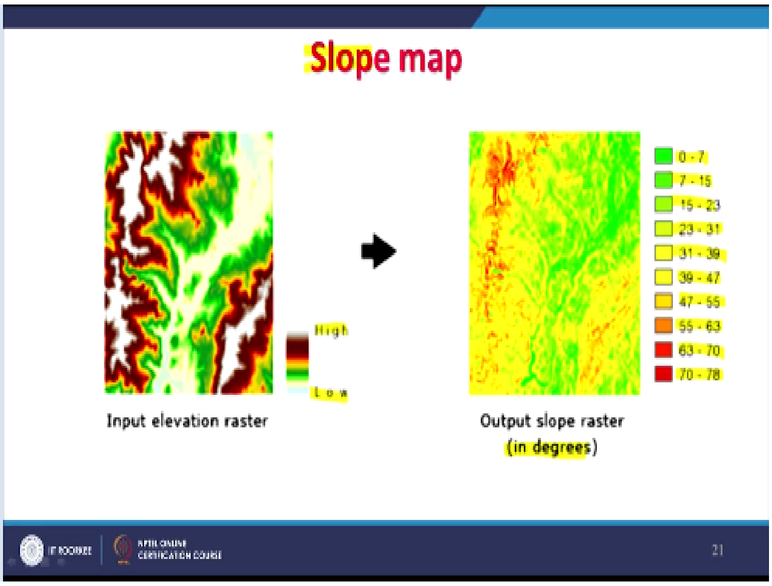
So, in that way, it can be done. Now, topographic functions basically calculate the value describing the topography at a specific location. Recall the roving window and moving window. That will be used to drive all these derivatives of a digital elevation model or characteristic like slope, aspects etcetera. So, that operation will be done in the neighbourhood to calculate the local terrain using for single. Slope calculation, aspect; these are the typical examples.

Now in all GIS, they will support or you can do the calculations in 2 ways. Either you can calculate in terms of percentage; that means we are keeping values between 0 to 100 or we can calculate in terms of degree; that means we are keeping values between 0 to 90. So, it's a nothing but rescaling, if we go 100. Now, it's a simple trigonometry that tanθ; this one is based on the rise and run. So, this is our rise; this is our run.

So, when we go for this, we can calculate. So, if θ is large like here then we are getting 45-degree here. Instead of θ is less here so we are getting 30-degree. And in terms of percentage of course, we can also calculate here. So maybe for 45, we are keeping 100% but that is not for some cases, we need not to do that. Generally, what we do? We always go for degree rather than percentage.

Percentage maybe sometimes misleading for many other users of the product. So, it is always better to go for degree because these degrees, we can understand even in the field very easily that some sloping surface is having 45-degree, 30-degree or maybe 75-degree or 90-degree. The vertical cliffs are also possible in terrain like Himalaya, almost vertical. So, their slopes are 90- degree or you can say 100%. So, the best way to use in degrees rather than percentage but both options are possible.

(Refer Slide Time: 33:57)

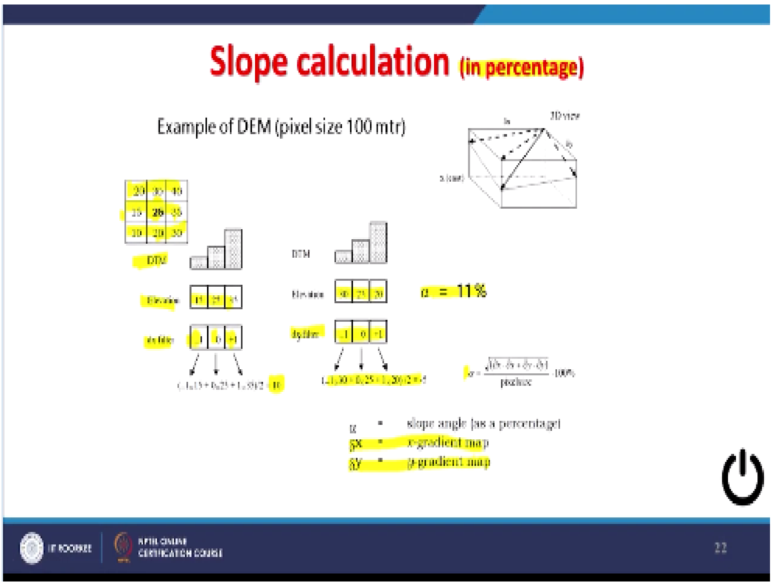


Example here, this is the input digital elevation model. As you can see, the values are here from low to high. And when we go for slope map calculation using neighbourhood operation functions, especially the topographic functions then we get a slope and here the slopes are shown of course in degrees. So, you get where you are having and this classification, again there are techniques how to classify these things also.

Whether you want equal interval or some other percentile or based on some other techniques; statistical techniques, standard division, all kinds of classification or reclassification can also be done here. But for timing being, if we take just this one so the slope map has been classified in these number of about 10 classes. The slope varying between 0 to 78-degree.

Zero or 7 is almost flat which is shown as green colour and the red one are shown between 72 to 78-degree.

(Refer Slide Time: 35:07)



Similarly, as you know that in slope calculation in terms of percentage, it is just rescaling the thing. So here, suppose this is my input roving window. Now for the centre cell, I am calculating. This is DEM or digital terrain model. Elevations are 15, 25 and 35. And then this filter, the weight in the surrounding in the neighbourhood in this particular example is given -1 and + 1 and of course, no calculation for the centre pixel. So, it is zero.

So, when this calculation is done, the new value comes is 10. So same way in case of other calculations, we can do but this dx and dy, dxy; all these calculations are done and then we get the value as 11%. So likewise, one can do the calculation also in the percentage. Slope angle in percentage or in degree and also the gradient map. So, you know gradient maps can also be calculated.

So, this brings to the end of this discussion that is part 3 of GIS analysis. We will continue on GIS analysis part 4 and others. So, for time being, thank you very much.