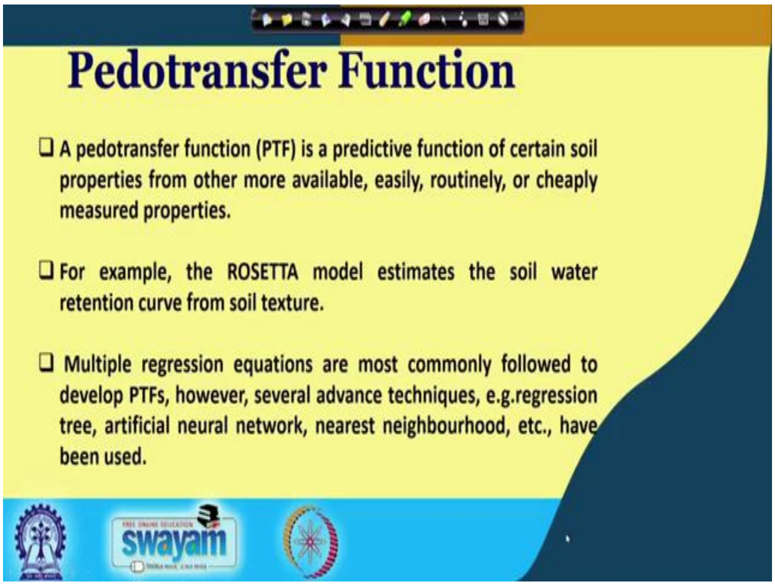
Soil Science and Technology
Prof. Somsubhra Chakraborty
Department of Agricultural and Food Engineering
Indian Institute of Technology, Kharagpur

Lecture – 60
Pedotransfer Functions and Uncertainty of DSM

Welcome friends, to this last lecture of Soil Science and Technology and in this lecture we will be consider will be covering 2 important topics one is Pedotransfer Function and Uncertainty associated with digital soil mapping.

(Refer Slide Time: 00:40)

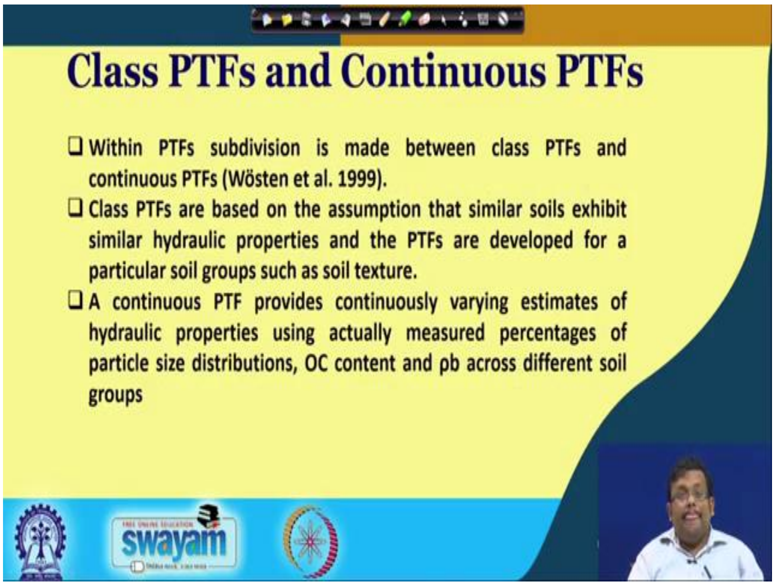


And, in the pedotransfer functions, remember that this pedotransfer function or PTF is a predictive function of certain soil properties from other more available, easily or routinely or cheaply measured soil properties. For example, the ROSETTA model estimates the soil water retention curve from soil texture. And, you know, multiple regression equations are most commonly followed to develop the PTFs, however, several advanced techniques; example regression tree, artificial neural network, nearest neighbourhood, etcetera have been also used.

So, in a nutshell this pedotransfer function is basically a mathematical relationship where we try to predict a certain soil property. For example, most of the time we predict different types of soil hydraulic properties; sometime we predict soil organic carbon and other properties from more available, easily, routinely, cheaply measured property. For example we try to measure the organic carbon based on the soil clay and texture analysis and also the other properties which are easily measured then we fit a relationship regression relationship then it is called the pedotransfer function.

And so, you know, for this fitting regression relationship we can use either linear regression techniques we have discussed about the simple linear regression, multiple linear regression and then we can also use some advanced method like artificial neural network, nearest neighbourhood method etcetera.

(Refer Slide Time: 02:12)



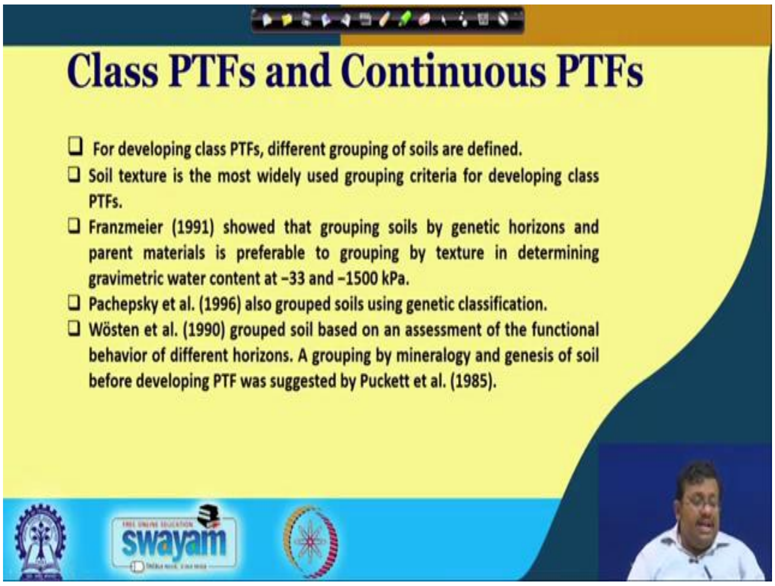
So, another important term is class pedotransfer functions and continuous pedotransfer functions. Remember within pedotransfer functions subdivision is made between class pedotransfer functions and continuous pedotransfer function. Class pedotransfer functions are based on assumption that similar soil exhibit similar hydraulic properties and the pedotransfer functions are developed for a particular soil group such as soil texture.

And continuous pedotransfer function provides continuous varying estimate of hydraulic properties using actually measured percentage of particle size distribution, organic carbon content and bulk density across different soil property across soil group.

So, again this class PTFs based on we generally assume that similar soil will exhibit similar hydraulic properties and then we develop the pedotransfer functions for a particular soil groups such as soil texture. For example, we assume that, you know, clayey soil will produce clayey soil will exhibit similar hydraulic properties then we fix a pedotransfer function or produce a pedotransfer function for that clayey group. So, this is called class PTFs.

In case of continuous PTF we disregard this broad classes we actually, you know, we take, you know, this PTFs we create the PTF provides continuously varying estimates of hydraulic properties using actually measured percentage of particle size distribution organic content. So, we cover we consider all the values across different soil group and then we create a universal model.

(Refer Slide Time: 03:47)

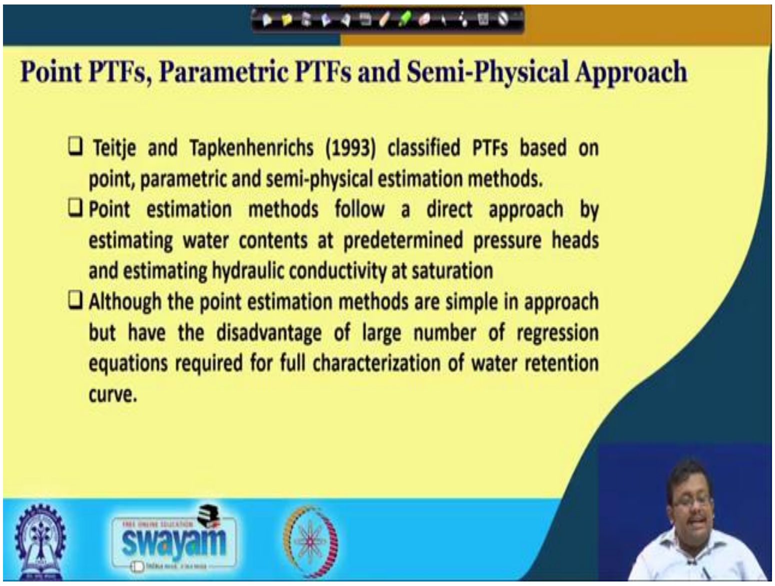


Now, class PTFs and continuous PTFs are also for, you know, it is also you need to remember that for developing the class PTFs different groups of soils are defined. Like soil texture, is a most widely used grouping criteria for developing the class PTFs. And this Franzmeier in 1991, shows that grouping soil by genetic horizons and parent material is preferable to growing to grouping by texture in determining the gravimetric water content.

We will be discussing, mostly focus on the soil hydraulic property. So, in this case you can see that this scientist showed that the grouping by soil by genetic horizons and parent material is preferable to grouping by texture in determining gravimetric water content at minus 33 to minus 1500 kilo Pascal. And Pachepsky et al. in 1996 also grouped soils using genetic classification and Wosten et al. grouped soil based on an assessment of the functional behavior of different horizons. For example, a grouping by mineralogy and genesis of soil before developing PTF also suggested by Puckett et al.

So, you see that different scientist have proposed different methods for producing this class PTFs and they have differentiated, you know, the soil into different classes based on either genetic horizons or parent materials and, you know, genetic classification and functional behavior of different horizons and by mineralogy apart from the soil texture broad classes.

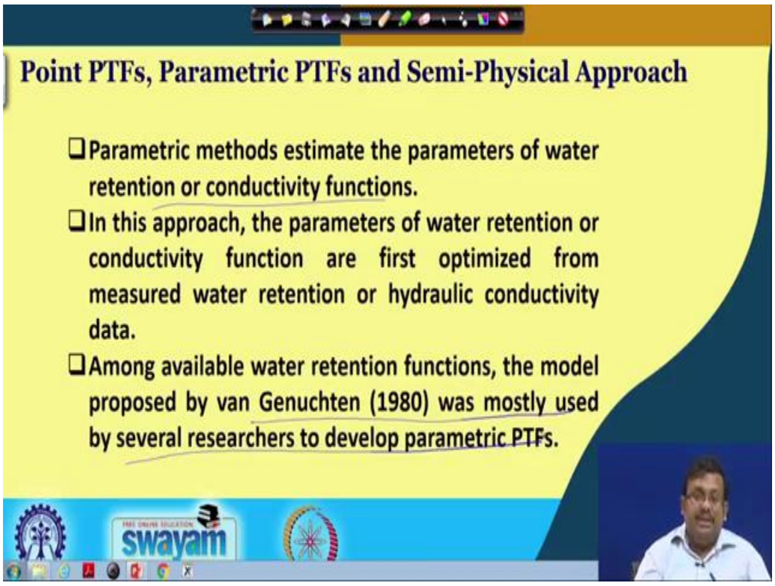
(Refer Slide Time: 05:27)



Now, this scientist Teitje and Tapkenhenrichs classified PTFs, you know, based on the point, parametric and semi-physical estimation method. Point estimation method follows a direct approach by estimating water content at pre-determined pressure heads and estimating hydraulic conductivity at saturation. Although the point estimation methods are simple in approach, but have disadvantage of large number of regression equation required for full characterization of water retention curve.

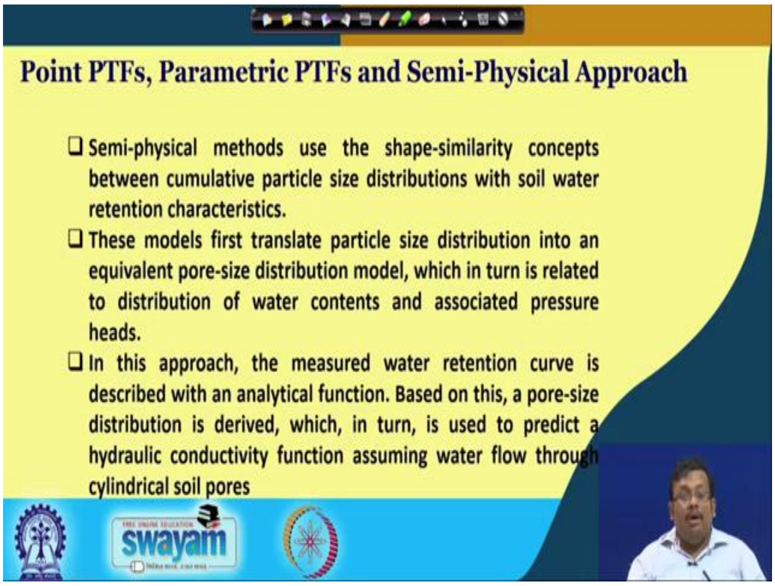
So, again this point estimation methods follow a direct approach by estimating water contents and predetermined pressure heads and estimating hydraulic conductivity at saturation.

(Refer Slide Time: 06:15)



So, another method is parametric method. So, parametric method estimate the parameter of water retention or conductivity functions. So, in this approach the parameters of water retention or conductivity function are first optimized from measured water retention or hydraulic conductivity data. And among available water retention functions the model proposed by van Genuchten was mostly used by several researchers to develop parametric PTFs. So, this is about the parametric PTFs.

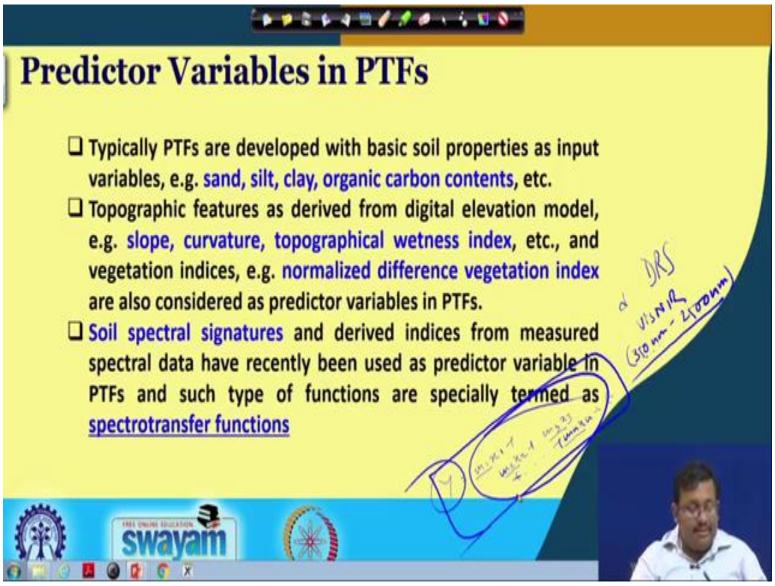
(Refer Slide Time: 06:49)



And, another one is the semi-physical approach. Now, semi-physical methods use the shape similarity concept between cumulative particle size distributions with soil water retention characteristics and these models first translate particle size distribution into an equivalent pore-size distribution model which in turn related to distribution of water contents and associated pressure heads.

And in this approach, the measured water retention curve is described with analytical functions. Based on this, a pore-size distribution is derived, which in turn, is used to predict hydraulic conductivity function assuming water flow through cylindrical soil pores. So, this is about the semi-physical approach.

(Refer Slide Time: 07:31)



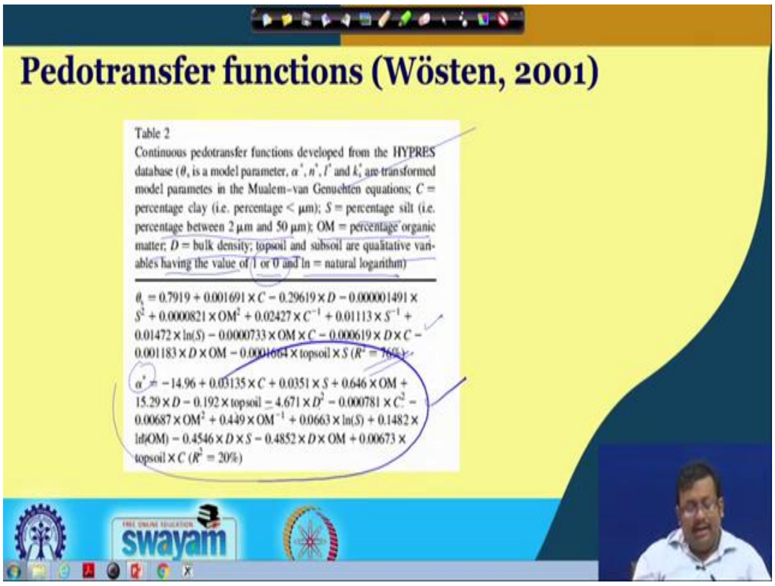
Now, predictor variables in PTFs, this is very important. You may ask ok, what are the predictor important predictor variables in PTFs? Typically PTFs are developed with basic soil properties at input variables which you can easily quantify; for example, sand, silt, clay, organic carbon contents, etcetera. Sometime topographic features as derived from digital elevation model like slope, curvature, topographic wetness index, etcetera and also vegetation index like Normalized Difference Vegetation Index or NDVI are also considered as predictor variables in PTFs.

So, apart from this sand, silt, clay, organic carbon content, we also use this slope characteristics as what we extract from digital elevation model I have already told you in the overview of digital soil mapping that this, you know, this DEM can be used to deduce several topographic as well as, you know, several types of topographic features and this topographic features like slope, curvature, topographical wetness index we can generate from DEM. And we also can use them for as predictor in this pedotransfer function and also vegetation index. This normalized difference vegetation index we can get it from remote sensing data.

And, you know, also we can get soil spectral signatures and derived index from measured spectral data. We have already discussed about the soil spectral signatures in our previous week of lectures where we when we discussed about diffuse reflectance spectroscopy and in the diffuse reflectance spectroscopy, I showed you how we use the individual reflectance values as a predictor for predicting a soil particular soil property and this type of function is also known as spectro transfer function. So, the in the spectro transfer function we basically try to predict a soil property based on numerous, you know, predictors where for example, here you can see m 1 x 1, m 2 x 2 plus m 3 x 3 up to m n x n and here x 1, x 2, x 3 and x n are considered as individual reflectance values.

In our case, in case of Diffused Reflectance Spectroscopy or DRS in case of visible to near infrared, DRS this x 1 to x n will vary from 350 nanometer to 2500 nanometer for individual wavelengths we will get the reflectance values and we will fit them into this regression model and ultimately we will be trying to predict a particular soil property say soil organic carbon and these function we call it spectrotransfer function.

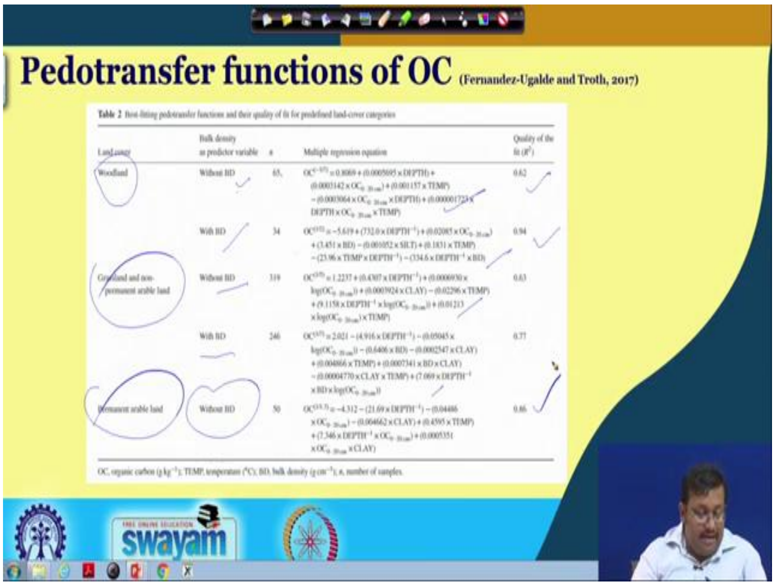
(Refer Slide Time: 10:40)



So, let us see some example of pedotransfer function. here you can see the continuous pedotransfer function developed by this Wosten et al in 2001 from this HYPRES data database where theta s is a model parameter and then alpha n, l and k are that are, you know, are the transformed model parameters in the Mualem van Genuchten equations, C is the percentage of clay, S is the percentage of silt. And then, you know, percentage of silt that is percentage between 2 micrometer to 50 micrometer, organic matter is percentage of organic matter, D is the bulk density and topsoil and subsoil are qualitative variables having the values of 1 and 0; that means, they have coded them topsoil and subsoil and ln is a natural logarithm.

So, you can see they are predicting these theta s by using this formula. So, and they are getting an R square values of 0.6 0.76. And also they have modeled this alpha asterisk alpha by using this model, so, this is another pedotransfer model, this is another pedo, two different pedotransfer model and here they are getting an R square values of 0.20. And you can see in this pedotransfer model they have incorporated all different variables which they mentioned here as predictors and ultimately they have generated a particular model and this called the actual pedotransfer model or pedotransfer function.

(Refer Slide Time: 12:12)

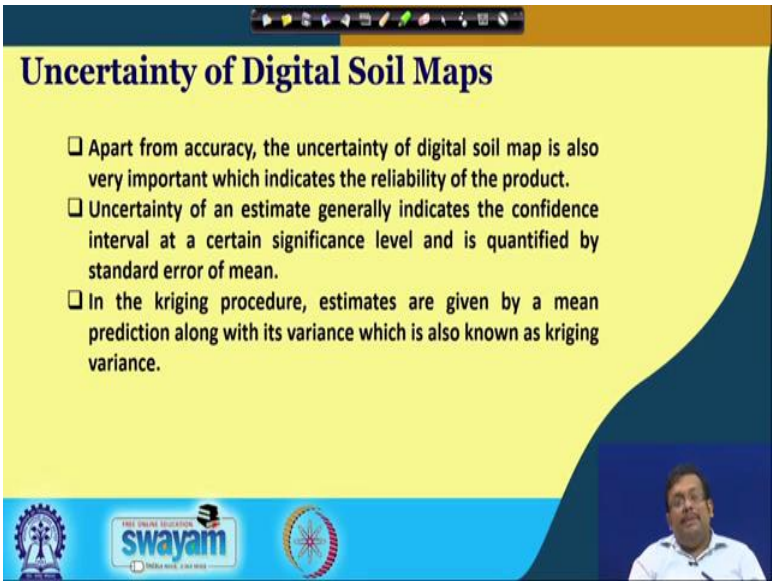


Another pedotransfer function example here you can see Fernandez-Ugalde and Troth in 2017. So, they have created different pedotransfer model for quantifying their, you know, soil organic carbon you can see in different condition without when they are did not considered with a bulk density they have created a pedotransfer function; by considering the bulk density they have created another pedotransfer function and then Grassland non permanent arable land they have created this pedotransfer function, with bulk density they have created this pedotransfer function.

So, for different a land used land cover types they have created two pedotransfer function without bulk without considering the bulk density and with considering the bulk density. So, these are two pedotransfer functions and for permanent arable land they have created one pedotransfer function without considering the bulk density. And these are the number of samples and you can see quality of the fit is basically calculated based on the R square values or coefficient of determination values. Here, you know, it varies from 0.62 0.94, 0.63, 0.77 to 0.86.

So, these are some examples of pedotransfer functions for predicting organic carbon in different land cover by considering bulk density or without considering the bulk density of the soil. So, we have covered this pedotransfer function. The final topic we will discuss is accuracy and uncertainty of digital soil mapping.

(Refer Slide Time: 13:46)

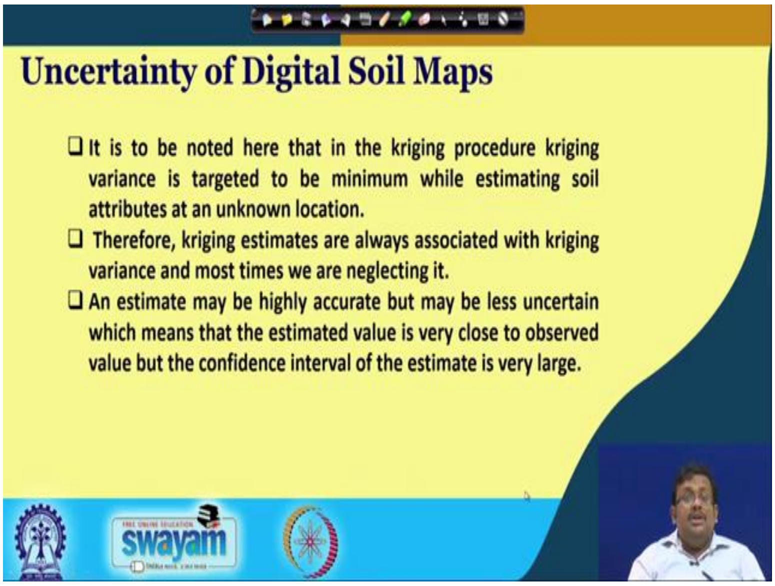


So, we have already cover the accuracy of digital soil mapping in our previous lectures where we, you know, when we discussed the validation of different types of models and we discussed about RMSE, we discussed about Lin's concordance correlation coefficient, then bias and all this things.

So, let us talk about the uncertainty of digital soil map. One of the best, you know, one of the major advantage of digital soil mapping is digital soil mapping not only produce the digital soil map of spatial variable and their spatial variability map, they also produce side by side the associated uncertainty map. So, apart from accuracy the uncertainty of digital soil map is also very important which indicates the reliability of the product. Uncertainty of an estimate generally indicates the confidence interval at a certain significance level and it is quantified by standard error of mean.

So, in the kriging procedure, estimates are given by mean prediction along with its variance which is also known as the kriging variance.

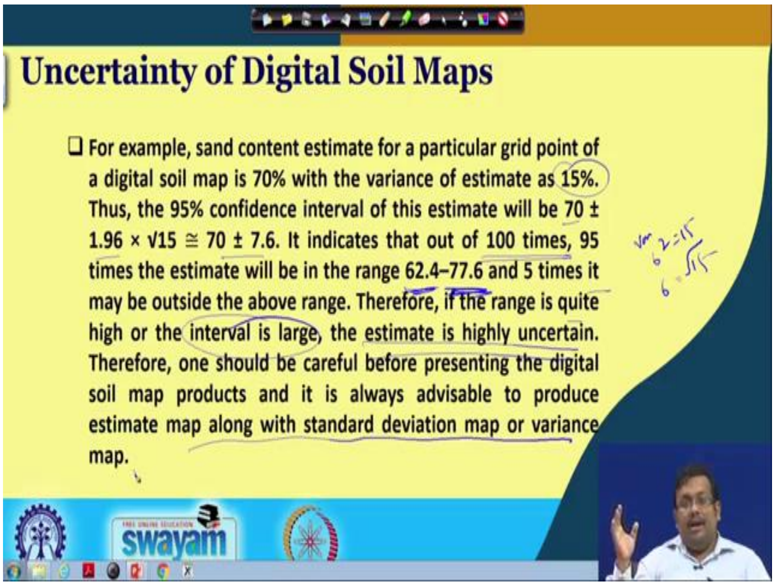
(Refer Slide Time: 14:54)



So, it is to be uncertain. So, it is to be noted here that in the kriging procedure kriging variance is targeted to be minimum while estimating soil attributes at unknown location. Therefore, kriging estimates are always associated with kriging variance and most time we are neglecting it. So, an estimate may be highly accurate, but maybe less uncertain which means that the estimated value is very close to observed value, but the confidence interval of the estimate is very large.

So, this is the practical implication of a particular situation. Again, when an estimate maybe highly accurate, but it is less uncertain which means that the estimated value is very close to observed value, but the confidence interval of the estimate is very large.

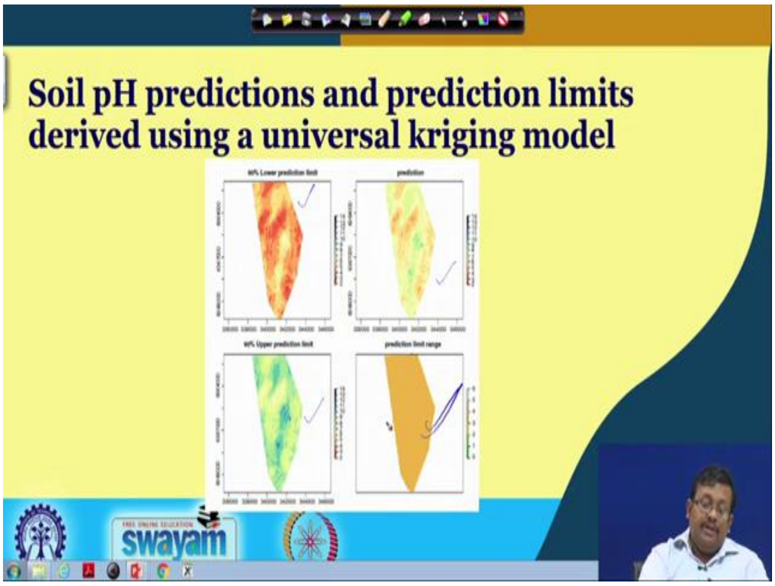
(Refer Slide Time: 15:42)



For example, let us see one example. Sand content of, you know, sand content estimate for a particular grid point of a digital soil map, let us see it is 70 percent with the variance of estimate of 15 percent. Thus the 95 percent confidence interval of this estimate will be 70 plus minus 1.96 plus standard deviation of 15 because, here we are seeing 15 percent. So, basically here variance or sigma square is 15. So, standard deviation will be root 15. So, basically we are getting an interval, 95 percent confidence interval, you know, the standard normal variance value at 95 percent confidence interval 1.96 multiplied by this sigma, so, ultimately 70 plus minus 7.6.

So, it indicates that out of 100 times 95 times the estimate will be in between the range of 62.4 to 77.6 and 5 times it may be outside the above range. Therefore, if the range is quite high or the interval is large the estimate is highly uncertain. Again, if the range is quite high or the interval is large as you have seen our as we have discussed in our last slide the estimate is highly uncertain. Therefore, one should be careful before presenting the digital soil map products and it is always advisable to produce estimate map along with a standard deviation map or variance map. So, that, you know, the client maybe aware or should be aware of the related uncertainty of the products of DSM.

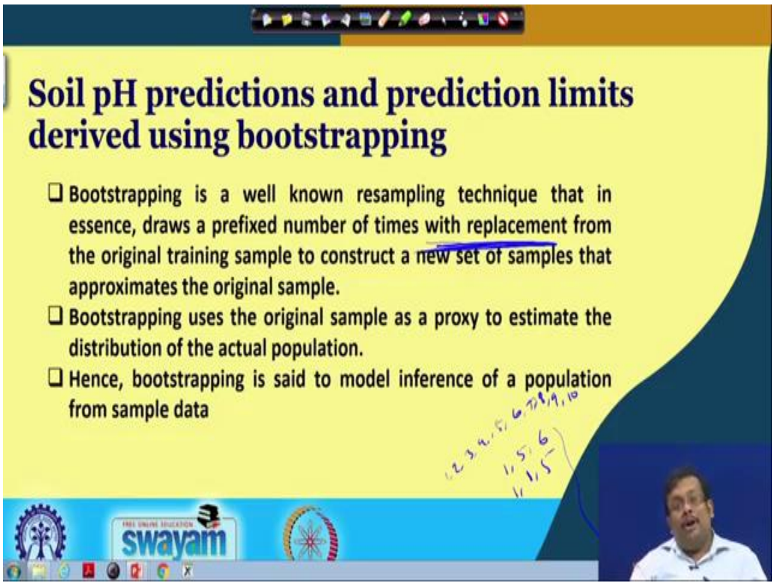
(Refer Slide Time: 17:36)



So, let us see some examples of, you know, uncertainty levels using different methods of digital soil mapping. So, it is a soil pH prediction and prediction limits using a universal kriging model. So, we are using a universal kriging model to produce the prediction of, you know, prediction of particular, you know, soil property here soil organic carbon and we are also having the 90 percent lower prediction limit and 90 percent upper prediction limit and this is basically it is a difference between the 90 percent upper prediction limit and lower prediction limit. So, it is a prediction limit range which is the difference between these two ranges.

So, not only we are getting the prediction value, but also we are getting the prediction limit range also. So, this is an example where we used the universal kriging model.

(Refer Slide Time: 18:35)

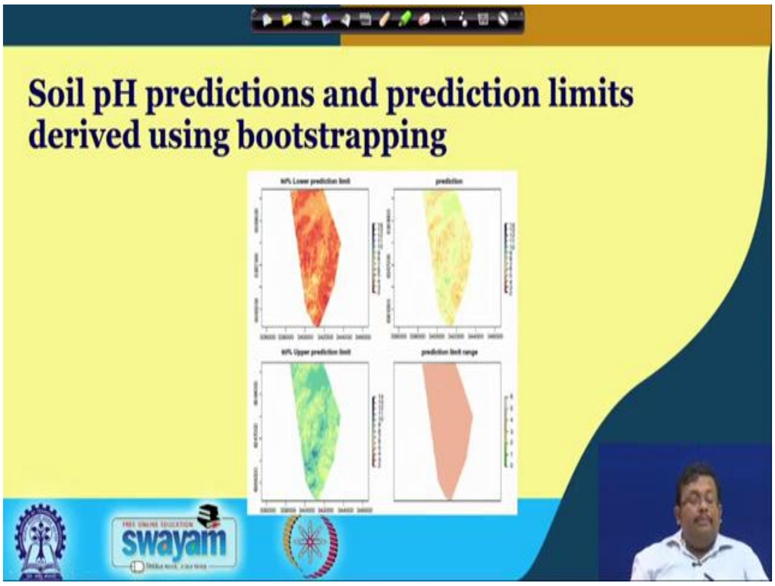


Another example; so, soil pH prediction and prediction limit using bootstrapping method. Now, bootstrapping is a well known re-sampling technique, there is essence or draws a prefix number of times with replacement from the original training samples to construct a new set of samples that approximates the original sample.

So, for example, if there are, you know, ten samples 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and we are drawing a bootstrap sample consisting of four samples. So, we can take 100 bootstrap sample which consists of each and each bootstraps sample may consist of 4 samples. So, when we draw the samples without replacement, it is very important resampling technique that is essence draw prefix number with replacement sorry, with replacement from the original training sample then we call it bootstrapping.

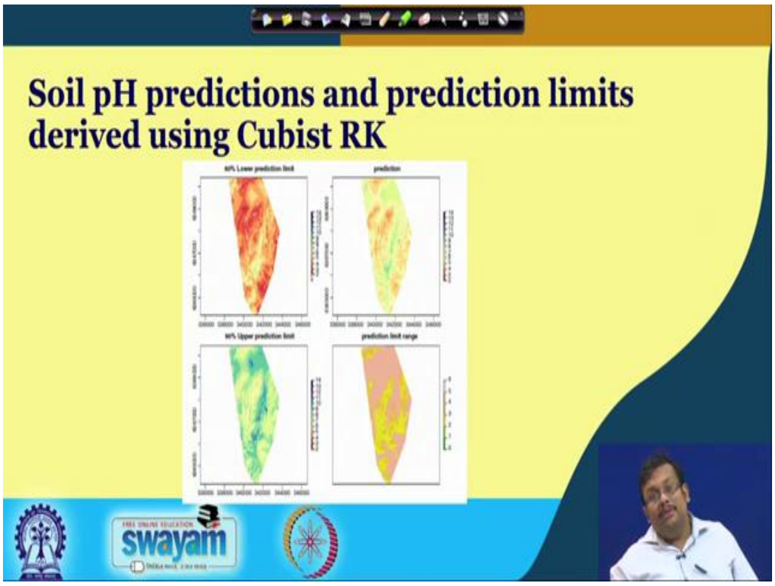
For example, if there are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and we are taking three sample as a bootstrap sample. So, in the first we can take 1, 5, 6. Randomly in the second we can take 1, 1, 5 again. So, you are seeing we are doing with replacement sampling. So, that is called bootstrapping. So, we can take number of bootstrap samples and we can estimate the original sample, approximate the original sample based on this bootstrapping samples. So, bootstrapping uses the original samples as a proxy to estimate the distribution of the actual population. So, hence bootstrapping is said to model inference of a population from sample data.

(Refer Slide Time: 20:36)



So, let us see one example that bootstrapping is used to create the digital soil map. You can see here the digital soil map is created. Actually these are digital soil map I am sorry these are the digital soil map of soil pH and this soil pH map is created you can see the ranges from 2 to 14. This soil pH map is created using the bootstrapping and again this is a original prediction and this has 90 percent lower confidence limit and 90 percent upper confidence limit and this is the prediction limit range. So, this is an example of, you know, producing the estimate as well as the uncertainty using bootstrapping method.

(Refer Slide Time: 21:18)



This is another example of producing the digital soil map of pH and they are related and they are; and their associated prediction limit by Cubist based regression kriging where in the regression kriging we have use the cubist model as a deterministic model and kriging with residuals.

(Refer Slide Time: 21:42)

So, guys I have finished all this uncertainty and let us wrap up this digital soil mapping and associated things I am this will be this is our last slide. And thank you for being with me for this in this journey of this course and we have covered several aspects of soil and I must mention in this slide you can see I must mention the special acknowledgement to my students without whom this course would not be possible.

They have constantly helped me to create the materials and they are also taking care of the other things and Miss Swetha and Miss Swagata they are my PhD students and they helped me for creating different slides and also, you know, creating the assignments and creating the assignments.

So, I hope that this course was helpful for you, you have learned something new and you have we have we tried to cover all the basic aspects of soil. Obviously, we could not cover all the aspects of soil because of time limitation, but I tried to give the basic overview of the important aspects of soil and you have learned those things. Obviously, I would also encourage you to go ahead and discuss and consult different books you know several books are available and I have shown you couple of reference books also which might be very much helpful if you want to gather a more detail information about soil science.

And I hope that you have enjoyed this course and if you have any queries regarding the aspects of soil which I have covered or any other aspects regarding the soil please feel free to e-mail me and I will be more than happy to answer your queries. And thank you very much for taking this course and I hope all the best to your future endeavors.

Thank you guys.