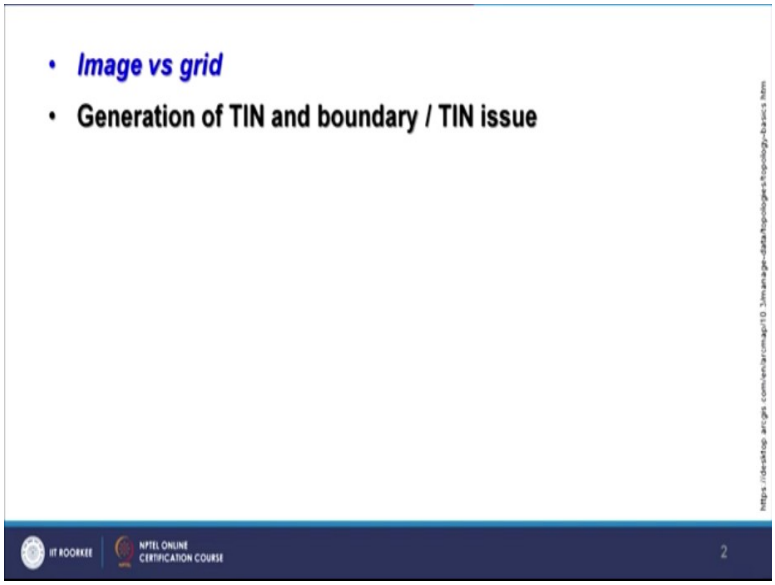
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology – Roorkee

Lecture - 10
Demonstration through GIS software

Hello everyone and welcome to new discussion. We are going for a demonstration on GIS software. And today, I am going to use ArcView GIS which is very old version but it is very convenient. And this I have also discussed in the previous demonstration class. But the same which I am going to show on ArcView can also be done, maybe even in much better manner in ArcGIS as well.

But since it is very fast so, many times for demonstration purposes, I prefer this one. What we are going to discuss here?

(Refer Slide Time: 01:02)



And will be showing that image versus grid differences. And early part of our demonstration would be spent on this. And then we will go for the generation of TIN and also boundary issues which we have discussed in the theory class. So, these two things we are going to discuss. So, let us go for now software demonstration.

(Video Starts: 01:38) So, in this ArcView GIS software it is much simpler. And what I am going to do, I am going to open an already existing database which I also showed you through that ArcGIS so that you would have the feeling that both really works in the same

way but it is very convenient. That is why, many times I prefer this one. Of course, currently we are only having in this project the vector data but we can add the raster data as well.

Here first thing is that you have to add extensions. There also in ArcGIS you have to add extensions. So, if I add an extension which is called Raster or Image Analysis also. Like Spatial Analyst extension and thus with the same name it is also called a Spatial Analyst extension there in ArcView, ArcMap or ArcGIS. So, once I have added these extensions, I will have now much more capabilities to do it.

And if I go for this Image Data or Grid Data Source. So, what I am going to do? I will bring some sample data which is of ESRI and which is very useful data for such kind of things. And, I display here the elevation grid like this. Now this is our grid data and currently it is being classified into different classes having these ranges which you can see in this TOC.

But if I choose here the option Image then same data I can display as image as well. And in that way basically system is taking care about the data in both formats. The original data will remain the same but only for display purpose it is now showing as an image rather than as a grid for the same data. So, many times depending on our requirements, we can play with these things and can have this kind of display.

Now, the grid can also be displayed as image but image cannot be easily displayed as grid. Now, when I say about the grid, now let me zoom some part here and explore what are the values here. Because what really matters the difference between image and grid is that if I click here, this is one cell of my grid and this is more distinct for example if I click here, I get the value.

And that is the real value I am getting in the floating and the precision is very high though up to 6 decimal places. But many times, as you know this is the elevation value and elevation value cannot be stored. So, this is showing some redundancy in the elevation value. But does not matter. Sometimes if things have not been properly defined then redundancy may come even in this. So, unnecessary it is there.

But we are not going to reduce that one. We will discuss these things related with raster data compression techniques. So, anywhere within that cell if I click, I get the same value but as

soon as I go beyond this cell and click here, I get a separate value. Now, here I want to bring a very important discussion that within the cell everything is having the same value.

Wherever I click, I get the same value because this cell is representing an aerial average elevation value. Because this is a digital elevation model so, this is an elevation value, an aerial average elevation value. And I have zoomed this grid up to unit level that means I cannot see anything within this one. Whatever is there, it is there and single attribute has been assigned to each cell of my grid.

And same also happens in case of image, both are raster. So, Because of definition, the major difference in the raster and the grid is that in image these values will not be real values, they will be integer value. That thing we will also see little later. So, what I can do? I can bring an image here. The true image not a grid being displayed as an image.

And then I can show the image data also. So, if I choose an image like this one which is a sport satellite image. And of course, this is quite dark in that sense so what we can do? We can enhance or stretch. So, if I go for Standard Deviation stretch, it will look like this now. But about the stretching and other things that is not for discussion today.

Again, exploring the value so that we can understand the difference between a grid and an image. So, if I zoom in this part, now I start seeing each individual pixel. So, if I zoom further this is what I get. So, if I go for the exploration because this is colored image, it is made from three colors and we are using an additive color scheme that is RGB or which are primary colors also.

So, for each pixel, I am having three values against three different channels or bands. So, even whenever I will click anywhere, I will get the same value against each channel. If I change my opposition and go to the next pixel, here we call as pixel because it is an image, I am getting these values. If it would have been a single band image then I would be getting values which would be only against one layer rather than three layers in this example.

But here important point is that the values are being displayed as integers in all 3 channels. So, wherever I click, I get the values like 124, 124, 97. There might be some very different

cases where we may get all values for all 3 channels. And sometimes, you may get same value like here for against 3 channels for the same pixel, we are getting 3 different values.

So, what we saw? We saw that in case of grid; the cell value is coming in decimal or in real numbers whereas, in case of image these values are coming as integer. And when this situation is there, in case of image we call the unit as pixel. So, automatically means when I say pixel, I am talking about an image and my pixel value is going to be an integer value or whole number.

And when I say cell, I am talking about the grid and that means the cell value maybe positive integer, negative integer or positive real number, negative real number. So, all 4 possibilities are there in case of a grid or in case of a cell of a grid. But in case of an image, only positive integer values are assigned or having with pixel. So, that is the major difference one has to remember about image versus grid, though both are raster.

And many processes are going to be same with that. So now, I go for the next thing and I make this thing clean first. Now, I add my data to prepare a TIN and as you know, that TIN can be prepared using point data. So, I am going to use this point data, let me change the color here so that it becomes very clearly visible.

And that is not a big problem because this is only the symbology or data and I can change the symbol as well. So, instead of that, I can choose this one and also color I can choose like this. So, I should be able to do it like this and apply and this, that is it. Now, this is the attribute table, non spatial data associated with each point which are being shown.

So, there is a column which is called pH. And as in the previous discussion I was mentioning that generally pH is recorded up to only 1 decimal place. And that is why one should keep it as it is. It should not have 4 or 5 decimal places for pH values. Same generally with the temperatures in degree centigrade or Fahrenheit, they too are recorded up to 1 decimal place.

So, in our database, you should not go for higher places after decimal or higher precision because neither instrument is measuring nor anybody is expecting that I would be doing that kind of analysis. And if results I show to somebody that average temperature for this area is

23.5678 then immediately people may ask the question that how you have measured? how you have achieved to this figure? So, maximum I should say 23.5. That is, it.

So, precision is very important here. I have selected this. I will be using this field which is pH field to generate a TIN. And instead of elevation, I can use any other field and numeric field to create a TIN also. So here, what I need to do? I need to add an extension which is 3D Analyst extension. Same extension is also used in case of ArcGIS with same name 3D Analyst.

And once it is done, now I am having an option in this surface because TIN is also a surface. So, create TIN from features. And also, in the bottom of this software, a brief help is available to me which comes here, just in this part. So, when I put cursor here say build TIN using selected features in the active feature themes. If I do not select any points, then all points would be considered to create a TIN.

So, that is what I am going to do. I will go for all points and the height field or source of height is going by pH as discussed. And these are my mass points and value field, I am not going to change anything whatever in the default just to keep things simple at this stage and go for interpolation or creating TIN surface. So, once I go for this, of course, it will ask where to put your data.

So, if I say I want the data to be put here, no problem now the TIN has been created. And let me bring these point data on the top and also change my color here. So, that point data is better displayed here like this. Now if I want to explore this TIN then for that I have to zoom it. So, let me zoom it. Now this is a regular network of triangles as you can also feel here. And all these triangles make a facet for us or a plane and there how you handle.

So, if I explore anywhere else like here, then I get different elevations because it is representing a facet or a sloping surface or plane. Therefore, I am getting different elevations, instead of elevations, these are pH values, remember this I have taken pH values but rest of them are accordingly slope and aspect. So, when we were discussing about TIN in a theoretical part of the TIN, then I said that whenever you generate TIN, 2 more important information are generated.

And that one is slope and another one is aspect. Whereas in case of raster, these has to be generated through further processing but here simultaneously once TIN is generated, both this information's are also generated and kept with the data and very completely intact. So, that is the biggest advantages of TIN. And each triangle is representing different elevations.

So, it is noted in that way exactly like in case of raster that if I am talking about grid within one cell every value will remain same. And whereas in case of a TIN within a facet or a plane or triangle, I may get different elevation values, slope values and of course, the aspect value will remain same because that is the orientation of that sloping plane or facet with reference to north as you can also see.

So, whenever I click here, I am getting the same aspect. Whereas the elevation keeps changing like this, see here. So, it depends in which facet I am. On the screen you may see the similar colors but in fact, in the data itself there might be different adjacent triangles. Another point which we discussed in theory is that wherever more observations and less variations are there, if this is the situation, larger triangles may be formed.

But generally, if density of observations are very high, smaller triangles are formed. And wherever you are having less number of mass points, larger triangles may be formed. And that is what you are seeing here, that in this part of our TIN we are having points which are very closely associated as compared to these 4 points. So, if I zoom, I am getting smaller triangles here.

But if I zoom in this part, I am getting larger triangles which have been formed. So, as discussed earlier while discussing the TIN in theoretical part that TIN is adaptable to relief roughness or changes whereas raster cell size is fixed. And within one cell everything is an aerial average. But in case of TIN, it is not like that. So, TIN is basically a completely different kind of data model to represent a surface.

Now another very important point which we also discuss associated with TIN is that if you see this border, the boundary issue is here. As long as the data is there, it will form a triangle. That means it cannot do extrapolation. But in case, if I use the same data set which I am having that is the point data set which I use to generate TIN. Now, what I am going to do?

I will use the same data set and generate a raster and then see what really happens? So, for that I can create a raster. And that raster we can create using interpolation. Interpolation discussion will come little later so that I will not spend much time on interpolation theoretical part but I can go for that. So, currently what I am going to do? I am saying that the extent of my output is going to be same as display.

And when I choose this one that means this entire windows area is taken as an extent because I have said that output has to be the same as my display. And now output, grid decides because in case of raster I can decide the spatial resolution but in case of TIN, I cannot decide spatial resolution. Currently it is deciding spatial resolution. So, I will accept as default values no problem.

But I can change these values to some other spatial resolution. There will not be any problem at all. Because if I change a value say here and make it as 2.0 and enter then you would notice that the number of rows and columns will also change accordingly. See it has happened. So, let us keep this one there is no harm. And then we will go to the next screen which says next option that which method of interpolation would you like to adopt.

There are my various methods but whatever default is coming IDW and IDW stands for Inverse Distance Weightage method, I am going to use and the same field I will use like in case of TIN we have used the pH field for Z value. And again, the search area and other things whatever in default I am going for that thing. And the barriers or brake lines here also in raster I am going for and for time being, I am saying no brake lines just taking the default values and go for interpolation.

Now, interpolation has been done and because of small set of data set that is good for demonstration and quickly it can be done. Now, when I will display you would notice that it has covered the entire display area of us which was available before. Now, let me bring first the input point data they are having. So, wherever we had the points between points interpolation has been done.

And wherever we did not have observations that means points, extrapolation has been done. So, if I take this part, for example then between these points' interpolation is done inside all these, on my right side. But as soon as the margins of point comes then this is all

extrapolation. And as you know that if we do not have observations on the left side here, except for that point data then of course, it is going to be very wide values also.

As you know that overall raster can only be stored in 2 shapes either in rectangular or in square shape. So, this current one is definitely in rectangular shape and If I display now TIN over it but I am seeing that TIN was restricted only wherever I had the points and beyond point because in TIN there is no concept of extrapolation, not at all.

Each point mass is used to create a triangle or a facet, instead of calling triangle we can call as facet. Even for in this case, like here, if you see this is one point, this is one point and maybe little lower. So, here no extrapolation, this point, this point and this point has been used to create a triangle. Recall our Delaunay theorem which says that a circle should be drawn in such a manner that only 3 points it should connect, not anymore and then that triangle is formed.

So, this is how these facets are formed. So now the issue is about the boundary, the TIN will not make rectangular or square shaped boundary as in case of raster. And therefore, you are having boundary problems. I can make subset of my raster very easily but I cannot make subset of my TIN. So, when discussing about TIN, I said that if at all I need a subset of that thing then first what I should do?

I should create a subset of input point first and then use those subset points to create a TIN otherwise later on a subset of TIN cannot be created whereas a subset of raster can be created very easily. So, this can also be demonstrated for which I need an extension and which I will add here very quickly and which is many years back it was developed by myself only.

So, here we can extract some data, one more extension would be required. So now, I am having this extension which is called Grid Analyst extension which was also developed many years back by myself. So, what I am going to do? I am going to create a subset of the raster which we have just created. So, first thing what I need to do? I need to draw a polygon over it. And once I have done it like here.

So, I will deliberately draw a polygon which is not really rectangular or square in shape. Therefore, I can demonstrate some other interesting thing about GIS capabilities. I have

created a polygon which is not really a rectangular or square but remember that a raster can only be stored in only 2 shapes either a square or rectangular. So, what would happen in this case?

Two things are going together. One is making a subset of raster which we have just created and see what happens to this scenario when the polygon is not really fitting in a rectangular or a square shape. So, I will say extract grid using selected graphics. This option I will go for. And as soon as I do it, it has done it. The colors may change, do not worry, colors we can bring the same as the previous one. There is no issue at all.

Same colors, I can go for that. And now see this is a subset of that original one. Values wise everything is same except that these colors have changed. But I can bring the original colors as well. But just to see both differently, let us keep this scheme of colors or enhancement. Now, what I am seeing here that this is my original data and the shape was completely different of that polygon. Exactly using that polygon, it has extracted.

So, when I see this one, this is what I am getting now and it is also satisfying the condition that overall shape here is now rectangular. So, what happens to the area which do not have any value? And in modern GIS these values are stored as no data, as you can see here. So, I am not getting any value if I click on this black part. As soon as I go on the digital elevation model part, I get the values, like here no value, like here no value.

I can also play little with this and say that these no data values should not have any value whatsoever. So, do not feel that now raster has become another shape than square or rectangular, not at all. Only for display, what I said that no data values should be displayed without any color. That means they would mix with the background of the screen currently the background is white.

So, it is mixing with that and I see a grid. I see a digital elevation model having an arbitrary shape. But in fact, the system is storing it in a completely rectangular fashion. But in case of TIN, I cannot make any subset. There is no such tool exists. This tool is of course not going to work because this only says that grid. So, you will not find any tool anywhere which can make a subset of TIN. So, that is one of the limitations with the TIN.

As I have said if you are ultimately going to make a subset better first make the subset of point data and then create TIN. There will not be any big problem in that way but the boundary issues will always be there. Now, let me also prepare one other way, little different, instead of saying same as display for the output extent, I will say same as my data. And what my input data is the soil map.

And here, I am keeping the same resolution as I did previously. So, I get the same resolution and then I choose here same method of interpolation and same field, rest of the things are remaining same. Now see what happens here? It has created raster out of my input data and see what is happening now? That it has only created raster till where it got the extent covering entire input points.

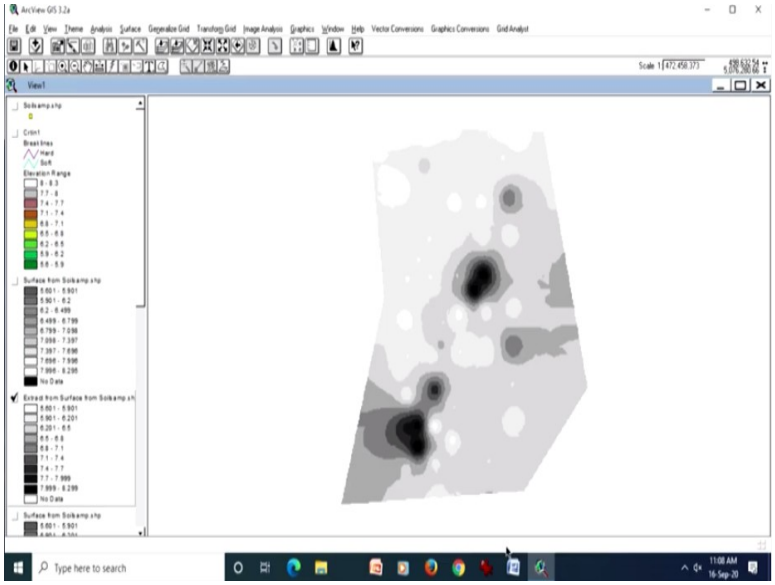
And if you compare this one with the TIN then TIN will also go. Because theoretically or practically the raster has to be stored either in rectangular or grid form, it is a two-dimensional matrix and number of rows and columns need not to be the same. So, 2 shapes are only possible rectangular and square. But in case of TIN, this condition is not there.

See even the border areas if I zoom this part, first let me get rid of this polygon which I created earlier. Now, see this border area that TIN is just touching those points and creating triangles but here in case of raster beyond these observations it has done the extrapolation. So, whatever be the area which you are seeing beyond the points, is all extrapolated but overall shape of raster is rectangle.

So, the border issue is always a big issue in case of TIN but boundary issue in case of raster is not an issue at all. Because if I extract then using this concept of no data which we will have separate discussion, I can still see an arbitrary shape. But at the same time, our system is storing data in two-dimensional matrix. So, this brings to the end of this demonstration.

(Video Ends: 35:21)

(Refer Slide Time: 35:22)



And I hope that by now, you must be able to get either of these software's or even QGIS or GRASS which are the open-source software. And start doing the same exercises which I have done on this software's. And get the feeling of the data, how system stores the data? How you can generate the data? and how you can compare different data sets, vector, TIN and raster?

So that, then you will understand the efficacies or advantages associated with each type of data model. And of course, disadvantages of all types of data models including your non spatial data. So, this brings to the end of this discussion. Thank you very much!