# Introduction to Geographic Information Systems Dr. Arun K Saraf Department of Earth Sciences Indian Institute of Technology, Roorkee

# Lecture - 16 GIS Analysis – 2

Hello everyone, we are going to now discuss in this particular lecture about Analysis and part 2. In the previous lecture we have discussed the basic selection retrieval functions and later on the classification especially the reclassification techniques of both vector and raster data as per our requirement based on their attributes. In this particular discussion we will have a more advanced function of a GIS analysis. As you recall that we have completed one and two in the GIS analysis in one lecture. Now, we are a we will be looking more on the overlaying operations.

(Refer Slide Time: 01:12)



Initially when a GIS started developing as you can recall by Roger Tomlinson. And later on when then commercial software started coming in the market, because there were very strong on vector data and people developed some perception they do not perception that it can only perform overlaying operations. So, overlaying operations on vector data as well as on raster data it is possible. But, when we are discussing about the comparison of vector data and raster data at that time we have seen that overlaying over using the vector data to some extend is alright, but if you known many maps vector maps and then overlay then the output will have so many polygons that then interpretation or usage of such maps become very difficult.

Whereas in case of raster data there is no limit and an overlay operations can very easily be performed on raster maps and lots of modeling and analysis functions can be done over raster data which we are focusing this particular one on overlay operations. In later part we will also discuss regional transformation and measurement in this lecture and also neighborhood operations we will be discussing in this lecture. So, let us start with overlaying operation.

As you know that we have gone through the mathematics which is based on the set theory and in our ten plus two classes and which is based on the Boolean logic or Boolean operators and this is also called Venn diagram, where we know that a intersection b how in a simplified form.



(Refer Slide Time: 03:15)

This is how it would look A union B or A exclusion B, A negation B this is how different scenarios will come. These are the logical operators which are used through softwares and you get different results. But this examples are very simple example where say two maps are having one polygon each, but in real operation say if I take given example of two maps, two maps might be having polygons say one map might be having 20 polygon another map might be having 30 polygon. So, this operation becomes complicated.

However, the fundamentally it is just based on the same set theory or a Boolean operators.

So, the mathematical which we have learnt earlier is now going to be used. However, when we were going through all this set theory or Venn diagram Boolean logics at that time we will never thought that this is how the mathematics which we are learning that that time we will use later in GIS. This is very power full tool within GIS, this is one of that become a function which we put them in under the function of logical functions.

(Refer Slide Time: 04:36)



So, these overlaying operations as I already mentioned can be performed on both types of vector data as well as raster data. Overlaying operations are difficult to perform on 10 data.

## (Refer Slide Time: 04:49)



Now here a very simplest example has been taken here that if map are having just two units another map is having two units then when we overlay or intersect this or making a union of these maps then four units a turn up. So, you know it is basically multiplying here. And if this would have been many, many polygons in one map then the output map will have might be few 100 polygons. In case of vector this is the difficulty, but in case of raster this is not a difficulty.

However, even if I have to handle in vector and the hundreds of polygons comes and the my map become complicated as in previous lecture we had seen that we can resort before we present these maps to say decision makers we can resort to the reclassification and based on certain attributes we can reclassify such output maps and then I can reduce the number of polygons and can make those maps very very useful. No problem even if you have to handle with vector data and there are other analytical tools in combination you can apply and achieve a good result.

# (Refer Slide Time: 06:10)



Similarly, a vector overlay here is a basically a table get a or attribute table get appended and this how we address at that this map A is having 2 units, map B is having 4 units. We end up with more number of polygons here and then your table becomes further enrich. And this is how you start getting thing there. So, you can then later on classify or reclassify your map you can reduce the number of a polygon and still you can use those maps and make a sensual interpretation of those maps. There are some off suits of the basic four a Boolean logical function.

impose

impose</td

(Refer Slide Time: 07:01)

So, you can have even other functions like, impose, stamp, join, compare.

(Refer Slide Time: 07:09)



There are other functions which a very common function which also used is a clip function or extract function. As shown that there is input coverage which is having this kind number of polygons and there is clip coverage, the coverage which we are going to use to clip the map 1 with the input maps. So, when you overlay on this then you get this intersects map is like this. These basic functions are being used, but their new name had been given so that they become much easily understand.

The similar clip or extract function can also be performed on your raster data. Suppose, you are having digital elevation model and shape of your overall shape of your digital elevation model of an area is rectangle. Now you want to extract that digital elevation model only for a shape political boundary say based on certain district boundary. So, you want to show only the elevation model of that district. What you need to have as input, you need to have a clip instead of a circles here you can have the polygon of that district.

Then once you intersect or use this clip or extract term a function then you can extract only for this area. But remember that the other out of that a circle you digital elevation model is still will have data, but it will be has a no data. So, only the elevation model below that polygon boundary would be visible nothing else will be. Similarly, instead of polygons you can also perform on line data. Again in clip coverage is same as in the previous example this was the polygon data, here is the line or polyline data and you can extract even line data. Similarly, you can also extract in the from point data.

So, I earlier gave example in GIS analysis 1; that if you are having blazes of India there will be thousands of blazes. And you want to extract or make a subset of those blazes, so you need to have a clip boundary and that has to be the polygon. If you use the polygon and then through this clip extract or intersect functions you can extract only blazes which are falling below that polygon. Maybe of a taste or maybe a district or block whatever is the requirement accordingly this function can view. So, this is based on simply on Boolean logic.

(Refer Slide Time: 09:53)



Also you can have other functions like erase function. So, again input map is same this erase coverage is same. Now, below this everything is gone and only you are having now and the remaining part. Same with the line data and same with the point data, so should depending on your requirements you will use the appropriate tool and perform the analysis.

# (Refer Slide Time: 10:19)



Union also the same thing that you want a merge everything whichever is available in the two maps or more than two maps then you can use the union function. Input union this one so output will have everything. Another example here that this input is this a map here and the union cover is this so you end up with this. This can be use has a merging of polygon as well, merging of two maps. It may be on polygon maps it may be of your polyline map or even point data.



(Refer Slide Time: 11:03)

Again union functions for polygon are shown here, and that everything in whatever was present in these two maps have been unioned here. The operator which you will use or so whatever present or input means map A or map B then you end up with this kind of (Refer Time: 11:21) and your attribute table will become further rich. And it will have whatever is available was in map 1 or map 2 that will be included in your attribute table as well.

(Refer Slide Time: 11:40)



And then intersect which is end function, end operation of a Boolean logic when two coverages are overlayed. Only portion of the input coverage that falls inside the intersect coverage will remain in the. This example we also seen in detail and the example are here also.

### (Refer Slide Time: 11:58)



Identity functions also be used which is everything located when the boundaries of input coverage is collected into the output coverage. The boundary of the output coverage is identical to the input coverage. Then identity procedures applies to point line and polygon coverages also you can apply on the raster data. But, remember that on TIN you cannot perform these operations.

(Refer Slide Time: 12:22)



Now raster overlay examples are here and you can perform on raster layers all arithmetic operations. Four basic functions are given here; plus minus multiply and division. Then

relation functions greater than equal to less than functions. And then logical operators based on Boolean logics and or x or not, all these functions can also be used. Also in addition to this conditional functions if then else another important thing that all these four basics kinds of functions can we use in combination as well. So, you can make a complicated syntax or complicated query to the system for overlay.

But one thing one has to remember that if you make a very complicated query or such overlaying operation during overlaying operation the output may which will create generate you have to understand or someone has to understand that one. So, it is generally in the good practices and do not bring too complicated syntax or this overlaying operation at the very first go, you can do it in this stages wise. So, that a few have involved two maps you use certain (Refer Time: 13:37) arithmetic operations.

As said also after this first step instead of going in one go you can do it in four steps and each step you can check for errors. Therefore, you would more comfortable and more confident in your output once you achieve and the results. So, a step by step is much better approach then going in just one go, because then things may become a complicated to interpret such maps.

(Refer Slide Time: 14:10)



And arithmetic operations examples are here. If there is a map 1 map 2 they are having all different attributes here. If I say that map C should be equal to map A plus 10. So, simple arithmetic operation and I want to add 10 values to each of these cells so this has

been done here very easily. Similarly I want to add map 1 and map B. So, I say map C should equal to map A plus map B, so this is you can check it like the top left value was a map A was 5 this is map B 4. This say this should be done because this is a very simple example where you are having this 4 by 4 cells in real one you might having thousands of rows and columns raster map.

And then values are not visible as such they will be visible through a color or a range of colors or in a continuous section. So, once you do a simple operation first check the value because if I have instructed that add 10 to everything then check at least one or two points whether things have been done accordingly or not. Similarly, if I have added two maps A and B then I should check whether it is giving correct result are not. And like I can also perform a little complicated one that map C2 should be equal to in map A minus map B divided by map A and B multiply by 100. This is kind of creating a indexes or an index like this.

Now you have put a little complicated syntax, so at least for few randomly selected locations you must check whether the results are coming correctly. Otherwise, determine if I use this map as input insert some other analysis and it is carrying wrong result then I will end up with a really confusing outputs. So, that the best approach is go one by one step by step.



(Refer Slide Time: 16:21)

Now an example of relational function here is the example is given for a greater than and this map A and map B are the same example in case of previous example. Now instruction is output map should have map A if map A cell value is greater than B then if it true then (Refer Time: 16:37) 1 otherwise 0, so let us interpret this map. See map A here the 5 value is definitely greater than 4 and therefore it is true and it is getting value 0. Let us see the example of a bottom right corner that is the value 6, 6 is definitely less than 8 and that is why it is false so it is assign 0. Now the output is in binary form.

So, once output comes you should be ready because condition which you have put forward here then according the map will be generated. And this is a simple binary map to understand. Likewise, such relational functions in combinations can also be used with other functions.

Boolean Logic : Truth Table							
it and ou	tput (trut	h table) for	Boolean o	operators.	1 = true, 0		
ATTRIBUTES/INPUTS		OPERATIONS/OUTPUTS					
A	B	A AND B	NOTA	A OR B	A XOR B		
1	1	1	0	1	0		
1	0	0	0	1	1		
0	1	0	1	1	1		
0	0	0	1	0	0		

(Refer Slide Time: 17:36)

Now Boolean logic or we also sometimes use a truth table, because which is false which is true that can be done. So, here the examples are shown that attributes input and output here and operations are here, so if A and B are there then A and B and intersection union all this are there and then you can create a table to understand, that if it is true it will get value 1 if it is false it will get value 0. And accordingly you can create a table for your understanding like map which will output map which will going to be in binary form.

## (Refer Slide Time: 18:19)



Another example the real one example is that here of end statement is that that here the land use equal to forest the condition is that a in this land use map search the land use which is showing the forest or which is having attribute forest. In slope map search the slope which are having steep slopes and when these two conditions will satisfy only for those areas the selection has been done, and rest areas are getting one single color. Likewise, you can put conditions.

So, these are telling that here the land uses forest and slopes are very steeps here and in the remaining area these two conditions are not getting satisfied and therefore nothing is there. Why we are getting wide because here there is a blank data that is why we are getting wide otherwise there should have been only two colors. This is the basically where the condition are have not been applied and remaining areas the condition have been applied.

## (Refer Slide Time: 19:35)



Now, relation and logical operations together you can put together and like here that relational and logical operation here is the map A and map B and different categories of say forest and their values are given here. So, we declare that map D should be equal to map A when the forest is the category in map A and in map B it is less than 500. Now the two operations; one is the relational and another one is the logical, both are there and you know in combination so then your output map would be like this. Again all will come in binary form, so 0 for false 1 for true.

Likewise you can create as complicated as you can think. However, as I was mentioned earlier that if you create a complicated overlaying operation the output map you have to understand first. You have to satisfy that whatever I have put here whether it is coming correct or not. Secondly, in real operation one has to also remember the bracket, where the brackets is getting opened and closed. Because, if you do not take care then some map will get their calculation elsewhere and then output map would become very difficult to use or interpret, so these cares has to be taken. Therefore, it is better to go step by step.

#### (Refer Slide Time: 21:09)



Now, conditional functions are here that map C here should be equal to if map A equal to forest true or say undefined has well, because here we do not have any information. So, that can also be used as undefined. Likewise in map C we are getting areas which are undefined. This thing can also be used as in earlier examples the white areas were coming in a map, so these are the undefined areas. Rest is satisfying when because of this map is all is having forest it is satisfying and therefore the first condition and therefore it is true and everything is getting 1. Remaining areas are undefined therefore they are question mark values has been assigned.

Similarly, here that map C if map A is forest and map B equal to 700 101 for true, 0 for false so it is coming accordingly. So, on 1 map 2 map multiple maps you can perform similarly.

#### (Refer Slide Time: 22:15)



And there are also cross tables we call or cross operations are possible. Like here that we can involve that in cross table we are multiplying forest multiply by alluvial or else multiply by shale and so on so forth. Therefore, the cross map will have this attributes as well. Basically this operations will, of course you will get a map because there is a dynamic linkage, hot linkage which been your special objects as well has attribute data. But, after these operations the major changes will occur mainly in the attribute data or tables, that is why this are also called cross tables. So, depending on your requirements you will go for these operations, all these four major types of operations which are available during your overlay analysis.

#### (Refer Slide Time: 23:06)



Now you can also go for two dimensional tables like here the categories are alluvial and shale, here the categories are forest grass and lake. This kind of matrix can also be developed that this is land use map, so land use map is having forest grass and lake. And geology map is having just two types of geology alluvial and shale. So, when you can create output map which are say if this condition is like if it is alluvial and forest then it is suitable, if it is alluvial and grass not suitable, if it is not suitable.

Similarly, for different type of geology shale and forest not suitable but shale and grass it is suitable. So, these in this two dimensional table to at two condition at two places suitable things are there and then you assign that one when you go for this. So, these two dimensional tables an initial over this matrix can be developed and this matrix will be used in your analysis, in your overlay operations and the input would be land use and geology. So, likewise lot of different types of analysis using those four basics overlaying operations can be performed.

### (Refer Slide Time: 24:36)



Matrix overlay that I have already given examples involving two map layers at a time. That two dimensional table is formed with attributes values of the two maps placed along row and columns and then output values of each combination of values is controlled by the operator. And so that another example is given here that there is a land use map, there is a soil map, there is a two dimensional table and here the attributes are A B C another attributes for land user 1 2 3 and you end up with this similar kind of a map.

It becomes much easier because here you can develop this two dimensional table and think before you submit for overlaying operations.

### (Refer Slide Time: 25:22)



Index overlay is also possible like here in this overlay operations one can overlay up to 15 20 layers in a single operation. The user can assign ranks or weightage to individual maps as well to their attributes. Suppose, you are working for the ground water exploration you know that lithology is very important, so can assign more way to lithology. Whereas, you may say the soil is not that important, so you can assign for a soil layer less important.

Similarly for maybe slope maps, maybe for linear maps or geological structures depending on their influence in particular area you assign the weight in this way you can create also output. So, you know the different weight for different categories of different types of map and for different categories within each maps accordingly you can assign and then final map can be created which will give you the better understanding say about the ground water condition of that area. So, all these are also possible.

### (Refer Slide Time: 26:47)

(through g	eometric co	ordinates and	topological re	lationship)
	N	leasurement		
Length	×	Area	the 3	
simp	x-x/	shape	(implies	
00	mplex L	both p	èrimeter Larea)	
Density	TELES	Unit conver	sion	
	# pointe /	Feet = Me	eters / 0.3048	
	Nr% Col%	Area&	Peri&	Name\$
	1 1	70897947 20	45240 33	NoName
	3 3	222894910.92	91584.20	NoName
	1 1	516260563 77	-1 00000E+038	TOTAL AREA

And now we come to the fourth analysis operations which are regional transformation or measurement. In this particular operation what basically is a measurements are done and then we add a in our attribute table or do some other operations. So, like measurement simple line and if it is polygon then we can measure the perimeter area and if there are points then different type of analysis the distances between points or pattern distribution pattern of points that kind of analysis plus random selections, random generation of points can also be done in this types of operations. Similarly like here, that we are having different categories so we can do the calculation of area, we can do the calculation of perimeter we can assign different attributes there.

Now, the last in this particular lecture about analysis operation is neighborhood operations. And these neighborhood operations will look the values in the surrounding as per the given condition by the operators or the GIS users.

#### (Refer Slide Time: 27:54)



Basically, these evaluate the characteristics of an area surrounding a specific location and the size search area and other thing is depends on the user requirements the project requirements. In this one; one is interpolation functions, so interpolation techniques we have already discussed. In this particular lecture I am not going to in detail about the interpolation. The topographical functions which we will discuss, and of course the search function we have been anyway discussing all these.

(Refer Slide Time: 28:29)



So, neighborhood operations as involved the creation of new data, based on the consideration of the roving window or moving window or neighboring points about selected target locations. The similar technique also exists in digital image processing or remote sensing data, where we call as a special filtering. There also a two dimensional matrix maybe of 3 by 3 pixel may be by 5 by 5 any odd number or 7 or 9 9 by 9 are used. These roving or moving windows moved through out your data raster data and calculate the value depending what kind of weight you have assigned. For the (Refer Time: 29:07) to calculate the center value.

Now whenever you use a neighborhood operation using roving or moving window on the margins; that means, one row first column and last column no calculation will be done because if it is 3 by 3 then you will have first row and last row and left most column and right most column no operation will be done. If it is 5 by 5 then there will be two rows on top and bottom two rows on left and right. So, the number becomes larger, and then number of rows and columns which are will not be involved for analysis will be accordingly.

So, in 3 by 3 roving or roving or moving windows you will have a just one top line and bottom line and left column and right column no calculation will be done. And they evaluate these roving windows may be 3 by 3 will evaluate the characteristics of an area surrounding a specified target.



(Refer Slide Time: 30:19)

Let me give you an example. Suppose your window is a 3 by 3, so what you do basically you can assign a weight for this. So, if I say the weight is equal all should 1, so this is one type of design of window. If window of different type can be designed which will have some directional weight like I can say that here is minus 1 here is minus 1, this will have 1 2 and say 0 0 0 something like this. So, you can assign even directional weight as well using this rowing window.

So, they will evaluate, so evaluate characteristics of an areas surrounding a specified target location here, target location is 3 by 3. If any my input raster is larger, so for the central pixel it will search this 3 by 3 one row one column all around that area and depending on the weight which have been assigned to these cells of my matrix there the new value will be generated. This is what it is says the neighborhood operations involve the creation of new data. It will create the new cell value or pixel value depending on the weights which I have given in my matrix which is a two dimensional may be 2 by 2 has may be 3 by 3, 5 by 5, 7 by 7 any odd number.

(Refer Slide Time: 31:57)



Now, in all neighborhood operation it is necessary to indicate one or more target locations. Generally it is signal target location that is a central cell or pixel, the neighborhood operation considered around each target and the type of function to be executed. Now type of function these means we are talking about the weight may be minus may be plus or you will not to multiply so on, on the attributes within the neighborhood.

And in the raster one single attributes is there so this will be done for all those wherever this will move one by one, from starting leaving in case of 3 by 3 will leave the first row and first column and then that means it will start from the pixels of a second column and second row. Likewise, it will go all around on the entire raster data set. And then the typical neighborhood operations in GIS are interpolation. So, we have seen in interpolation there we have discussed the search window; we have discussed the weight assigned to the search window and so on so forth. So, in interpolation we also use this neighborhood operations over a say digital elevation model to drive certain topographic features or topographic outputs from a map like a slope map (Refer Time: 33:16) map gradient map.

So, all these will consider under topographic functions which we will see very soon in little detail about those. And then finally the search functions, they are also will search the neighborhood. And this is how the neighborhood functions. Interpolation we have already covered that in order to for completeness within this a module of this lecture I am just very briefly going to interpolation that it estimates the unknown value at a unsample sides.

(Refer Slide Time: 33:50)



Using known values of existing observations and there are different types of interpolation you can use the point data as input or you can do the line interpolation. But basically line interpolation what it would be done, first in a within a given GIS software the polyline is first converted into points and then points are used for interpolation. But certain software will allow you to input as a polyline, but some software may not allow. If they do not allow there are tools available you can convert any time a polyline data into point data and then you use for interpolation.

(Refer Slide Time: 38:38)



Now, topographic functions, before that we have to little bit understand what is basically topography. Topography is the surface characteristics of terrine and these characteristics can be slope, relief, and form of an area which are referred as topography. So, when we say that if a area is completely flood we do not see any features, so that we call has a flood topography. When topography is having lot of undulations valleys hills and ridges we call as ragged topography.

Because, in the case of whether you will have slope relief aspects and all those things where there, and topography of a surface can be represented very easily in a digital elevation model. And what are digital elevation models, what are what are different types of digital elevation models maybe available, what are how to evaluate the quality of digital elevation model which we have discussed in some previous lectures. As you know that a DEM represents a topography surface, if it is a carrying the elevation value a typical DEM in terms of a set of an elevation value measured at finite number points and contain terrain features of geomorphological important; such as valleys, ridges, peak, pits etcetera.

(Refer Slide Time: 36:00)



And DEM are commonly they are organized in the grid form they are typical raster. And uses of such thing become much easier in GIS software using our computers because they use all the functions which are available for matrix analysis.

And due to their fixed spatial resolution regular grids are not adapted to changes in relief roughness. So, this in this here our raster is fixed the resolution is fixed; that means the cell size is fixed, so whereas the TIN is adaptable to relief roughness. And use of smaller sizes one option is that we go for higher and higher spatial resolution DEM, but there will be a one limit also. So, that also we have discussed that what should be the higher spatial resolution, where we should stop, so that we have discussed in the previous class.

#### (Refer Slide Time: 36:55)



And this triangulated irregular network which also a one way of representing terrain; that too we have discuss in detail. And this a definitely TIN is having advantage because it can accumulate your relief roughness and it will adoptable the unit which are triangles in case of 10 they will change the size and shape according to their requirements. But in case of a regular grade or raster the cell size are fixed.

(Refer Slide Time: 37:35)



Now, that to calculate topographic function we use the elevation values and typical topographic a transformation or functions which are used like slope which is rate of

change of elevation. Gradient maximum slope is called gradient and then the aspect the direction of a slope with reference to north. These are the typical topographic derivative which we drive from using a digital elevation model in GIS. Very easily to drive all these, there are already standard functions are available and these can be used and immediately you can drive.

The one thing you have to remember, whenever you are calculating the slope because the slope what it says the rate of change of elevation, so that means that change of elevation as you move horizontally and their and this scale both your scale that means the horizontally scale and vertically scale has to be same. If it is not same then you have to use a factor multiplying factor which will change one scale. So generally what we do, we will change the elevation to the mapping units, because this x y scale horizontally scale is generally in mapping units. So, we change that vertically scale into mapping units because there is a one to one relationship, this has to be remember.

So, a factor has to be used according to the location of your as per the different latitude because as they not perfectly squared therefore this complication is there. Plus the scales are different horizontally scale are generally in will be in geographic coordinates, whereas vertically scale that means elevation value and most common is a might be in meters. So, you have to convert. If you are having your horizontal scale or maps say in utn in meters then this will not come.

(Refer Slide Time: 39:42)



So, topographic function to calculate values they describe the topography at a specific location. And neighborhood is used to characterize local terrain. And typical examples are slope calculation and aspect calculation. Now using this GIS you can drive slope in two ways; one based on the degree and one based in the percentage. Very simple concept here, say tan theta that is rise over run. So, that a distance theta this is calculated, so as you move that is why these two scales has to be same.

So, here I can drive here the degree of slope is 30, whereas a in percentage it would be 45. Because, in case of degree it is 0 to 90, in case of percentage it is going to be 0 to 100. And therefore for different like here it is a 30 percentages 58 and here 45 percentages 100 and so on so forth. Depending on your requirements you can either choose degree or percentage. Later on also one can transform from one to another, but this is very easy calculation in any standard GIS software.

Now, aspect is the orientation of a slope surface with reference to the north that can also be calculated. Generally in by default you will get the output which will show aspect and eight directions starting from north and to north west.



(Refer Slide Time: 41:27)

Likewise, let us see you the example of a slope map in terms of degree it is given so this is my input digital elevation model which a showing a continuous values low to high. And when it subjected to slope map as per the concept which we have just discussed it has calculated the slope, so red areas are showing slopes between 70 to 78. Now depending on your requirements as we have discussed in GIS analysis 1 that reclassification can be done and you can say classify between 0 to 20, 21 to 40 and so on so forth. So, number of classes in your output map can be reduced if at all those are required.

This brings to end of a GIS analysis 2. In a next lecture we will discuss further GIS analytical operations.

Thank you very much.