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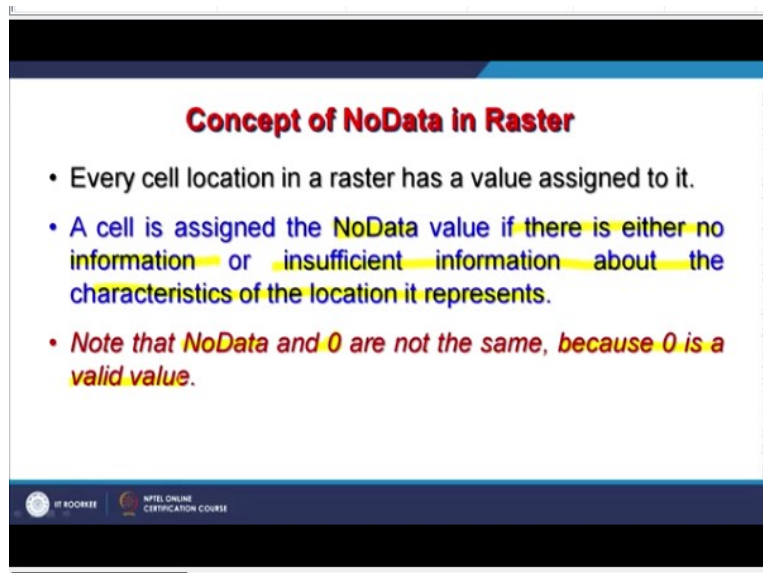**Prof. A. K. Saraf**
Department of Earth Sciences
Indian Institute of Technology-Roorkee

Lecture-32
Concept of NoData in Raster

Hello everyone! and welcome to a new discussion which is the concept of NoData in raster and if you see the title the NoData something is awkward might be for many people but in GIS this concept is very much required. So, initially when we started using GIS, it was mainly focused on the vector data handling an analysis.

(Refer Slide Time: 00:57)

Later on, then we got the capabilities in GIS to handle raster data but then problem came about that how to keep information about the outside of my area of interest. So, for that, this concept was developed which is called NoData. NoData means there is some value but it is not 0. So, NoData is not 0; this is the first statement here.

And then we will see that how it is important in GIS. As you know that in raster, every cell is having its location and a value and that value is assigned depending suppose if I am having 8bits digital elevation model or 8-bit satellite image then my values might be varying between 0 to 255 that means I can have maximum 256 values but at the same time, we also know that raster data can only because raster is a 2 dimensional matrix.

And therefore, raster data can only be stored in 2 shapes; either square or rectangular that means the number of rows and columns need not to be the same. If they are same, it is fine. But if they are not same then overall shape of the raster is going to be rectangular but sometimes when we are dealing some arbitrary boundary suppose district boundary or a state boundary, country boundary or a watershed boundary; hydrological boundary.

Then how I will have these things because these boundaries are not in a square or rectangular shape. So, how these will be handled in raster especially outside of my study area so that the entire data becomes either rectangular or square. So, keeping this thing in mind, this NoData value is assigned if there is either no information or insufficient information about the characteristics of location, it presents.

And that is what the outside because inside we need to have information but outside, till it gets the overall shape of a square or rectangular. We will see further details also with examples. And as I have already mentioned that NoData does not mean the value is 0, NoData 0 is a still a value. I will give you an example that if I am working in a coastal area and as you know the mean sea level near the sea or beach, I will have a 0 elevation.

But if I go towards the sea side then I should get minus value as well. And if I go towards the land, I will start getting elevation values in positive. So, 0 has to be kept for elevation and therefore 0 is a value. And NoData is also a value but it is very awkward value which is unimaginable in a data set and that value is assigned as NoData set. And what happens then? Once we assign that this is my value for NoData that is the information which is outside or insufficient information.

Sometimes within our study area also, NoData will get there; I will be showing examples. So, then this value is recognized by the processing software that whenever they encounter that value, just avoided. Do not consider in any kind of calculation or enhancement or any other thing. So, that is why it is NoData, just ignore it but it has to be kept in the system because overall shape of raster can only either square or rectangle.

(Refer Slide Time: 05:08)

Concept of NoData in Raster

- Cell values can be either positive or negative, integer, or floating point.
- Cells can also have a NoData value to represent the absence of data.
- Sometimes there are homogeneous areas / outside of a boundary in a raster dataset that the you do not want to display.

Now, these cell values you know that either can be positive or negative or also can be integer or floating points, depending if I am handling satellite image then pixel values will be only positive integer value but if I am handling a grid in GIS then I may expect positive values, negative values, I may get also integers or floating point or decimal values. And these cells can also have NoData value if there is NoData or absence of data.

That means because I have to even if data is not there but still overall that raster has to be presented; so, how it will be presented because I cannot assign 0 values. So, I assign NoData value and NoData then can be handled separately through these programs. So sometimes there are homogeneous areas outside of boundary of our area of interest in a raster data set that you do not want to display.

(Refer Slide Time: 06:14)

And therefore also, there we use. See this schematic; this is what the raster data will look. Only thing which you have to assume that units are not rectangular, they are a square but in display you are seeing this one. So, overall shape of this raster can be either a square or rectangular, when we do not have any information like here, I do not have any information. So, these things I can declare as NoData value. I cannot declare as 0 because 0 is considered as a value especially in your digital elevation model. So, this concept was very much required.

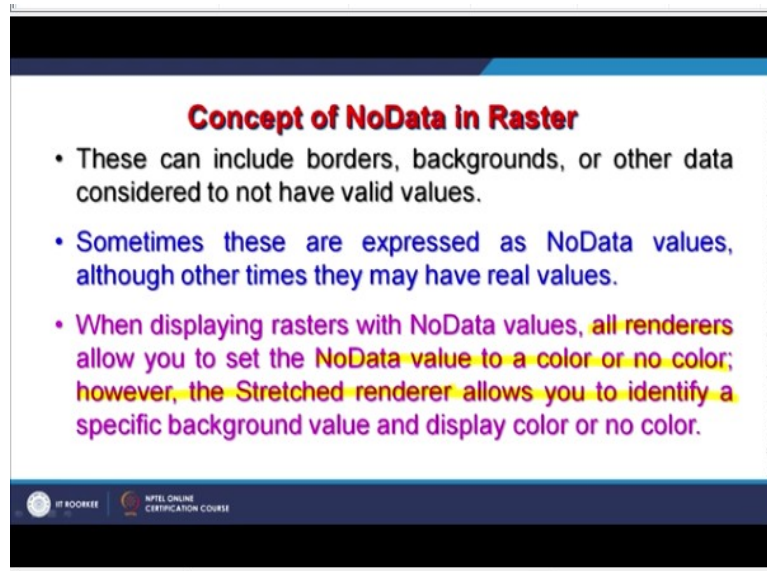**(Refer Slide Time: 06:57)**
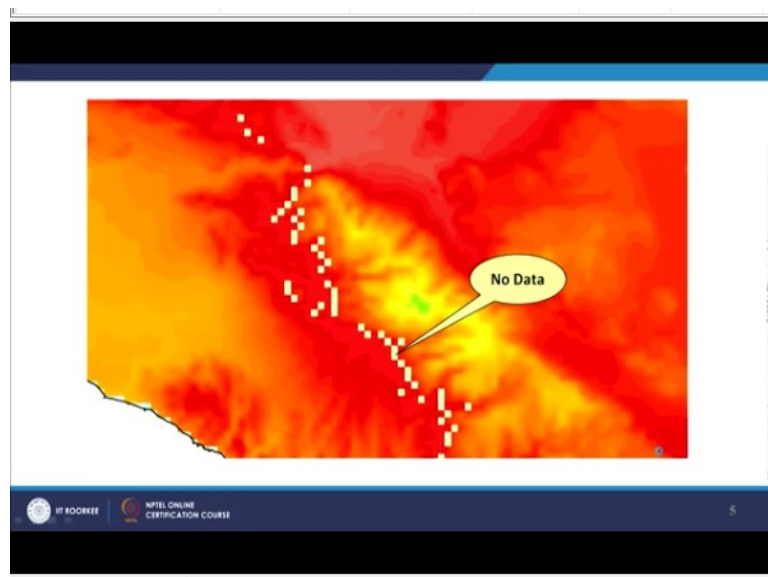
As you can see this is the real data. And this is the example from SRTM DEM; the SRTM stands for Shuttle Radar Topography Mission which was a 16 days mission of NASA to generate a global scale DEM at different spatial resolution. So, when this mission and it was based on SAR interferometry techniques.

When we will discuss about how digital elevations can be developed then we will be discussing this part also that how using SAR interferometry, we can generate digital elevation model. But before that, when this mission was completed then during processing, it was realized that there are various cells in the digital elevation model which are not representing that true elevation values. So, these cells were assigned as NoData value.

And these areas were generally where the snow peak or because of some other errors or reasons during data acquisition, insufficient information was there or absence of data and therefore these cells were assigned as NoData so that the system would know. And whenever there is a calculation, suppose I develop a slope map from this digital elevation model then

during that calculation, these NoData values will be ignored. Now, this concept of NoData also includes these borders because in border areas when we are handling an arbitrary boundary; raster then in border areas in the background or other where we can have the NoData value.
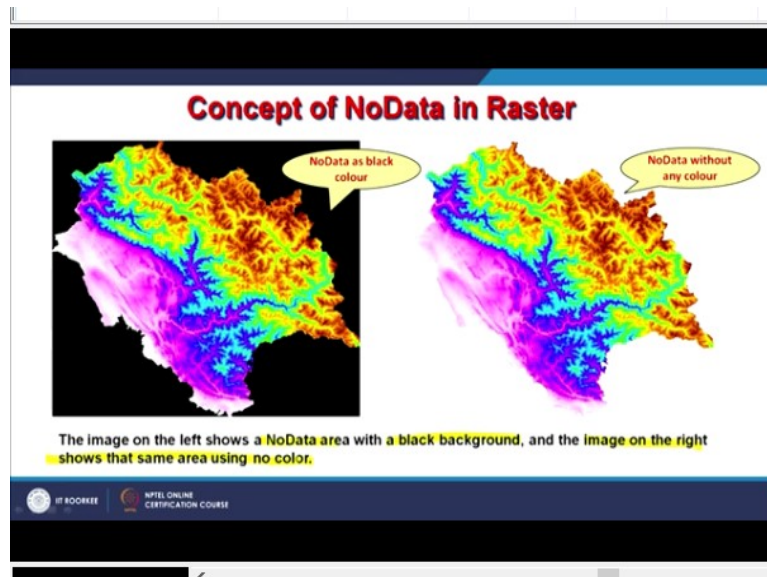
(Refer Slide Time: 08:45)

And sometimes these are expressed as NoData values, although other times they may have real values. So, in early GIS systems when NoData handling capabilities was not there or was not implemented or rather not known, we had great difficulties for areas for outside and we used to do some a completely different way to handle this data. So, when we display raster with NoData values, sometimes the software automatically assigned NoData; they assign no colour basically.

So, it mixes with the background of your screen and therefore you do not see anything outside of your arbitrary boundary. However, when we do the stretching over the data then it has to be remembered by the system or should be identified that these values should not be considered in any of these processing. So, nowadays it is handled very efficiently. Now let us take an example.

(Refer Slide Time: 09:58)

Concept of NoData in Raster

The image on the left shows a NoData area with a black background, and the image on the right shows that same area using no color.

Like here this is a Himachal Pradesh SRTM DEM and using the Himachal Pradesh political boundary, I have extracted from that tile. So, when you extract that data for the outside of Himachal Pradesh because overall raster has to be either square or rectangular. So, what happens to the outside area? So, for outside areas, here we are assigned black colour and therefore you are seeing like this.

If background of my screen is also black then I will not see any problem but if it is not, like white then I will be seeing these black area and black areas are nothing but NoData. So, now as I have said the colours do not matter in GIS. What matters is the value? So, we can change that value and we can say that if value of NoData is like this, assigned white colour.

So, it mixes with background of my screen like on the right side which you are seeing. So here, as a user or whoever is visualizing such data set will not experience that black areas and will not feel that it is either rectangular or square; he would only see the arbitrary boundary. So, the image as you saw that NoData area is assigned black colour and whereas the image on the right the same without no colour or white colour.

So, both options are available if background of your screen is white. Generally, we assigned white colours because it has to be printed on white paper then again there will not be printing like left image, it would be printing like right image and no black colour will be used. So, it looks more realistic also.

(Refer Slide Time: 11:53)

NoData examples in ArcGIS

- For example, NoData is stored natively inside the Esri Grid format.
- However, this concept of pixels with value-based NoData is foreign to most other raster formats.
- Because ArcGIS treats nongrid raster formats the same as grids, NoData is allowed in any supported raster format.

Now, how we handle in different software's, example I am taking from ArcGIS. NoData as you know is stored natively inside the grid format. And not only in this grid format but most of the grid formats which are being used in GIS. all can handle very well nowadays. So, this concept of pixels in case of satellite images or cell values in case of a digital elevation model, that values based on NoData is foreign means it is unknown for other raster formats.

So, sometimes if a software has not very well designed to handle NoData then you would have real difficulties. Initially we used to have these problems but now good software's, all of them have implemented concept of NoData. Now as you know that like ArcGIS treats non grid raster formats the same as the grid and NoData then is allowed and supported format.

(Refer Slide Time: 13:04)

NoData examples in ArcGIS

- Changing a raster's geographic location often involves rotation, and when a raster is rotated, wedge-shaped areas are created.
- The wedges are not filled with data and are unknown, so these areas become NoData.
- With nongrid formats, the areas of NoData are stored in the auxiliary file (.aux) next to the raster itself.
- Generally, the default values of NoData are displayed as transparent.

So, this is very standard practice that if data is coming from another format, still it will handle if it is having NoData, it will handle as NoData. And changing a location in the raster like geo referencing, when we do the geo referencing and after resembling when there is a change in the data means there is a change in the scale, there is a rotation and there is a warping, or it is fitting the surface.

Then there will be on the sides, there might be some background or wedges which might be created which will be bringing this NoData zones or NoData area. So, by default if we assign no colour then it will mix with the background and there will not be any issue. So, wedges on the sides of an image or raster are not filled with data and they are basically unknown values.

So, they become automatically the NoData. And with you know like non grid formats, maybe the satellite images or other data, the areas of NoData are stored in auxiliary file that is .aux next to the raster itself. So, when we store on our hard disk or any storage devices then you would see many times that a file having the same initial part but extension is aux. So, here the information about NoData is stored.

So, whenever that file is used, suppose I am having a file, India.grid then I am will be also having a file India.aux and whenever I will use India.grid then this file will also be used to help the system to ignore NoData for any kind of calculations. And generally, the default values for NoData are displayed as transparent as I have said.

(Refer Slide Time: 15:00)
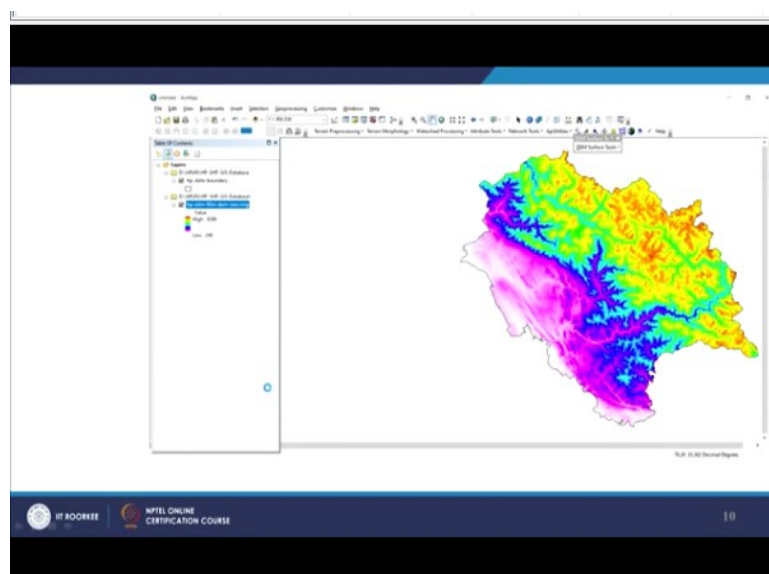
Now here, this is again example from ArcGIS and what you are seeing here that the same digital elevation model of Himachal Pradesh is shown here. And here the point which is important to note is that what it says display NoData has no colour. So, here no colour has been assigned that means whatever the background of the screen, it will mix with that. Though the overall shape of this data is rectangular. So, remaining areas are having NoData values. And which you can also check that what the NoData value each layer is carrying, that we will also see little later.

(Refer Slide Time: 15:50)
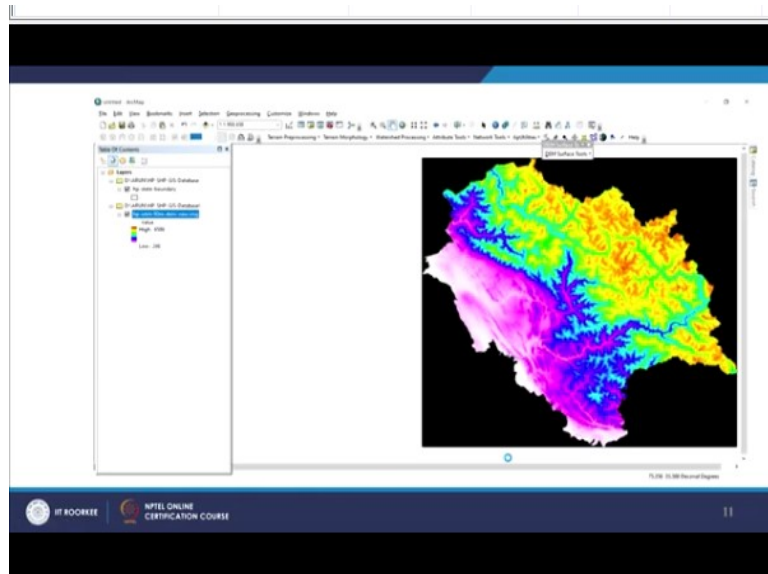
Now here, if I use some colour instead of no colour then this is what the result is that I see the background. It does not look nice either. So here for the NoData value, a black colour has been assigned and this is what the result on the right side you see.

(Refer Slide Time: 16:11)

So, when basically doing any kind of calculation or statistics operations on raster data set then this NoData are ignored. Any cell if it is carrying NoData value that is A node. And when performing operations on raster data which is containing NoData values, there are typically 3 ways that NoData is treated for each cell. NoData can be returned for the location no matter what. And NoData is ignored and the value is computed using any available values. And the third one is a value must be estimated and NoData cannot be returned.

(Refer Slide Time: 16:55)
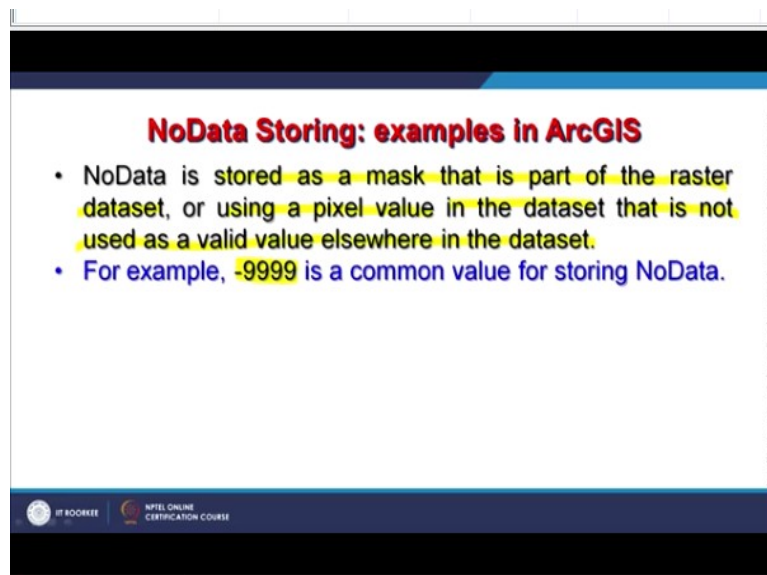
So, likewise the calculation is done or software's are designed to handle NoData. Now, as you know that NoData like in Himachal Pradesh example that NoData is stored as a mask that is the part of my raster data set in a .aux file in ArcGIS format. And using a cell value or pixel value in my data set that is not used as a valid value elsewhere in the data set so that means in every calculation, it is ignored.

For example, I may be having NoData value like this, -9999. Now I am having a digital elevation model or part of Himalaya and I have extracted for Himachal Pradesh like I have shown. Now I know that there in real digital elevation model or in Western Himalayan terrain, there cannot be any elevation value which would be -999. It is impossible and therefore for the system, it is declared that this is that value and whenever it is encountered, please consider as NoData and ignore in all your calculations.

So, basically this is how the NoData is handled. It is very-2 important to understand the concept and utilize whenever we require because during display sometimes, we may encounter problem and the background might be coming black and which is giving a

completely wrong presentation. So, in order to avoid that one, NoData can we assign either no colour or the colour of the background of your screen and you can get rid of outside area of interest.

So, this brings to end of this discussion. Before I close, I would urge you that try to use these whatever the discussion which I am having and giving example of ArcGIS. If you do not have access to ArcGIS or ArcView, ArcView also handles NoData without much problem or ERDAS software also but if you do not have access to these softwares but you can have definitely access to QGIS which is an open-source software and try all these things explore that how QGIS handles NoData. With this, I end this discussion. Thank you very much.