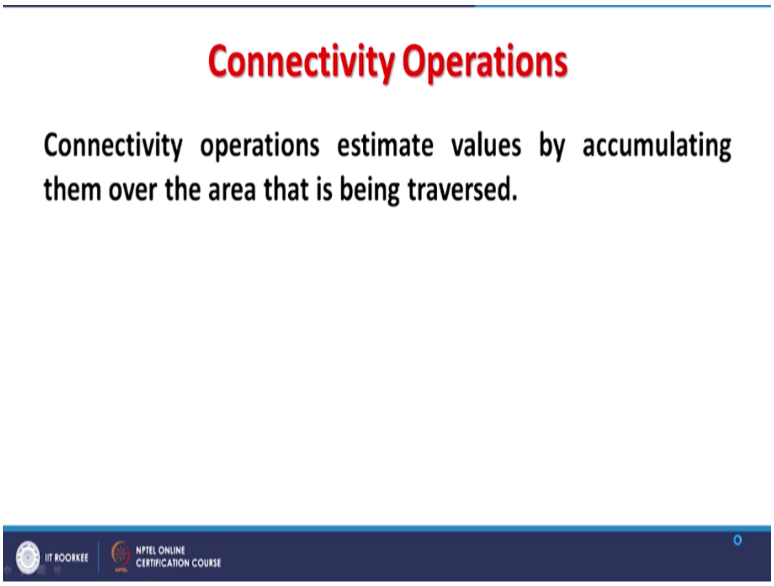
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Sciences
Indian Institute of Technology – Roorkee

Lecture - 24
GIS Analysis – 05

Hello everyone! and welcome to this geographic information systems course. And today we are going to discuss GIS analysis Part 5. And in this as we have been discussing earlier about the neighborhood operations and before that also, the buffering and other things. Now in this one, we are going to discuss this connectivity operations which are again very important especially related with the various analysis for finding suitability for certain projects and also from network point of view also.

Because, that is another very good application of GIS in network related thing. So, we will be going through this discussion about the connectivity operations.

(Refer Slide Time: 01:10)

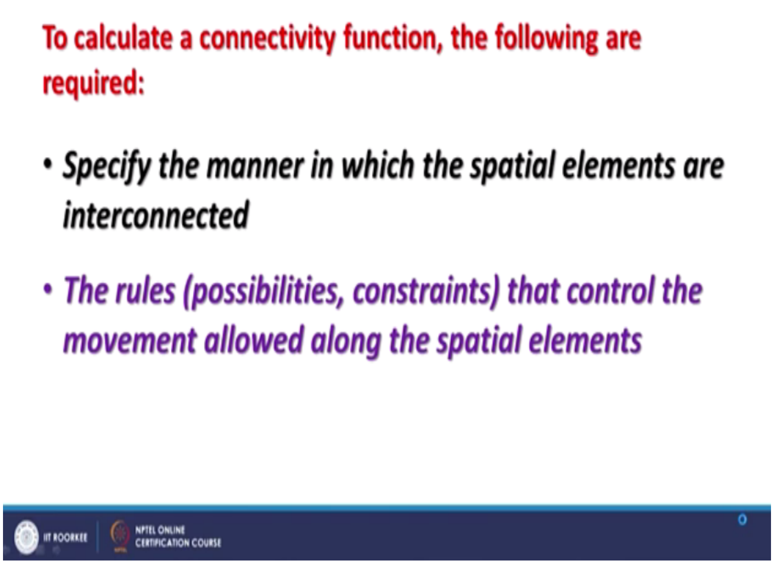


Basically, here also like in interpolations, the estimations are done. So here, connectivity operations estimate values by accumulating them over the area that is being traversed. And you know if I give you an example here that like if I have to climb a mountain then initially, I might be going at a little higher speed and suppose I have spent 1 hour time.

Now, it will take more time to travel to the same distance so that the tiredness is getting accumulated. Similarly in a plain area, if I walk for first an hour then I may complete 5 kilometers. But in next hour, I may complete 4 kilometer and likewise my traveling or traverse reduces because the tiredness gets accumulated. In case of like pollutants and other studies where dispersion occurs.

So, it’s a reverse way but again, you know it can also be estimated using such operations in GIS. We will see further details on this with the examples also. So, in order to calculate this connectivity function, what are the requirements? What are the inputs?

(Refer Slide Time: 02:38)

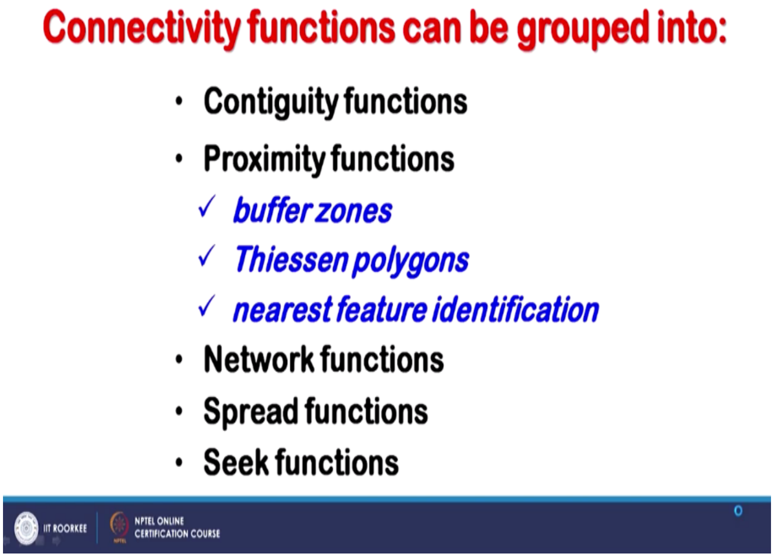


So, first the specify the manner in which the spatial elements are interconnected. Because this is based on the connectivity and therefore, this interconnected information is very much required among the spatial objects or spatial elements. And second thing that how things move. You know what are the constraint, possibilities that means the rules and that will control the movement which is allowed along the spatial elements.

So, suppose there is a pipeline or a river system, a stream or maybe a road or a power line. So, what are the possibilities and what are the constraints that we have to also provide. Because recall the initial discussion on GIS; GIS is providing various software tools. It is up to us to utilize them judiciously and very efficiently. Though, most of the tools which we can imagine are available nowadays with good GIS software’s.

But only it is our capabilities, how many we can exploit. So, these rules or this understanding about that particular phenomenon is very much required. And of course, third one is the unit of measurement. This we have to always take care while doing such calculations or estimations in GIS. So, connectivity functions can be grouped or can be divided into various sub functions.

(Refer Slide Time: 04:17)



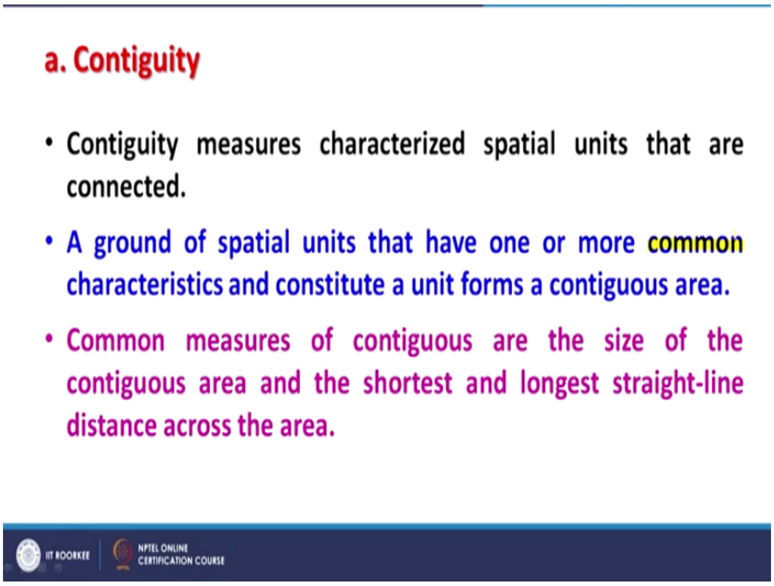
The first one is the contiguity. We will see details about these functions little later. Then proximity and proximity part, we have already discussed like buffering or buffer zones; creating buffer zones. So, it is directly connected within a spatial object. A point along which, we create a circle so that is proximity directly connected. Similarly for polyline and polygon.

And this proximity functions also comes with the Thiessen polygon which we have discussed while discussing interpolation techniques. So, that is also a proximity function under this connectivity function. And nearest feature identification, that we also put in this proximity functions. Then the third sub function of connectivity is the network functions which are very important.

And they all come under this category of network functions. And then we are having spread functions. I will be showing some examples. How this spread function can be exploited for various applications, related with hydrology or planning of certain reservoir or even for groundwater. So, this is spread function. I will be also dealing in demonstration which will be the next after this lecture.

And the last sub function under this connectivity functions is called the seek functions. Seek functions is based generally on the surface hydrologic modeling. And when we will discuss in detail about the derivatives of digital elevation models then at that moment, we will be discussing also the surface hydrologic modeling and in that, we will be using this seek functions also.

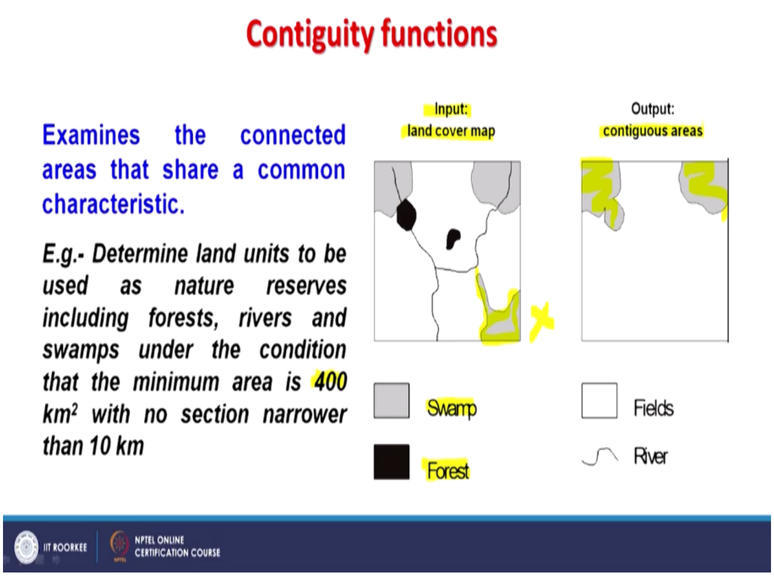
(Refer Slide Time: 06:14)



So, let us start with first contiguity. The contiguity basically measures characterized spatial units that are connected. Because overall it is a connectivity function so things have to be connected then only such measurements or estimations can be performed. A group of spatial units; those may be point, line, polygon that have one or more common characteristics. Now, common characteristics is another essential requirement and constitute a unit forms a contiguous area.

Now, common measures of contiguous area are the size of contiguous area and the shortest and longest straight-line distance across the area. So, these conditions will come under this contiguity function.

(Refer Slide Time: 07:03)



So, in this one, what we are basically doing? We are estimating the connected areas that share a common characteristic. For example, we might be doing in suitability analysis. We might be determining a land unit which will be used as a nature reserve. So, this is our requirements. This is what we want to extract from or query from the system through this contiguity function. And what are the conditions?

The condition that it should include forest, it should include reserves, it should also include swamps under the condition that the minimum area. Now, further constraints are there. Further rules, we are putting before the analysis that the minimum area which is being looked for this natural reserve is 400 square kilometers with no section narrower than 10 kilometers.

So, many conditions have been put. The only requirement would be after these conditions that this information must be available within our GIS database. So, like including forests, rivers, swamps; this information can come from 2 layers. One is land use and other one is drainage network. And we can have then this condition about the 400 square kilometers that minimum area and no section narrower than 10 kilometers.

Like for example in this schematic, what we are having? This is our input map and land cover map are there in which we are also having our drainage lines and other connectivity things so. So, there is a swamp, there is a forest and after this iteration through this contiguity function or after this estimation, this is what we get that these are the areas which are satisfying all your conditions.

That you know they are having area larger than or equal to 400 square kilometers. These are not narrow. Like here, this is narrow so that is why, it has not been chosen. This is not falling equal to 400 or more than 400 square kilometers. So therefore, this area has not been selected. And also, it is including forests, rivers and swamps.

So, in that way, our output is the contiguous area like this. So, in this way several conditions, if our data is having or our layers are having that kind of information then we can estimate or we can determine land units to be used as a natural reserve in this example. But the similarly, like if I give the example for municipal corporations, many times they are looking for suitable sites for solid waste dumps.

Now, there will be some condition because there are environmental constraints and so many things are there plus, it should not be close to population. And you know many-2 things are there. So, if we put all those conditions to our database; if we are having all those details in our database then we can really find out very good suitable sites for waste disposals.

Similarly, we have been involved in such projects where you know, this analysis was done to find out the suitable sites for this fly ash dumping. Because these thermal power plants which use coal to heat the water and that waste of that coal is fly ash. And that fly ash has to be dumped somewhere. So, wherever it will be dumped, it will destroy the soil and also to some extent the groundwater.

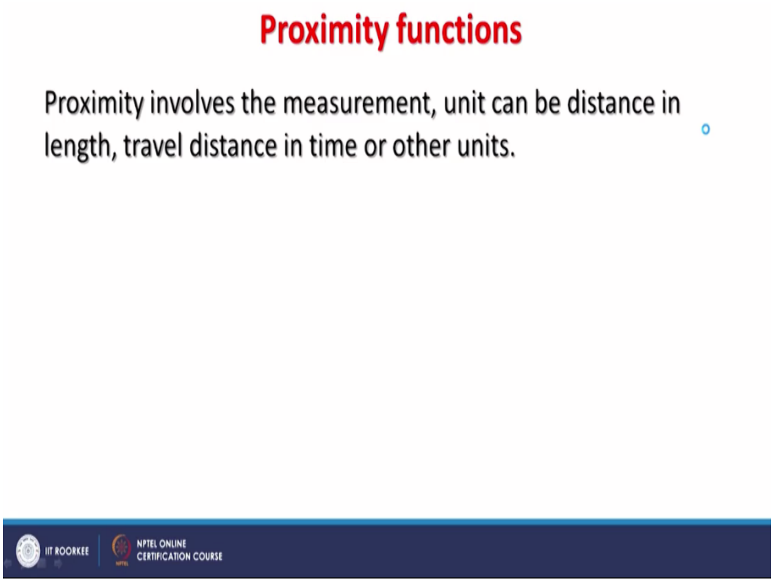
So now, if we know that how we can minimize this and we can also think that what would be the transportation cost and other things? So, all these constraints, conditions and rules can be put to the GIS system where we are having a very rich GIS database. And most suitable sites can be identified, maybe then 2-3 scenarios; that this is the most suitable site then next then next and then next. And then decision makers can make decisions.

This is what the purpose of GIS to provide different scenarios through our modeling, through our analysis so that decision makers can take very appropriate decision. If GIS user himself is a decision maker, then of course there will be one less step. But in many cases, decision makers are different people than GIS users. So, no problem! We can develop 3-4 scenarios that this is the most suitable site.

But if you know socially or economically, it is not possible then this is next then this is next. So, a simple function which is contiguity function under this connectivity function is so useful to solve various problems related with our real life in different domain, whether it's a waste disposal, in civil engineering or in earth sciences or electrical engineering and many-2 scenarios, these things can be used.

Now, we come to this proximity functions. Some sections of proximity functions, we have already dealt.

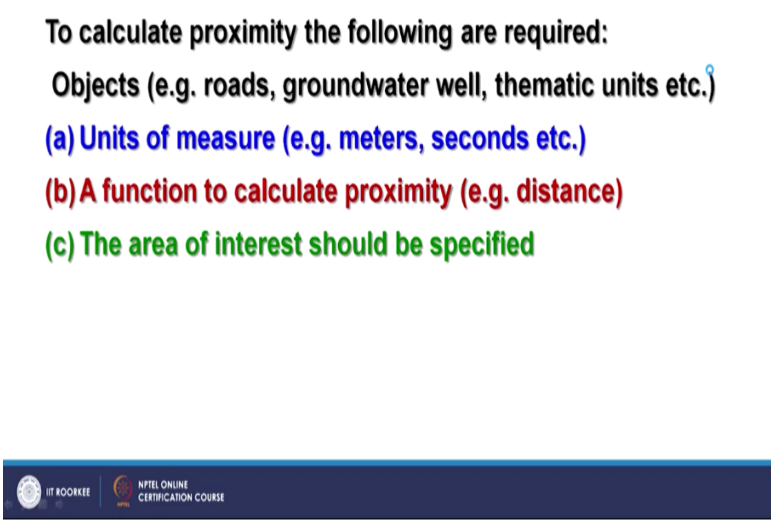
(Refer Slide Time: 13:07)



Nonetheless, the proximity basically involves the measurement as you recall the buffering in our previous lecture. And unit can be distance in length generally or travel distance in time or other units might be there like dispersion of pollutants concentration. Now, in order to calculate proximity, what are the requirements? The first requirement is object; that always means spatial objects.

For example, there might be roads, groundwater well, thematic units, land use, forest anything basically.

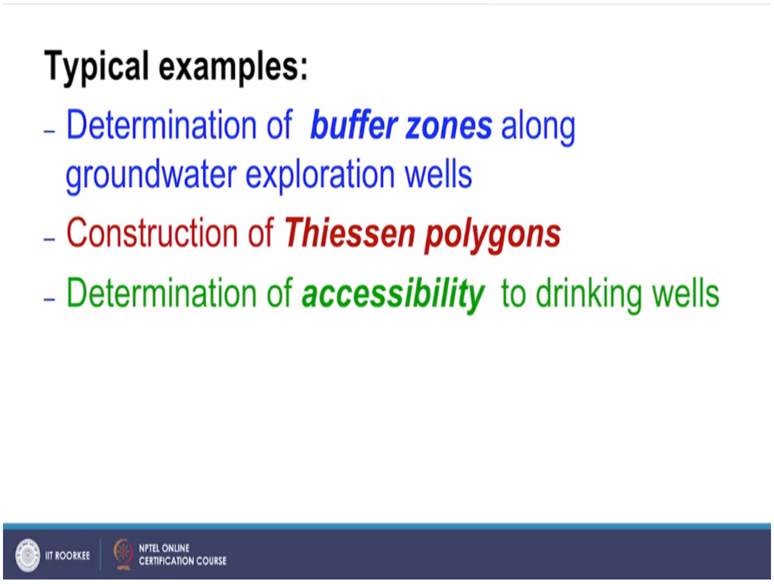
(Refer Slide Time: 13:36)



And then you require unit of measurement because if buffering has to be created then you have to provide whether it has to be in terms of distance or in terms of time or in terms of concentration. And then a function basically algorithm or knowledge about that phenomenon for which you are going to use this proximity function and that to calculate proximity. For example, that might be a distance but this is not always true.

There might be some other functions also. And the third one; the area of interest should be specified. In which area, you are going to do. If I am having a complete database say of Uttarakhand but I want to perform this analysis only for a small area. So, I have to provide that information also. So, it would be quicker and easier for me to create such scenarios.

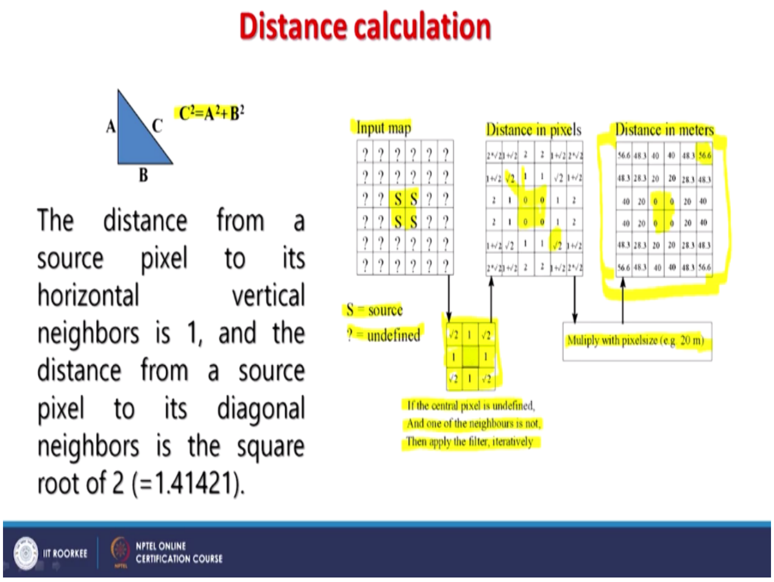
(Refer Slide Time: 14:40)



Now, example where this buffering is done. Some examples we have already discussed. So, determination of buffer zones along groundwater exploration wells; that is also one possibility. Construction of Thiessen polygons while discussing interpolations, we have already discussed. Then determination of accessibility to drinking wells. In many scenarios especially in villages, we have to do this analysis that what is the distance, people have to travel to fetch the water from a well.

So, that we can also analyze. Here first the distance calculation. So, basic trigonometric function will be used here.

(Refer Slide Time: 15:25)



The distance from a source pixel because here, it is not necessary that you would be performing only on vector data. So, it can be raster data to its horizontal, vertical neighbors is 1. And the distance from the source pixel to its diagonal neighbors is square root of this 2; that is equal to 1.41421. Now, while discussing this you know spatial filtering; the concept on neighborhood operations, there we discussed this moving window or roving window.

You recall this is 3*3 moving window or roving window in case of even many you know functions like calculating slope, aspect and other things also. Anyway, so this is our input map and there are 2 categories in our input map. So, you can say it is binary one that there is a source and this analysis being done for fetching water. So, the source of water is located here and rest areas are having no water.

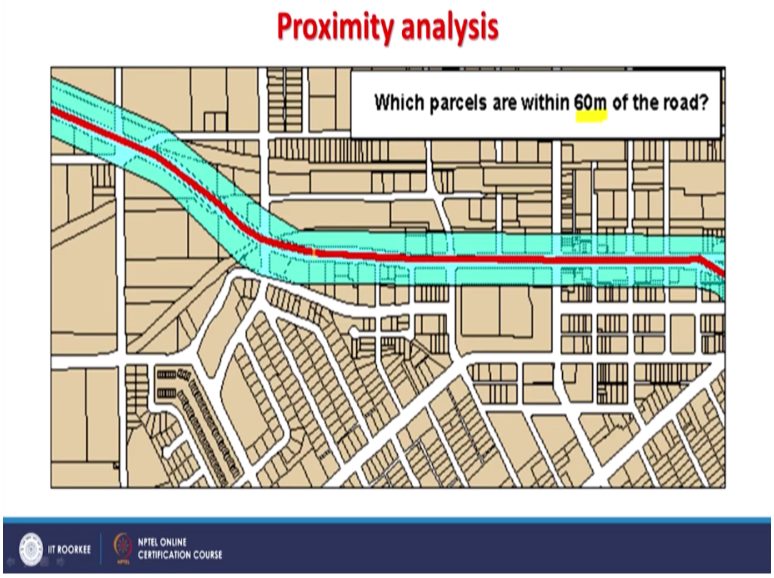
So basically, they have been given as unidentified or undefined. Now, this is the distance in pixels. So, what we have discussed here a diagonal distance is calculated from these 4 cells. So, if you go just vertically upward then it will be 1. But if you go you know diagonally upward then it is square root 2. Similarly, in downward and as far you go, the distance increases from the source of water.

So, when we use this that multiply with the pixel size, here is the spatial resolution of this input data. So now, we can get distance from water source to the different cells. So now, we are having distance. So, if the central pixel in roving or moving matrix is defined here, one of the neighbors is not then apply filter iteratively.

That means that we will move this filter or this matrix roving window throughout the data and then calculation will be done. The only problems in such scenarios comes when we use roving window or moving window which is not shown in this schematic is on the borders. So, one thick pixel on right side, one thick pixel on left side, same in the bottom and top; there will not be any calculations because beyond this, you do not have the input data.

So that way, you can find out what is the distance. Now, if you are having travel time information also coming like this 56-meter to this place, how much time it will take by walking or by cycle or some other mode then the time can also be involved or can be estimated in such calculations.

(Refer Slide Time: 18:54)



Now, we come to this proximity analysis. During buffering I gave this example. But here again, the red line which is shown here is the road. And now, the buffer has been created on both side of the road is 60-meter. Suppose, now there is a plan to widen this existing road which is going through a habituated area. Now, whose house or whose plots will get affected?

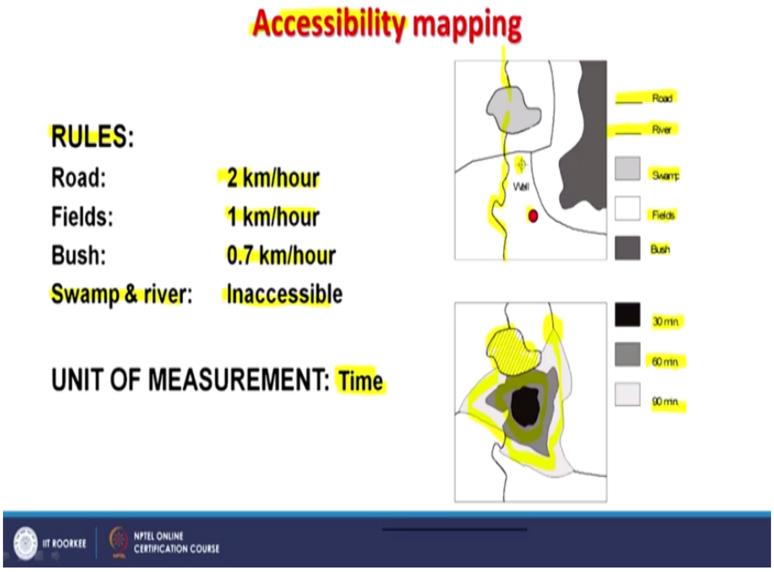
Immediately, if you are having a layer which is showing the revenue records or the plots of individual houses and if you overlay this buffer then immediately, you would know that whose houses or whose plot would be taken by this widening of this road. Similarly, you can do analysis like if a stream is flowing, now I want to calculate the flood plain area. If there is a flood say maximum flood which I get, generally hydrologist use the 100 years time period.

So, how much area would get inundated? That too can be analyzed using this proximity analysis. So, various such applications are there. I was involved in one of the projects in which there was a mineralization. So, one economic mineral thick vein of few meters was going through an area. Now, if government give to lease for exploitation of that mineral, then they have to give some extra area so that the mining can be done.

Now, that means simple buffer analysis. And after that, then whose land will go in that lease? If we want to extract that information then we need to have a revenue record. Now, here the example I am just taking is involving only 2 layers. But sometimes in our proximity analysis, there can be more than 2 layers and still we can perform such analysis.

The only things are that our rules or constraints can become little complicated as compared to just having 2 layers.

(Refer Slide Time: 21:22)



Now in this proximity analysis, we can also do a mapping which is called accessibility mapping and that accessibility mapping basically, it is also under connectivity function. So, the rules what are being put here for example if by road, it has to be travelled, now time is coming here so then 2 kilometer per hour. And if one has to travel through fields, not by road then the distance it will cover is 1 kilometer per hour. So, it is just half.

And if one has to pass through the bushes then this would be 0.7 kilometer per hour. And of course, the swamp and rivers inaccessible that nobody will go and swim through that and unit of measurement in our analysis of accessibility mapping is time. Now, let us take example. Here, there is a groundwater well. And what we are having; a land use map where we are having the swamp, where we are having fields.

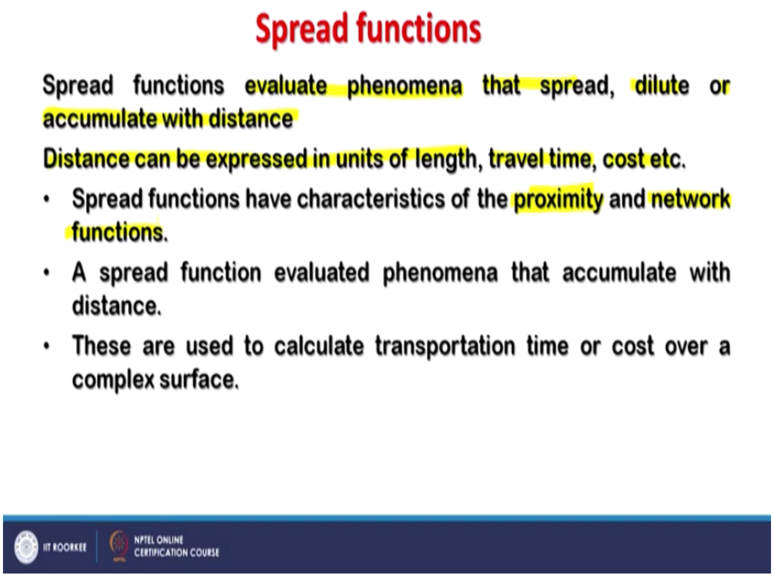
Because in fields, the travel time is 1 kilometer per hour. Then we are having information about the swamp and we said these are inaccessible so nobody will travel through. We are having then river network, a stream network and road network. So, this is just to make clear that this is the river and rest polylines are roads. Now if for this location, I want to prepare an accessibility map that how much time, it would take to fetch water from that well?

So, here now these buffers are being created, not circular buffers. More realistic buffers where these conditions are there; these constraints have been put. So, like here, what we are seeing that if we close to this place then it will take only 30 minutes to reach to the well and fetch the water. Whereas if one is located on this area; second sort of polygon, just outside the 30 minute then it would take 60 minutes and then finally the outermost buffer is showing.

That it will take 90 minutes. Because we have constraint that the swamp and rivers are inaccessible. So therefore, we do not see any accessibility here. So likewise, lot of things can be planned before anything really happens on the ground or lot of things can be analyzed to provide certain benefits to the community by this. So, if a person who is here say for example, then he has to spend 90 minutes time, if it is possible.

And if it is possible then a new well can be created; a new hand pump can be installed so that the people those who are living will fetch the water in only few minutes, not in hour. So likewise, we can do these accessibility mapping involving overall connectivity functions under which we are discussing proximity and accessibility. Now spread function is having lot of applications, especially in flooding or reservoirs planning and other things. So, what a spread function does basically?

(Refer Slide Time: 25:19)



Again, it evaluates the phenomena that we are trying to spread maybe water or nature oil, dilute or accumulate with distance. So here, we will be putting then other conditions also that the distance can be expressed in unit of length, travel time or cost also. That can be calculated, no issue. So, the distance either unit of length, the travel time, also in the cost form. Now spread functions have characteristics of proximity and network functions.

It's a mix up of these 2 functions; hybrid kind of function where the proximity is also maintained because things have to be connected in case of say flooding, a lower ground maybe having water but in between, we are having a ridge and then on the other side, we are having a much lower ground but since it is not connected, water cannot reach to that area. So, there has to be a connectivity and proximity and of course, network is also there.

So, that we will see. And spread function; basically, evaluate phenomena that accumulate with distance. And what are other things that these are used to calculate transportation time or cost over a complex surface. Because many times especially like in mining, you might have seen some opencast mining. These are not underground but open to sky so they are called opencast mining.

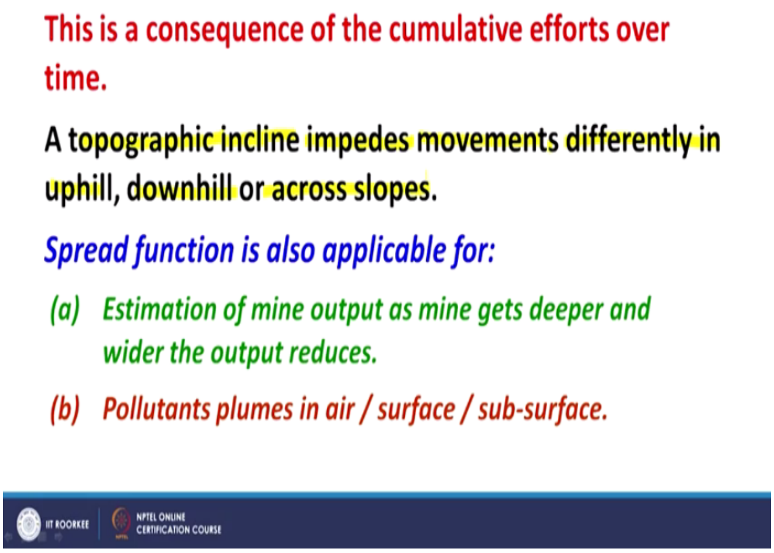
As we know that as mine gets deeper and deeper, it involves more time and ultimately more cost. Whereas when mine is shallow, only the travel time reduces or travel time is less and therefore, cost is less. So, basically what it is accumulating? It is accumulating the travel time to fetch a mineral or ore or coal from that place or deeper from the mine. And ultimately, it cost not only time but it cost in terms of money also.

Because the material has to be transported. Say a mine is already 100 meters deep now so that much is there. Initially, it was on surface or near surface; 10 meters. So, the cost was nothing and the production was very high. So, if we want to maintain the same production as we having in initial stages, that means we have to put extra cost by putting extra trucks, dumpers or many such features.

Then only we can maintain the same output otherwise output will also reduce. So, the same thing; this spread function can also be utilized in even in case of you know underground mines. As mine gets deeper and deeper, the travel time increases and cost. So, this estimation of the time taken to walk in a relatively rugged terrain. This example I have given that the more you walk, longer it takes to walk over the similar terrain.

That the first 5 kilometers, you may walk in 1 hour; the first hour. And then next 5 kilometer, you may take one and half hours and maybe next 5 kilometers, you may take 2 hours. Because your tiredness is getting accumulated, it is a spread function. And dispersion is just opposite of spread function. So, concept wise, it is almost same.

(Refer Slide Time: 28:59)

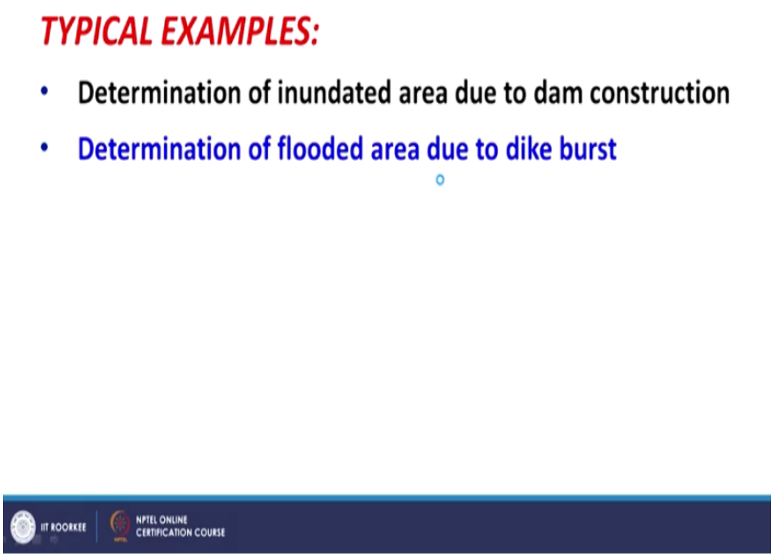


Now, this is a consequence of cumulative efforts over time as we know. So, if there is you know topographic relief or undiluted terrain because this incline or anything which impedes the movements. So, if I have to go uphill, it will take more time but while going downward, it will take less time. So, while doing such estimations, we have to take care about this that a topographic incline impedes movements and differently in case of uphill, downhill or across slope.

So, in a good planning even in case of mining or road construction or for many other transportation purposes, these things can be planned. These things can be estimated or modeled before really anything happens on the ground while employing such functions.

So, spread function is also applicable for various other things like estimation of output as mine gets deeper and wider, the output reduces; this example I have just given. Then pollutants plume in air, surface and subsurface. As more pollutants will cover the distance, it will disperse or the concentration will reduce. It is a reverse but the same way.

(Refer Slide Time: 30:26)



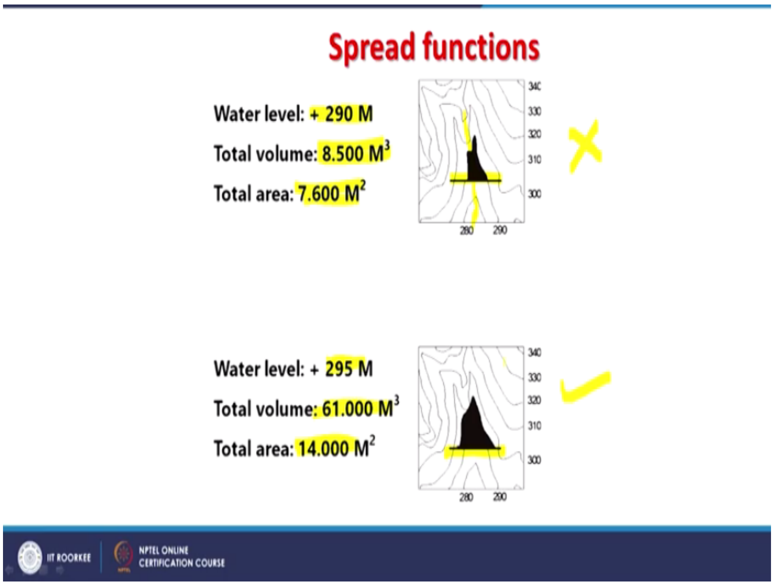
Now, other examples are the determination of inundated area due to dam construction. This is a very much requirement. And nowadays like in India, lot of a small sized dams or recharge structures are created or check dams are created for groundwater recharge. And all these things before really anything happens, we can model not only the location but the height of the dam. And once it will come on the ground, how it will inundate the area; how much area it will get inundated?

And if you can have a land use map and revenue records map also then you can know that which type of land will be inundated or whose land will be inundated. So, that much modeling can be done before really anything happens on the ground. And as I promised that I will be also showing a beautiful demonstration of dam. How dam can be simulated? DAM; not D E M. Dam can be simulated or a reservoir can be simulated on computer or can be modeled.

Now, determination of flooded area due to dike burst. Sometimes they are embankments dikes so if anything happens, how the flooding will occur that can also be modeled through this tool. Then spreading of pollution, dispersion and the construction of canal is related to accumulation of cost. So, we know that with every meter by which the construction of canal proceeds, cost accumulates because now more traveling is involved.

This is about the dam construction. So, first I am showing through a schematic. Then later on through demonstration, I would be showing how to you know simulate or how to model on the GIS software Here in the background, what you are seeing contours.

(Refer Slide Time: 32:28)



But basically, we are using the digital elevation model. But for simplicity, the terrain is shown through topographic contours. What are the conditions that water level is plus means above this? So, 290 meter above mean sea level so that is why plus is there. Now total volume is you may set the target or while some modeling, you may get that calculation also and the total inundated area in the upstream.

Suppose a river is flowing like this. Now a dam is constructed; this is the dam axis and having this height of 290 meter then how that area will get inundated in the upstream? Now, since it is being done on computers on a GIS platform so we can create various scenarios for making better decisions by decision makers. So, instead of if I just increased by 5 meter the height of the dam like here, the location remains same here. The height has increased.

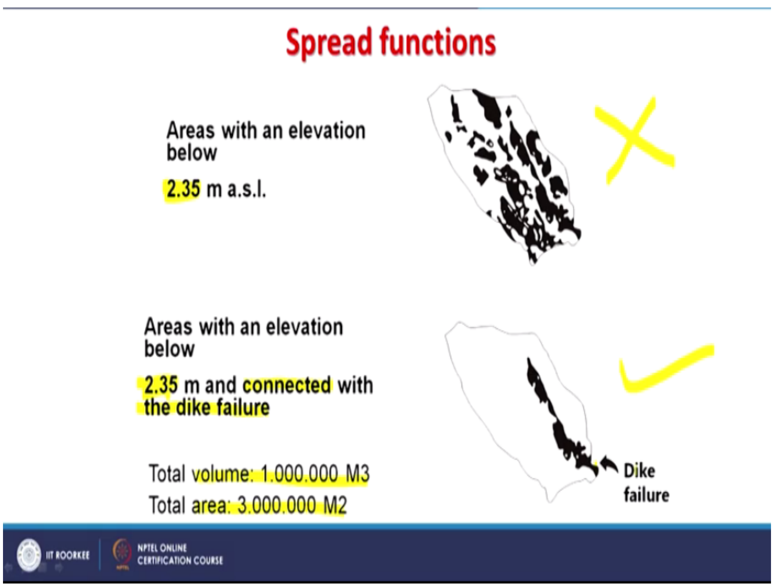
Immediately, you would find that the area has increased tremendously and also the volume has increased. So, if this first option is not you know providing sufficient water say for irrigation or for groundwater recharge then we can drop that idea and we can work on this idea that I want this much of water. Now immediately, you will know that how much area will get inundated in upstream.

Now, next question for decision makers is whose land is going to be inundated or which type of land use is going to be inundated? Whether it's a reserve forest, it's an agricultural land or wasteland or any other type of land use. So, if it is agricultural land then whose farms or whose agricultural lands will be inundated, that can be found out very easily by involving a revenue record map.

In northern India, it is called Khashhara map. So, if we involve that map, immediately I would know whose land will be inundated. And for compensation purposes or for other environmental clearances or other such processes, such informations can be very useful. And remember, nothing has happened yet on the ground.

Everything in the planning stage and various models or various scenarios have been generated so that the decisions makers can take the most judicious and appropriate decision. Similarly in case of flooding like here, this example is the real example for Bangladesh.

(Refer Slide Time: 35:24)



And because we know that most of the land of Bangladesh is not having much height above mean sea level. So, it is just 2.35 meter. Now, the area with elevation below 2.5 meter and connected because overall recall this is connectivity function. And if dike fails then which are the areas; this is the failure of dikes so which are the areas will get flooded? These are the areas.

Why other areas will not get flooded because they are not connected. Like here in a crude way of doing analysis because here, we said any ground less than 2.35 meter shown as a flooded area. But this kind of estimation or mapping is completely wrong because different areas are not connected. Though different areas are having lower grounds but in between, there are higher grounds.

So, for a more realistic analysis of this spread function or flooding, we need to know where exactly the failure or a scenario and only the areas which are connected with dike failure then they will be flooded. And once you have got this because in the background, you are having a digital elevation model. So, definitely you can calculate the area and also the volume of water.

And now, this water has to flow then also you can bring the discharge and other things. And if you know the velocity, you can also calculate how much time it will take to come out of flooding situation. So, all these things once a more realistic scenario has been presented then further analysis can be done. And here I would like to bring a very important point here is that from a GIS user or a specialist point of view, more you deliver; more you are asked to deliver.

I am repeating. More you deliver; more you are asked to deliver because you have given some very realistic say in this one an example that for your higher ups, you say sir, this is the best scenario which I am going to present to you that only these areas will get inundated. This does not mean that any area which is lower than this area will get inundated because it is not connected.

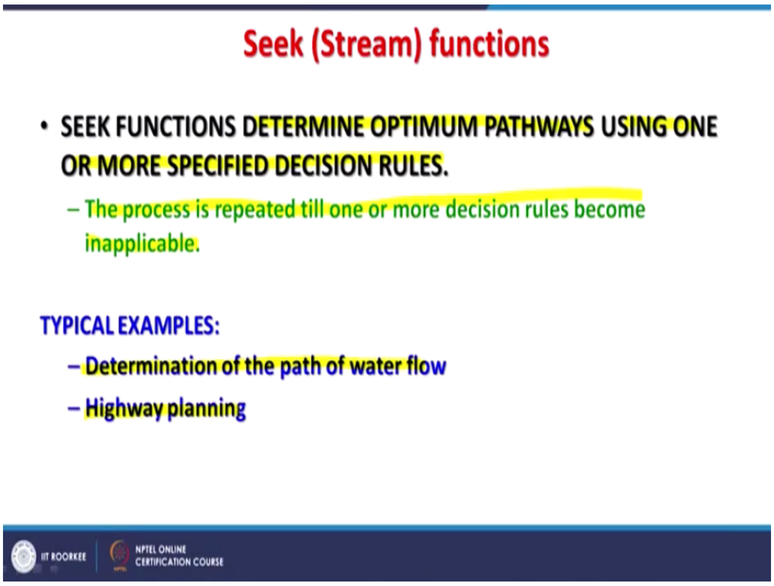
So, your higher ups would be very happy that yes, you have presented a very good you know realistic scenario. Next question he would say, can you calculate how much area will get affected? Can you calculate how much water will be involved there for this inundation? So, that can be done very easily in GIS. Now, once you have done then he may say, can you calculate that how much time it would take to drain out this water?

And what would be the you know the speed of the water or velocity of water from that area? So, more you deliver; more you asked to deliver. Now this do not take in the bad sense. It means that when we are asked to deliver more, we work more. We develop new tools in GIS. And this compulsion has really allowed us to you know develop GIS further and further. And it is making inroads in all real ground related problems in the world.

Whether it is related with pollution, it is related with mining, it is related with civil engineering, road construction, dam construction, everywhere or in natural disasters. Almost every domain of our life, GIS can play some role to large extent. Now, the last one before I close that the this seek or stream functions. As I have said that full length discussion, we will have when we will be discussing surface hydrologic modeling using digital elevation model.

But anyway, before that, the seek functions basically determine optimum pathways using one or more specified decision rules.

(Refer Slide Time: 39:51)

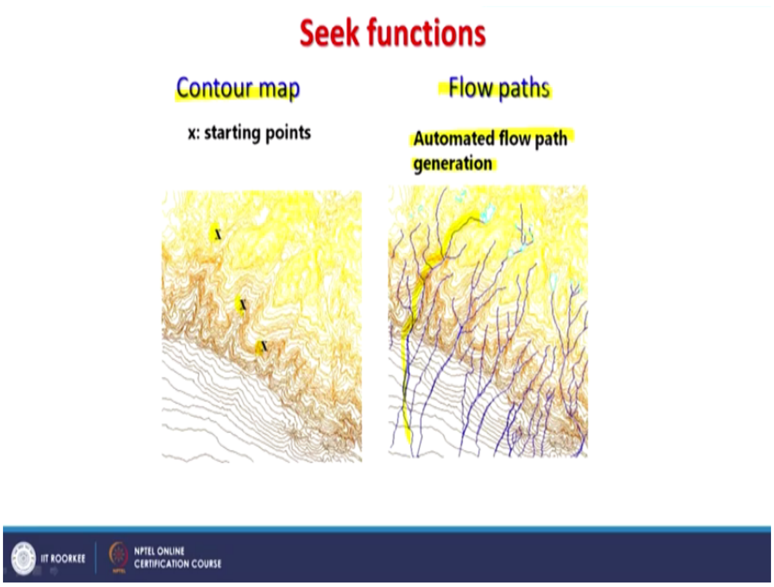


So, rules have to be there. The process is repeated till one or more decisions rules becomes inapplicable. And examples are determination of path for water flow. So, using a digital elevation model that means one can determine even a drainage network. And sometimes very accurate drainage network or a stream network, you can do it. Once the stream network is there then a person who works in hydrology will ask can you develop a stream ordering?

Which are the first order stream, second order, third order based on 2 very popular you know schemes; one is Strahler and Shreve. And you say yes, I can do it. So, once you do that one then they can say can you identify the individual watersheds? Of course, you would say yes, I can do it. So, more you deliver; more you are asked to deliver and as you keep delivering, people will keep asking.

And they will compel you to develop more and more your expertise as well as more GIS subject itself. In highway planning also, this seek functions can be used to large extent.

(Refer Slide Time: 41:14)



And also, like example is given here that these are my starting points over my terrain which is shown in form of contours and automated flow path generations can be done. So, if there is a point source pollution. Now I would know exactly how it will travel through these streams and how it will get dispersed? If that information that knowledge is available, that too can be modeled here.

So, using simple contour maps, I can develop flow paths as well using this seek functions. So, to that extent, it is very useful. So, this brings to the end of this discussion. After this, we will be having a demonstration on software. Thankyou very much.