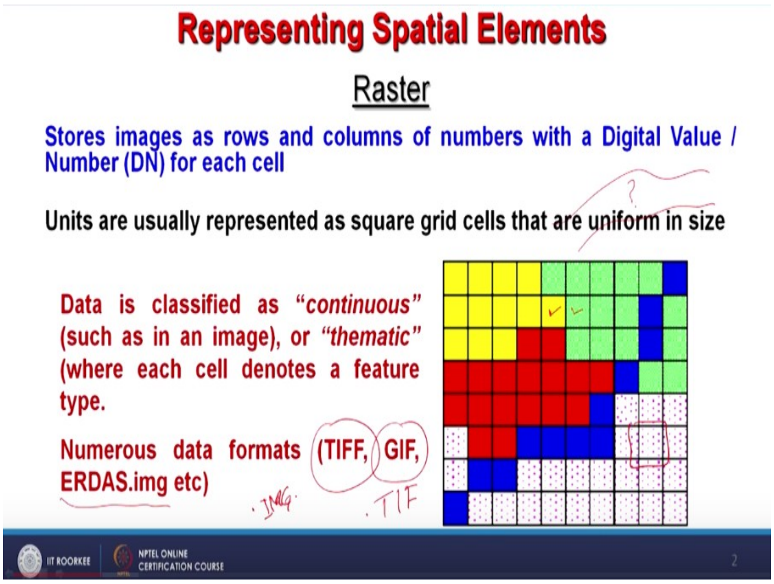
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology – Roorkee

Lecture - 06
Raster Data Model and Comparisons with Vector

Hello everyone! And welcome to new discussion. In this lecture, we are going to discuss about raster data model. And we will also try to compare with vector. We will see that what are the advantages and disadvantages associated with both types of models. As you recall that when we have discussed the data in GIS, at that time I said that there are 2 basic types; one is spatial data and other one is non spatial data.

And in the spatial data; I mentioned first the vector- 3 types of vector entities; point, line/polyline and polygon and which we have discussed. We have also seen demonstration also. Topology part; we have also discussed. And next spatial data was which I mentioned in that one was the raster data. So raster data; again having 2 types that we will also see and then try to compare with vector data.

(Refer Slide Time: 01:25)



Now, as you know that raster data is little simpler compared to what a vector data is. And here you are having basically rows and columns. So, mathematically if we look this one as a 2 dimensional matrix. And one thing also very important like in matrix that number of rows and columns need not to be the same. That means if they are same then we will have a squared matrix.

And if these are different rows and columns then overall shape of our matrix is going to be rectangular. So, same with the raster data that rows and columns need not to be the same. However, the cell which is the unit of a raster has to be always square in shape. No other shape is allowed in raster data because each cell is representing a ground area and both x and y of cell is also same.

So if I say that though unit is in square in shape, however the overall shape of a raster can be either a square or rectangular. But the unit has to be always square. This one has to remember. This very fundamental and this plays very important role in later analysis and other things. As you also know that raster data is a continuous data. As you can also see here that one cell ends, another starts.

So if I take like this yellow cell here then next adjacent cell is green cell and there is no gap between these 2. But if there are 2 polylines like this or like 2 contours; in between we do not have any information whatsoever. So, that is why we call vector data is a discontinuous data or discrete data. And whereas for raster data we call as a continuous data; one cell or one pixel ends another starts.

So, within that raster file like here which you are seeing a 2 dimensional matrix. Each cell is having its own color or pattern. But basically is a value. And that value is assigned certain pattern or color. So overall in the view, you see like this. So, the overall shape can be square or rectangular like here; it is a rectangular. But the shape of a raster unit like here, it always is square in shape.

Both x-y dimensions will remain same. This is very-2 important from unit points of view because the third type of spatial data which we are also going to discuss later on. Not in this lecture, maybe next lecture about the triangulated irregular network in which the unit is always triangle of varying size and shapes. There are advantages having such kind of model data model to represent the surface.

When I say continuous data or representing a surface; all 3 models- vector, raster and TIN can do it. But this vector is being discontinuous or discrete; only we can use the Iso-heights or contour lines. But between 2 contours; we do not have any information. But there is no information gap in case of raster and neither in case of TIN as you would see later on, when we will discuss about TIN.

So, all 3 data models can represent a surface especially like terrain surface or topography of a terrain can be represented very well. But in case of vector, it would be discrete. In case of raster it is continuous throughout your data file. You are having some information on other. And same would be in case of TIN also. So, this is another very important point related with. As I have already mentioned that unit has to be square in shape in case of raster.

Various formats for image and data are available. Depending sometimes, big software people have developed their own format. Like ArcGIS having their own format which is called GRD- grid format. TIFF- tagged information file format, this one is a very common format for image data. Because raster is having 2 types; one is grid and image. So for image format, TIFF- tagged information file format or in short we also write sometimes only TIF.

Because this was earlier in the MS DOS systems after in the extension space, there were only 3 digits or 3 places we are allowed so that became still popular. Another file format especially for image data is very popular is called GIF- graphic interchange file format. Now, this format is used for animated files. So whatever the files which you see or some cartoon kind of thing on your internet or on your mobile through WhatsApp or other social media platforms, they use this GIF.

And if you want to create, there are tools available by which you can create a GIF file that would be animated file whereas TIFF will not allow you to store animated files. Another file format very famous for raster data, especially the image data is ERDAS imaging or .img. .img is also very popular and compatible with many-2 GIS and digital image processing software.

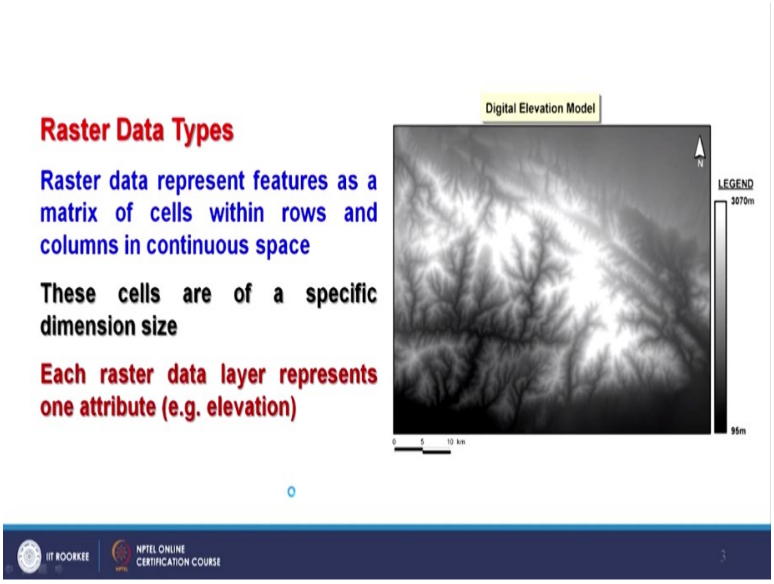
So, not a problem and there are tools available in all good softwares or separate tools are also available where you can change from one format to another. So suppose if I want to change from TIFF to ERDAS imaging, I can do it within ERDAS imaging. I can do it in ArcGIS or separate softwares are available where I can change from one format to another. But one important point I want to mention here that transforming from one format to another, always it is not fully transparent.

Transparent means that there might be introduction of some errors. So, one has to be very careful. Let me give you an example here like if I convert a file which is originally in TIFF to JPEG. That is JPG or in short we call as JPEG. Now this format JPEG format is a compression format. And therefore, it will change the quality or deteriorate the quality of a TIFF file permanently.

So, by mistake if I have deleted original TIFF file, I got converted into the JPEG and later on I want to restore the quality of my image then it becomes impossible. So, that is why I said this the conversion from one format to another, one image format to another or one grid format to another before you delete the original file, one has to be very careful about this whether this transformation from one format to another is fully transparent or not.

If it is not, then please do not delete the original file because that may be a very valuable data for you. So, we have to keep this thing.

(Refer Slide Time: 09:41)



Now raster data types; I have already mentioned 2 main types- your grid and image. Example on the right side you are seeing the grid example and a very common grid data which we use in GIS. And later on in you know in the last 10-15 lectures, we are going to have lot of discussion about digital elevation models and their derivatives. And we will be using the same digital elevation model as shown here.

So, this is grid data as you can see. It is continuous. And here in this case in particular this example; number of rows and columns are not same. Therefore, the overall shape of my raster is rectangular in shape. But data is continuous. There is a question will come about the spatial resolution. And of course, we will have more discussion on this but as the dimension equivalent ground area if it is very small then we call as a high spatial resolution.

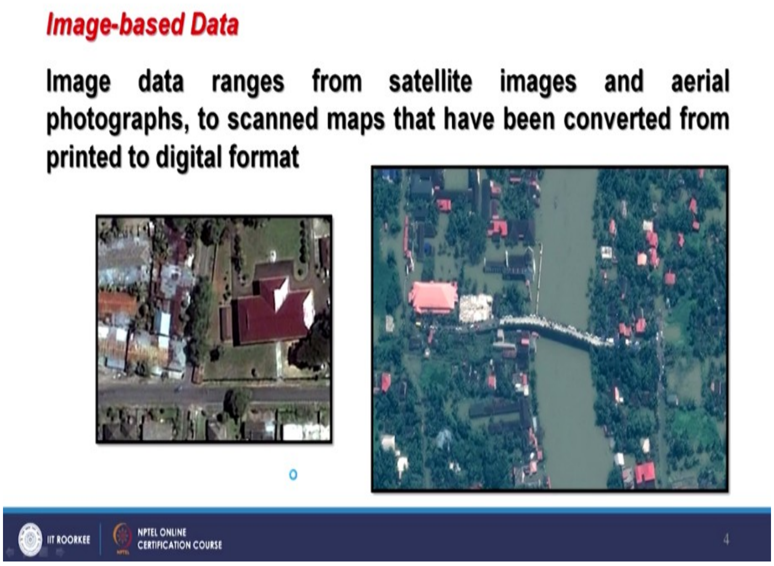
And if ground area coverage is large by one single cell then we call as a coarse resolution. And this in relative terms; for someone, 10 meter maybe higher resolution but for another person 10 meter maybe coarse resolution. So it depends on the work or project in which you are going to imply a particular resolution data. So for those who are working in a large area, say continental scale; even 100 meter resolution or 1 kilometer resolution may be sufficient for them.

But if I am working in a 10 square kilometer watershed or a district then maybe I require very high spatial resolution data. So, this dimension of the cell which is equivalent to ground area decides basically what spatial resolution is. We will have a complete separate discussion also on this. Now as you have seen in case of point data or in our previous discussion related with the demonstration of softwares that each vector entity whether it is a point, polyline or polygon can have theoretically n number of attributes.

But in case of raster each cell can have only single attribute. That is the limitation. Though it is continuous, the limitation is you cannot have multiple attributes associated with each cell, only one value. So if it is a typical elevation model that cell value would be elevation value. But if I have created a surface for say pollution then that would have particular pollutant concentration.

So in a typical digital elevation model, that value is elevation value. Now many times though we may have a surface which is representing pollution or any other thing but still we call as digital elevation model. So do not get confused in that. This is a common term which is used for grid files which is representing surface in a continuous fashion.

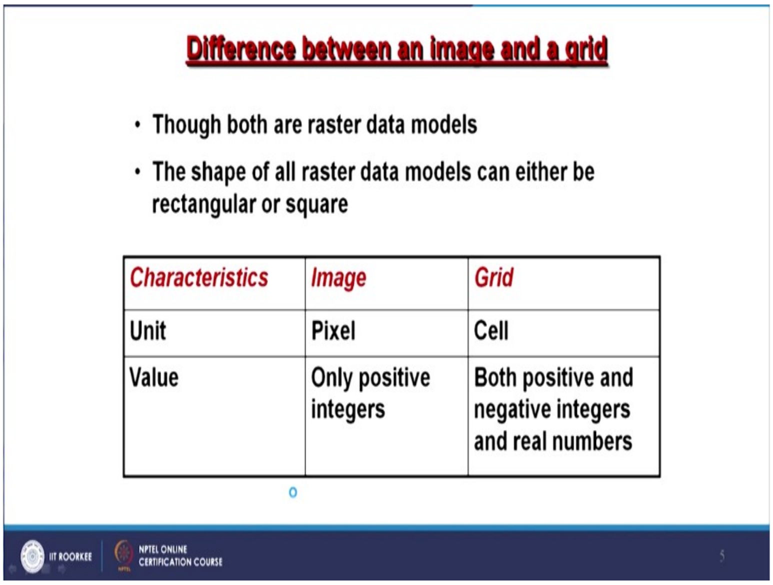
(Refer Slide Time: 13:07)



Another type of raster data is the image data. And this image data which you are seeing; 2 examples of images are here. Both are having rectangular shapes, though they can have also square shape. And one is aerial photographs; the left one and the right one is your satellite image of very high resolution. As you can see that all details are available here and almost these 2 images are having almost the same colors as you see in the real world also.

Because of very narrow bands, it is possible now to get images almost in the real color. If we are not using infrared or other channels then it is possible. So, your raster grid will look something like this and your raster image may look like this. Both are raster. Here the cell instead of calling cell, we call as a pixel which is a short form of picture element. And the cell values; how they are stored in the system are also little different which we will discuss later.

(Refer Slide Time: 14:33)



So, the major difference between an image and grid is; though both are raster and shape of raster model can be either rectangular or square. However the units in case of image, we call as a pixel- picture element and whereas unit for a grid, we call as a cell. So whenever I say cell, I mean I am talking about the grid. And whenever I say pixel, I am talking about an image.

So, that is why these 2 different terms are used for units of these 2 different raster datasets. Another thing is the value. What should be the value of pixel and cell in a grid? So the value is always in case of image or in case of a pixel value; it is always positive integer value. So if I recall the discussion which we had related with our TIFF file; so typical you know, TIFF file will always have values in whole numbers; in integers.

If it is suppose a TIFF file is 8 bit file then the pixel values can be 0 to 255. Maximum variations I can have only 256 and all values our whole number. So, that is why we say positive integers. We cannot store in case of image, any negative value or decimal or floating point values. So, the values have to be positive integer values only in case of image pixel.

However in case of grid; the cell value can be either positive or negative and also can be either an integer or real number. That means all kinds of values are possible. So if I go back to a slide here and about this digital elevation model or a grid; now someone might be working in an area which is land part and sea part also or beach area. So, there I will have some positive values above mean sea level, I may have surface below mean below the water.

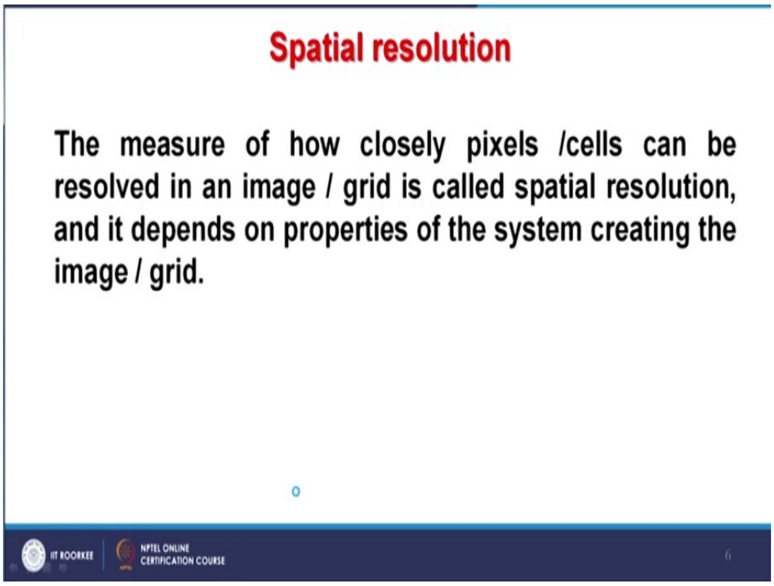
So, minus values also are required to be stored. But in case of image, all values will be positive values because these values are not elevation values. These values are you know either reflecting or reflectance values or emitted energy values. So, these values cannot be in decimal. So that is major difference and whenever you store in TIFF format, TIFFs will force the cell pixel values to become integer values.

So if you convert a grid to TIFF, you have to be very careful that the precision and the flexibility for cell values which is available in case of grid is not available in case of image. Though overall raster data is much easier to handle on either digital image processing or on GIS platform because of its simplicity but image especially of one of the raster is much easier as compared to grid.

Because of it handles only the positive integer values. And the life becomes much easier for processing person. So, these 2 major differences are there between image and grid. Otherwise concept wise both are 2 dimensional matrixes. And again if these are mathematically 2 dimensional matrixes, so all tools for analysis of matrix which have been developed in mathematical domain through computer science, they have come into the GIS.

So, all these digital image processing and GIS raster data handling; all this has percolated into GIS. And therefore GIS is becoming further powerful. Any development which takes place in related with 2 dimensional matrix is beneficial in GIS also.

(Refer Slide Time: 19:02)



Now, I mentioned about the spatial resolution. So, this is a common definition for image as well as for grid. That the measure of how closely pixels or cells can be resolved in an image. That means how I can identify different objects in 2 adjacent cells or pixels. And it depends on the basically properties of the system which is creating image or grid.

Here in case of a grid if 2 cells are having a large difference in the values then they will come very distinctly whereas if they are having almost similar values then you will not see. So, these resolving things will be a different problem there. So, higher the spatial resolution; when we say higher the spatial resolution that means that we are talking very small size of the cell or pixel in comparison to ground area.

If ground area by cell or pixel is being covered very small then we call as a high spatial resolution. For example if a pixel is representing 1m*1m ground area whereas in another image a pixel is representing 10m*10m area then in relative sense, I would call the 1 meter is high spatial resolution and 10 meter is a coarse spatial resolution.

Poor spatial resolution, I should not call because 10 meter if I compare with 100 meter then 10 meter is better. So, relatively is a coarser resolution if I compare with 1 meter. And higher the spatial resolution, I can distinguish different adjacent features more clearly. That means the resolving power improves. And it is only possible if I have the better spatial resolution data.

So, this is very important because people talk about spatial resolution. Especially nowadays when anybody who goes for buying a new mobile always you know, earlier we used to have 5 megapixel camera then 12 mega, 15, 16, 18. Now, people are talking about 48 megapixels camera. So that means the number of rows and columns if you multiply that will come somewhere about 48 megapixels.

So a one image taken by a 48 megapixel camera would have these number of cells present there or pixels present there. So, higher spatial resolution sometimes may not be required for particular application. So, it is not necessary that we should always go for higher and higher spatial resolution. It depends on our requirements. It depends on what kind of project I am doing.

Because higher the spatial resolution, more the storage it will require and more the processing time it would require. Unnecessary it may occupy my hard disk space and even my RAM, though I do not require that high spaces. So, optimization in a spatial resolution must be thought before we imply a very high spatial resolution data. And that you can decide based on at what scale, you are going to produce your results.

For example I may be using a 10 meter spatial resolution satellite image but overall if I am going to create output on A4 size for a district level analysis then my scale in ultimate printout map is going to be very small. So, why do use 10 meter spatial resolution data if I cannot show it to anyone. So optimization has to be always thought about this. For another way, we say higher spatial resolution means clearer image.

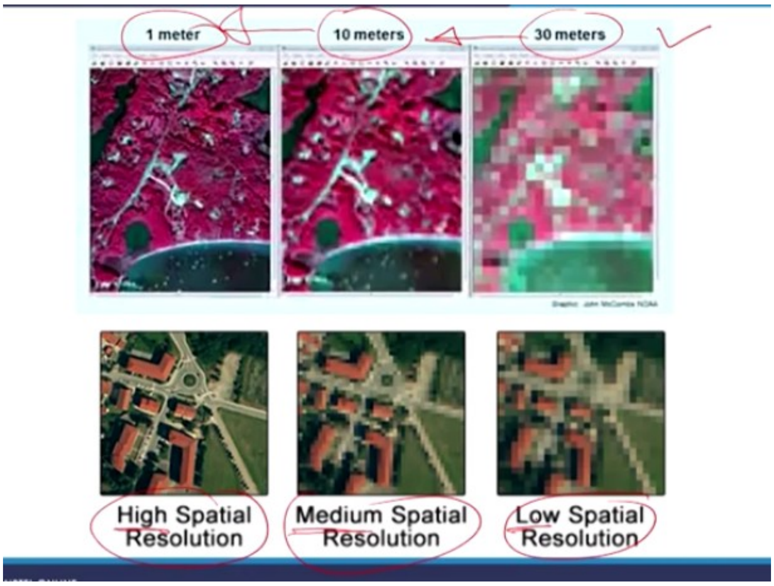
And you know this is how in case of image or grid. Now, one another very interesting thing may not be directly related with GIS but we sometimes you know handles that format also, I have already mentioned that is about JPEG. And we will have a separate discussion about data compression where this JPEG reference will come again. Now JPEG is a not a constructive image file format.

That means it compromises on the quality of an image. However it provides better compression. However there are various image compression tools which are available which are non-destructive. Non destructive means I can restore to original image quality but JPEG is not. Now what people are committing mistakes? Even they are buying a camera or mobile having 48 megapixel capabilities but the same time their image file is being stored in JPEG format.

So, you are employing a very high resolution camera but at the same time, you are destroying the image quality in order to save the space of memory. So, one has to be very careful. If you are really looking for very high quality images; change the option, store in TIFF which is very common format. And there are compression tools available to you compress your TIFF files which are non-destructive in character.

And you can still compress it not to the level as JPEG provides but you can anytime restore to the original image quality. But in case of JPEG, an image was acquired by a 48 megapixel camera and originally it has been stored in JPEG. Now, no image processing techniques can ever improve the quality of an image. So this is very important here.

(Refer Slide Time: 25:36)



Just to continue on this spatial resolution thing, I am having few examples in 2 rows. On the top right corner, we are having a spatial resolution of 30 meters. And you can see that the image is having, people call as a pixelized or jiggered image. And each cell or each pixel here is clearly visible. And it is relatively difficult to identify different objects which are present in this image.

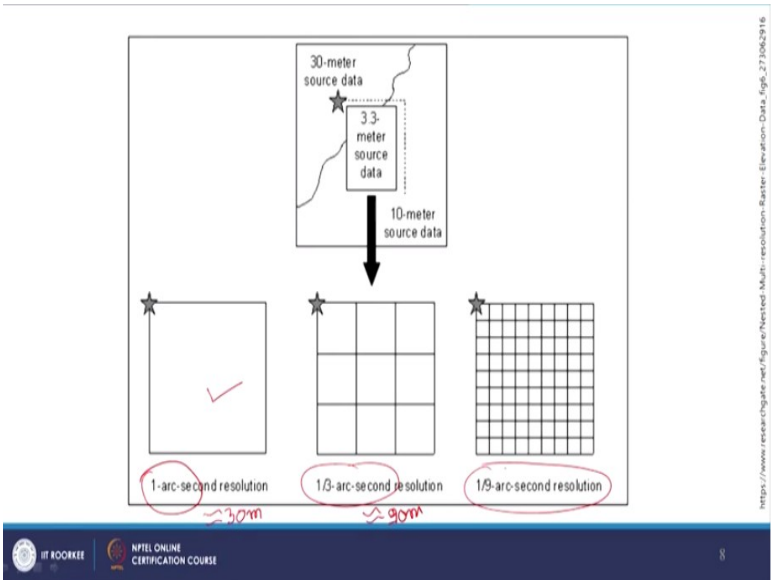
But if I move from 30 meter spatial resolution to more higher spatial resolution that is 10 meter, now things have become quite interpretable. But if still my application requires much higher resolution, I can also resort to a resolution which is 1 meter. And now things are really very clearly. So, the resolving power is much-2 higher in case of 1 meter as compared to 30 meter.

Similarly here, again low resolution image, medium resolution image, high resolution. These are relative terms. So, whenever we talk about spatial resolution, we always talk in relative terms. Otherwise none spatial resolution is the highest and none is the poorest. Because if I talk about remote sensing images, now you know, even 30 centimeter spatial resolution satellite images are available.

And the same time people are extensively using satellite images which are having 1 kilometer spatial resolution. So, 1000 meter versus 0.3 meter. All for different applications; different spatial resolution data are implied. So, optimization has to be there. Thinking is required before I imply that.

Though the data may be available to us as a higher spatial resolution but if my output does not require, why to imply, why to use it; use lower one because ultimately I am going to produce say 50,000 scale that does not require because in that I will not have much details available so why to do it and then reduce the size.

(Refer Slide Time: 28:09)



Now here for you know, for further discussion on this; sometimes you get the data in this kind of wordings- 1-arc second or 1/3-arc second, 1/9-arc second, 3-arc second. And what does it mean in terms of meters and other things? So, what happens? That lot of digital elevation model data i.e. DEM are available in these terminology. So, when they say 1-arc second, it basically means they are talking about roughly 30 meter.

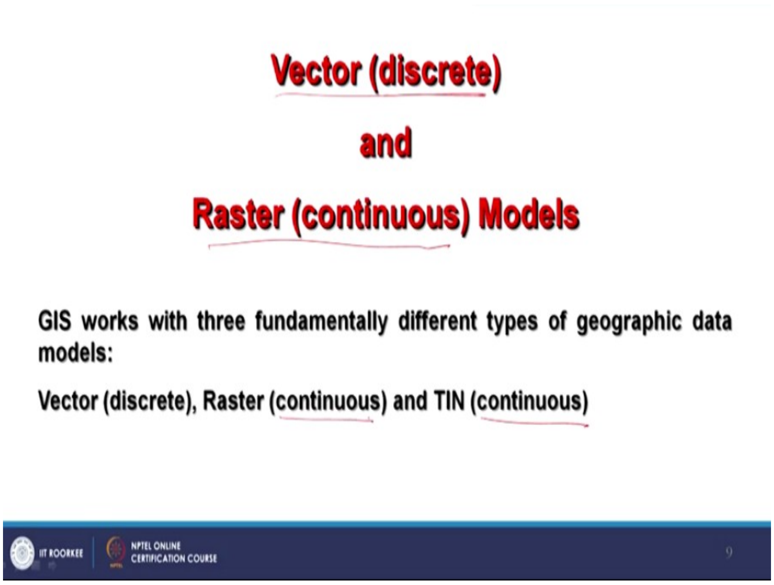
And when they say 3-arc seconds then they are talking roughly about talking about 90 meter. And how this arc second is coming? Basically it is in terms of latitude, longitude because the earth is a 3 dimensional body. So, these are arcs not like straight lines; though they are in maps, they are shown as arc. So that way also, we would understand what is the meaning of 1-arc second or 3-arc second.

Most of these digital elevation models are available in these terminologies. They use this terminology- 1-arc second, 3-arc seconds or 1/3-arc second depending on what kind of data is available. So as you can see that here, this is the data in 1-arc second. 1 cell will represent this much ground area but if I go for 1/3-arc second then I am now having for the same area in 9 cells.

And if I go for 1/9-arc seconds, I am having many-2 cells to represent the same area. And the area here maybe of like 1-arc second: only the 30m* 30m area. So likewise instead of simple meters sometimes it is also denoted in arc seconds. Now in this later part of this discussion, I would like to spend some time on this comparison. Indirectly we have been comparing.

But now we will have a complete dedicated discussion about the vector and raster data. How they are different? What are the advantages and disadvantages?

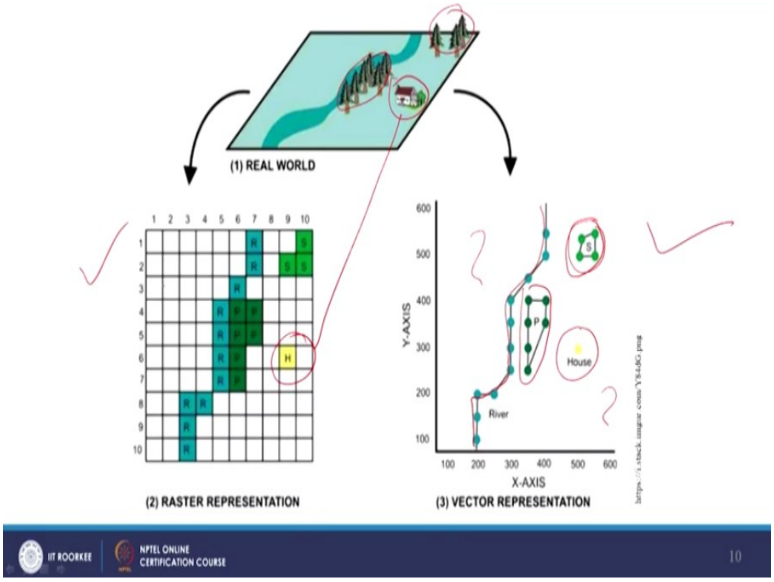
(Refer Slide Time: 30:48)



As you can see, we have already discussed that vector is a discrete data and whereas raster is a continuous data models. So, that is one major difference is there. And the good GIS softwares are capable of handling vector, raster and TIN and of course, non-spatial data as well. So, there is no issue about the data. But GIS handles these 3 data sets completely differently.

But, we currently we are going to discuss vector versus raster. Now TIN is also considered as continuous as raster. But the concept wise, it is different.

(Refer Slide Time: 31:22)



So if a real world has to be represented in these 2 different models; one is raster, another one is vector, how do would we represent it? So real world depending on the scale like house here in the vector data, it is represented as a point. A stream river though, it is having width in real world but in this particular scale, it is represented as a polyline which you can see here.

And a forest like here which is shown in the form of tree, it is represented in the form of polygon and same with this. So, all 3 vector entities can be used to represent the real world. And of course, when we model the things from real world to vector; this is discrete and this is abstract reality like river which have so smooth curves and width, all have gone except some lines.

So, it is a reduced version of the real world; abstract version of the real world. If I have to represent the same real world through raster model then the representation is going to be completely different. For example here for the house; now this house is occupying an area because if say resolution is 1m*1m or 2m*2m then I am 2 meter resolution then this house is being represented as a 4 m^2 area.

But as per definition of vector; point is zero dimensional entity. But this cell in this grid may not be representing the true area of the house; the constructed part of the house. So that means here also, this is not a true representation of the real world. In earlier discussion, I have already mentioned which I am going to repeat; real world is really real in one to one scale and which is cannot be represented by any models in true fashion.

Whatever the models which we are using are abstract version of reality and therefore errors will keep come in our data. So like comparing a point; a house being represented as a point data zero dimensional entity in vector whereas in raster, a cell will occupy some space. Similarly if I compare the row this stream river then here it is being represented but not as smoothly even in case of vector.

And it is completely in a stair steps case or jagged representation. And same with my forest; like here it looks little smoother but here it is just 3 cells which are representing them. So if I am having relatively coarse resolution data, my error component representing the real world would be very large.

But if I go for very high spatial resolution data in relative sense then my error component in terms of area and the quality; my error is going to be less. That is also true in case of high spatial resolution data. But if it is required then only and if that margin is available for errors, it is fine to even use medium resolution or coarse resolution data.

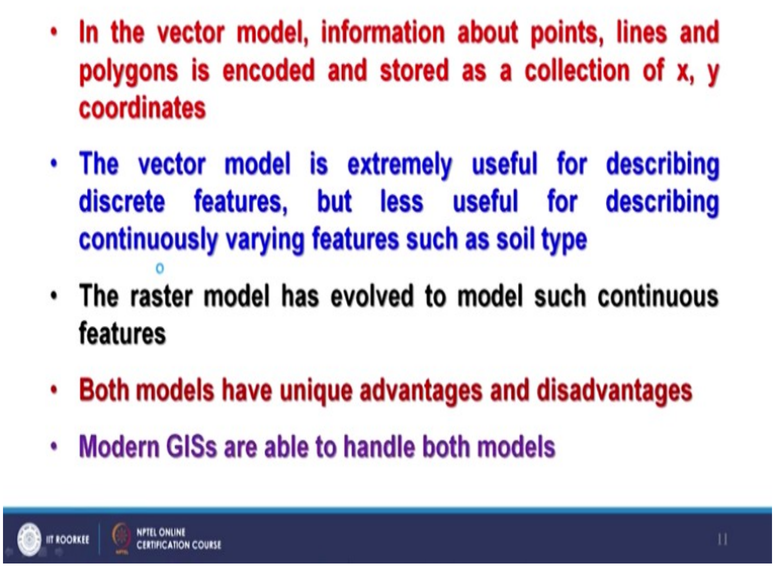
So, real world can be represented in 3 different spatial models; vector, raster and of course, later on in future lectures we will be also discussing TIN. You can see here that vector is a discrete and like in this area, you do not have any information. Though in real world, there is a land; that maybe agricultural land, waste land or any other. But there is a land. But in case of raster, each cell will have some value.

And that is why it is a continuous; no gap in the data. One cell ends and other starts. One pixel ends, another starts. No gap in between. So in vector data model, information about point, lines and polygons is encoded and stored as a collection of x, y coordinates. So for a point, you are having just one single pair x and y. For a polyline, you will have many and same with the polygon.

And vector model is extremely useful describing discrete features but relatively less useful for describing continuous. So if I have to describe a soil type as you know that in nature, the soils will not have a very sharp boundary. And therefore if I want to use that soil type layer or theme then it is better to represent the soil map through a continuous model that is raster rather than vector.

So, it depends on what kind of data I am going to handle, what kind of theme I am going to handle.

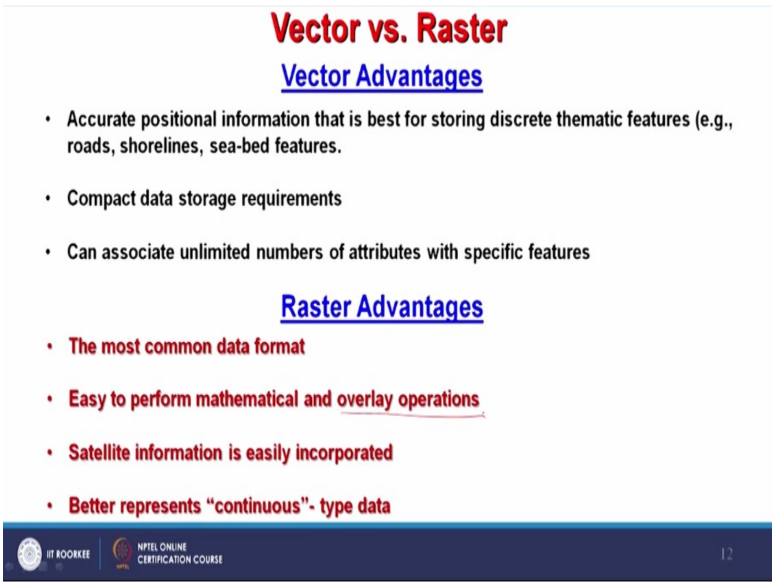
(Refer Slide Time: 37:00)



A raster model has evolved to model such continuous features. Terrain; another one, if I want to represent topography of a hilly terrain like part of Himalaya then vector in vector I will be using contours and it is not very easy. And therefore in order to represent terrain; variations of undulations and ruggedness then the best model to represent either TIN or raster.

And both models that mean vector and raster are having merits and demerits. That means they are having unique advantages and disadvantages. And all these models including your TIN can be handled with all modern GIS softwares.

(Refer Slide Time: 37:42)



So, very quickly we will go through these advantages and disadvantages. Like vector advantage is that position wise, it is quite accurate representation. But in case of raster that kind of accuracy may not be available. You have seen in case of house or forest or river stream or whatever. So, accurate positional information if one is required then go for vector representation.

Of course vector does not require much space. Memory requirements are very little compared to raster. Therefore, it is compact. Processing is compact also. And probably you do not require any compression techniques because there will not be any redundancy. Redundancy means unnecessary space, it will not occupy on your hard disk. So, redundancy is not present at all in vector.

And especially if you have built up a topology then the quality of vector data is going to be very high without any redundancy and you do not require compression tools either. That is why there are hardly any development or compression techniques for vector data. And vector data, you can associate theoretically unlimited number of attributes. Whether it is a point, line or polygon; each point can have hundreds of attributes.

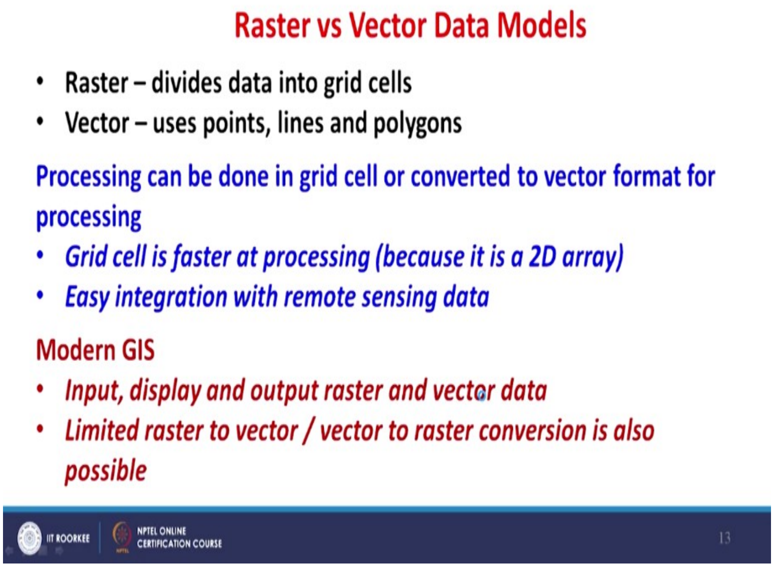
I showed example of those Himachal Pradesh hydro potential sites. I had about 20 attributes associated with each point data. But in case of raster, only one attribute is possible. That is the biggest limitation. So far, this is what it has been developed. Maybe in future, we do not know. But different layers you can have for the same. So, one raster layer you are having for elevation, another one you are having for pollution.

Third one is you are having for depth or any other thing. You can still have multiple attributes but it through different layers. But in case of vector, in single layer you can have many n numbers of attributes theoretically. Since raster is based on 2 dimensional matrix concepts and processing therefore mathematical and overlay operations are very-2 easy compared to vector data.

I will be showing an example where vector will find a limitation but raster does not have a limitation over overlay operations. And other is the satellite information is easily incorporated because if I am using grid, I want to also use satellite data. Both are raster; no issue, very compatibility is there. But if I am using vector data and I want to use satellite images then I may not have that kind of compatibility.

So for any kind of continuous phenomena, one should always resort to the raster rather than vector.

(Refer Slide Time: 40:58)

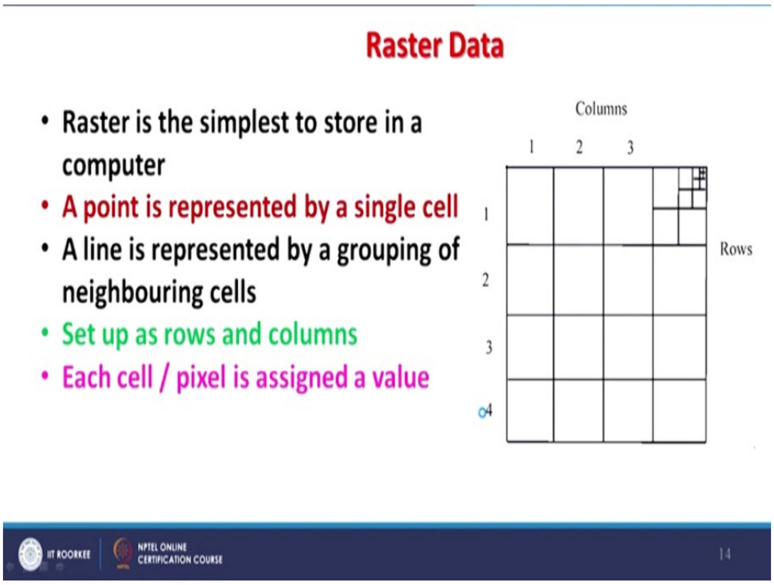


Further on discussion about the comparison of these 2 raster versus vector. Raster divides data into grid cells you know. Vector uses point lines and polygons. And processing can be done in grid cell or converted to vector format. That means the conversion from one format to another is also possible but it is not fully transparent. That means if you convert from raster to vector, you will introduce some errors.

And if you convert from vector to raster, again you will introduce because both models conceptually are completely different. I will discuss little bit about how you can convert from one model to another model. So the processing part is quite fast in case of raster if it is not very big dataset or a 2D array and easy integration with remote sensing data. Modern GIS; they all allow and support, handling of raster and vector data.

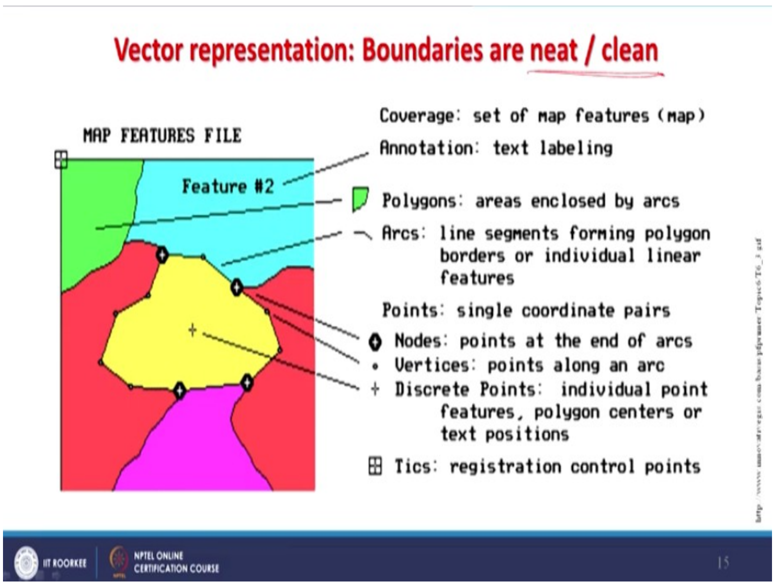
This conversion is quite limited. And like raster to vector conversion, it is not fully automatic yet. It is still semi automatic or you can call as a manual. But vector to raster, automatically it can be done.

(Refer Slide Time: 42:18)



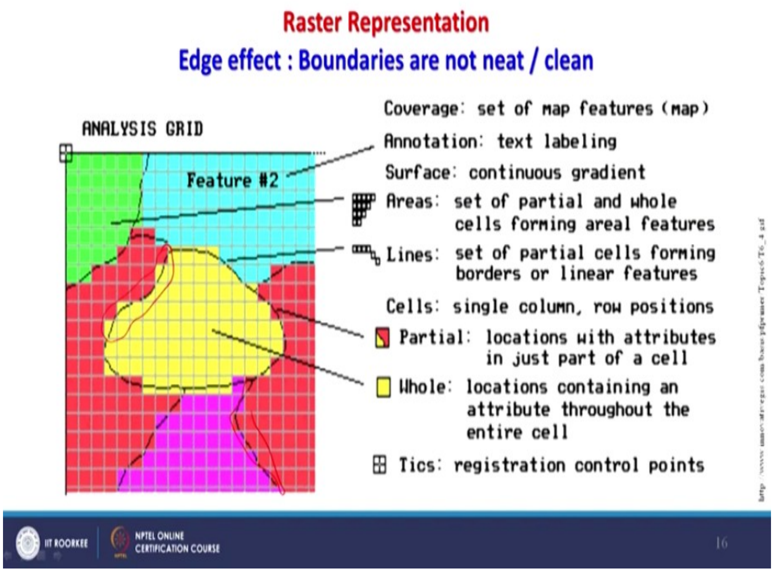
Raster is simplest to store in the computer. Point is represented by a single cell as you know. And a line representing by a group of neighboring cells and set up as rows and columns in case of raster. And each cell pixel is assigned a value. So, unit in case of vector can be point, line and polygon. But ultimate is the coordinate; x and y pair. But in case of grid or raster, you are having either cell or pixel as unit.

(Refer Slide Time: 42:51)



High spatial resolution; cells representing a smaller ground area will require more space to store and more processing time. Now from a visible quality point of view, how boundaries are represented better in which model. So if I see a vector map then I find that the boundaries are very well represented by a map in through vector representation. But if the same map has to be represented in this raster format then this is how.

(Refer Slide Time: 43:26)



So, you would notice that major problem is all along your boundary line. Your boundaries are not as smooth or clean as in case of your vector data. Because vector is you know, capable of representing very accurately the boundaries whereas raster will have a jagged kind of issues here. So, the boundaries are stair steps case or a jagged thing is there. Boundaries are not smooth.

However if we improve the spatial resolution or get high spatial resolution data then this much of error, you will not notice.

(Refer Slide Time: 44:13)

Sl. No.	Characteristic	Vector Structure	Raster Structure
1.	*Data structure*	Complex	Simple
2.	*Ease of learning*	Difficult – software is complex	Easy - functions tend to be more intuitive than in vector
3.	*Positional precision*	Can be very precise and thus accurate	Precision increased with increased processing time and data storage needs- accuracy. Limited by pixel size
4.	*Attribute precision*	Good for polygon, point and line data; not good for continuous data unless connected to TIN or similar technology	Good for continuous data; limited by size of pixels in representing attribute distribution in real world
5.	*Comprehensive ness of analysis capability*	Good for spatial query and relatively simple data, analysis-limited to Intersections	Not good for spatial query but very good for spatial analysis filtering, and modeling

So in a complete comparison through table, we will go one by one. These are the characteristics; vector data structure, raster data structure. And in summary if we go raster data structure is simple compared to that vector is complex. If I have to learn then of course, raster is simpler. Because it suggests 2 dimensional matrix; 2D array whereas our vector sometimes, it's difficult and softwares are also complex in that case.

If I talk about positional accuracy then more precision and accurate is the vector data whereas precision increased only if I increase the spatial resolution in case of raster. And attribute precision again; there are lots of tools I can store. And attribute precision in case of vector is relatively higher than what you are having in case of raster.

Now comprehensiveness of analysis capabilities, good for spatial query; I showed how to build a query a simple or nested query on a point data or any other vector data. And you can do it quite easily. However if I have to do the query on a spatial data, it is not that good capability currently of GIS software. So, that is one of the limitations of raster data.

(Refer Slide Time: 45:50)

Sl. No.	Characteristic	Vector Structure	Raster Structure
6.	*Overlay ability*	Good, but overlaying many layers can cause many splinters, etc. in the result which are difficult to eliminate	Because all pixels line up, overlay procedures do not create problems
7.	*Storage requirements*	Relatively small but complex	Relatively large and simple but may be complex
8.	*Ability to work with image data*	Poor - data must be vectorized first	Good - uses same kind of data structure
9.	*Conversion to other map projections*	Usually included in package and relatively simple to do	Difficult and quite often creates warped images which do not fill the raster, causing problems with neighborhood functions
10.	*Ability to work with network data structures*	Good - because system can handle lines	Poor - raster structure not amenable to network
11.	*Cost*	Expensive	Inexpensive
12.	*Output map quality*	Very good - looks like a map	Poor - doesn't look like a map to lay people

Now if I go for overlay ability then vector will provide overlay ability quite good but up to certain layers. And if you increase number of layers in overlay analysis, like A union B union C, D then your output would become very complicated. But in case of raster, all pixels or cells are lined up; though they may have different spatial resolutions but still overlay can be performed on tens of 20 raster layers in one go.

So, that is very big advantage with raster data structure. Storage requirements, we have already discussed the vector requires small but complex. Complex means here many files will be created as we have seen while discussing the topology. But in case of raster, sometimes even a single file will store the entire raster layer without any problem. Why I said sometimes?

Because sometimes there will be mainly one layer or one data like .TIFF file or sometimes you also have the compression information and other things. So 2 more files might be there. So, that is why I said that number of files in case of raster on a computer system against one layer is going to be less as compared to the vector. Now ability to work with image data; of course image is part of the raster, one type of raster so fully compatible with raster but in case of vector, it is very poor.

Conversion to other maps projections because as you know that each country is having their own map projection and everyone would like to represent their country in a true shape and size. So, there are hundreds of such projections have been implemented into the GIS. But all of them are not fully transparent again.

So if I have to convert my data from one projection to another especially the vector data then generally it is quite smooth. I can do it very easily. Most of the GIS software will provide tools to convert from one form one map projection to another. But in case of raster, it is having a little difficulty to convert from one format to another because their boundary areas; margins will have big problems there.

Now another characteristic is ability to work with network data structures; raster is very poor because it is continuous data whereas a vector is good because like stream network or power grid network, it is very good. Remember also, the first GIS software was also developed for handling network data. And therefore since early 60s till today, lot of development has taken place on this front.

So, most of the GIS softwares are very good to handle the network data; though handling the data, preparing the data for network analysis may be little challenging and difficult. But once you have developed a layer which is having complete network build up then there will not be an issue. Because there might be a flow direction maybe in two directions or single direction like case in case of natural drainage system, a stream network will have only one direction flow.

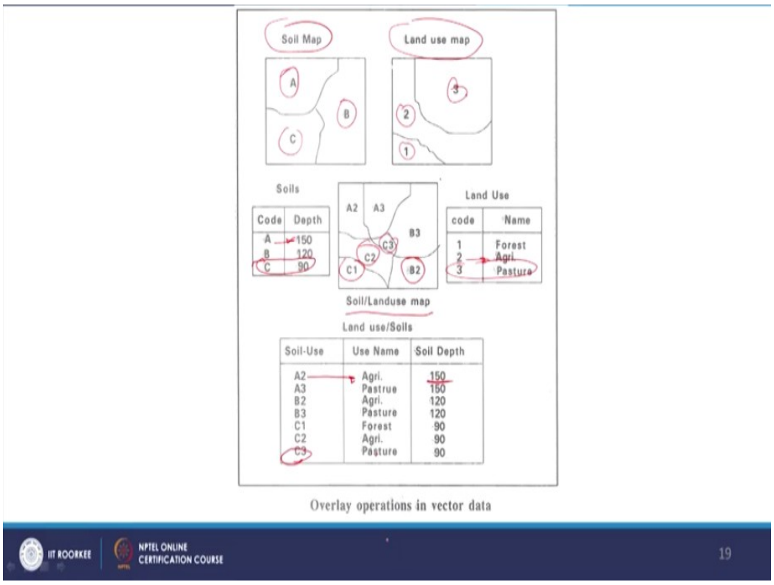
And therefore developing such network and analyzing you know, flow conditions or pollutant flow through these networks is easier. But if it is a 2 way flow like in road or other systems then there may be challenges there. But accordingly, the systems are developed. If we bring the cost; expensive and really very expensive compared to raster data. Nowadays lot of satellite images, digital elevation models; both are raster's are available almost free of cost on net.

But hardly vector data is available free because it has to be created. You cannot do it like satellite images which are acquired almost automatically via satellite sensors. And then using those images; you analyze and like you, develop a vegetation map or forest cover map or for you use for agriculture or for in our sciences, geology and you can also develop digital elevation models.

So many digital elevation models have implied different remote sensing datasets. And therefore only the processing is there but data becomes inexpensive. But in case of vector, each point, line, polygon has to be digitized that is why it is expensive. If we consider output map quality then of course, the vector is very good whereas raster is poor from map quality point of view; poor only when it depends again on a scale and depends also on the spatial resolution.

If raster data in relatively higher spatial resolution that may not be that poor and if you know high spatial resolution, you are representing in a map which is to say, in a full size of 50,000 scale or so then that may not be poor either. So these output map quality is also very important. Now, I mentioned about this overlay ability. So, we will spend few more minutes on this part.

(Refer Slide Time: 51:51)



And that is very simple example here is taken. I am having one map which is soil map and another one is land use map and both these maps are having only 3 polygons, each. Now when I overlay then I may end up with 9 polygons or more or less, depending how they intersect with each other.

And therefore not only the number of columns in my attribute table will increase but number of polygons has also increased in my map which you can see here which is soil land use map. But GIS are capable to handle these overlays of vector data up to certain extent. But if you go and overlay you know, 10-20 or even 5 maps like, this each is having 3-4 units in it then you would end up with 30-40 polygons of smaller and smaller size.

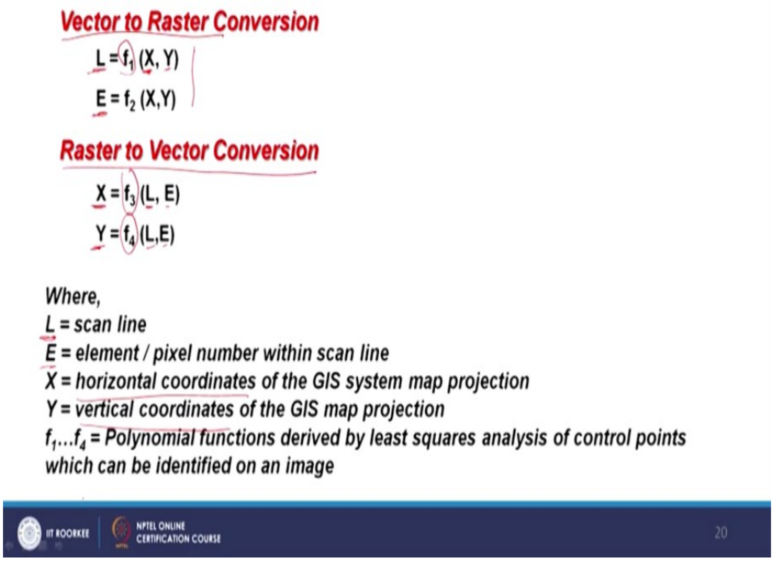
And the usage or interpretation of such maps would be very difficult; theoretically. Practically you can overlay but output we have to understand. That is the thing in GIS whatever the advanced analysis you perform but whatever the output would come; we have to understand. We have to explain. And the best thing as I told also in the demonstration, do it in step manner. Maybe 2-3 maps, first you only create an output.

If I have to do it say for 6; 3 in one go, 3 in one go and later on these 2 outputs of 3, 3; I can overlay. So I will have understanding in a step manner. So here if I take this example, here units are numbered for our simplicity, they have been kept A, B, C. Here they have been kept as 1, 2, 3 as ID of these polygons. And when I overlay then there is a A2, A3, B3, B2 something like this.

So here in you know, output table; A2 is agriculture in my land use map. So, A is agriculture in my map which is A2. So, 2 are here. And A soil depth is 150 centimeter so likewise. If I take say C3; C is having depth of 90 meter and 3 means pasture so C3; pasture 90 centimeter.

So likewise, a table can be developed or developed by the software or can be developed by ourselves and overlay to certain extent of vector data is possible. But if I have to do it for 20-30 layers, having all in vector then it would be very difficult but if I do it using raster data, it is no not that difficult. Still output from both models, you have to understand and use that one.

(Refer Slide Time: 55:09)



So, last here is the conversion; vector to raster conversion and raster to vector conversion. As I mentioned earlier, it is possible to convert from one format to another or one model to another. But always when you convert from one to another because these conversions are not fully transparent and therefore they will introduce errors. So here if I have to convert from you know vector to raster; the first one then what I am determining.

I am determining the column and row number. So, L- here is the scan line or row number, E is the pixel or the element; pixel number. That is the location, not the value. Now here, there is a function. So, these are my polynomial functions because in case of vector to raster conversion, I am having input X and Y; my coordinates.

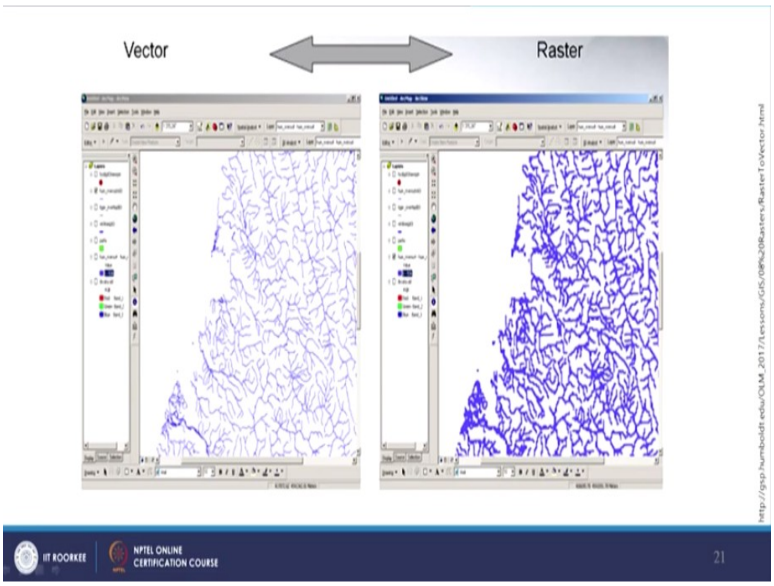
So for row and column, I will have the same equation and similar polynomial functions. In case of raster to vector, again I will be implying different polynomial functions. But here I am going to determine coordinates using input as row number and column number; L and E. So L stands as I have said, scan line or row in raster. E is the element; the pixel or column number.

X is the horizontal coordinate. Y is the vertical coordinate and so on. And the value if it is from vector to raster like first one, what would be the value of the cell is a separate issue. This is only conversion from there. The value has to be assigned also. And that is why because these 2 models are conceptually completely different. Even then we are trying to convert from one format to another. We are bound to have errors in it so one has to be very careful.

Again the same point I can bring that whenever you are doing conversion from one to another, do not delete original file especially the vector file because vector data is a very expensive. So keep the original file and you can have another file which is converted to raster and same with vector to raster and raster to vector.

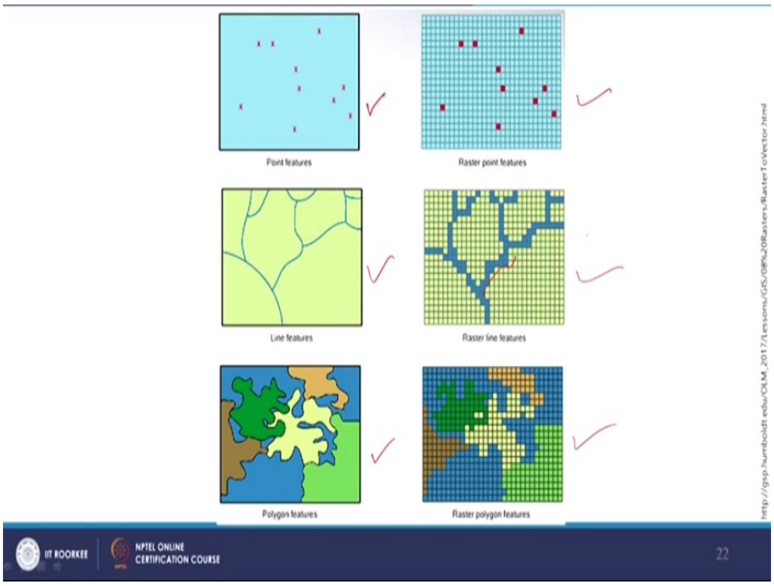
So this one has to remember very carefully because once I have converted from vector to raster and same time, I have deleted my original vector file. Tomorrow I required my original vector file, I cannot restore by any means whatsoever. No processing tool, you would find in any software which can bring you the original vector file except if you get the backup from somewhere. So, these are the very important point here.

(Refer Slide Time: 58:18)



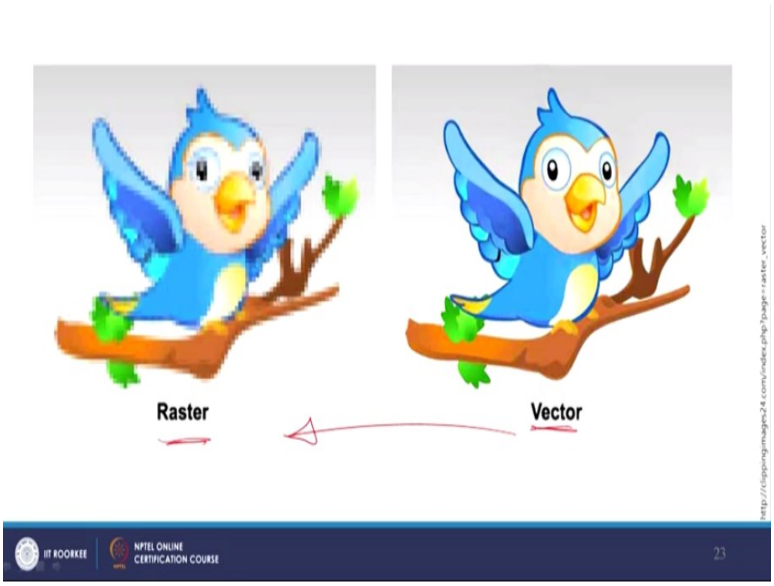
Examples are also here; this network which we are seeing if I convert into raster this is how it would be visible. But if I zoom it further then I will start seeing that jagged or stair steps case in case of raster but not in vector. So there is always a conversion issues from one to another.

(Refer Slide Time: 58:44)



For point data if I convert point to raster something like this will happen. If I convert line to line features to the raster line features, again I will have these jagged issue or stair steps case issue. And if I am converting from polygon to raster polygon; basically raster, I am having the boundary problems and spatial resolution problems with raster data. So both models; raster and vector are having various advantages and disadvantages.

(Refer Slide Time: 59:13)

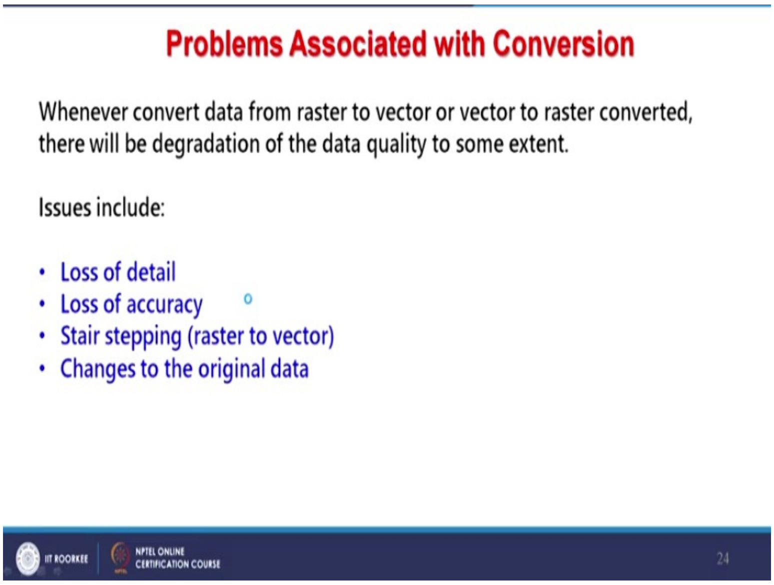


This is the last image just to understand in a much better manner. Because everyone is having a child in him and therefore whenever we see cartoons, it becomes automatically attractive.

So the same thing I am going to explain through that if you see that in vector; it looks very smooth, very clean. But in case of raster; the same cartoon is being represented here which is not as clean. If I improve the spatial resolution of this raster image, I can have better representation.

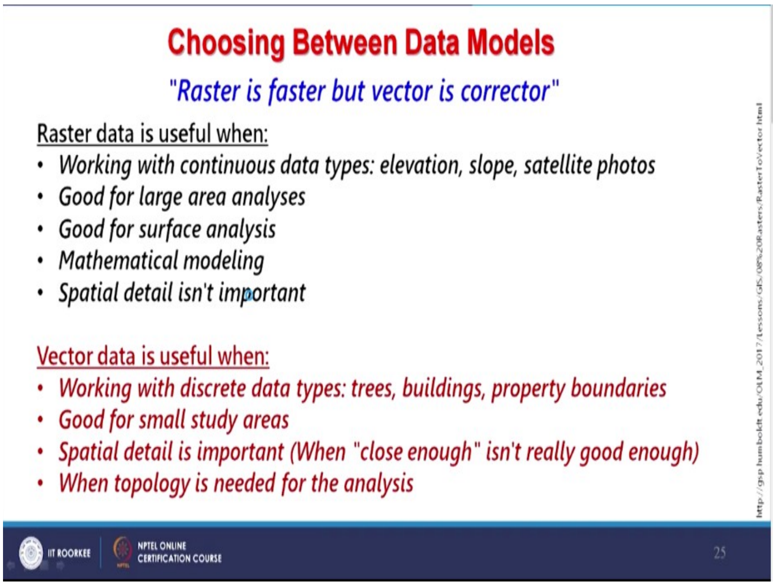
But that would not be close to my original vector file. So if I convert from vector to raster, this is what I am going to end up. And therefore the vector file should not be deleted without much thinking. Make sure that you are having backup because it does not require much space; memory. So do not unnecessarily delete vector files because it is not easy to create. So problems I have already mentioned. So, very briefly we will go.

(Refer Slide Time: 1:00:32)



So, whenever you convert from one format to another, there will be degradation of the quality of data. Many examples I have shown and issues include like loss of details, loss of accuracy and stair steps case in case from raster to vector and changes to the original data and so on so forth.

(Refer Slide Time: 1:00:52)



Now which model to choose? Of course, it depends. The best answer to this question is that it depends on your project. Also depends what is available. If a raster is available, my project is also requires raster data then why I should bother about vector data and same with vector. But sometimes we have to do it in this kind. So we know raster is faster but vector is corrector; it is having a more correctness, more accurate representation.

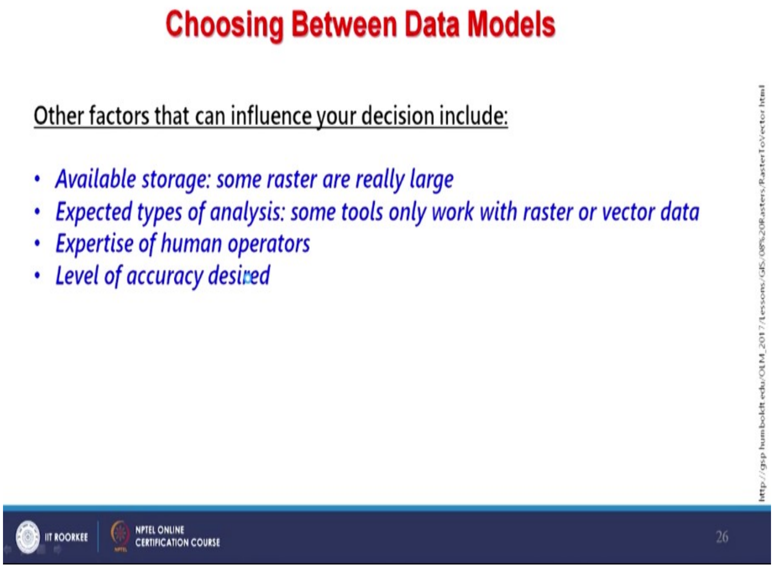
So raster is useful when working with continuous data types for example; elevation, slope, satellite images and so on. Good for large area analysis. Good for surface analysis because raster by nature is representing the surface. That surface maybe elevation surface. That may be a pollutant surface, soil surface and so on. And mathematical modeling because it is 2D or 2 dimensional array so all those modeling and other things is easy to perform.

Spatial details: if they are not important then it's fine. Vector data is useful when working with discrete data types like tree, buildings, property boundaries etc. If somebody is working say in revenue department then raster is not good because accuracy is very much required there. And you are representing various you know boundaries or land records vary in form of polygons then vector is good and good for a small study areas.

If a study area is large like continental scale then vector is not good. Spatial details are important when close enough, it is really not good. That means that if I am working for a terrain like Himalaya and I am using the vector data that means my contours would be so close to each other that sometimes it is very bad from analysis point of view. So they become so close because the ruggedness; the slope is very high.

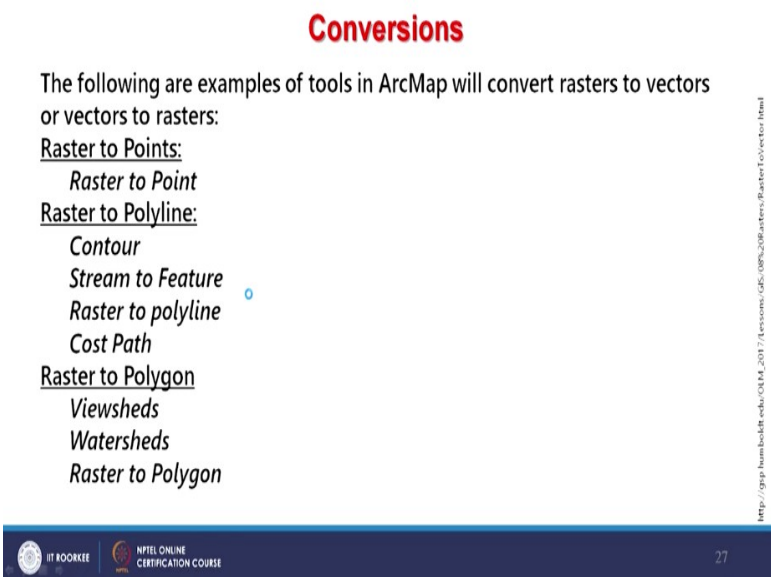
But in case of raster, the difficulty will not be noticed. And finally here that when topology is needed for the analysis; in case of raster, you do not have to bother about topology but in case of vector of course, topology is always required especially for the polygon data.

(Refer Slide Time: 1:03:29)



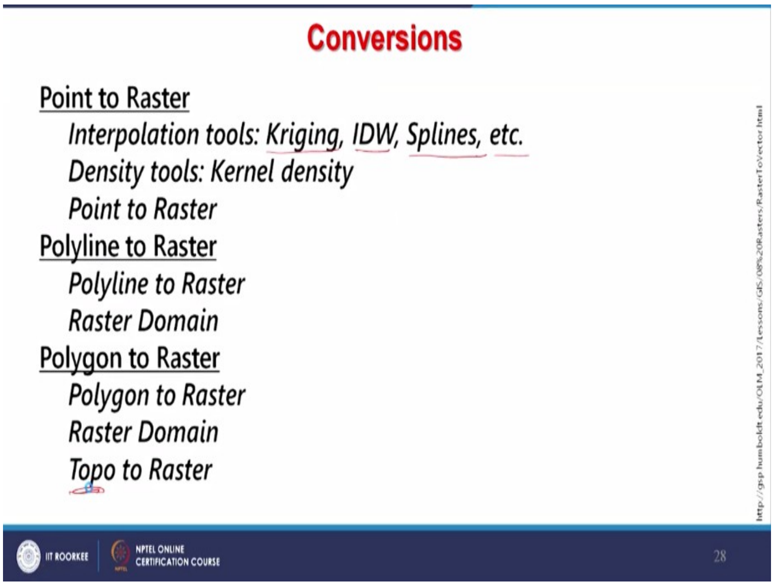
And other factors which can influence your decision which includes the available storage because raster requires a really large storage space. Expected types of analysis; some tools only work with raster or vector data. And expertise of human operators is also required. Level of accuracy desired that will make our decisions.

(Refer Slide Time: 1:03:52)



And some examples are here with ArcMap GIS software converting from vector to raster. So, there are tools available. You can convert raster to polygon, raster to point, contours, streamlines features, raster to polygon. And when you have to calculate cost part and other things then it is always go for vector. Raster to polygon also required when we are working for viewshed, watershed analysis. Raster to polygon is also required elsewhere.

(Refer Slide Time: 1:04:27)



Now point to raster data that means creating from discrete to continuous and the tool is available for us, is interpolation tools. And these interpolation tools include various interpolation techniques for example Kriging very popular one, inverse distance weightage method, Spline and many more. And variants of these tools are also available not only in this particular software but many types of software are there.

Density tools are also available because point data density analysis is required in many applications. Point to raster, polygon to raster; you convert raster domain and polygon to raster. So point to raster, polyline to raster and polygon to raster; all these things sometimes for certain types of analysis, we require. And like topo; topo to raster means here the topography or topographic map but there is no automatic conversion is available from raster to vector, this one has to remember.

So with this, we come to end of this discussion. Thank you very much.