Digital Elevation Models and Applications
Dr. Arun K. Saraf
Department of Earth Sciences
Indian Institute of Technology, Roorkee

Lecture – 06
How to assess quality of a DEM?

Hello everyone, and welcome to 6th lecture of digital elevation models and application course. And in this way particular lecture we are going to discuss how to assess a quality of a digital elevation model. This is very important. Because, a as a I have been mentioning, that there are lot of digital elevation models nowadays available free of cost on net at a different this spatial resolutions. And therefore, before we imply or before we use this digital elevation models, the quality must be assessed. There are a different ways of a assessing that quality, and while discussion about this spatial resolution, indirectly we have gone through a some products and derived from digital elevation model. And I have seen that howthis spatial resolution when it varies how the quality of a product changes.

But, here is specially we will be discussing about how to assess the quality of a specific DEM which we have downloaded for a particular area. Because, digital different digital elevation models which are available, have been prepared using or together different techniques. As you have seen that when we discussed availability of a different digital elevation models; especially, those who are available which are available free of cost, like U S D E M which was derived using survey to proceeds.

And then interpolation we have done. And therefore, there were problems about the seams at a one-kilometer resolution and so on; so forth. But, a when we go for 90-meter resolution a S R T M D E M, then we what we see is that, these are not good for hilly terrain; however, they are very good for plain areas. For example, they are not, may not be good for a small area of a Himalayan terrain.

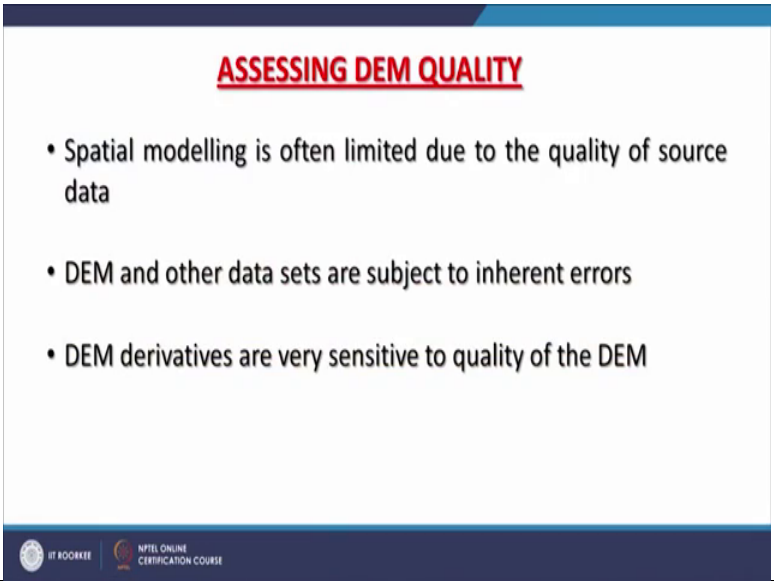
But they may be good for a flat area of indo gangetic plain. And contrary to this say your aster D E M, may be good forhilly terrain area, because of the technique is different, it is a stereo pair data, and may not be good for plain area and so on. So, these are the issues which are there associated with each types of elevation models, which are available and we will be seeing how to assess the D E M quality.

So, a spatial modeling is often limited due to quality of source data. If a, the data itself is having poor quality, that means, it has having lot of errors may be vertical errors, horizontal errors. Then, the product which we are going to generate using these models may be and slope map aspect map, may be indo gangetic network or may be some modeling for flooding or for a dam simulation or other things.

Then we are going to have problems. So, therefore, this is very important to assess. And a these are the D E M and other data sets are subject to inherent errors 2 types of the errors we can classify them. One is inherent errors which are errors which got a incorporated while data or digital elevation was being created and where are the errors which are operational errors which may erupt while doing the processing choosing a wrong processing step and not controlling the errors.

So, inherent errors here are will assess the quality of a digital elevation models whereas, operational errors will allow us to assess the quality of a product which has come through digital elevation model. But understanding about inherent errors is very much required. And that is why, in previous lecture I have mentioned, that we must go through the metadata, which is available through for each type of a digital elevation models withwhich will give us lot of details about these errors accuracy part and spatial resolution and other things.

(Refer Slide Time: 04:44)



So, D E M derivatives are very sensitive to quality of the D E M. If a digital elevation model is having lot of errors, and if you use the D E M to derive a product like slope map or denied network or any other thing, then the, it quality may further deteriorate, because these process at some stage may introduce some more errors.

So, if the raw product in this, cases digital elevation model itself is having errors, then these may propagate. So, one has to really control keep control over these errors. And errors in either remote sensing data or in G I S or a processing steps or in digital elevation model are generally least discussed. And therefore, we can mention that these are the errors are still ignored. But these are very, very important because they provide a statement about how accurate the data is or how accurate a particular product which one has generated using a digital elevation model.

So, that is statement or the confidence in our output will completely depend on the quality of a D E M, in which in also includes the error components. So, errors in D E Ms are likely to be larger on the shaded slopes. Because, if I am using the say radar data of for S R T M then, that is why, I am mentioned earlier, that for hilly terrain sometime we find especially the hilly rugged terrain like Himalayan terrain, we find that the area which got completely in shadow areas, because of oblique angle of radar waves or radar pulses. So, we may get errors above elevation values in a shaded slopes.

So, that is the errors which may there in case of a S R T M data. A smooth and featureless terrain, now this can also come in a optical data stereo data like extra D E M. So, this one has to take care about this one, that what kind of terrain this D E M is going is representing of which area this D E M belongs to.

And depending on that we should keep eyes on error part error component. D E M errors are especially variable and especially auto correlated; that means, they keep changing within an area, within a set of or one single file it may even change. So, that one has to remember. Another term which we comes while, assessing the quality of a digital elevation model is that uncertainty; that is, a model is an basically model when we say digital elevation model or any other model. So, it is a abstract form of reality or an approximation of reality, reduced reality; it is not truly a real; that is why, spatial resolution is there. So, courser special resolution we are reduced the reality much more.

When we are having a finer spatial resolution, we have not reduce reality to that extent are in previous example.

But any model once. We use the term model automatically means it is a reduced reality or a freak approximation or abstract form of reality. And it is a basically is down scale representation of the real thing. And as we know that D E M is also a model of real world features or phenomena, which is represent. The phenomena in this case is the terrain or undulation or you know a (Refer Time: 08:35) which are present ups and downs on a terrain.

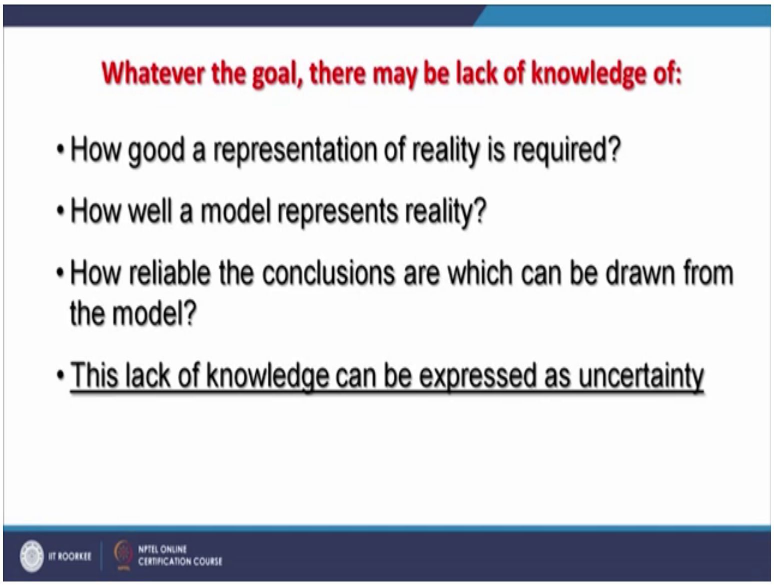
And the goal of a spatial modeling using digital elevation model is to produce close approximation of reality; that means, the products should not represent entirely a different thing then the real the terrain. For example, if I have derived a slope map, then slope map should look realistic to the real terrain conditions. And a if I have not taking care about the quality of D E M or the say z factor or a scale vertical accuracy horizontal accuracy, then with the product may be having lot of errors and may be much more away from this approximation of reality.

So, whenever whatever the goal, there may be the lack of knowledge of how good a representation of reality is required. Now, again it is depending on our output. As have been mentioning that they if we are covering a very large area, then we go for a may be a relatively courser resolution data. And they are the representation of reality is much more, approximation of reality or abstract form of reality is much more.

So, this has to be seen from that a scale point of view. And how well a model represents reality, that how good we have seen even the same resolution digital elevation model at drive from 2 different sources; say for example, a S R T M at 30 meter and G D M at 30 meter, but both are representing the same terrain a quite differently. Because the way they have been generated the quality got involved in it. And therefore, this a, these models which one is more accurate to reality of a particular part of the globe is more important.

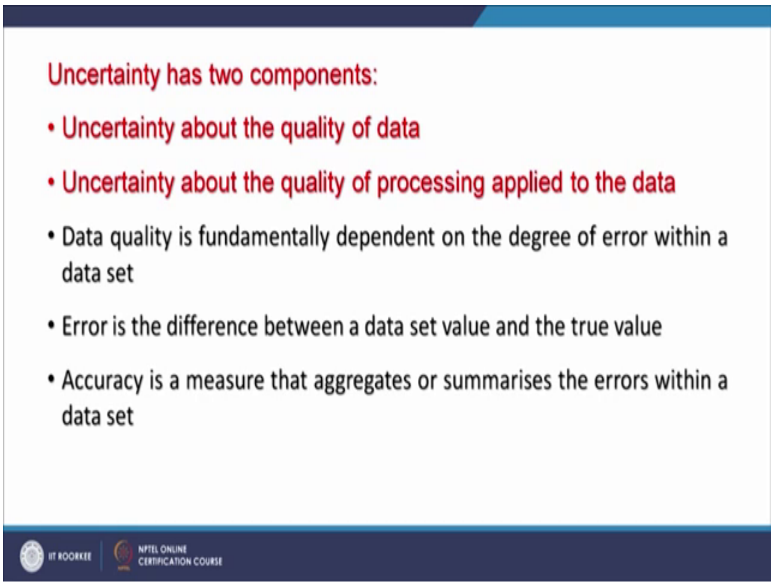
So, this say this has to be also understood very clearly. And how do I will the conclusions are which can be drawn from the model. So, if I am using a digital elevation model, example, I have been given is a if I have derived the slope map how close it is to the real conditions of a terrain. So, that a our confidence or reliability of that product will totally depend our about our understanding about the quality of a digital elevation model. And they if may be lack on of this knowledge can be, and this say can be expressed as uncertainty.

(Refer Slide Time: 11:15)



So, if we do not know much about the quality and error component and other things, then we put them as an uncertain. So, this is how it is expressed as uncertainty. Now, this has got uncertainty has got 2 components; the first component is a about the quality of data and second is the uncertainty about the quality.

(Refer Slide Time: 11:43)



Of processing applied to the data and earlier I have used to and term for errors inherent errors and operational errors.

So, the first uncertainty or a incorporates the inherent errors that is, the quality of data and they in the second uncertainty that is, the quality of processing, incorporate the operating errors. So, these 2 are important component of uncertainty and data quality is fundamentally dependent on the degree of a errors within a data set.

If we are having an large errors, then level of confidence goes down uncertainty increase and our quality become very poor, but contrary to this if we are having a small error components, then uncertainty is very little and we get high level of confidence in our data set and the errors are also very less. An error is the difference between a data set value and the true value.

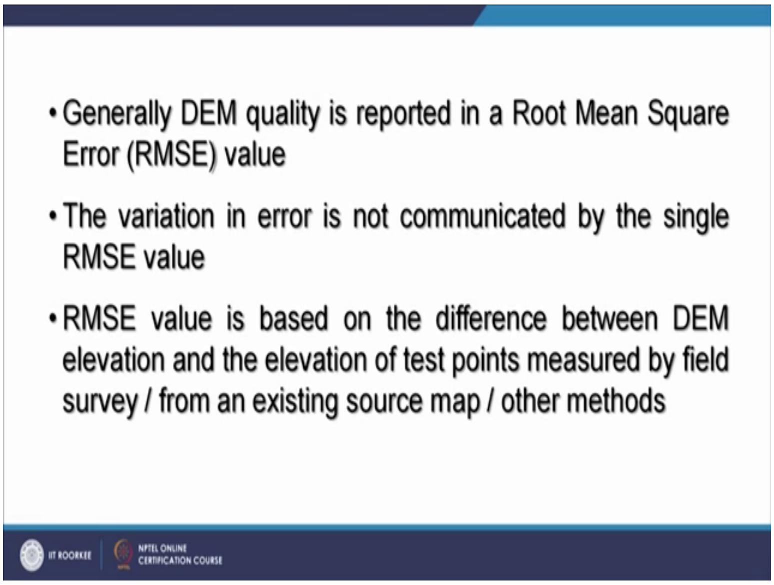
So, because they in this in our case is, if we want to really assess the quality of a D E M then we need to have to collect a near truevalue. Why I am using word near true and why not true value? Because, true for every location it is not possible to collect the true value.

Because, if it would have been possible then, there is no requirement of any model; that means, at one to one scale and model is approximation of realities. So, we have reduces the risk scale. We have reduced the reality. So, that is why, this error will be there in most of the data sets, but how it is small or large that is more important here.

So, the error if we want to define is the difference between a data set value and a the true value. For example, if a cell in a digital elevation model is showing elevation which is 100 meter above means 11, but if a the true value that is on the ground at that particular location, the average aerial average of elevation is a 90 meter then, we can say the error component is 10 meter. So, this is that is why it is important here to understand error.

In this sense especially related with digital elevation model and now that at the term which we have used earlier in the previous lecture, is the accuracy this is a major that aggregates or summarizes the errors within the data set. So, the accuracy is a statistical statement remember this thing.

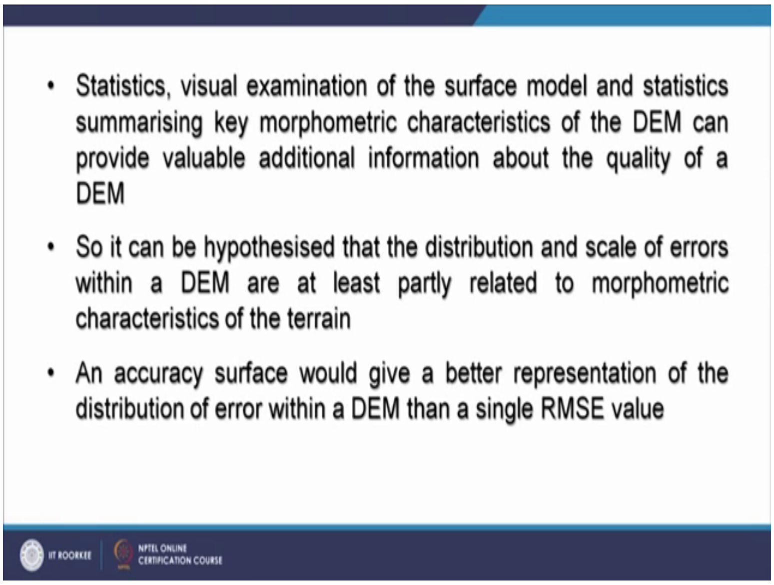
(Refer Slide Time: 14:33)



Now, generally D E M quality is reported in a root mean square error or R M S E value and that variation in error is not communicated by single R M S E value. So, that has to be also one single value because, as we have seen that a the quality of a digital elevation model very specially at as well as auto it is auto correlated. So, these that is why, a single R M S E value may not be a true representation of error component in the data and R M S E value is based on the difference between digital elevation and the elevation at test points measured by field survey or from existing map or other methods may be implying a differential glow on a G P S or similar techniques.

So, a different locations we may get R M S E differently because, especially these things vary within a data set depending on the terrain conditions. And now, when we see this statistics, the visual examination of surface model and a statistic summarizes key morphometric characteristics of a D E M and can provide value will additional information about the quality of a digital elevation model. So, a statistics visual examination, these will play very important role while assessing the quality. So, it can be hypothesized that the distribution and a scale of errors within a D E M are at least partly related to morphometric characteristics of the terrain.

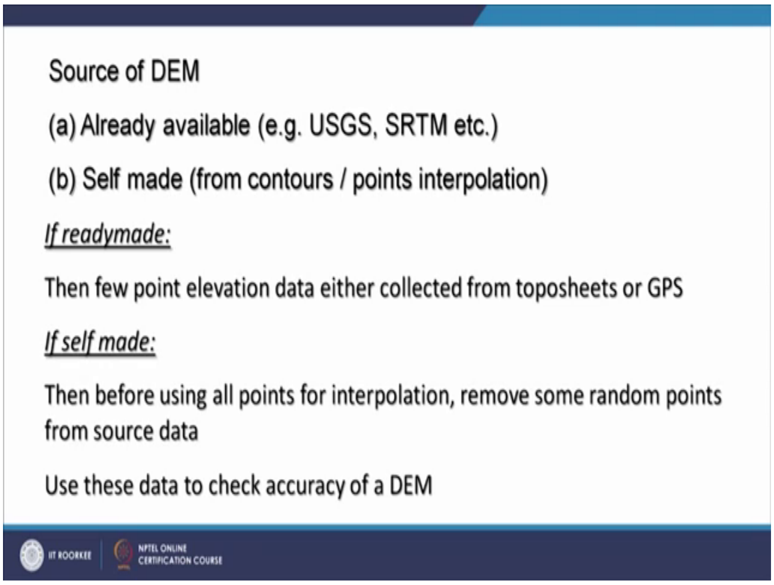
(Refer Slide Time: 16:09).



So, the features which are present will tell us that a how a the errors are there, what are the scale of errors and so on, so forth and an accuracy surface would give a better representation or distribution of error within a D E M then is single R M S E value. So, this is a equally important in case of a digital elevation model.

And while assessing the quality, as we know that a there is a the inherent errors, when we want to assess the quality of a digital elevation model, then we have to assess the inherent errors and before that we have to see the sources of D E M. So, for example, this is not a exhaustible list and the sources of D E Ms which we will be discussing in a separate lecture, but here just in a very brief some example 4 examples we have taken sources of D E Ms like already available.

(Refer Slide Time: 17:19)



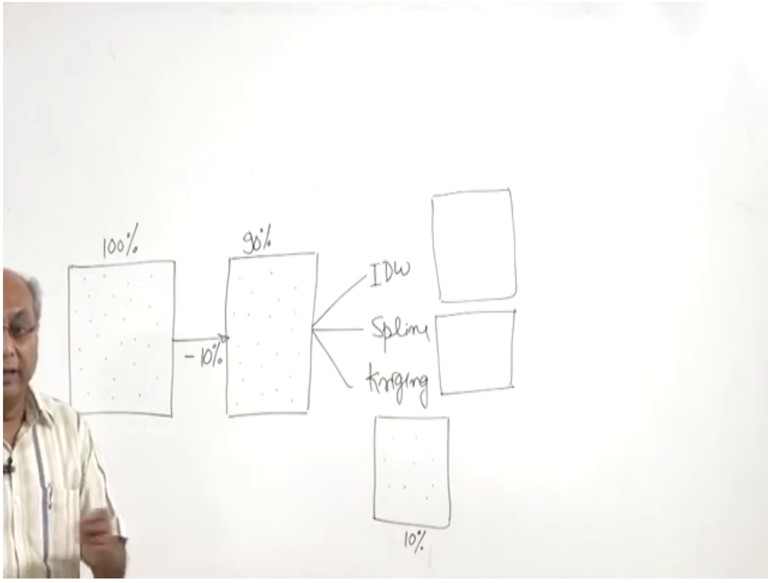
D E Ms for example, U S G S or S R T M or G T M etcetera and we may be having self-made D E M driving from contoured point data or may be collecting elevation values of point heights using a D G P S or different real G P S technique and then creating digital elevation model.

So, we will put them all in self categories if we use readymade, a digital elevation models or freely available digital elevation models then what we need to look in to they then few point elevation data either collected from toposheet or G P S. So, if we use these data sets, compare with the real values or the values which are being represented corresponding values, which are being represented in a digital elevation model, then probably we will have a idea about the quality of error component, especially vertical component of a digital elevation model.

So, how close it is, but if it is self-made if be our self are made, then we can control these inherent errors in a much better way. So, that then before using all points for interpolation remove some random points from the source data and I will explain on this if we are making our and that example will fits with the, it is this is spatial interpolation techniques.

So, what we do basically in a spatial? What I am suggesting through the quality? How to assess the quality of a D E M and how to choose a particular a spatial resolution and a this a spatial interpolation technique? We can devise a very simple method that, suppose a I am covering an area

(Refer Slide Time: 19:10)



And in which I am having 100-point heights for which I am having elevation values. These a I was supposed to create a digital elevation model implying a particular spatial interpolation technique, but a before that, what I can do that, I take a you know 10 percent points randomly selected points out of this data set, initially. So, what I will have, here I am having 100 percent points, I have randomly selected 10 percent of points and quarter and saved it elsewhere for further use. So now, I am having left with 90 percent of my elevation points for which I will be using for different interpolation techniques.

So, I can imply different interpolation techniques like maybe I D W, may be spline or may be kriging. Now what I will have I will have in this particular example, I will have 3 surfaces. So, I will have 3 surfaces using 90 percent of the data set; one surface of spline, one surface of kriging and one surface of I D W.

Now those points which I did not involve in interpolation the 10 percent randomly selected points are plotted here and the corresponding elevation values, 10 percent points are plotted here. Their corresponding elevation values are checked, how far they were away from an these values and you know out of say, in this example, the 3 interpolation techniques I have used. So, if I find that I D W values or the surface which has been generated using I D W is giving a more close values corresponding to these 10 percent randomly selected points,

Then what I will do, I will in the final step what I will do I will take 100 percent points and directly go for I D W interpolation and can create a surface. So, through this process 2 things I have achieved one is which interpolation I should imply for my set of data set.

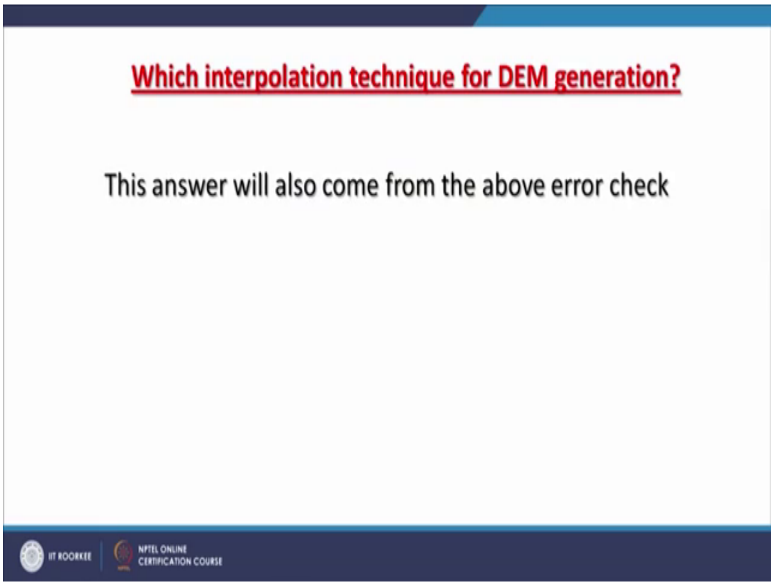
So, that is a the set of data that I can decide based on this exercise one and. Secondly, same time I can also access that what is the quality of the D E M which I am going to generate. Because, this will by when will compare these 10 percent random points with corresponding cell values digital elevation values then I will know exactly what is the error component here. So, this ray while doing this exercise, we can achieve 2 things; one is which interpolation techniques should be implied one and secondly, what is the inherent errors which I am going to have with my data set.

So, this a this is a rather very quick method how were one has to remember that for each data set we had to do this exercise again and again. Because, a one particular data set suppose a data set instead of elevation values, I may be having a groundwater surface for which I am having observation wells for which I am having water level data. So, for that set I D W may not be suitable.

I may find end of this exercise may be spline more suitable. So, for different data sets different interpolation techniques may be implied one and there may be different data quality. So, this is what it here means, then that if we are going for a self-made digital elevation model then, we can first perform this exercise, achieve 2 things together and then go for use the entire 100 percent data set and generate final digital elevation product.

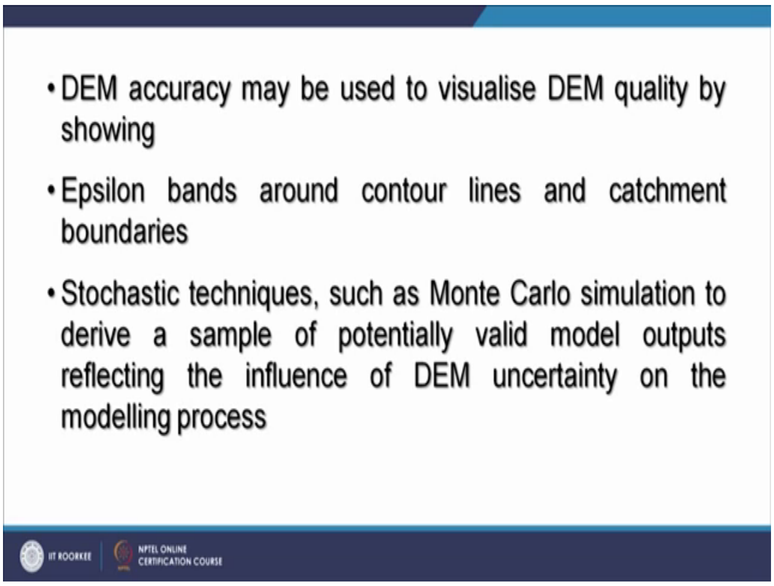
So, in this we will have full control over the quality of the data set, but if we are using readymade data set then we do not have much choice we have to live with that errors, but we must know what are the inherent errors and that how that can be assessed that we will discuss little later. Almost implying the same thing, but the data has to come from some other source, because if the data is coming from the same source then it will give it will not reflect about the errors. So, we will see little later. So, use these data check and to check accuracy of a D E M this these kind of checks or processes can allow us to check the accuracy of a digital elevation and now this I have already mentioned that which interpolation techniques for D E M is and.

(Refer Slide Time: 24:45)



This answer will also come from the check or on the process which I have already discussed so.

(Refer Slide Time: 24:48)



D E M accuracy may be used to visualize D E M quality by showing epsilon bands are around contour lines and catchment boundaries that is also.

Because, these boundaries can tell us that, how accurate a digital elevation model, this is another way stochastic techniques, such as Monte Carlo simulation to drive a simple sample of potentially valid model outputs, reflecting the influence of D E M, uncertainty on the modeling process. So, that also can be implied and quality of a D E M can be assessed using G C Ps ground control points or point heights here requires little discussion that if I am having digital elevation model.

I want it is a readymade digital elevation model I want to assess the quality of a digital elevation model and as mentioned just few minutes back, that we need some external source of data, may be a point elevation data derived from say digital these G P S or differential G P S or may be some other survive techniques if that a set is available with us the G C Ps ground control points on which we have got the high reliability, we feel that they are more accurately collected elevation points.

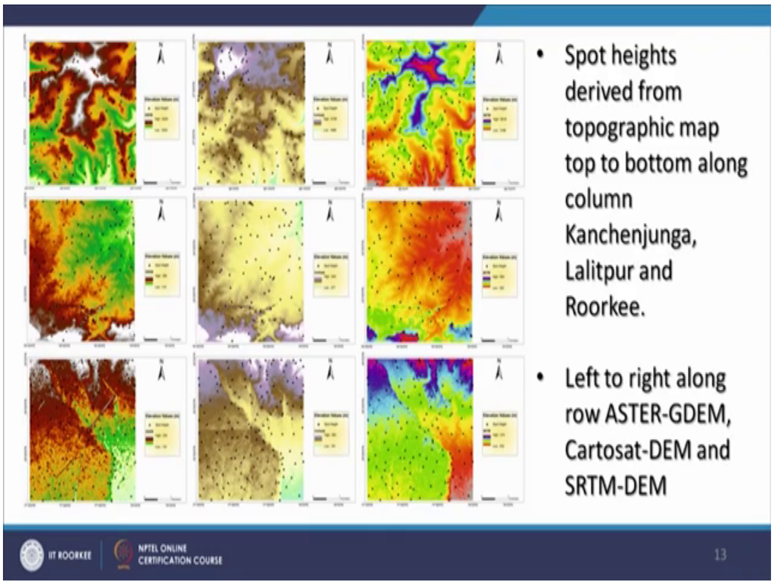
Then we can put again as we did here in this exercise that a those points and check the corresponding elevation of a digital elevation model and how close the elevations of digital elevation model are coming close to these observations which are coming either through differential G P S or some other survey techniques and by which this exercise this can be done as I understand that N R S say ISRO India has developed a large library of G C Ps. They have collected using differential GPS.

So, if a the this large library is made available then it becomes much easier for any one of us to use those corresponding point heights values derived from differential G P S and then check with the elevation of a different area or parts of India. So, because, then we can use these library of G C Ps as a standard one assuming that, they have while collecting these G C Ps they have taken utmost care and maintain the highest quality standards, but unfortunately these are not available.

So, we have to generate our own or bring from some other source these points heights value; the other source can be a toposheet, but when we are using a very high relatively higher special resolution digital elevation model and we are using point heights of 50 thousand a scale toposheet, some times that may not be very appropriate. So, here the scale or the sources scale of the point heights from a toposheet will matter most. So, one has to take care about this a aspect as well or relatively medium resolution or course resolution data and corresponding point heights coming from 50 thousand scale may be good, but for higher spatial resolution data and these heights for which these G C Ps are coming from 50 thousand this scale may not be good.

So, one has to remember when example here which is shown that a I am having a digital elevation models and derived a from variety of a sources here and a.

(Refer Slide Time: 28:42).



Corresponding point heights are also plotted here and what we see that a when we compare these with the point heights and the corresponding elevation values which are available and then when we compare then we find results. So, there are a different results are shown here against the different point heights. And so, what these spot heights derived from topographic map top to bottom top to bottom here which is of different areas this is along Kanchenjunga one set then, this is a Lalitpur area which is a less or very smooth terrain, note a much a undulations and then 3 bottom 3 images are of Roorkee area which are near flat.

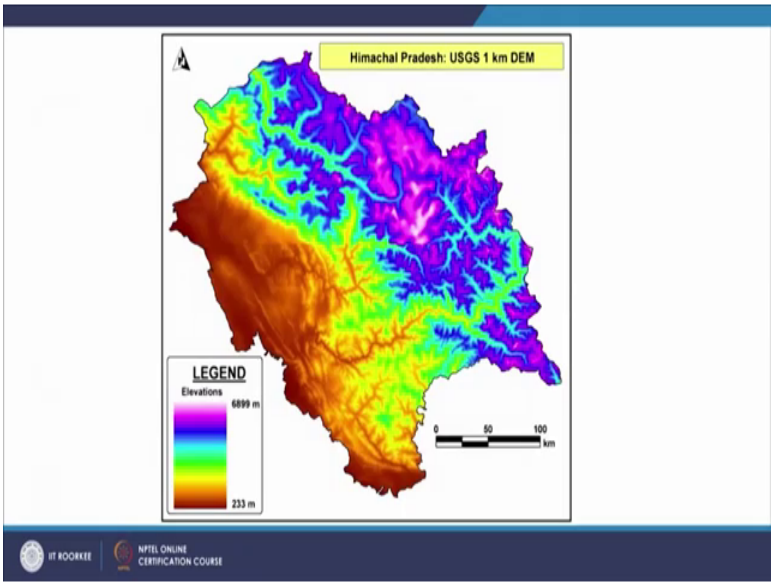
So, what we see is a highly rugged terrain and less rugged terrain and come almost a flat terrain and when we use the different for these use to assess the quality, then this is how this can be done that you get the spot height in in these 3 examples, the spot height have come from these topographic maps was basically survey of India 50 thousand scale toposheets and when we compared we got different results for different elevation models.

So, what we here, the first one at the first row is the aster G D E M for all 3 sides this is the (Refer Time: 30:27) D E M and this is the S R T M D E M. So, 3 different types of D E Ms at 3 different resolutions are here for 3 different entirely 3 different areas representing 3 different types of terrain. And so, as mentioned earlier, depending on the terrain, the errors would be there. It is under it is generally it is formed that the for highly rugged terrain the error component may be little higher.

But it depend also will depend what is the source of data, how the data has been generated, if it is remote sensing based digital elevation model, then which technique has been employed, but generally this is a this is has this has been observed that areas of hilly terrain rugged terrain, we may have large errors as compare to area which are less rugged or smooth you know terrain or completely flat terrain like indo gangetic plain part of India.

So, this a this a this way we can compare the point the point heights with corresponding elevation values from digital elevation model and can give can have a feeling about the quality of a digital elevation model. So, this way we can assess and quantify the errors a present, if at all inherent errors in these readymade digital elevation models.

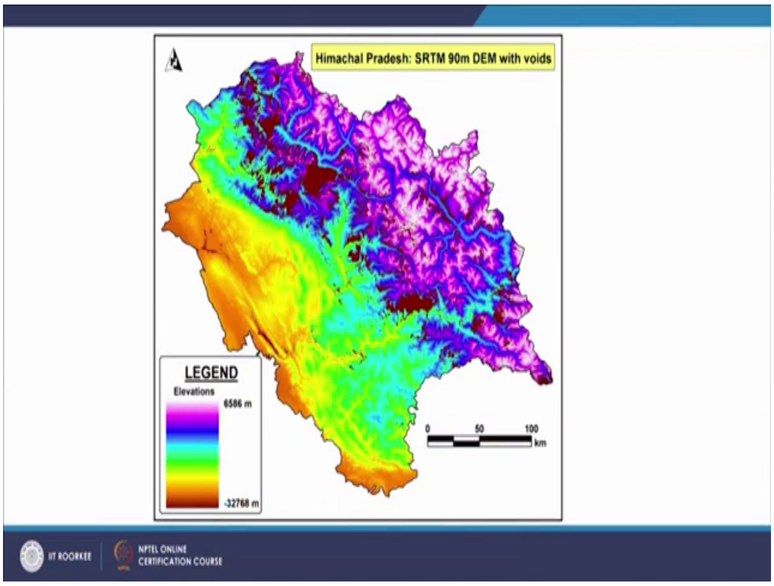
(Refer Slide Time: 32:04)



Now, one example we have already discussed earlier, but a since this comes here.

In this reference that if this is a readymade digital elevation model of one kilometer resolution and if we see very carefully what we find that we may get some seams present in such type of digital elevation model when we didn't have these S R T M and other digital elevation models at 90 meter this thousand meter or one kilometer resolution D E M was very useful especially of hilly terrain, but a it had the inherent error that was the scene issue apart from the interpolation and other errors were there and the another.

(Refer Slide Time: 32:52)



Problem came initially when we started using S R T M D E M with voids. So, voids means no data basically; that means, while these images radar images were collected for a snow covered on glaciated areas, no data was collected because of some shade over some and dielectric constant and they are would this voids which are represented in this particular example in dark maroon color that is why, the value is very void and ah.

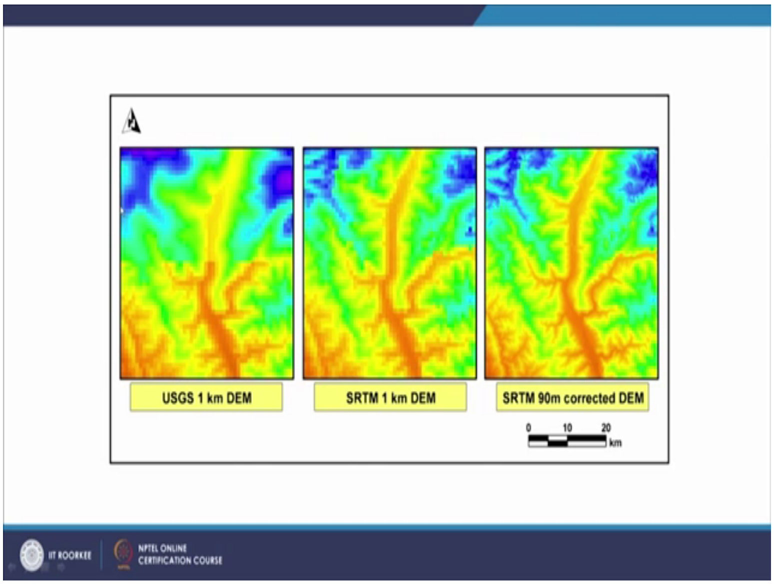
So, that we can declare to the system that when whenever it encounters this value in this particular case the value was assigned minus 32,768 meters so, whenever this value is assigned, this constant value then consider as no data and do not involve in any calculation whenever this value is encountered within a data file, but later on implying a simple G I S techniques overly in.

So, you in the background you are having relatively courser resolution digital elevation model for example, may be S R T M one-kilometer resolution which was also available or may be U S G S D M one-kilometer resolution and you can pick the relatively courser resolution digital elevation values from the low resolution digital elevation model for this 90-meter resolution and you can fill the voids and this is what the exercise was done. So now, you are seeing a completely void free digital elevation model at 90-meter resolution.

Now, later on for entire S R T M digital elevation model of this 80 percent of the glow the same exercise was done and the voids where removed, but remember that a these voids we are present in especially a snow and glaciated part. So, and the data these voids have been filled with relatively courser a spatial resolution cells and therefore, they are suffering from certain types of error.

So, if I imply such digital elevation models where voids have been filed with courser resolution data then, we have to keep in our mind that whatever the products which we are driving, will also have these problems or these errors in inherent errors, but visually some time we do not see these errors, but we go if we zoom it, we may see these errors very clearly as this is a zoom part.

(Refer Slide Time: 35:36)



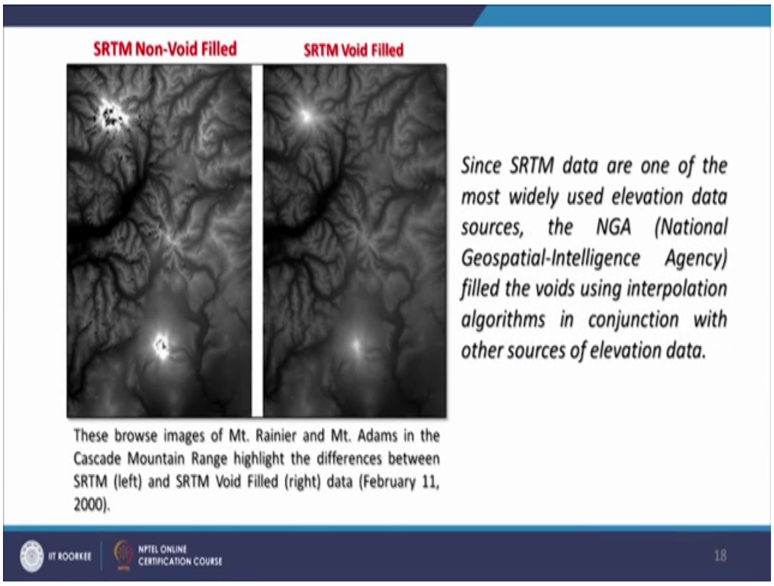
So, I am this what we are seeing the left image is U S G S one kilometer D E M and here the seam can be seen very clearly why seam is coming because 2 different toposheets may be of 2 different scales or may be surveyed in 2 different years and when these where Moser Baer made interpolations were done using contours these seams got developed and this is inherent error with U S G S one kilometer D E M is still these D E

Ms are been used because for different modeling we may not require a very high resolution relatively high resolution D E M.

So, be may resort to even one kilometer resolution another example at the same resolution one kilometer resolution S R T M here we do not see the seam luckily in this particular example we do not either seeing any voids as well, but when we see the S R T M 90 meter which is a void free correct D E M then as you can come you can see that a compare to U S G S D E M this is highly reliable.

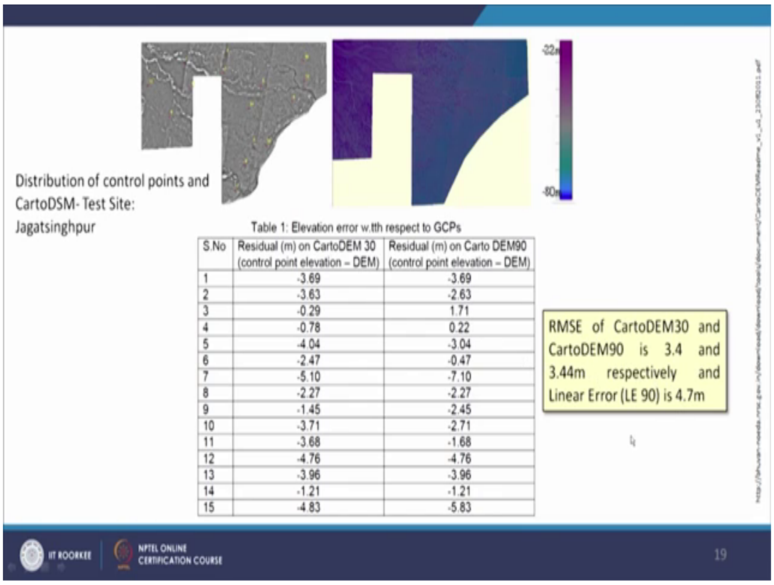
Means the quality of D E M is much much better than a one-kilometer U S G S D E M or one-kilometer S R T M. So, there are various ways either you can involve the visual inspection which can give you a you know a quick assessment of the quality of a digital elevation model zooming at and zooming out at different locations especially of heli terrain and implying some point heights and then checking the errors or some statistical methods by which one can do this thing. This is a again example of none void field digital elevation model of a hilly terrain and then S R T M void field D E M this is a Mount Rainier and 2.

(Refer Slide Time: 37:36)



Mounts are there Mount Rainier and Mount Adams are given here and this is how the after processing the D E Ms have been improved.

(Refer Slide Time: 37:45)

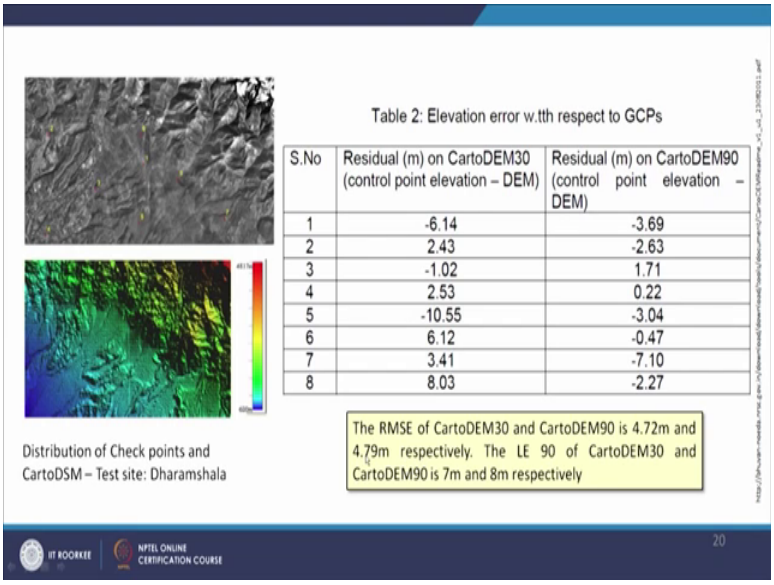


These control points as I was mentioning that a carto digital surface model this exercise where done using the this G C Ps library of N R S A.

And, this is one area where you can see at different spatial resolution what are the residuals or R M S E which you see for one set of data set. So, 30-meter resolution this is what these are the values which you are getting almost you are getting ah. So, you are underestimating the height value and whereas, in case of 90-meter resolution, you are again getting values some are varying for different.

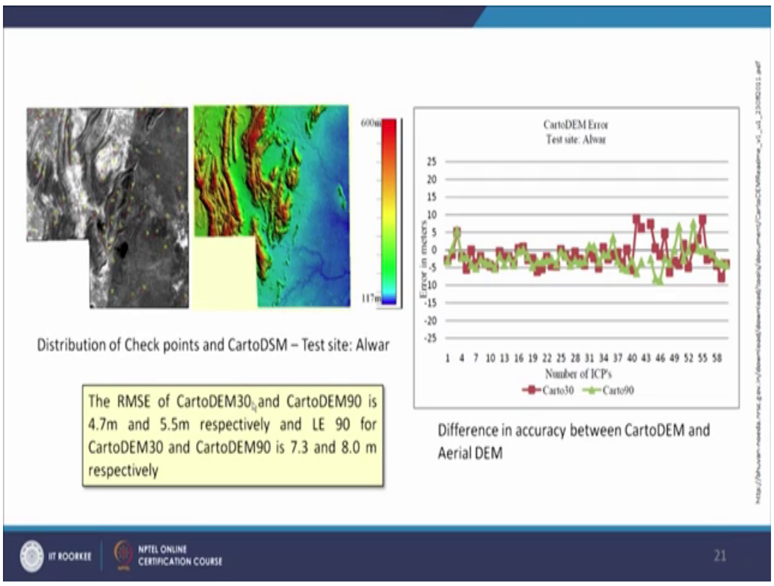
So, since the resolution has change and the therefore, these R M S E values have also change for with respect to different G C Ps as well. So, this is one way one example is here that R M S E of carto set D E M which is this one and R M S E is this is 3.4 and whereas, in this case this is 3. 44 respectively and a linear error is 4.7.

(Refer Slide Time: 39:02)



And other example of Dharamshala which is a hilly terrain and what you are seeing is again same comparison cartosat D E M 30 meter carto set 90 meter and this is what you see the error now as I been mentioning that in hilly regions you may get all together different values and here what we see that a this is the value in 30 meter R M S E is 4.72 whereas, here it is 4.79 say generally it is larger in of a heli terrain one more example of a of Rajasthan not as raised as Dharamshala and these a corresponding check points are also there.

(Refer Slide Time: 39:49)



So, this is the difference accuracy of cartosat D E M and aerial D E M which have been derived and again R M S E 30-meter resolution and 90-meter resolution 4.75 point. So, courser the resolution the error component may be higher R M S E may be higher of a digital elevation model of a hilly terrain again R M S E error may be higher.

So, that has to be kept in mind. So, this brings to the end of this assessment of the quality of a D E M and there are various methods which we have discussed in this one and a depending on the area on which you are working depending what is the source of digital elevation model and depending on the this special resolution over digital elevation model one should try to first assess the quality and then later imply to derive various products using that particular digital elevation model. So, this brings to the end of this presentation.

Thank you very much.