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

Lecture – 30
Demonstration Through GIS Software

Hello everyone! and welcome to this demonstration of few things which we are going to do. The first one we are going to see through these two softwares ArcGIS and ArcView that is buffer creation. How buffers differently can be created. And we will also discuss in between during demonstration how these can be utilised. And second demonstration which we will see in this discussion about attribute classification in GIS.

So, these two main discussions or demonstrations, we are going to have through softwares. **(Video Starts: 01:02)** So, in ArcGIS, what we are having one point file which is soil samples which we will be using for creating buffers. And we can choose a field or some other you know fields or many fields together where you would find this tool in this particular software is in geoprocessing. So, when you go for this geoprocessing, there is first option which is called buffer.

As I have been saying whenever you use any software whenever help comes, at least first or second times, you must read that information. So, like here, it is saying geoprocessing tool that creates buffer polygons around input features to a specified distance. In our case, the input feature is our point data. Alright. Now, once we go then there new window will come.

And there we have to declare few things before we go to create the buffer. So first which file? So, I am going to take this soil sample file here. If more than one theme is displayed then there will be a list. If single themes are displayed then there will not be any other option. Now after this, we will go for where to keep the data that the output feature class, that will come very soon. And then we will have other details request like here.

So, whatever is in default, it is creating. No problem. We will go about that. Now here, linear units; these are in our mapping units and mapping units as you can see are in degree decimal. So, these are in decimal degrees and therefore, we must take care while creating buffer about

this units. It has to be in the mapping units, whatever the value you provide in mapping units. Suppose I am giving 30 here in decimal degree in the mapping unit.

And what I am going to say that I want no dissolved type initially. I can do like this and in next round, we will do a dissolved type and then see what is the difference between these two. Basically, when we go here dissolved type as you can see that if we choose none then an individual buffer for each feature is maintained. And when we say that dissolve all then all buffers whichever having, overlap will be dissolved.

So, time being, I am just going for the default that is none and then let us do this buffering. It may take some time and then we get a buffer layer as per the given units or distance. So, around these input points, buffers would be created having this 30 decimal degrees linear units which are quite big in that sense but depending on your requirements. I have just taken one value.

And deliberately, I am about to create a big circle. So, they become obvious and visible to in this demonstration. See here now these buffers around each point has been created. But since we have chosen option not to dissolve and therefore you will see that there are many buffers around many points which are overlapping to each other.

So, for real analysis, we need to you know resolve or dissolve these boundaries of individual overlapping buffers so that we get a good real picture. As you can see that here, a large overlap is there. So, in the background, we are seeing 6 buffers and on top of those, 2 buffers are there. In next processing what we are going to do? We are going to create buffers with dissolve and the rest of the thing, we will keep as it is.

So, I have taken that one and here, I am again taking 30 units. And here, only the option which I am going to change that dissolve all. All buffers are dissolved together into a single feature removing any overlap. So now, there will not be overlap in these things. It may take a little more processing time because not only the system has to create buffers around each point.

But later on, it has to dissolve those buffers as well. Now here, what we are seeing now these dissolve buffers. So, buffers earlier and these dissolve buffers are having quite good

difference. So, if I give a different colour to the previous one and if I bring top of this one, you can visualise that what I am talking. So here, what do you see that it looks more realistic.

But this is about the soil sample data. And if this would have been a different thing like suppose it would have been a network of mobile towers. And around each mobile tower, suppose buffer has been created. Now by doing this thing, we would know immediately that where are gaps in that area where no one mobile signals can reach. So, such kind of analysis can be done involving buffers very easily.

And what I am going to do? I am going to overlay the previously created buffers top of this so you can see that what I am talking basically. See these red circles are showing those overlapping buffers but the new buffer which has been created is without any overlap and all has dissolved. So, it looks more realistic and more useful than without overlap. So, this care must be taken.

Now, sometimes there might be requirement of creating multiple buffers around a point. So, this can be done very easily in ArcView. Also, I am sure it can be done in ArcGIS in multiple stages. But at least, the tools which are visible to us does not allow us to create multiple buffers in one go. Though multiple go, obviously it can be done.

So, what we will see? The same data file we are going to take in ArcView and here also, we have to add an extension buffer wizard. And this will allow us to create buffers. So now, this create buffer option is there. Once we have set the units. So, wizard means that basically, these are sequential processing and much easier way, you can do it very quickly.

Of course, it is asking that what is going to be your input file? So, we are seeing since in ArcView where we are having only one file so automatically is there. It is also declaring that there are 74 points so should number of features selected. Should I create buffer on selected points or on all? So, since selection is zero and therefore, it is going to do on around all input points.

Now once we go for the next, it asks few things that what is the specified distance? So, suppose if I give a specified distance 20 here because we are going to create multiple buffers. And I would like to give a field here say for example pH. And then we can choose for

multiple buffers also. So, multiple buffers if I do then also. So, first we can go for say this kind of buffering that giving 20 metres in distance and can create buffers likewise.

There is some issue with the number, okay! This exercise basically, let us do it for multiple buffers. Just for that purpose only, we were doing. And multiple buffers; instead of 1000, we are keeping here 3 and 10 and of course, it is in metres. And this will work. We can keep little larger value since it is asking so no problem.
Now again should it be dissolved? So, we say yes, it should be dissolved.

Either you can store as a theme or you can store as a graphic. So, we will do as a graphics. Later on, you can convert as a theme also. And we finish it. Now here, buffers should have been created. Now see, the buffers have been created. And we zoom because it depends on the units you know. So, if we zoom, we should be able to see like here. See buffers have been created like this.

So around different points, we can aa create buffers, multiple buffers also. More practice would be required to have a full control over these softwares. Otherwise in one go, you can create multiple buffers like here, we could do it multiple buffers just for a few points. But if we want to do it for many points, again we can try and run that one, that here.

And then here, I choose a field like this one and then say okay! Again, problem about the distance. So, okay! Maybe a little larger distance we give and then multiple buffers and we are going for a new theme. Again, multiple buffers have been created around each file. Depending on what distance we are giving, your multiple buffers around each input point which you are seeing here.

So, each point is having multiple buffers because earlier we have created 3. Now giving the different possibility, we have created 3 so they are more now buffers have been created. So multiple buffers can be created. What is the use of multiple buffers? That if we know that something gets disperses with distance like pollutants then we can create some multiple buffers though, they are having little harder boundary.

So, in order to have a softer on realistic boundaries, what we can do? We can first create a surface using interpolation techniques. Create the buffers. And for buffers, we can extract

from the background layer that is interpolated layer. And we can get a very smooth multiple circles showing the conditions which might be there or would be there. So, it depends on which application you are trying.

Then as per that and depend on distance, these circles or buffers can be created. I have taken the simplest vector feature that is point. Similarly, the line features or polygon features can also be taken and around which, the vectors can be created without any problem. So, this is how the buffers can be created. Now, what we will do? We will also see one more thing here that is the classification of attribute data.

So, what we are going to do here? We are going to you know add one file which is a raster file and will try to perform this classification. So here, the classification is going to be on raster data. Similarly, you can do also on the vector data as well. There will not be any problem. Now here, this file which I have added is a digital elevation model of the same area where we were handling the soil sample data.

Now for timing, it is displaying in a continuous fashion because it is raster. So, it is a continuous data. The lowest value is 1489.33 and the highest value is 1532.21. This much information we are having but if we go for this classification then as discussed in the theory, we get lot of options here. And this also as I told you when we have been discussing classification of attribute data.

At that time, I told you that there are 7 types of classification techniques which have been implemented into GIS including that manual classification technique. So, where we get these options? When we go for this symbology. First layer properties then symbology and then when we choose this classified then we get option about the classification.

So, when we press this one, what we get here is a histogram. As I was saying that the histograms should be studied first before we choose any kind of classifications. So, see the distribution of the cell values between within this digital elevation model are varying. It is not a bell shaped. Generally, it is true that most of the data sets may not show the distribution in bell shaped.

So therefore, now options are there. By default, in this particular software, it always goes for the natural breaks which is based on you know this Jenks optimization. Though, the one which we have discussed first by discussing the types of attribute classification was equal interval. Though manually, we can also define intervals. Manual classification is also there. Then after equal interval, we also discuss quantiles then we discuss natural breaks.

We also discuss statistical ways that is standard deviation and geometrical interval. So, any of these methods of classification can be chosen. By default, it always does it in natural breaks. Suppose, if I go and say do it you know equal interval. Here, I can decide how many classes I would like to create? By default, it is giving 5. I say I want 7. Now, as you see that when I have chosen 3 classes, it is showing this because it is equal interval.

So, it will cover all the distribution of cells between having this range into one class then next and then next. If I choose 7 then there will be these 6 vertical lines dividing the whole histogram into 7 categories/7 classes. And the breaks are going to be like this break values, it is going to be like this. Now, I can also choose a percentage or a normal one. So, I do not do any further thing.

But before that, if you want to see the standard deviation of this input data then you get immediately by choosing this or you want to see the mean; mean line is here and that is giving your further information. So, study of histogram will make us understanding of the distribution of the data within one file will make us to choose most appropriate classification method of our attribute data.

Here, in case of raster data which is a digital elevation model which I have taken is a continuous. And therefore, it will have only one attribute that is the cell value or elevation value here. Here also, you get simple statistics including a standard deviation, mean values, sum value, minimum and maximum and total count; everything is there. So, if I classify this one using equal interval and 7 classes, this is what the result I am going to have as you can see.

Also, during discussion, I said the colours do not matter in GIS. What really matters the value; value of the cell or attribute of your feature or your object that maybe point, polyline polygon or a raster. So here by default, the grey ramp has been chosen. We can choose some

other other coloured ramp also like this one, I can choose and then say okay. So, if I say okay then I get this thing varies without any problem.

Now, if I change my classification method. Instead of say using equal interval. if I say I want natural breaks. In this exercise, I will keep all classes 7. Total number of classes to divide in 7 and then say okay. And when we say apply, let us see what changes brings. Here you have to focus in this part. So, see slight changes has brought. Now what the point here for discussion is that different classifications should create always different results.

But all these will depend on the distribution of the data or that simple statistics, we are seeing. It is not necessary that all classification will produce all distinct differently results, not necesa1rily. It will depend on the distribution of data. Now this time, if I choose some other classification like a standard deviation and I say there will be only five classes. This option is not open then so 2 in minus side, 2 in positive side and 1 is in centre and the mean one.

And then when I say this one and apply. See, it has completely changed the appearance of my classified map. So, if this map is satisfying my needs, I can accept this classification as well. But one important thing which I would like to highlight here that there is no change in the data. This is what the word manipulation comes in our discussion. What we are manipulating only with the classification and trying to display new things on screen.

But the real data is intact all the time. There are no changes whatsoever in the real data. So here, you see the 7 options but defined interval means this is also manual one. Here I can define interval or equal area. I can choose equal interval. You can define also or manual where you want so manual, you can also choose. By which depending on your requirements, you can choose a classification technique like here I've just chosen geometrical interval.

And see this distribution; this is based on also frequency. So, the width of classes here is not same. So, wherever the values are having less frequency, the classes are wide here. The distance between these two blue vertical lines is big here. Whereas wherever you are having more you know frequency then it is less. So, it is basically also in frequency area-based kind of thing.

And when we go and say okay classified, just focus on this one. The entire appearance will change. Nothing has happened to the data. Only for our visual inspection, things have changed. So, different classification methods can be implied not only on the raster data sets but on vector data sets as well. As you choose the theme accordingly, this will be done very easily.

But the main point here is that one has to keep in mind that before choosing a classification; this I am repeating again before choosing a classification, first study the histogram. That means the distribution of your cell values/pixel values or other values. Once that you have understood then you would get some idea that which classification method, you should adopt.

Simply statistics will also allow you to go through that one and then you choose appropriately the classification method and number of classes as per your requirement. And then you can go and perform the classification. It should produce desirable results. Now suppose in this one, I have found that geometrical interval is giving good results. But that does not mean that if I am having another set of data for maybe same area or different area, the same classification method having seven classes will give the same kind of result.

May not be necessary because it depends on the distribution of the data. So, that is why I am emphasising again and again that histogram; simply statistics will let you know that which is the best way of doing classification. Otherwise, methods are always there. Nobody will tell you that which classification methods you use. But whenever we go for equal interval which is the simplest one, many softwares are having in general in the default.

But you know that is not very scientific in sense that what is the basis of classifying everything in equal interval. Because it should be classified based on certain statistical or distribution or shape of histogram likewise. But sometimes, even equal interval is adopted just for simplicity. People will easily understand definitely equal interval. They understand equal interval classification method more easily than any other.

Because first, they are seeing how things are distributed. And what you are seeing here, this will become also part of our legend. So, they will see the colours and try to relate things with the map and the legend. So, as I mentioned that equal interval is the most popular or most

easy to understand but sometimes for many types of analysis, that may not be most appropriate.

This we also discuss while discussing the theory part of attribute classification. So, this basically brings to the end of these two demonstrations about buffer and classification. See, the theory generally is you know will tell you what are the things, what are available and how things are done? But when we come for implementation or application or using some softwares then various options are available to us.

Like when I was saying about buffers; you can create multiple buffers. You can create multiple buffers in ArcGIS also but maybe not in direct one goes but in multiple go. So, basically there is no instructions or nobody will tell you that how much rings, you should create? How many buffers, you should create? what should be the distance among these buffers?

So, all these when we come for the real practicals in our projects or studies, these things have to be decided by ourselves. Softwares are providing software tools to us; a set of tools, hundreds of tools. Which tool to use and when? This would depend definitely based on our requirements. And once you know that what my targets are then you can decide how I can achieve my targets or how I can process the data which will create or which will satisfy my targets or my higher ups in the system?

So, for that, there is no straightforward thing but in theory, things are straight that I have discussed that there are 7 types of classification; one by one and advantage, disadvantage likewise. But here, when we come for the real practical things then I cannot force you or can tell you that which one is the best, not at all. It will depend as per your requirements; one and in particularly in case of classification, it will depend on the distribution of the data. So, with this, I end this discussion or demonstration. **(Video Ends: 30:40)** Thank you very much