Geographic Information Systems

Prof. A.K. Saraf

Department of Earth Sciences

Indian Institute of Technology – Roorkee

Lecture – 5

Demonstration Through GIS Software

Hello everyone! and welcome to our new discussion. And today we are going to spend some time about the demonstration of GIS software. I will try to do on 2 software's, I am just using these software's to demonstrate you, I am not promoting by any means these software's to you. I am using these software's to demonstrate a few capabilities of GIS which we have discussed so far in previous lectures because I use personally these software's.

And I have found quite interesting and sometimes convenient also. So, that is why I am using but there are a lot of options available for us to use a GIS including open source software's also. So, very briefly, I will discuss this part also. As you know that the first GIS software which was developed as CGIS that was somewhere in 1962-63, early 60s in Canada and though it was not in commercial software.

And after about 20 years, from 62 to 1982, the first commercial GIS software came and that was Arc-info. So, that was on UNIX platform and later on, it became available on Sun Solaris and later on also on windows as PC ArcInfo and parallely the older UNIX based or Sun Solaris based also continues and then with the PC Arc Info when we became this very familiar with windows because PC ArcInfo was command based.

So, when this windows became very popular and stable then we got a software ArcView. So, today I am also going to demonstrate to you on ArcView as well. Because I find very convenient and I will give you a little comparison also and the latest ArcView and the latest one which is called ArcGIS. So ArcGIS also, I am going to use today and before that I would like to mention few other GIS software's which are very popular.

And like if I talk about open source GIS software then the top ranked open source GIS software is QGIS; a very popular one and quite capable of doing this your processing part or analysis part in GIS then there is a commercial software like GeoMedia, MapInfo, Global Mapper etc. So, there are maybe 5-600 GIS software's from commercial costing few lakh rupees to free like QGIS.

And another GIS software which basically started in this open source domain was GRASS, is still continuing anyone who is interested to use the open source GIS software's then these 2, I would definitely suggest either you start using QGIS is very powerful. And also GRASS GIS, because initially GRASS was developed by US corps of engineers. And of course, then it came in public domain.

And still many people will prefer that software because it is open source, they can do their own developments and another thing. But in commercial software also, you can do your own development using either Python, Visual Basic or C++ languages. And you can develop your own tool. If I compare Arc Info and ArcView and just because these 2 software's, I am going to demonstrate today.

ArcGIS is quite heavy on the computers because it is very powerful, have got hundreds of tools available and anything which can you imagined that should be done by GIS software. You would find a root in this one. You would find may not be a direct tool but sometimes if your understanding is good then you can find a very good tool or steps using multiple tools to achieve your results.

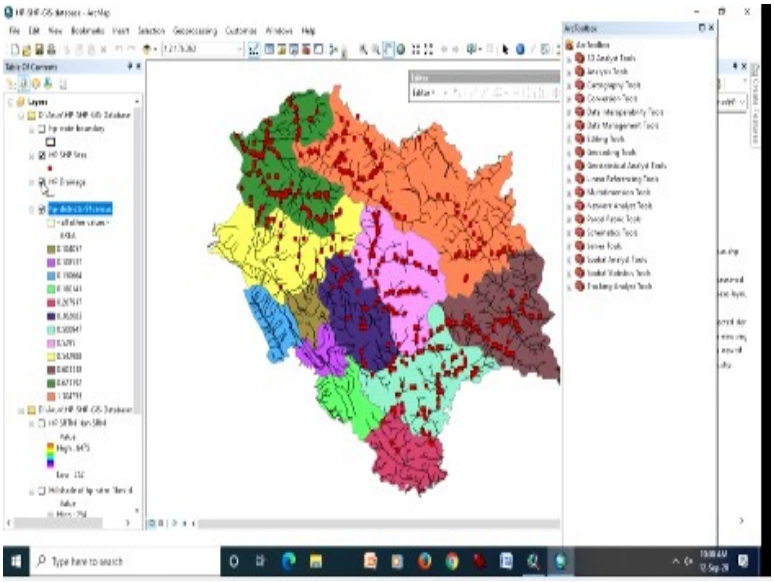
Second part, that is since it is own means upgradation is continuous by the company which is ESRI. And therefore, full support is also available for this software. Contrary to this ArcView is a very light on the computers. ArcView, we used to come in just 6-7 megabytes of memory and still it works on Windows 10, which I will show you. And it is a very easy to use anyone and you can find even today hundreds and 1000s of users have ArcView.

They are very convenient, they are very happy to use and if you can get, you can also get free of cost on net. But now company that is ESRI has stopped developing ArcView further because all the capabilities which ArcView had; all have been incorporated in ArcGIS along with many-many additional capabilities. But as I said ArcGIS is a little or quite heavy on the computer resources then ArcView GIS.

We are going to demonstrate about basically I will be start with simply point data. And then I will also see different other vector data entities and finally, we will also discuss something about topology, how topology is handled in the GIS software. So, these 3 things basically I am going to do on 2 different GIS software's and one is still continuing very, very powerful that is ArcGIS.

Another one which is not supported anymore by the company but very easy to use, and very light on the computer resources. So, let us say start seeing things in real about the demonstration of these GI software.

(Refer Slide Time: 07:18)



(Video Starts: 07:18) Here I am going to demonstrate on ArcGIS. ArcGIS is having 3 major components and the first one is a of course, ArcMap where you make the display of the different layer's data. Second is Arc Catalog. So, if you see here on the right side; the catalog word is

written and it is something like quite similar to your windows explorer. And the third one is called Arc toolbox and it is having hundreds of tools organized under different categories.

So, these tools we can use for different kinds of processing and analysis. Currently, what I am displaying here is a district map of Himachal Pradesh and top of this; I am having some point data. Here for example, I am taking potentially small hydro sides of Himachal Pradesh, which we developed for a project and this we can use as a point data for demonstration purpose.

So, 2 layers are there are on this table of contents side, you can see multiple layers also. Many layers are there but currently I am displaying only 2 layers just to keep things simple and all 3 types of data I can also show you hear that red boxes which you are seeing are the point data and in the background you are seeing the polygon data and also there is a HP drainage that is Himachal Pradesh drainage lines.

And that data is also available with us. So, this say HP and this one is also here drainage lines. So, I can show you one by one, like a point data as you know that in GIS, data is organized in different themes or layers. And anytime we can bring any layer, remove any layer, edit any layer and analyze any layer along with other layers. So, like here in the background, I have put the district's boundaries, political boundaries of Himachal Pradesh and also the drainage lines.

So, if you want to see only drainage lines, no issue. You can also see and of course, there are many-2 other associated things are here. You can zoom in or zoom out. And also, if you see here and when my cursor is moving on this map that is in ArcMap view then the same time the latitude, longitudes are also changing because as we have been discussing that every data is organized in geographic coordinate system.

And therefore, though the data in degree decimals but system can display very easily in form of degree, minute, seconds so that we can compare with other data sets. And the top of this; this is what I was mentioning earlier about the Arc tool and there are hundreds of tools, within tools and sub tools are there. And so, if I go for a Spatial Analyst tools now, within the Spatial Analyst tools; there are many-2 tools.

And if I go say in hydrology, again there are sub many tools and tools means; suppose if I open a tool like this and then it will ask for input, it will also ask where to store the file and also if there are options available and you also provide those options to the system. And by default, things will also be available and against each option, there will be help file on the right hand side.

So, here this is the tool and I have chosen in that hydrology analysis tool. I have chosen fill. So, input raster which is must mandatory, Output raster where your data will be restored, and that is z limit. And anytime you need any help just see on the right hand side, its detail will help, and generally available against all options, all tools. And further if you want more help, you can go to tool help option and do it.

Almost all software's are having similar help files and these things. You see when we venture in serious analysis of data, many times because of maybe an issue related with data or many other things, we need help and today's the best resource is either use the online help which is updated one, through the net or help within available within the software. Earlier we used to have manuals. Now online manuals or online help or when you install the software, the help is also installed.

But sometimes that helped maybe a little older. And so for latest help, one can always resort to online help for the same software and you not only get the help, but also theoretical background or all details even algorithms you may find on these online resources. So, that is a very good thing which one should always resort whenever it is required. Now, I will go one by one again on the point data which is being displayed.

Point data as I have told you that it is a zero dimensional entity and therefore even if I keep zooming this because I in the symbology; I have chosen and this is as a box, a square filled with red color. This symbology I can change to anyone but fundamentally it will always remain a point and point means a zero dimensional entity. So, no matter how I zoom it, I will not be.

So, here you can change the symbols. There are hundreds of such symbols are available, some are standard one also and as you can also design your own symbol, you can edit existing symbols as well and this is all allowed in almost all software's. So, symbol is just for display for our eyes only. But the data remember; it is the point data and no matter how much I zoom.

The size of that point is not going to increase because point is a zero dimensional entity and therefore no measurements like perimeter, area or centroid is possible. Same with the line data if I start thinking about the line data and then if I start zooming again my line does not become thick that drainage line does not become thick. It will remain the same. The reason is because line is a one dimensional entity.

And only the length I can measure along the line and not the width though the width information may be stored as attribute. So, that is why it is like this only. However, if I keep zooming a point polygon data like this and then of course, area and perimeter is not going to be changed; only the scale is changing basically. The scale is changing and that always you can also check what is the scale when it is being displayed here.

So, these are the very basic things about any software so, in ArcGIS. Similarly, as the same project which I developed on ArcGIS, the same project originally it was developed on the ArcView software. I told you ArcView is very easy, very user friendly and it is very light on a computer resources. So, it works very well. Again two things are here; one is district boundaries, political boundaries and another one, the point data.

And same way you can bring layers on the top of another; you can change the priority and other things also. Now, I would like to show you about all 3 data; I have just mentioned point data, line or polyline data and polygon data. How it is created that will come in our discussion maybe a little later but how we can analyze or take the advantages of this thing and that I can demonstrate also on this software.

So, what I am going to do; I am going to select based on the criteria. Recall the definition of GIS what it says as per user defined specifications. So, today I am the user and I will define the

specification and I want to retrieve the data as per the specification which I will give from this 0.5 which is Himachal Pradesh potential small hydro sites are there.

So here like in ArcView, you are having a query and by which you can be query builder. On the left hand side you are having all attributes information and then you can get these functions are available for you, which you can use and build a query but I will show you this the same thing here and in different name, it is called as select by attributes. And I can do it but before that I would like to show you like against each point as you know, we can have n number of attributes at least theoretically. So, if I open the attribute file here and then what I will see here and that there are various columns in which the database was developed, the spatial database was developed and each against each point and there are information's available to us.

So if for example, if I want to have the information about one particular point and if I click here, immediately information about this point will can be retrieved very easily from our attribute table like this one. So, and all these fields are column here are also available to us against that particular point. But if someone would like to do the analysis for many points together, then that is what we can do it.

Also if you see that, I can select a few points like this one. And once these points are selected, they will be shown in a different color and even in different symbol. But the same time, I can also see these points in my attribute table and I can bring those points, are here. See all these points which are there, have been only selected points are being shown here. And all points are shown here.

So, this is how all points are shown including the selected one. So, if I want to see only the selected one then I can get only that one and get rid of. Another thing is that if I make selection here as you have seen that if I make a selection in my attribute table then same time, I will also get the selection elsewhere. So, there is a dynamic link between my table and my map.

Map is here the point data and in the background I am having boundaries, district boundaries of Himachal Pradesh. If I do any operation on my table; the same time it is selected, they are also

see this point was selected or if I do an operation here then automatically that thing is also selected which I can see that where is that one or if I want just that one, it will come automatically.

So, it is very convenient to see what are the things which has been selected or not but it is a very-very simple operation but very frequently use operation on GIS platform. Now, as we are discussing about so we will go for this selection. First, we have to choose the layer. So, I am going to choose HP SHP. Only two layers currently are being displayed; one is HP SHP sides and another one is the district.

The initial query which I am going to make is only on this point data because this is the first simplest vector entity and I am going to raise a question to our GIS software that I am looking those catchment areas which are having value say more than equal to 100 square kilometers and that means I am having one field in my database or one column in my table which is having catchment area CAT underscore A, Where A is for the catchment area.

So, if I apply this one, these are the sides so these are the points which are satisfying this condition and they have been selected here. Similarly, they must have also been selected in the attribute table. Now, I can make little more complicated search there or we call as nested search then it is based on the Boolean logic if you recall which we will of course, discussing overlay operations in GIS in much more detail.

But based on set theory; veins diagram, Boolean logic that A Intersection B, A union B and so on. So, same those operators are also here. So if I say, I am raising a query to the system that if catchment area is more than equal to 100 and let we take 1 more here and head is more than equal to 50 meters because different fields are having different units there.

So, it is for query building, it is not an important thing. Whatever is the units are there, those will be there. We do not have to type or say anything about here and when I say Apply now, only those sides out of my point data which are satisfying these two conditions have been selected. And just to make sure that the selection is correct or in order, as I've been saying and keep

repeating this thing that one of the first rules of GIS is that after each and every step, check for errors, and check for your results.

And if you are satisfied then you go for the next step. Otherwise, correct it and then move forward. So, what as I have done the selection based on the nested search that if catchment area is more than equal 200 square kilometers and the potential head for the generation of hydro power if it is more than equal to 50 meter, then select those. So, if I go for only to show me the select things.

Now, we will concentrate on these two columns; one is the catchment area which you can also see here and the head. See, all these selected sides are satisfying all this condition even if it is more than equal to 50 is there; 50 head, it has been selected and here you may find maybe very close to 100, may not be exactly 100.

So, all above 100 or equal to 100 all have been selected that means if my database is having say 20 columns or 20 fields; I can make a nested search, a complicated search using all those fields like here these fields are available to me roughly 20. I can make a search and can use all these operators without any problem. However, only thing is that if you go for a very complicated nested search, first of all, you have to understand the results.

Here, nested searches are still simple that if this x is equal to more than this and y is more than equal to this, please select. But if I use 20 fields and make a very nested search, we have to also understand the results. So, what is the solution to this problem? So, the best thing is you do it in a stepwise manner that initially you can involve 4-5 columns or fields of your database. Do the search and on that you can do further search.

So, may not be in one go but in multiple go and once you are satisfied then in future, you can preserve that query language or whatever query you have created here and that you can go for in just one run also. So, initially better is to go step by step, to do the query. Similarly, here also same way query can be done. I will do this time little differently. If I say that districts equal to Kangra and say head equal to more than 60 meter.

The only thing which is required here to take care about the parentheses because these parentheses sometimes create a problem. So, only thing you have to take care about these parentheses, otherwise it works fine. So, these are the sides which are shown here in yellow color are satisfying these two conditions that; these belongs to the district Kangra and head is more than equal to 60 meter.

Now, let me show you the results. Here, these are the results for in the attribute table and you can bring all results together also like this, like in ArcGIS and here is the boundary between selected and so what we did? We put a condition that if district is Kangra and head is equal to more than 60 then show me or select all those points. So, you can see that head is here of course, more than 60 and all these sides which have been selected belong to Kangra district.

So, both software's; for basic operations are having same capabilities because while developing ArcMap, or this ArcGIS, this is the most modern version. Developers have seen that what were the best tools available either in Arc-info, PC Arc-info or in ArcView so all those best tools have been incorporated. However, when I moved from ArcView to ArcGIS, I had too many times search the same tool which I was using very easily in ArcView.

But much more powerful and advanced tools are available in ArcGIS. Only thing, it is heavy on the resources and it requires more learning time then ArcView. Otherwise, you know that ArcGIS is definitely much more powerful as compared to ArcView but if we have to do regularly very simple GIS operations then the best way to use these very lightweight GIS software. So, both options are available that is not a problem at all.

As I was saying also earlier that, here on this corner, top right corner, both scale and latitude, longitude in degree decimal. What in ArcGIS, we have been seeing that these are available in degree, minutes, seconds format here they are available in degrees' minute's seconds, though you can change this option and you can also see in degree but just wanted to show you that computer stores information in DD that is degree decimal.

And you can display in both formats either in DD like here I am displaying or in degree' minutes' seconds like in ArcGIS, we are displaying. It also shows the scale at this moment. So, if I zoom it like this, a scale will change. If I zoom further, scale will change and that means we are going for towards larger scale.

So, if the ratio value is smaller; we call as a larger scale but if I zoom out then the scale will keep changing and if I go like this, it has gone further smaller. So, the entire Himachal Pradesh data is being displayed here. Now, for the similar way, you can do a query against different themes in all your vector data. When the discussion on raster data will come, we will have another demonstration on that.

But today, I would also like to discuss about the topology part. So, with this demonstration, I will do it here also, in this one. So, what we need to do is first to create a theme and because the topology is built mainly for polygons that is the most useful one. So, I am going to create a polygon file and say that where to store so, I will target directory. For time being, I am putting here in temporary folder and then I give the name like poly1.

And now; though the file has been created but from data point of view, it does not have any polygon or anything neither attribute, or nothing. So, what I have to do, I have to create the polygon. So, I will choose the tools to create a polygon like this and I draw, see that when I start creating a polygon, automatically it is drawn; by definition, for the polygon origin and destination has to be the same.

So, that is why it is keeping origin and destination since the beginning and as soon as I double click here, like 1 polygon has been created. If I open the attribute table, I can see that is also here and I can assign an ID or automatically you can also assign ID later on. Now, the part about the topology so if I go for say this option, which says the append another polygon along with this and there you would see the really good part, good feeling about topology, why topology should be constructed.

So that, if I go and append another polygon along with the existing one on the neighborhood or side of this then what I have to do? I have to overshoot for the next polygon means I will start within, inside of the existing polygon and then I will keep doing like this and see the cursor and it is not now origin and origin point and end point is same. Here it will construct so, that option is different and therefore, the beginning and end is not the same.

And again I will come back inside the existing polygons. As soon as I finished, it has created a polygon and let me open the attribute file. Now, it is another polygon and I give the ID as 2. I can also display, this one and change the colors. So, I want to give unique colors like this for ID 1 and ID 2. Now, the main part of this topology or advantage of topology is the common boundary between this brown one and yellow one; these two polygons.

So, if I zoom any part of the common boundary, I will not find any gaps or artifacts or any other polygon. Total numbers of polygons are 2 which you can check in attribute table. If one is not building polygon topology and keeping data in a Spaghetti model then you would find a lot of gaps along this. So, the common boundary between polygon; left polygon and right polygon has not been drawn or digitized twice. Only once I have done it.

So, the common boundary between two polygons has been digitized only once and that is the basically major advantage of having polygon because if we follow for polygon topological model then the quality of data would be very high and of course, in a spaghetti model, there is no polygon topology and the line and the common boundaries between two polygons has to be digitized twice and there will be gaps and overshoots and all kinds of problems are there.

One more thing, I want to show you here that if I cut a polygon existing polygon suppose I cut from here to here then see what happens? And this is another way to understand how topology is understood by these GIS software's. This is, I just create simply a line, straight line or I can create a polyline as well like this one and as soon as I finished, now another polygon has been created.

And that means, the line which I have just digitized has become common boundary between these two polygons. And this here, I have to give now a different ID so that I can add display in 3 different colors which I am going to do it. So very quickly, it creates the 3 polygon and again using common boundary. So, one common boundary, initially was between these yellow and central polygon was from this to this. Now, another common boundary has been created like this.

So, lot of such tools would be available on GIS platforms or any software which you can exploit and create own polygon map. As this also discussed while discussing polygon, this topology part in a theoretical class, I told that earlier software's, we are not capable of creating this topology simultaneously. Now, this software's have become such convenient and powerful that they keep building the polygon as soon as you finish the digitization work.

One more thing, I would like to show you that if later on, I have decided that I am not interested in this polygon which I have just selected. This one which you can see this dark brown one so, I selected and deleted. See the common boundary is still, is there. But since one polygon, the leftmost polygon has disappeared but it has not harmed anything to the existing polygon which was, in its neighborhood.

Let me give you an analogy that it is something like 2 good neighbors are having two adjacent plots. And the first one construct his house and later on, the another neighbor is starts constructing the house. So whatever the money, the first one has spent for the common boundary or common wall that can be shared between these two in a manner that tomorrow another; any one neighbor or neighbors decide to reconstruct the house that the common boundaries should not get or common walls should not get disturbed.

And if such understanding can be developed like in polygon topological model then lot of disputes in real life, can be avoided. So, that topological concept should also come in our day to day life that common things should be shared without any condition between neighbors and whenever one neighbor would like to change the things, the other ones would not get affected. And if this thing can be done that would be really very good.

So, that is the advantage of having polygon data in typological model, not in a spaghetti model. No other software, no other concept like CAD CAM design, these graphics software's. They are not going to support topology. Only the GIS; topology is supported and as you know that of course, it has come from mathematical domain and first, it was developed there the concept prints, 200-300 years back.

And now; we are exploiting this one in our GIS software, the similar way I can also do it here, in ArcGIS also. Because both are at this stage, for demonstration purposes; both are having equal or you know almost equal capabilities. **(Video Ends: 07:18).** So, as I have shown to you the demonstration of some basic GIS capabilities through 2 software's.

You can try the same thing, having some point data or some line or polygon data which you can download also from various places and can try this thing on your GIS software. Suppose, somebody is using QGIS software, you may find a lot of sample data available here in there which you can use it, try it and also start editing that data whatever is available sample data. So, you would have the confidence how to use such software.

As the more time, you can spend on the software's, you learn more, you develop more confidence. Like this ArcGIS, I have been saying it is very powerful tool and therefore a lot of time is required initially. And once you have learned, you know, which menu and which options and which tools are available where and once it is mapped in your brain then things become much easier.

And one can really exploit these powerful tools for various types of analysis, which involves spatial data; data which is having geographic coordinates. So with this, I would like to end this demonstration. And in future, whenever we will complete some theoretical part, we will also have similar demonstration on those topics. Thank you very much.