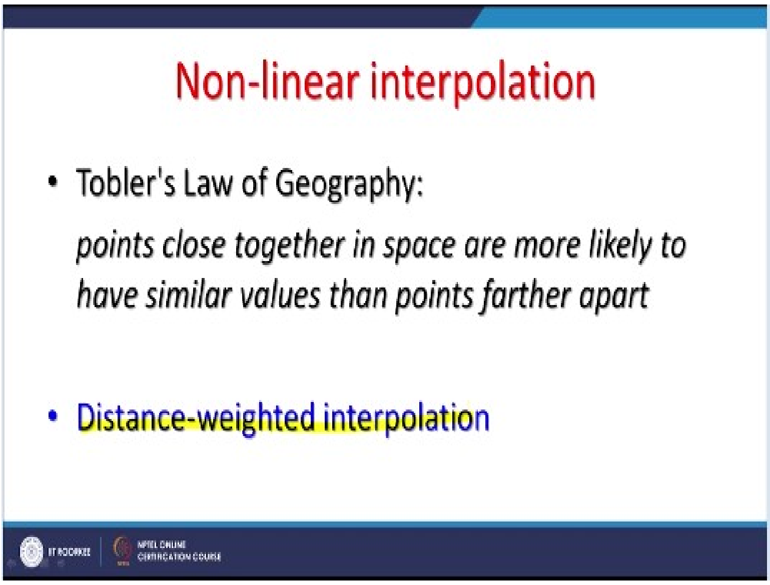
Geographic Information Systems
Prof. A. K. Saraf
Department of Earth Science
Indian Institute of Technology – Roorkee

Lecture – 19
Spatial Interpolation Techniques - 02

Hello everyone! and welcome to this discussion which is second part of our spatial interpolation techniques. So, in the first part, we have discussed interpolation up to the linear one. The linear interpolations are not that popular but sometimes for certain requirements, one can definitely go for linear interpolations. Those options are available in the softwares and one can go even for Thiessen polygon.

One can create Thiessen polygon but it's an old method when we did not have this GIS software or any such softwares then people were only resorting to Thiessen polygon.

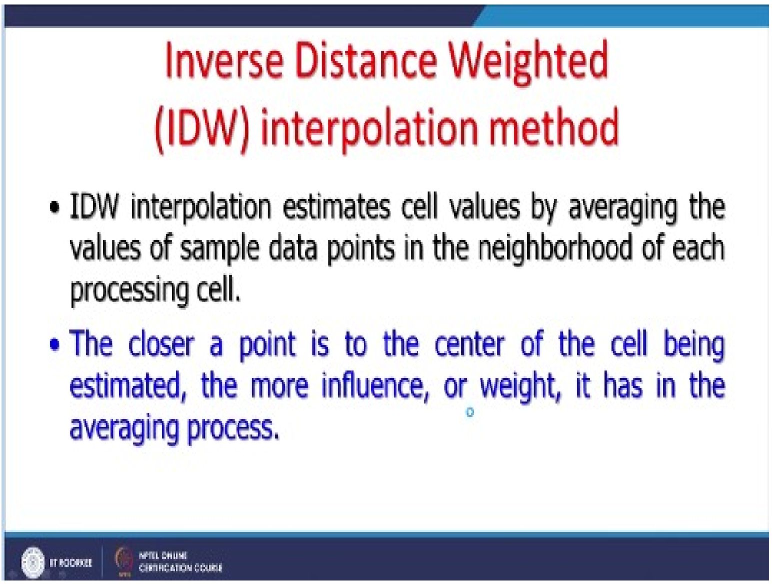
(Refer Slide Time: 01:12)



Anyway now, we come to this other type of interpolation techniques which is being considered as nonlinear which is again based on the Tobler's law of geography and which says that the points close together in space are more likely to have similar values than points further apart. And using this concept of Tobler's law of geography, there are you know interpolation techniques which have been developed which is the distance weighted interpolation techniques.

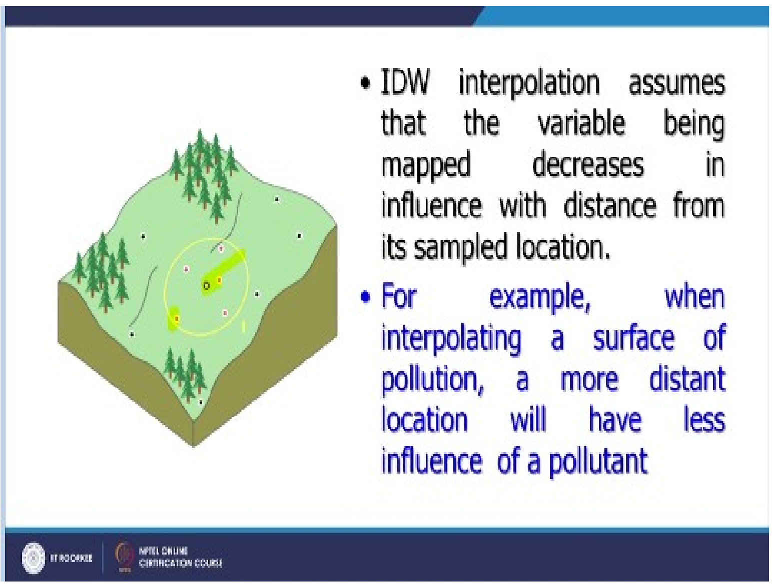
So, for unsampled site, interpolation has to be done then among the input points, the point which is closest will have the highest weightage and the point which is having further will have less weightage and this is what is called distance weightage interpolation. Exact interpolation in that way is IDW which is inverse distance weighted interpolation method in which points which are very far from the cell which is going to be estimated then their weightage will be less.

(Refer Slide Time: 02:19)



So, this interpolation estimates cell values because after all going to create a raster by averaging the values of the sample data points in the neighborhood of each processing cell. And as per Tobler's law of geography that closer a point to the centre of the cell being estimated, the more influence or weight, it will have in averaging process.

(Refer Slide Time: 02:49)



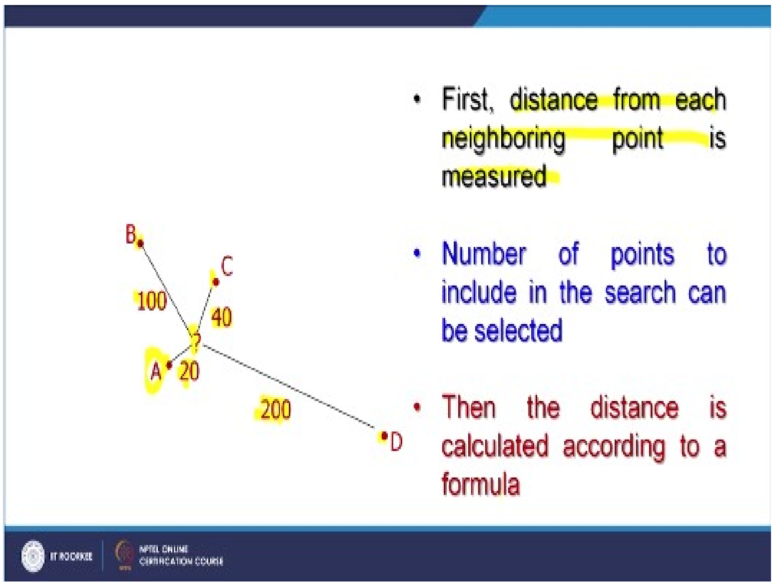
Now for example, if I take this schematic which is showing a terrain and some input points and also showing a search radius in a form of circle. So, for this unknown point which is in the center whatever these small red points are there, the point which is very close will have the maximum influence while determining the value for that one and the point which is very far like this one, will have the less influence.

Now, we can fix this radius also. So, we can fix it either by number of points for this kind of search and estimating the value or maybe a radius for a circle. Now IDW interpolator assumes that the variable being mapped like here in this figure, decreases influence with distance from its sampled location that is why it is called inverse distance weight age.

For example, when interpolating a surface of pollution, as we know that a more distance location will have less influence of a pollutant. So, if a chimney is emitting certain pollutants, if you go away from that, it would be less concentration of that polluted but if you want to do such a prediction about this, you know pollution which is getting dispersed in the 3D space or in the air or in the atmosphere.

Then for realistic prediction, you require not only the location of that chimney in geographic space but you also required the wind direction and wind speed as well or any other impedance, there might be some tall trees or other things. They may again create. So, if you want to create a real good prediction about such plumes then the best thing is as best inputs you can provide, better the interpolate surface is going to be.

(Refer Slide Time: 02:49)

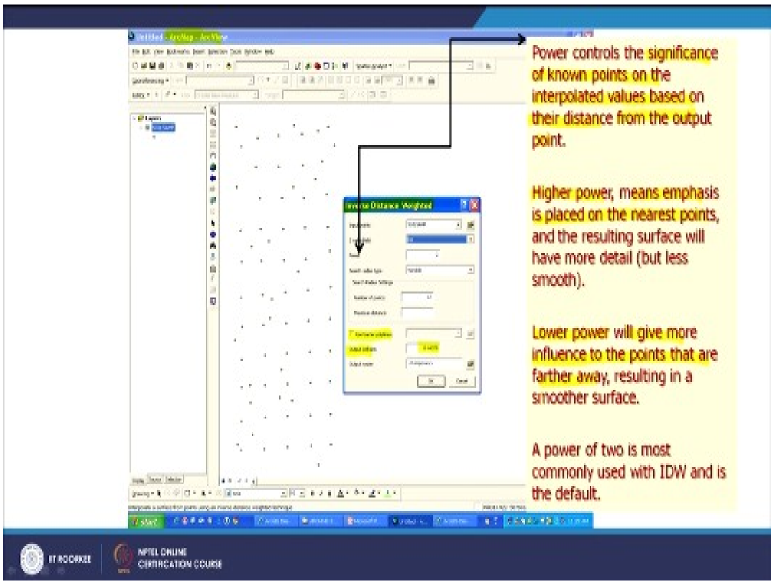


Now, for a linear example as we have seen that first you know if it is falling whichever the point is closed will have the maximum influence. If I take a little different example, here for this unknown location where question mark is there and these are my input information which I am having and these are the distances which are marked as like 40, 20, 100 and 200. And my points are these A, B, C, D.

Now for this unknown location, I want to predict that means this nonlinear interpolation techniques or especially this IDW technique; inverse distance weightage technique, the point which is having the minimum distance, the point A is having minimum distance to unknown location will carry the maximum weight while calculating the value for this unknown location.

So, in IDW, what is the first step? That distance from each neighboring point is measured. Like here in this schematic, it has been shown. Now number of points to include in the search can be selected either in form of number of points or in form of radius of a circle. We will see some examples also then the distance is calculated according to the formula.

(Refer Slide Time: 06:40)



And how these options are available? So, this is an example from ArcView GIS. The green dots are my input points. If I choose this inverse distance weightage method then what I am seeing and the z value, I am taking here as a pH value and I am providing power and other things which are different options are available. Search radius I am saying variable and it should search minimum 12 points around my unknown location for that cell.

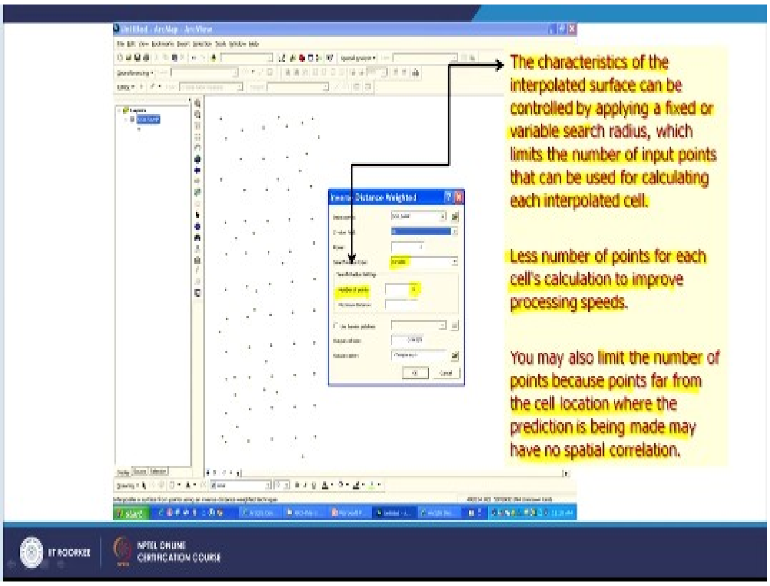
And further details are there and I can also control the spatial resolution of my data and if I am having prior knowledge about the terrain of that area then I should also provide the barrier information which will bring abruptness in my data and that would be more close surface representation then a surface representation without barriers. So, if we want to create more realistic surface, all these information's especially about the barriers is very much required.

Now when we say about the power, what does it mean by power here in this particular method? So, power basically means that it controls the significance of known points on the interpolated values. So, you provide high power instead of say 2, you provide 4 that means the input point will have more weightage while determining point in the surrounding for unknown location. So, this based on their distances of course, this is inverse distance.

So, everything is decided based on the distance. Less the distance, more than influence from the output point. Now when we provide higher power means emphasis is placed on the nearest points. When we provide low power, it will give more influence to the point that are further away and that means smoothen surface. But in case of high power, we get less smoothen surface. Though, the overall technique is inverse distance.

So, a power of 2 is most commonly used and that is why it has been kept in default here with the IDW and therefore, it is default here.

(Refer Slide Time: 09:14)



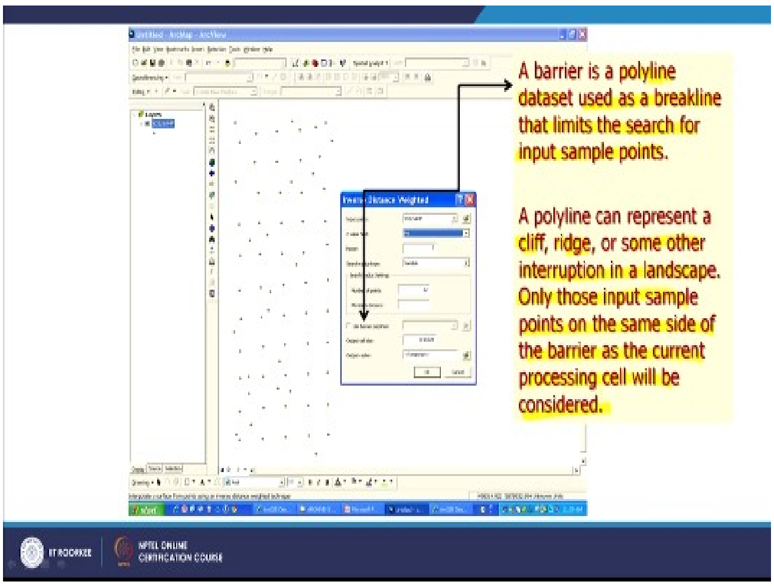
Now further if we go for this search radius value here in this example, we have kept as variable but when we keep a variable or there are some other options are available. So, these characteristics of the interpolated surface can be control by applying a fixed or variable search? So, we can choose 2 options which limits the number of input points that can be used for calculating each interpolated cell.

So here in this example, I kept variable but if I fixed then only those number of points in the neighborhood would be searched. So, again here if I want more smoothen surface then I will go for variable but if I want more you know, rigid surface; less smoothen surface then I will go for fixed radius. So, a smaller number of points for each cell calculation to improve processing speeds.

Of course, if the search for a smaller number of points is involved then the processing time would be less. Otherwise, software point of view, they have kept a value 12 as a default value. One can also limit the number of points far from the cell location where prediction is being made have no spatial correlation. That means the points if they are very-2 far, though we may keep a large number.

But as you know that they should not have any influence for that cell for which we are determining the value. So always there has to be optimum number of points or size of the radius which we choose. Considering very far points, basically you are creating a predictive surface which may be very far away from the real situation or reality. So, one has to be very careful while choosing these options.

(Refer Slide Time: 11:20)

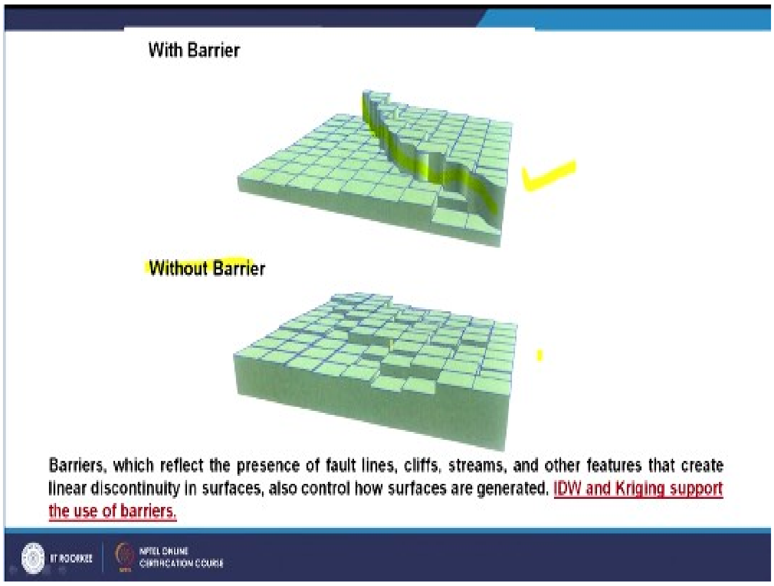


Now next is the barriers. Barrier's part, we have already discussed a little bit but anyway, it is a polyline dataset. It is not a point. Barrier always has to be a polyline input file based on the brake-line, also called brake-line which limits the search for input sample points. Basically, during the interpolation if you have provided a polyline theme which is representing say, geological fault then beyond that line, no search would be made.

And this polyline as we know that can represent in topography cliff, a ridge or some other interruptions of landscape. Even a stream in case of groundwater or surface water can be considered as a brake-line and only those points samples on the same side of the barrier as the current processing cell will be considered. So, this controls your surface.

And when you create a surface using barrier lines and you are having prior knowledge about that then your surface would be closer to the reality then without using barrier lines. So, in any application if you are using such surfaces whether for groundwater, hydrology or pollution or any other thing if you are having that information, bring that information in form of polyline theme and provide that input and you can create a very good, more realistic surface.

(Refer Slide Time: 12:59)

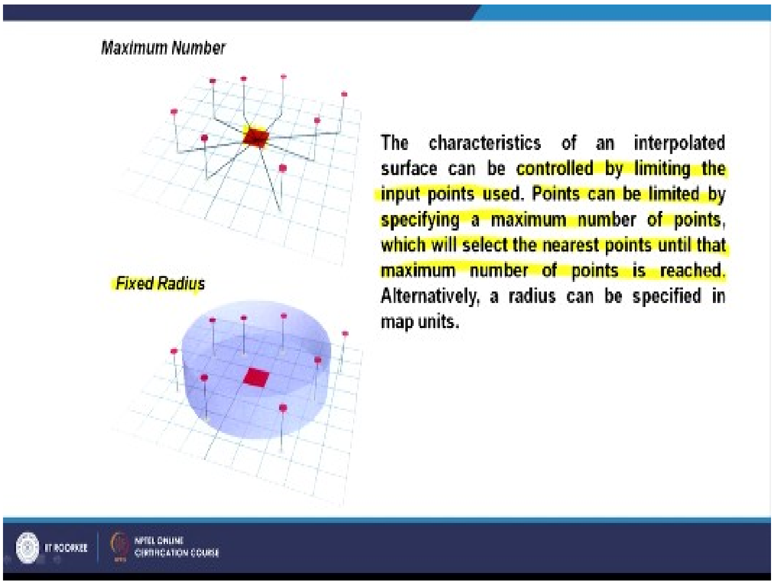


This is how it influences the brake-line. If I provide the barrier information then this is how my surface in 3D would look and without barrier, this is would be a more or less quite a smooth surface. So, it is more real, closer to reality than this one. If they are in nature, in a topography there might be a vertical cliff like this and if I am not using barrier, I would be predicting their surface like this which is grossly in accurate.

So, in order to have better accuracy of our surface, it is always if possible, provide the barrier information while doing the interpolation. So, from geological point of view or from civil engineering point of view, fault lines, cliffs, streams and other features which are discontinuous lines polylines and which controls how the surfaces are generated. Now, there are various interpolation techniques.

So, IDW and Kriging will support the barriers. Like if somebody is trying to use barriers in case of linear interpolation like Thiessen polygon then it is not possible because Thiessen polygon by its characteristics itself is a abrupt interpolator and therefore, such options will not be there in the softwares either because theoretically it does not fit there.

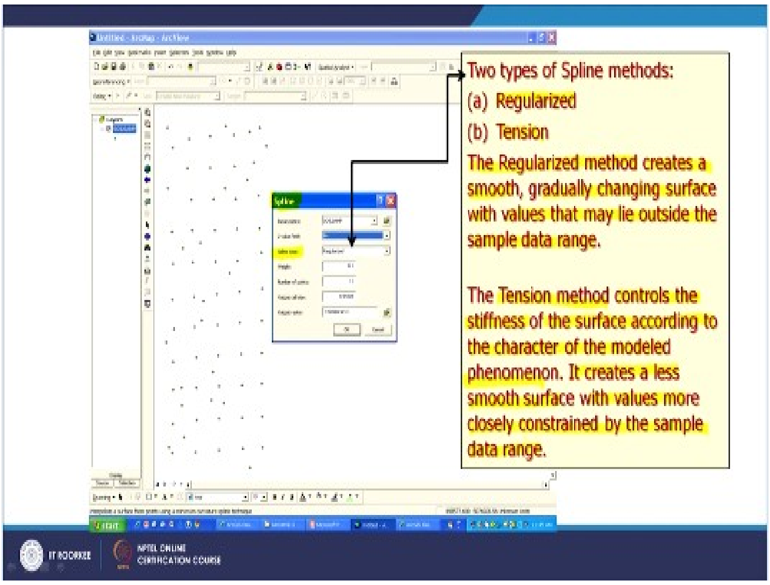
(Refer Slide Time: 14:35)



Now about the search radius, so if we provide the number of points that search 12 surrounding points or 5 surrounding points to determine the cell value of this grid then it will do the search up to the number of points. But if I have provided a fixed radius then it will search based on the radius rather than number of points. So, both options are available. And now these characteristics about the search radius is an interpolated surface can be controlled by limiting the input points used.

Points can be limited specifying the maximum number or points which will select nearest points until that maximum number of points is reached. Alternatively, a radius can be specified in map units also, as you can see in the bottom figure. So, both options are there.

(Refer Slide Time: 15:29)



For example, here this is you know which type of your interpolation is. So, this search radius can be controlled either with maximum number or giving a search radius. Now, there are some other parameters which are also considered and if I go for another type of interpolation technique which is called Spline; of course, this also falls in the category of nonlinear interpolator. Here, we are taking the same input layer of points.

Now here, options would be a little different because it is a different interpolation technique. So, first option is about which type of Spline because as I mentioned earlier that each interpolation technique especially these non-interpolation techniques will have their own variant also. Now which is most suitable for your dataset is a very difficult question that we will be discussing later.

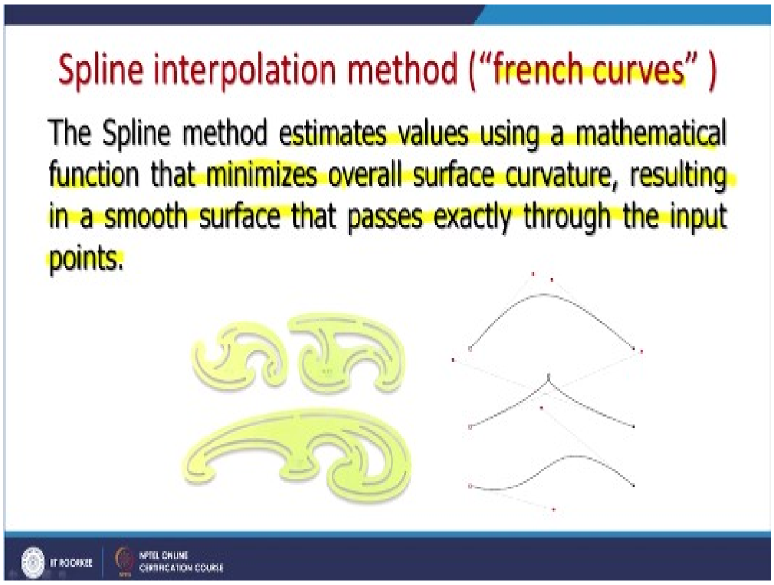
But for the time being if I say that I am going to use Spline instead of IDW interpolator then there are various Spline. So, first if I say 2 types; one is the regularized and another one is tension. So, 2 options would be available. So, if I choose the regularized method which creates a smooth, gradually changing surface values that maybe lie outside the sample data range.

So, regularize is option in Spline will create a very smooth surface relatively compared to the tension method which controls the stiffness or abruptness of the surface, according to the character method of modeled phenomena, whatever whether say elevation or concentration of any other parameter. And it will create of course, less smooth surface as compared to regularize with values more closely constrained by the sample data.

So that means that each interpolation method whether it is a IDW or Spline or Kriging which we will be discussing, all will have their own variants. And while choosing these variants, one should fully understand what are there and what are in the defaults. Because this software people may put something in default but that may not suit for your dataset. So, one has to be very careful while choosing.

Now this Spline interpolation method based on a concept which is called French curves, say older technique when we did not have basically the digital techniques then at that time people used to have these French curves.

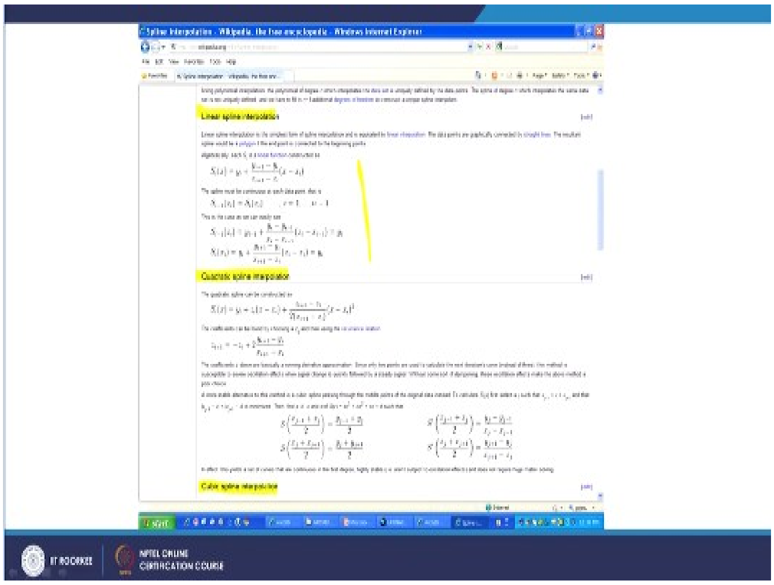
(Refer Slide Time: 18:27)



Still, you might see that these tailors when they cut you know, cloth to fit a pant or trouser for a person, they use these curved scales, you might have observed that one. Or while preparing the coat, they use curved scale to bring that curvature in the cloth, so that it fits very well with the body of a person. Same concept here about these French curves, so in Spline method, estimates the values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points.

Like here, these are the examples of French curves. Earlier architects when they did not have the softwares or civil engineers, they used to imply these French curves use to made of acrylic seeds or wooden seeds, you know. So here, what they doing Spline that it is minimizing overall surface curvature and what it doing is creating smooth surface. Here we are seeing in 2D but this is how the smooth surface can be created like this also here.

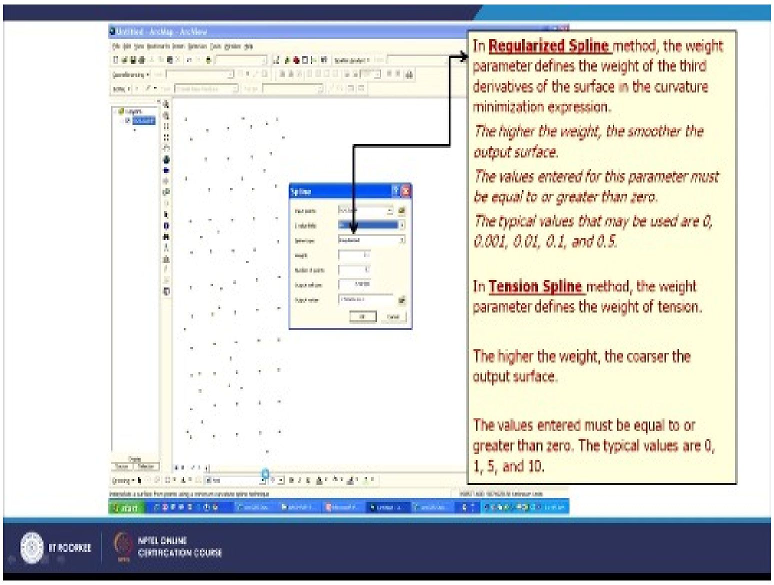
(Refer Slide Time: 19:37)



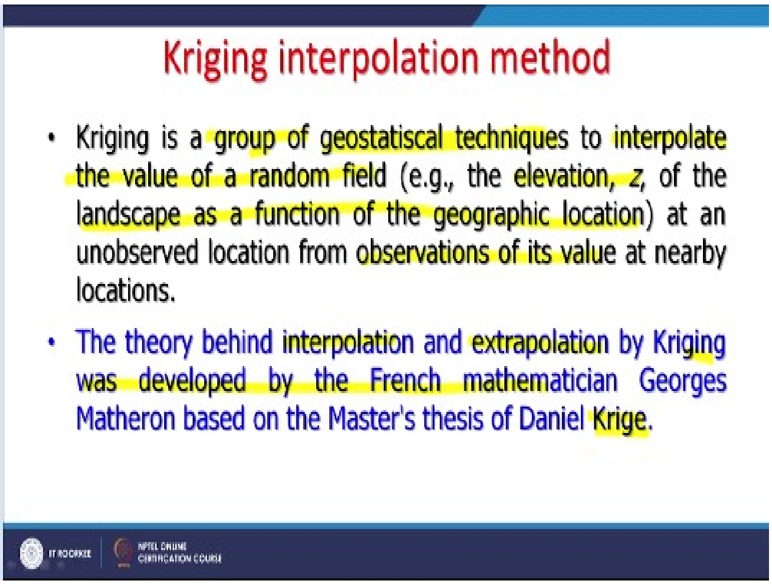
Now if one would like to learn what is mathematics behind against these all methods and their variations or variant then all that is available in standard softwares or elsewhere on the literature. For example, if I choose the linear Spline interpolation then this is the mathematics which is involved there. If I go for other options like quadratic Spline or cubic Spline interpolation, I am having complete mathematical background available to understand or to develop further.

Many of such options may not have been implemented today. So, we can further develop individual interpolation techniques to suit our datasets.

(Refer Slide Time: 20:33)



(Refer Slide Time: 20:43)



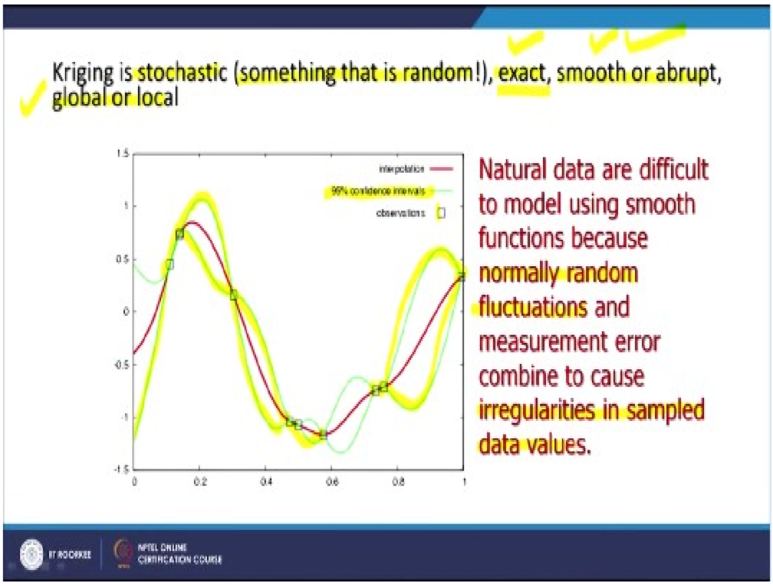
Now, there is other this regularized and this tension method. Tension apply; we have just discussed. So, we will go for the next interpolation technique which is called Kriging; very-2 popular and when we did not have the proper GIS software, we used to resort to a software which is called surfer. And this surfer is very powerful and very good for Kriging interpolation.

But now, GIS softwares are also capable of producing raster surface using Kriging and their variants versions of these Kriging are also available. So, basically Kriging is a geo-statistical technique which is a group of geo-statistical technique that means various variants of Kriging are available, to interpolate the value for the random field or for an unknown location.

And that value can be elevation value like z value of the landscape as a function of geographic location at that unknown unobserved location from whatever the observations which you are having in the nearby locations. So, it is more or less in the same way as other interpolation techniques but it is based completely on geo-statistical techniques.

And there are certain advantages with Kriging interpolation which we will also see. So, what theory says that the interpolation and extrapolation by Kriging which was developed by a French mathematician whose name was George Matheron based on the master thesis of Daniel Krige. So, this term Kriging is coming from a master thesis and it was a really great contribution in the field of mathematics, especially in geo-statistical techniques and spatial interpolation.

(Refer Slide Time: 22:40)



Now Kriging, if you recall that we had different kind of categories or different way of you know, classifying different interpolation techniques. So, in that sense the Kriging is a stochastic and which is incorporates the randomness in it and in nature, many surfaces may have that phenomena and also, it is exact, smooth or abrupt depending on what choice we go and global or local, those options would be available.

So, smooth and abrupt option would be available and global and local would be available but exact is very important here that is why a lot of people prefer. So, that it provides both exact, smooth or abrupt, global or local and is a very good plus it also estimates the error. For other interpolation techniques, it is not possible easily to estimate errors but in case of Kriging, it is possible.

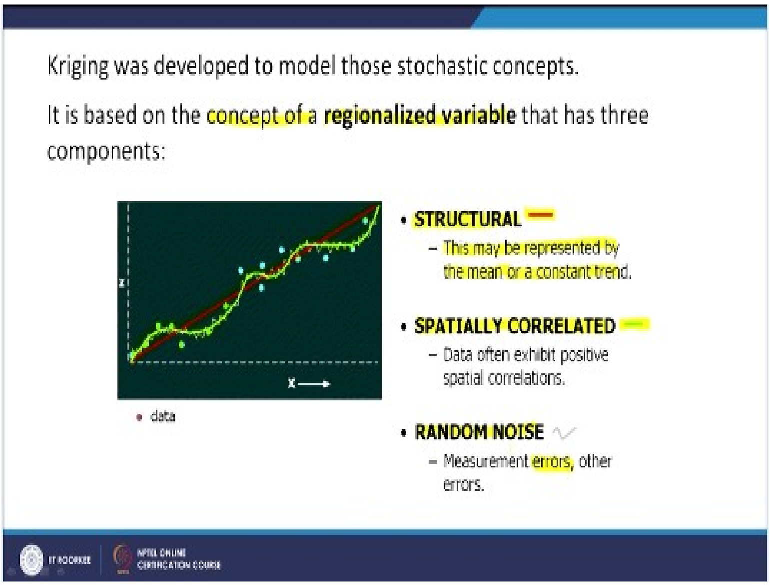
This is what this figure is departing that interpolated surface is shown in the red curves or red lines and then you are having input points and of course, this is a 2D linear or 1D representation like this, not really as the surface but in a linear fashion. So, these are the observations in blue boxes. But if I keep 95% confidence interval then there may be a surface which is going like this Kriging surface.

Or there may be a surface, though they are exactly passing through our observation points, that is why it is exact. Same time, it is smooth as well. So that between these 2 points, it is following some smooth path. So, it is a smooth. If we can constrain this one then it can also become abrupt. And if we change the radius or other parameters then it can become either global or local. So, Kriging is a quite powerful interpolation technique.

So, as you know that natural data; maybe elevation or any other characteristics of nature data, which are difficult to model using a smooth function because nature is may not be like that only. Because normally random fluctuations exist especially the terrain like Himalaya, I have given examples. If I am predicting a surface for terrain representation then I cannot go for a smooth surface say, it is a random fluctuation.

And measurement errors will combine to cause irregularities in the sample data values. So, chances of errors are much more in case if this is not random fluctuations are there. And therefore, we have to be very-2 careful about this.

(Refer Slide Time: 25:55)

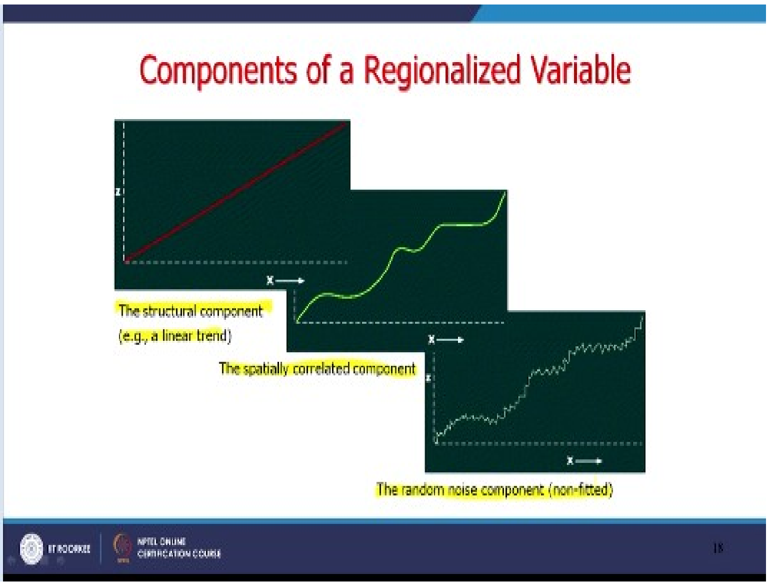


Since Kriging has been developed, as we have discussed is based on the stochastic concept and it is a regionalized variable say, based on the concept of regionalized variable which has got 3 components. The first one is of course; your data is there. Like here the data is shown in these dots. So, the first component is the structural one which is shown in the red line in this diagram.

So, this is a structural component of a Kriging surface which is completely a straight line or you can say linear. So, this may be represented by a mean or a constant trend. So, a structural component is a linear kind of thing whereas spatially correlated component which is represented in this diagram as a green line which you are seeing as a thick green line like this and the third one is the random noise.

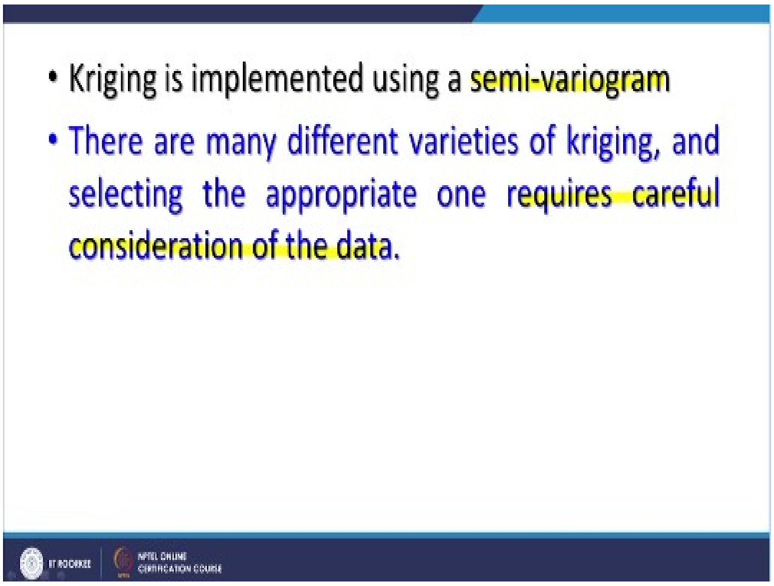
Now, this random noise is errors and this is what the advantage that it can also predict errors when you go or resort to the interpolations. So, we can separate out all 3 components or key components of Kriging; one is a structural, spatially correlated and random noise.

(Refer Slide Time: 27:32)



So, we continue with this that earlier figure, we can separate out as the 3 components. So that the first one is the structural component which as I mentioned is a linear trend; linear one. This is a spatially correlated component and last one is the random noise component which is non-fitted in that sense.

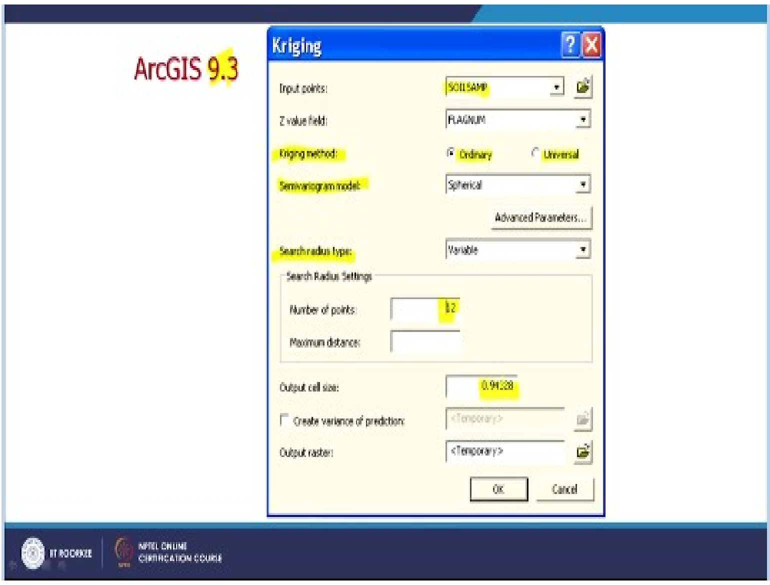
(Refer Slide Time: 27:57)



So, Kriging is implemented using a semi-variogram and there are many different varieties of Kriging means variant of Kriging exist. And many of such variants have also been implemented in many GIS software. Selecting the appropriate one, which one is the most suitable for my input dataset is may be sometimes very challenging. And therefore, it requires careful consideration of the data.

And if we are having prior knowledge of the terrain conditions or characteristics of my input data points then I can choose otherwise no literature or no software will ever let you know, that choose this variant of the Kriging or these options within the Kriging. No, nobody will ever tell you. So, it depends on your data type or data input or characteristics.

(Refer Slide Time: 28:56)

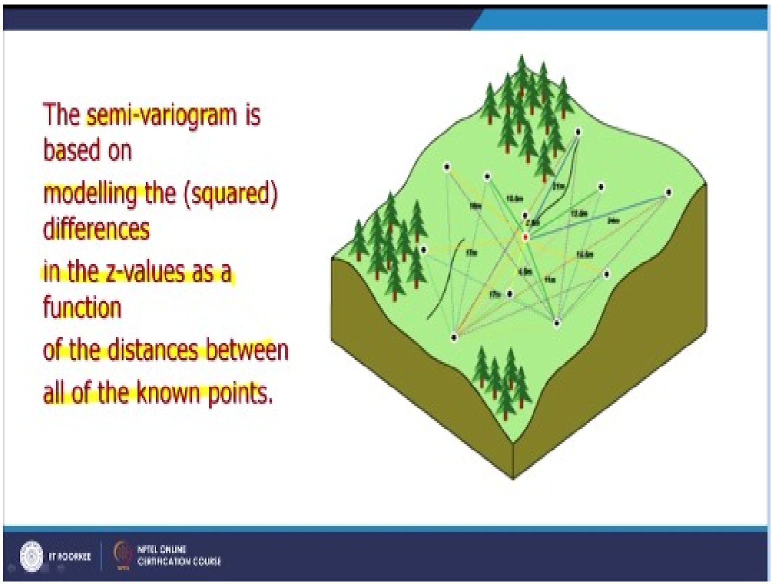


Now if I take example, though it is little older version of example but does not matter even in the latest version of ArcGIS, almost same options are available. So, here I am taking an input which is a soil sample data and I made take a value here say pH value; currently it is just displaying other value. Now 2 Kriging methods variants are there; either ordinary Kriging or universal Kriging.

And as name implies, universal will create more smoothen surface and if we go for ordinary, it will create a more you know local fitting surface. Now, semi-variogram models can be various options as spherical or circular or many other. And then of course, the search radius type which we have discussed in other interpolation techniques almost the same way, variable or fixed; 2 options would be there.

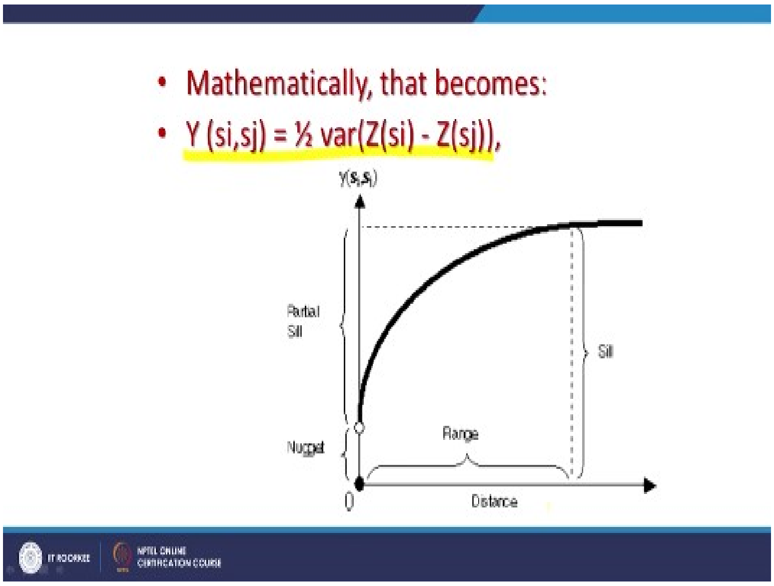
And of course when you are creating a raster surface from input point data, you can always control spatial resolution. By default, it may give some but may not be acceptable to you and then rest of the things you provide where to be stored and other things.

(Refer Slide Time: 30:17)



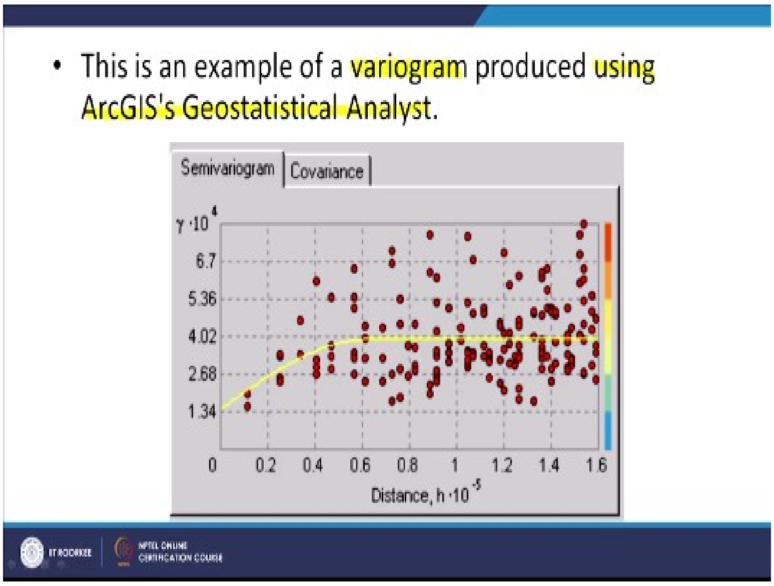
So basically, let us say denote some time about what is this semi-variogram? So, semi-variogram is based on the modeling that is that squared differences in the z values as a function of the distance between all known points. So based on this input, semi-variogram are modeled.

(Refer Slide Time: 30:41)



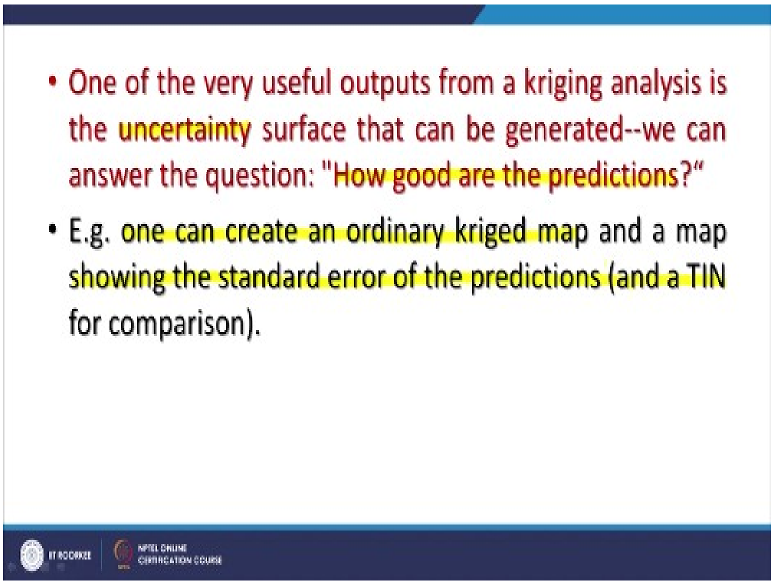
Now, mathematically, basically that becomes as per this equation and as we go from the distance, this becomes quite flatted. So, that is a Tobler's law of geography.

(Refer Slide Time: 31:00)



And if we see a further example on this variogram or semi-variogram produced using say this geo-statistical analyst of ArcGIS then here what we are getting a semi-variogram like this against these input points. You can also calculate other parameters which may be required to better understand the generated surface for that purpose.

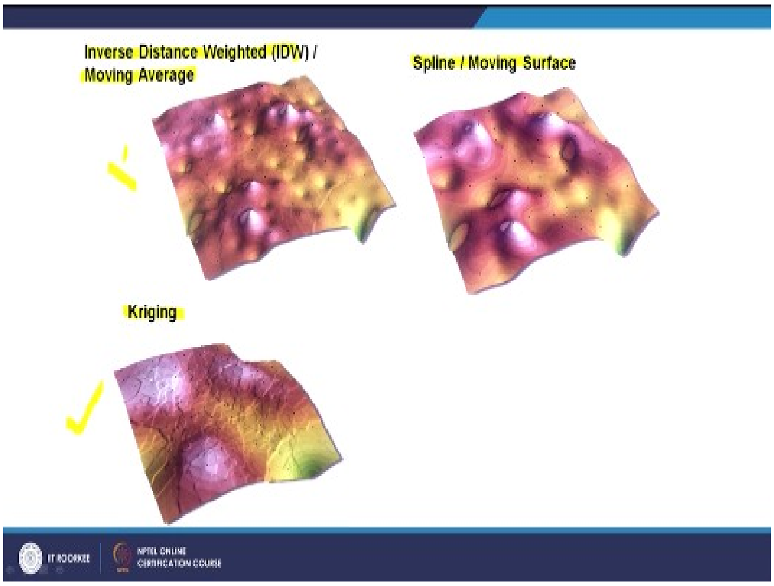
(Refer Slide Time: 31:29)



Now one of the very useful outputs from Kriging analysis is that error part or uncertainty surface that we can generate. And this basically will answer how good the predictions are or how good the surface which you have created which is closer to the reality. Though, there are various techniques like IDW, Spline, Kriging and their variants, Thiessen polygon but except this technique; Kriging, no other technique will let you know that what is uncertainty factor.

So, that is why many people resort to the Kriging. And one can create an ordinary kriged map and a map showing a standard error of predictions and you can compare with TIN also.

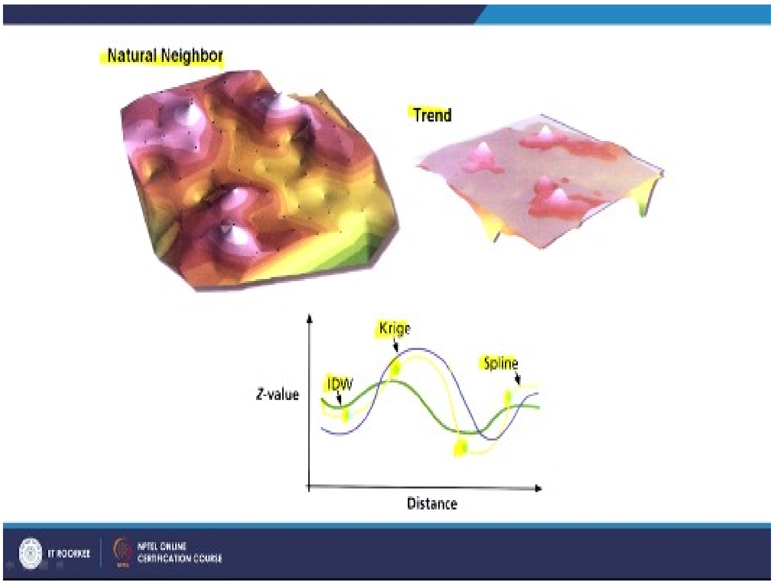
(Refer Slide Time: 32:24)



Now here, using the same input points 3 surfaces have been created using the 3 nonlinear interpolation techniques which we have just discussed. So first one, we discuss the IDW technique and the option we chosen is moving average and this is how the surface though, this surface is shown as a 3D surface, does not matter. So, here this is how the surface will look but if I choose the Spline and the option is moving surface; see the moving average, moving surface; different concept.

So, if I use this option in Spline, this is how my surface is going to be created. Remember that the input point theme is the same for all 3 surfaces and lastly is the Kriging. If I create a Kriging surface using the same data points, this is how the surface would be created. Now which one is closer to the reality that has to be assessed. In case of Kriging, uncertainty information would be available but in case of IDW, that information is not available to us.

(Refer Slide Time: 33:40)

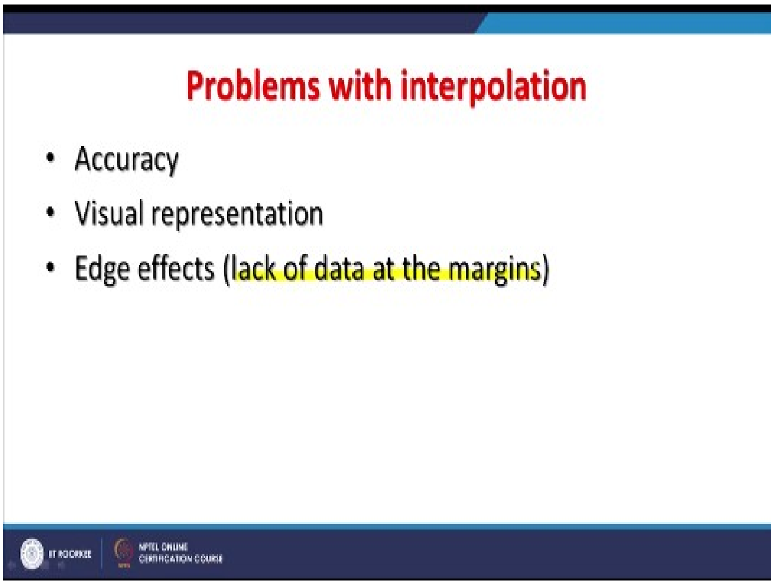


Now you know, there are various options like here natural neighbors then you can create a trend surface and if I say draw a cross section or profile for all 3 surfaces which we have just discussed and input points are also shown here as blue dots. What you are seeing that this blue one is the krige surface which is having more variations but sometimes, it is very well fitted between these 2 points.

But in case of Spline, as you can see this yellow one is exact fit. So, it is going and touching all input points whereas IDW is going in between points, not exact in this particular example because all these interpolations while doing through these softwares, you are having different options. So, if you do this, you may get various surfaces. And you have to be very careful while choosing which is the best for your input dataset. So, this is the problem.

The problem about the accuracy which surface or which interpolation is more accurate for my input dataset is very difficult or challenging but it is possible to assess to some extent.

(Refer Slide Time: 35:08)



Visual representations can be compared if you are already having some prior knowledge. Another problem will come about the edge effects that is lack of data at the margin or extent of the input data points. You do not have any input points and that means, it would be doing extrapolation. So, there might be some edge effects in the data. The best solution is, you define a large extent initially and later on you can extract for a smaller area and that will reduce your edge effects.

And basically, this brings to the end of now this discussion. So, this was spatial interpolation technique. We have discussed in 2 parts. This was the last part. Thank you very much.