# Machine Learning for Soil and Crop Management
**Professor Somsubhra Chakraborty**
**Agricultural and Food Engineering Department**
**Indian Institute of Technology Kharagpur**
**Lecture 24**
## Diffuse Reflectance Spectroscopy: Basics and Applications for Crop and Soil (Contd.)

(Refer Slide Time: 0:21)

NPTEL ONLINE CERTIFICATION COURSES

Machine Learning for Soil and Crop Management
Prof. Somsubhra Chakraborty
Agricultural and Food Engineering Department
Indian Institute of Technology Kharagpur

Week 5: DIFFUSE REFLECTANCE SPECTROSCOPY: BASICS AND APPLICATIONS FOR CROP AND SOIL

LECTURE 24

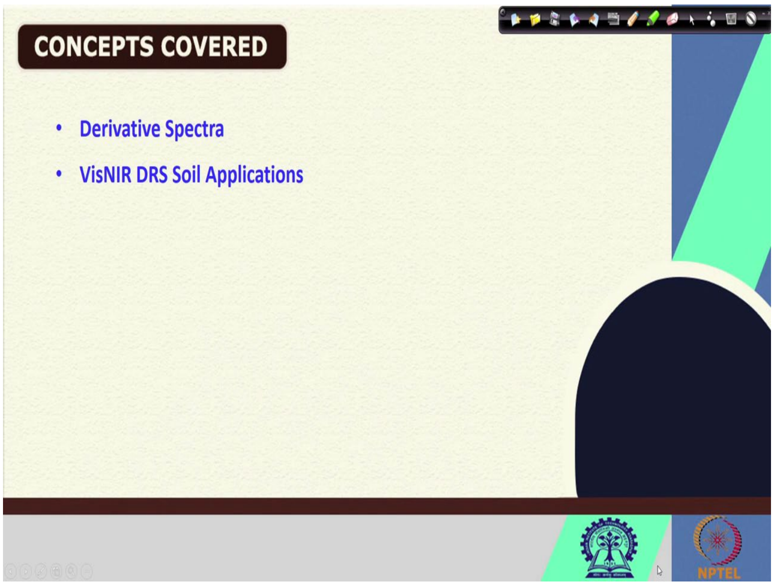
Welcome friends to this fourth lecture of NPTEL of week 5 of NPTEL online certification course of Machine Learning for Soil and Crop Management. And in this week, we are discussing or this is week 5 and in this week we are discussing Diffuse Reflectance Spectroscopy; its basics and application for crop and soil. So, in our previous 3 lectures, we have discussed some of the important aspects of diffuse reflectance spectroscopy, we have seen the hyperspectral remote sensing, the hyperspectral data cube, and then we have seen that resolution of varying resolution from panchromatic to multispectral to hyperspectral and ultraspectral data.

We have seen what is soil spectroscopy, how we can generate the soil spectral library, then we have seen different spectral pre-processing, different spectra of spectroradiometer commercially, available spectroradiometers, we have also seen several foreoptics like different probes of commercially available soil spectrometer. We have also seen the spectral resolution sampling interval all these definitions.

So, in the last lecture, we have started discussing about spectral pre-processing, which is the first step of the spectral analysis and we have seen how to convert from wavelength from nanometer to wave number and then spectral trimming is we have seen, we have seen the moving window average and then standard normal variate, continuum removal, detrending, multiplicative scatter correction. So, all these different types of spectral processing we have seen and we have seen what are their benefits also. Also we have seen Savitzky-Golay filtering for smoothing the spectrum.

However, one of the major pre-processing for contemporary spectroscopic application is derivative spectroscopy. So, today we are going to discuss the derivative spectroscopy and its application, also the advantages of derivative spectroscopy. Why we go for derivative spectroscopy, we will see what are the reasons and finally, we will see some examples of soil based we will see some DRS application for soil property prediction.

(Refer Slide Time: 3:20)

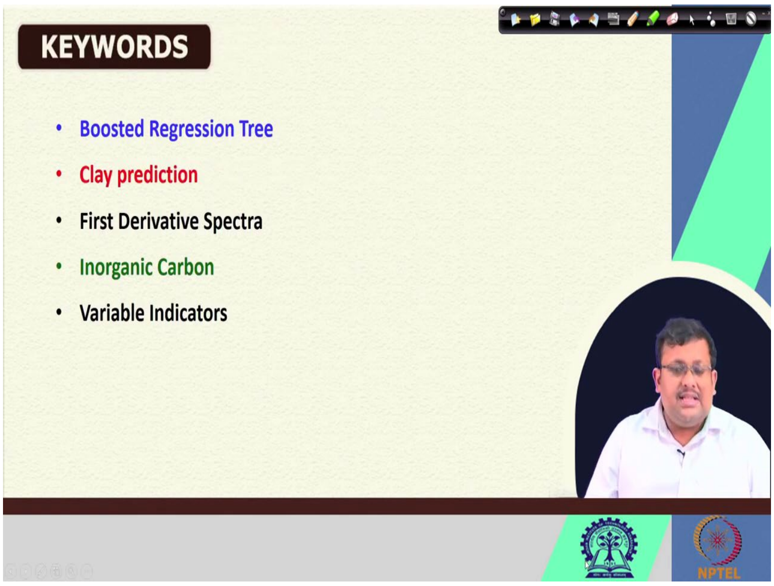


CONCEPTS COVERED

- Derivative Spectra
- VisNIR DRS Soil Applications

So, these are the concepts which we are going to cover in today’s lecture in this lecture, lecture number 24 and in this lecture, we are going to first cover the derivative spectra and then we are going to cover some of the classical application or classical paper or VisNIR DRS soil applications.

(Refer Slide Time: 3:43)

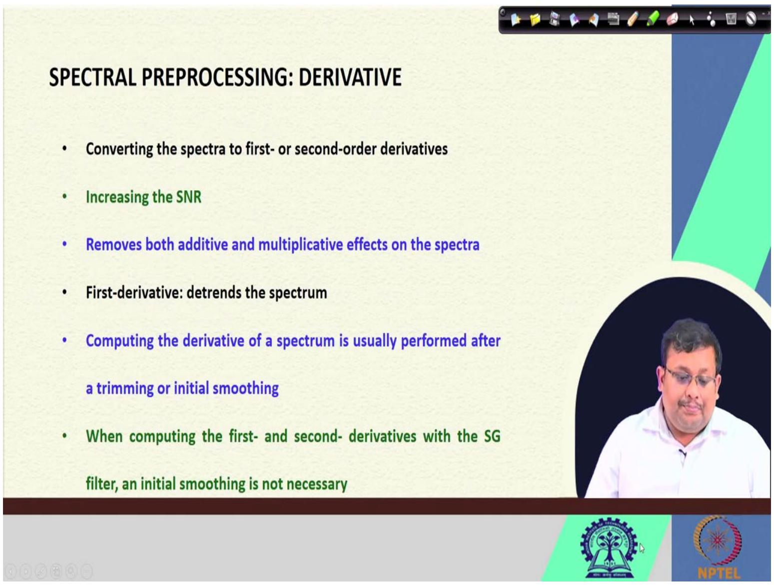


KEYWORDS

- Boosted Regression Tree
- Clay prediction
- First Derivative Spectra
- Inorganic Carbon
- Variable Indicators

These are the keywords which we are going to discuss today. We are going to see the application of Boosted Regression tree, then clay prediction, first derivative spectra, inorganic carbon and also variable indicators.

(Refer Slide Time: 4:01)



SPECTRAL PREPROCESSING: DERIVATIVE

- Converting the spectra to first- or second-order derivatives
- Increasing the SNR
- Removes both additive and multiplicative effects on the spectra
- First-derivative: detrends the spectrum
- Computing the derivative of a spectrum is usually performed after a trimming or initial smoothing
- When computing the first- and second- derivatives with the SG filter, an initial smoothing is not necessary

So, if we see that derivative spectroscopy, now, why we go for the derivative spectroscopy? So, derivative spectroscopy is basically converting the raw reflectance spectra to the first order and second order derivatives with respect to the wavelengths. Now, why do we do that, we do that basically to increase the signal to noise ratio or to remove the noise. There are certain other benefits of using the derivative spectroscopy which we are going to discuss. But the basic fundamental assumption behind doing the derivative spectroscopy is to increase the signal to noise ratio.

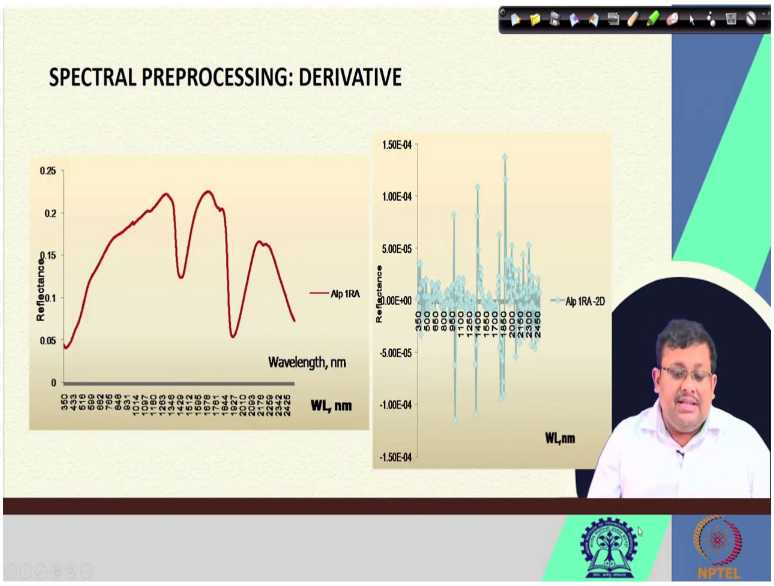
So, also when you go for the derivative spectra, it removes both additive and multiplicative effects on the spectra. So, when you go for the first derivative of reflectance spectra, it detains the spectrum. So, first let us discuss why you go for the spectral derivative. So, generally spectral derivative is generally, derivative spectroscopy relies on developing the first and second derivative of the raw reflectance spectra and then question comes why we go for this?

We go for the first the major reason for doing the derivative pre-processing is to increase the signal to noise ratio. So, apart from increasing the signal to noise ratio, the derivative spectra also removes both additive and multiplicative effects on the spectra. So, in case a first derivative, when you go for the first derivative, it detrends the spectrum and we have seen what is the benefit of detrending in our previous lecture. So, first derivative helps in detrending the spectra.

Now, computing the derivatives of a spectrum is usually performed under trimming or initial smoothing. So, you know what is the initial smoothing, so initial smoothing generally we can perform with the Savitzky-Golay filter and then we can trim the data from 350 to 490 nanometer and then from 2451 to 2500 nanometers. So, we can trim these to extreme and then we can go for subsequent machine learning models. Sometimes when you go for the first and second derivatives with the Savitzky-Golay filter itself, then an initial smoothing is not necessary.

So, generally we go for the first derivatives of the reflectance spectra then followed by the with the Savitzky-Golay filter then filtering then then initial smoothing we generally avoid. So, there are we can either take the first derivative, we can either take the second derivative; from my own experience, I can tell that in most of the application first derivative works better.

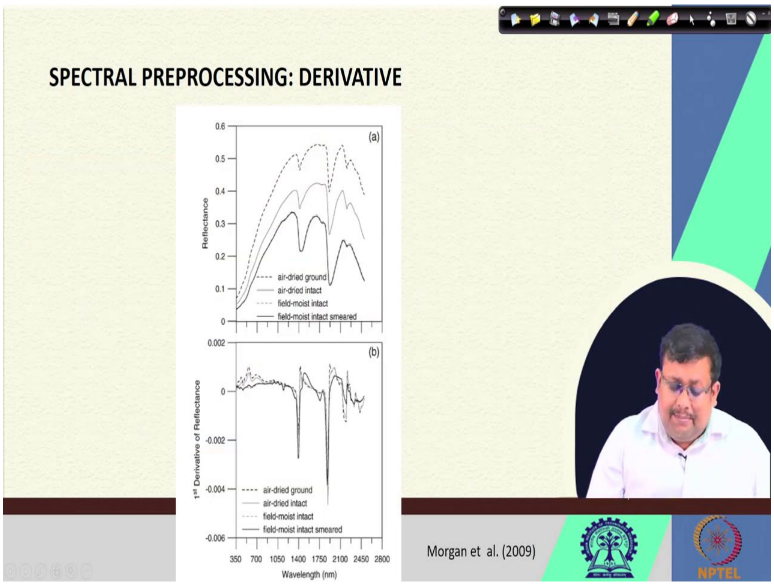
(Refer Slide Time: 7:39)



SPECTRAL PREPROCESSING: DERIVATIVE

Now, the question is how this first derivative increases the signal to noise ratio. So, you can see this is an original raw reflectance spectra of a soil and this is the first derivative spectra of the same soil. So, you can easily see that the broad spectral features are almost absent in the in the in this derivative spectra plot and we can see the signal to noise ratio has increased greatly. So, our idea of behind this derivative, generating the radiating spectra is to increase the signal to noise ratio.

(Refer Slide Time: 8:16)



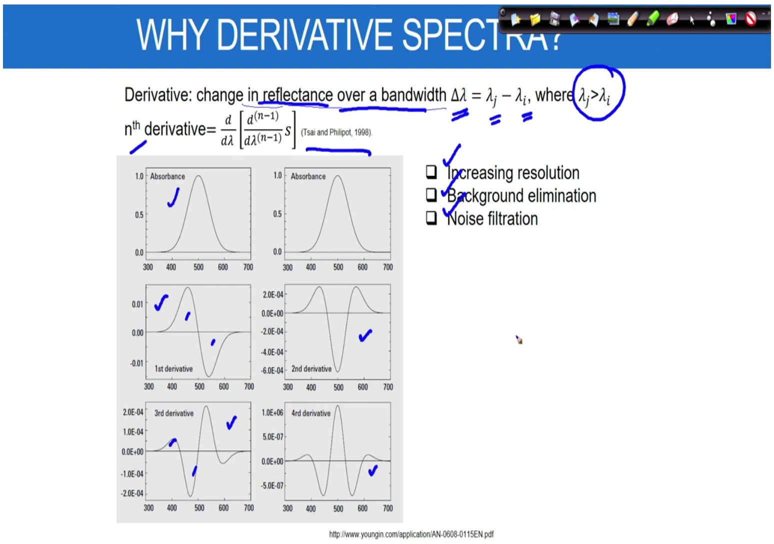
SPECTRAL PREPROCESSING: DERIVATIVE

Morgan et al. (2009)

A similar application you can see in Morgan et al, 2009, who showed that there are different types of spectrum. You can see this is air drying spectrum, this dashed line, then air dried intact spectrum and also field moist intact and feel moist intact smeared spectrum, you can see different types of spectrum from the soil, that basic shape is same. However, when you go for the generating the derivatives the first derivative of the reflectance, then we can see this is these are the first derivative of the reflectance.

So, you can see the spectral shape has been completely changed and we can see some of the important spectral features which are arising after this first derivative reflectance conversion. So, the idea is increasing the signal to noise ratio.

(Refer Slide Time: 9:07)



WHY DERIVATIVE SPECTRA?

Derivative: change in reflectance over a bandwidth $\Delta\lambda = \lambda_j - \lambda_i$, where $\lambda_j > \lambda_i$

n^th derivative= $\frac{d}{d\lambda}\left[\frac{d^{(n-1)}}{d\lambda^{(n-1)}} S\right]$ (Tsai and Philipot, 1998)

- ☐ Increasing resolution
- ☐ Background elimination
- ☐ Noise filtration

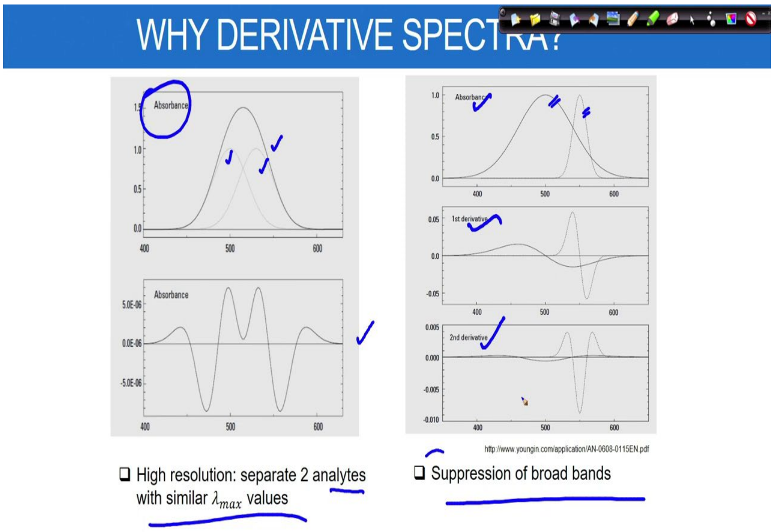
http://www.youngin.com/application/AN-0608-0115EN.pdf

Now, the question comes what how we can calculate the derivative. So, you can see that here the derivative is basically the changing in the reflectance over a bandwidth. So, just, so, increasing the band increase, changing the reflectance over a bandwidth where this bandwidth is denoted by delta lambda, where delta lambda equal to delta j minus delta i where lambda j is greater than lambda i. Now, this delta lambda could be any value and so, the nth derivative can be calculated by using this formula given by Tsai and Philipot. So, there are a couple of reasons for which we go for the derivative spectroscopy.

First of all, to increase the resolution, second, we want to remove the background and also noise filtration. So, you can see one example here, this is an absorbance peak and from this absorbance peak, you can see this is the first derivative, this is a second derivative, this is a third derivative and this is a fourth derivative. So, as we are increasing the derivative, the number of peaks are also increasing. So, in that way we are increasing the resolution.

So, here you can see this is one only one feature spectral feature, we can see in first derivative, there are 2 features. In case of second derivative there are 3 features, in case of third derivative 1, 2, 3, 4 features so and so forth. So, as we are increasing the derivative, we are having much more information gathered from the derivative spectrum.

(Refer Slide Time: 11:00)



WHY DERIVATIVE SPECTRA?

http://www.youngin.com/application/AN-0608-0115EN.pdf

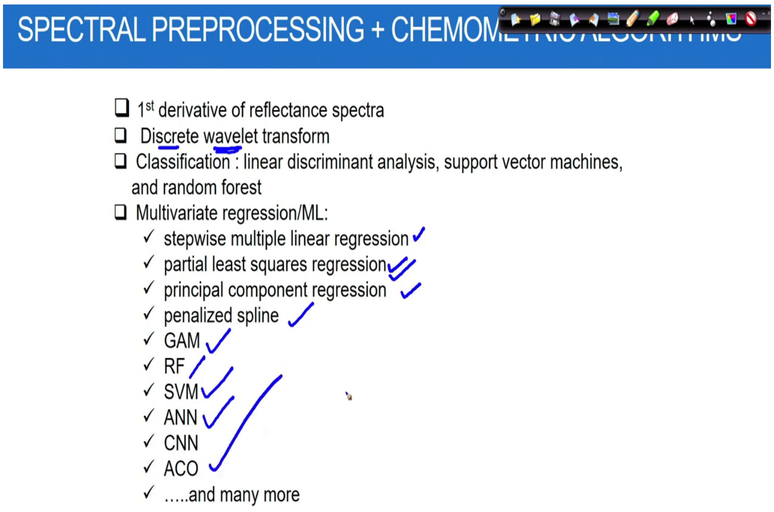
☐ High resolution: separate 2 analytes with similar $\lambda_{max}$ values

☐ Suppression of broad bands

Now, also you can see here another example, another important aspect of derivative spectroscopy is when you go for the, from the when we change it from the absorbance to derivative then you can see that it separates to analyze with a similar lambda max value. So, you can see here the 2 analyte with the similar, similar lambda max values were actually overlapped, but they are separated when you go for the conversion from absorbance to the derivative.

Also you can see suppression of broad bands. So, here you can see the broad band is suppressing these smaller spectral features, but when we go from the absorbance to first derivative between second derivative these smaller features is getting much more resolved. So, we can see these smaller features and then the suppression is also going down in each of these subsequent derivatives. So, this is how it is clear that why we go for the derivative of, why we go for the derivative spectroscopy for pre-processing the spectrum and why it is very helpful.

(Refer Slide Time: 12:32)



SPECTRAL PREPROCESSING + CHEMOMETRIC ALGORITHMS

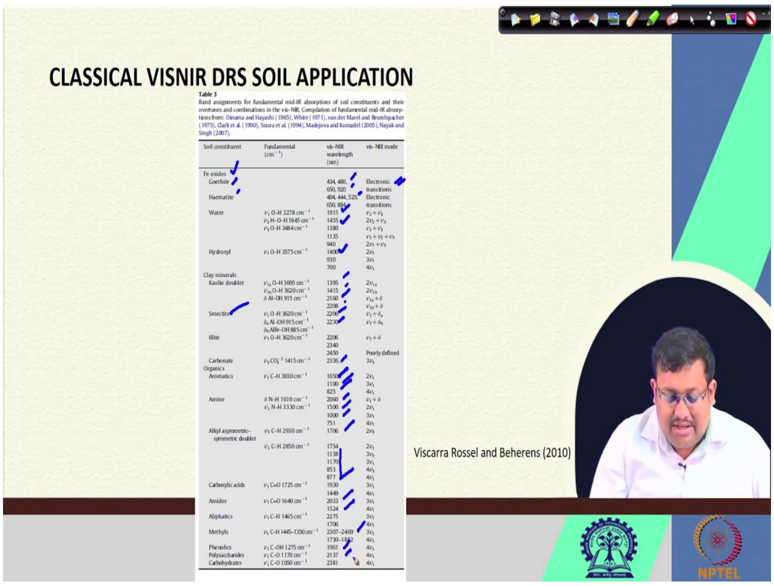
- ☐ 1st derivative of reflectance spectra
- ☐ Discrete wavelet transform
- ☐ Classification : linear discriminant analysis, support vector machines, and random forest
- ☐ Multivariate regression/ML:
  - ✓ stepwise multiple linear regression
  - ✓ partial least squares regression
  - ✓ principal component regression
  - ✓ penalized spline
  - ✓ GAM
  - ✓ RF
  - ✓ SVM
  - ✓ ANN
  - ✓ CNN
  - ✓ ACO
  - ✓ .....and many more

So, generally we take the first derivative or different types of other spectral pre-processing as we have discussed. Another one is discrete wavelet transform, wavelet is a small wave like function. So, these wave like function generally we move these any these wavelet function can be defined in terms of locational parameter as well as scale parameter and by changing the location and also the scale parameter, we can fit any minute change in the spectral data and we can match the wavelet with any minute change in the spectral data and then we convert it into a 2 dimensional plot we call it wavelet conversion plot.

And from there also it is when we go for the wavelet transform, we generally successively decompose the spectrum into a number of scales and each of these scales that have, are having a number of coefficients using these wavelet coefficients we can selectively remove the noise which is there in the spectra and we can go for the subsequent machine learning application.

So, what are the subsequent chemometric and machine learning applications we generally use? We generally use the stepwise multiple linear regression which is also partial least square regression, this is the most widely accepted method or utilized method to deal with the spectral data, then principal component regression, then penalized spline regression, GAM regression, random (())( 14:15), support vector machine, artificial neural network then also ant colony optimization and so on so forth. So, there are multiple ways through which you can deal with the spectral data.

(Refer Slide Time: 14:33)



CLASSICAL VISNIR DRS SOIL APPLICATION

Table 3
Band assignments for fundamental mid-IR absorptions of soil constituents and their overtones and combinations in the vis–NIR. Compilation of fundamental mid-IR absorptions from: Oinuma and Hayashi (1965), White (1971), van der Marel and Beutelspacher (1975), Clark et al. (1990), Srasra et al. (1994), Madejova and Komadel (2005), Nayak and Singh (2007).

| Soil constituent | Fundamental ($cm^{-1}$) | vis–NIR wavelength (nm) | vis–NIR mode |
|---|---|---|---|
| Fe oxides | | | |
| Goethite | | 434, 480,<br>650, 920 | Electronic transitions |
| Haematite | | 404, 444, 529,<br>650, 884 | Electronic transitions |
| Water | $\nu_1$ O-H 3278 $cm^{-1}$ | 1915 | $\nu_2+\nu_3$ |
| | $\nu_2$ H-O-H 1645 $cm^{-1}$ | 1455 | $2\nu_2+\nu_3$ |
| | $\nu_3$ O-H 3484 $cm^{-1}$ | 1380 | $\nu_1+\nu_3$ |
| | | 1135 | $\nu_1+\nu_2+\nu_3$ |
| | | 940 | $2\nu_1+\nu_3$ |
| Hydroxyl | $\nu_1$ O-H 3575 $cm^{-1}$ | 1400 | $2\nu_1$ |
| | | 930 | $3\nu_1$ |
| | | 700 | $4\nu_1$ |
| Clay minerals | | | |
| Kaolin doublet | $\nu_{1a}$ O-H 3695 $cm^{-1}$ | 1395 | $2\nu_{1a}$ |
| | $\nu_{1b}$ O-H 3620 $cm^{-1}$ | 1415 | $2\nu_{1b}$ |
| | $\delta$ Al-OH 915 $cm^{-1}$ | 2160 | $\nu_{1a}+\delta$ |
| | | 2208 | $\nu_{1b}+\delta$ |
| Smectite | $\nu_1$ O-H 3620 $cm^{-1}$ | 2206 | $\nu_1+\delta_a$ |
| | $\delta_a$ Al-OH 915 $cm^{-1}$ | 2230 | $\nu_1+\delta_b$ |
| | $\delta_b$ AlFe-OH 885 $cm^{-1}$ | | |
| Illite | $\nu_1$ O-H 3620 $cm^{-1}$ | 2206 | $\nu_1+\delta$ |
| | | 2340 | |
| | | 2450 | Poorly defined |
| Carbonate | $\nu_3$ $CO_3^{-2}$ 1415 $cm^{-1}$ | 2336, | $3\nu_3$ |
| Organics | | | |
| Aromatics | $\nu_1$ C-H 3030 $cm^{-1}$ | 1650 | $2\nu_1$ |
| | | 1100 | $3\nu_1$ |
| | | 825 | $4\nu_1$ |
| Amine | $\delta$ N-H 1610 $cm^{-1}$ | 2060 | $\nu_1+\delta$ |
| | $\nu_1$ N-H 3330 $cm^{-1}$ | 1500 | $2\nu_1$ |
| | | 1000 | $3\nu_1$ |
| | | 751 | $4\nu_1$ |
| Alkyl asymmetric-symmetric doublet | $\nu_3$ C-H 2930 $cm^{-1}$ | 1706 | $2\nu_3$ |
| | $\nu_1$ C-H 2850 $cm^{-1}$ | 1754 | $2\nu_1$ |
| | | 1138 | $3\nu_3$ |
| | | 1170 | $3\nu_1$ |
| | | 853 | $4\nu_3$ |
| | | 877 | $4\nu_1$ |
| Carboxylic acids | $\nu_1$ C=O 1725 $cm^{-1}$ | 1930 | $3\nu_1$ |
| | | 1449 | $4\nu_1$ |
| Amides | $\nu_1$ C=O 1640 $cm^{-1}$ | 2033 | $3\nu_1$ |
| | | 1524 | $4\nu_1$ |
| Aliphatics | $\nu_1$ C-H 1465 $cm^{-1}$ | 2275 | $3\nu_1$ |
| | | 1706 | $4\nu_1$ |
| Methyls | $\nu_1$ C-H 1445–1350 $cm^{-1}$ | 2307–2469 | $3\nu_1$ |
| | | 1730–1852 | $4\nu_1$ |
| Phenolics | $\nu_1$ C-OH 1275 $cm^{-1}$ | 1961 | $4\nu_1$ |
| Polysaccharides | $\nu_1$ C-O 1170 $cm^{-1}$ | 2137 | $4\nu_1$ |
| Carbohydrates | $\nu_1$ C-O 1050 $cm^{-1}$ | 2381 | $4\nu_1$ |

Viscarra Rossel and Beherens (2010)

Now, of course, the basic assumption of spectroscopic analysis, so as spectroscopic analysis comes from the fact that in the visible near infrared region, the most of the spectral features that their combination bands and overtones are appearing in the visible to near infrared range. So, let us see what are those spectral features?

Now Viscarra Rossel and Behrens, one of their classical paper published in Geodarma in the year 2010, they have compiled a table and from different sources, where they have mentioned what are the spectral signatures for different soil components, which can be captured by the visible near infrared spectroscopy.

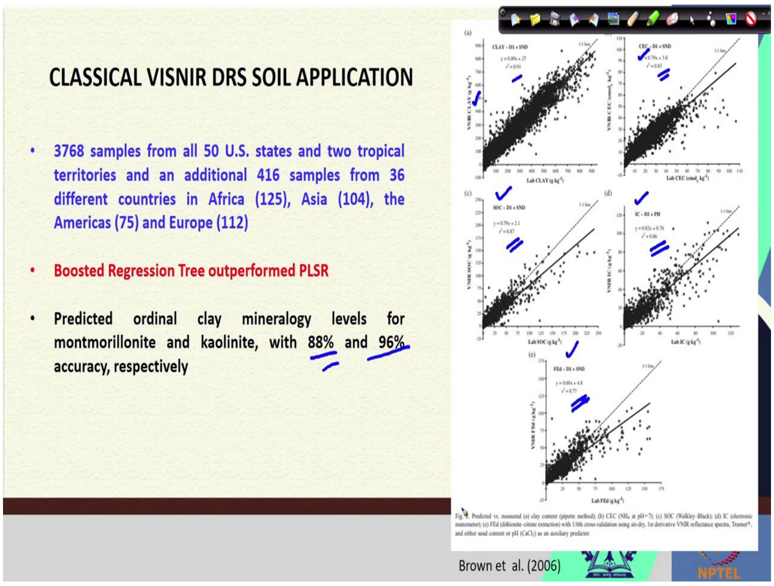
So, we can see here for iron oxides like goethite, hematite, you can see that most of these are can be identified in the visible range like 434, 480, 650, 920 and so on and so forth. So, basically the mode is electronic transition and so this, due to this electronic transition, these wavelengths are very much sensitive to these goethite, hematite.

Water, you can see that, the hydroxyl you can get the VisNIR wavelength at 1915 nanometer, then 1455 nanometer and these are the VisNIR mode. So, hydroxyl you can also get at 1400 nanometer. In case of kaolin minerals, kaolin doublet you can get 1395, 1415, 2160, 2208 and so on so forth. (())(16:30) you can see the smectite at 2206 and then 2230. In case of carbonates, you can get it from 2336. So, in case of organics like aromatics, aromatics you can get at 1650 nanometer and then 1100 nanometer, 825 nanometer.

Amine can be found in a 2060, 1500 and then 1000 nanometer, alkyl asymmetric, symmetric doublet you can see at 1706 nanometer and then so on so forth. Carboxylic acid, you can get at that 1930 nanometer. Amides you can get at 2033 nanometer. Aliphatics, you can get at 2275 nanometer, methyls you can get at 2307, phenolics at 1961, polysaccharides at 2137 and carbohydrates at 2381.

So, it is quite clear that it is quite clear that most of the inorganic mineral minerals as well as the organic components in the soil like oxides as well as the clay minerals and organics are heavy, they are defined spectral features in the VisNIR zone which can be identified by this the VisNIR DRS and mostly in the VisNIR DRS as I have already mentioned, they are not the fundamental vibration, they are basically the overtones and also the combination bands.

(Refer Slide Time: 18:10)



CLASSICAL VISNIR DRS SOIL APPLICATION

- 3768 samples from all 50 U.S. states and two tropical territories and an additional 416 samples from 36 different countries in Africa (125), Asia (104), the Americas (75) and Europe (112)
- Boosted Regression Tree outperformed PLSR
- Predicted ordinal clay mineralogy levels for montmorillonite and kaolinite, with 88% and 96% accuracy, respectively

Fig. 2. Predicted vs. measured (a) clay content (pipette method); (b) CEC ($NH_4$ at pH=7); (c) SOC (Walkley-Black); (d) IC (electronic manometer); (e) FEd (dithionite-citrate extraction) with 1/6th cross-validation using air-dry, 1st derivative VNIR reflectance spectra, Treenet®, and either sand content or pH ($CaCl_2$) as an auxiliary predictor.

Brown et al. (2006)

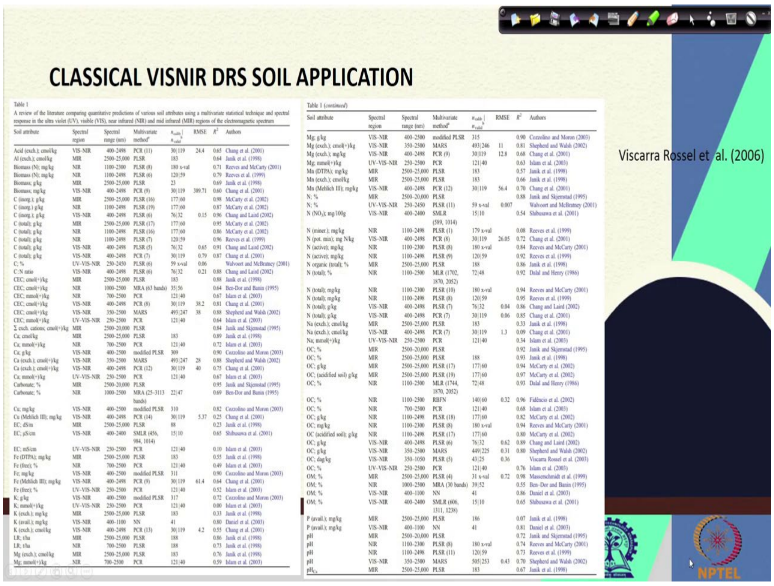
Now, although there are 1000s and 1000s of application for different soil properties. In this lecture, I will confine myself to discussing some of the important publications and some of the classical papers regarding the VisNIR DRS soil applications. So, one of these classical paper appeared in the journal Geoderma in 2006 published by Brown et al and they have collected around 3768 sample from all 50 US states and 2 tropical territories and an additional 416 samples from 36 different countries in Africa and the Americas and Europe.

So, after collecting the samples, then they have used the Boosted Regression tree model which is a tree base model, you know what is a tree base model. So, they have used Boosted Regression tree model and also the partial least squares regression model to model several soil properties and they have seen that the Boosted Regression tree perform better than PLSR in that case, in predicting the soil properties you can see here they have predicted different soil properties like clay and also is catenation capacity, soil organic carbon, then inorganic carbon and also the FED they have predicted using the VisNIR DRS.

So, the predicted ordinal clay mineralogy levels for montmorillonite and kaolinite with 88 percent and 96 percent accuracy, they have also used they have also got 80 percent, 96 percent accuracy for predicting the montmorillonite and kaolinite using the spectral data. So, you can see that prediction accuracy varies from one property to another property for example, in case of clay we are getting 0.91 R squared values. In case of CEC, we are getting 0.83, in case of organic carbon we are they have got 0.87, in case of inorganic carbon they have got 0.86 and in case of FED they have got 0.77.

Although the accuracy varies from one property to another properties. In general, it can be seen that the VisNIR DRS has satisfactorily predicted most of the tested soil properties. So, it is one of the classical paper and in the field of VisNIR DRS soil application.

(Refer Slide Time: 21:06)



## CLASSICAL VISNIR DRS SOIL APPLICATION

Viscarra Rossel et al. (2006)

Table 1

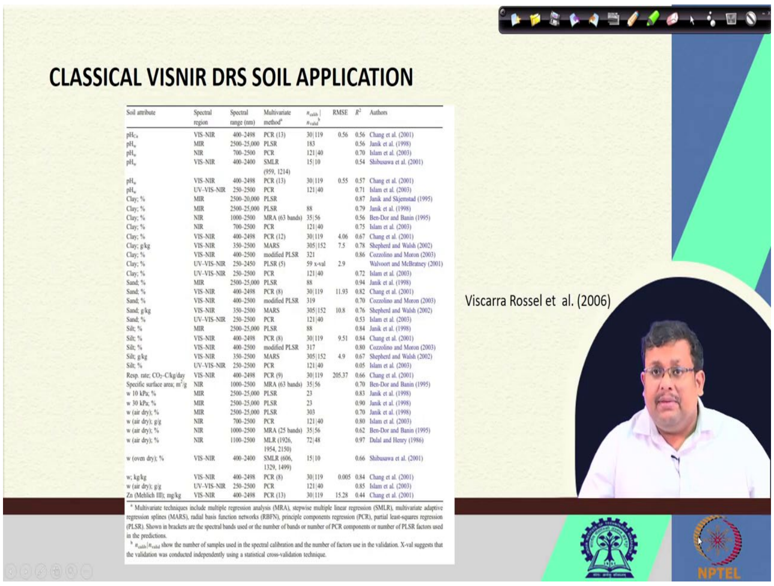
A review of the literature comparing quantitative predictions of various soil attributes using a multivariate statistical technique and spectral response in the ultra violet (UV), visible (VIS), near infrared (NIR) and mid infrared (MIR) regions of the electromagnetic spectrum

| Soil attribute | Spectral region | Spectral range (nm) | Multivariate method$^a$ | $n_{calib}$ \| $n_{valid}$ $^b$ | RMSE | $R^2$ | Authors |
|---|---|---|---|---|---|---|---|
| Acid (exch.); cmol/kg | VIS-NIR | 400-2498 | PCR (11) | 30\|119 | 24.4 | 0.65 | Chang et al. (2001) |
| Al (exch.); cmol/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.64 | Janik et al. (1998) |
| Biomass (N); mg/kg | NIR | 1100-2300 | PLSR (8) | 180 x-val | | 0.71 | Reeves and McCarty (2001) |
| Biomass (N); mg/kg | NIR | 1100-2498 | PLSR (6) | 120\|59 | | 0.79 | Reeves et al. (1999) |
| Biomass; g/kg | MIR | 2500-25,000 | PLSR | 23 | | 0.69 | Janik et al. (1998) |
| Biomass; mg/kg | VIS-NIR | 400-2498 | PCR (9) | 30\|119 | 389.71 | 0.60 | Chang et al. (2001) |
| C (inorg.); g/kg | MIR | 2500-25,000 | PLSR (16) | 177\|60 | | 0.98 | McCarty et al. (2002) |
| C (inorg.) g/kg | NIR | 1100-2498 | PLSR (19) | 177\|60 | | 0.87 | McCarty et al. (2002) |
| C (inorg.); g/kg | VIS-NIR | 400-2498 | PLSR (6) | 76\|32 | 0.15 | 0.96 | Chang and Laird (2002) |
| C (total); g/kg | MIR | 2500-25,000 | PLSR (17) | 177\|60 | | 0.95 | McCarty et al. (2002) |
| C (total); g/kg | NIR | 1100-2498 | PLSR (16) | 177\|60 | | 0.86 | McCarty et al. (2002) |
| C (total); g/kg | NIR | 1100-2498 | PLSR (7) | 120\|59 | | 0.96 | Reeves et al. (1999) |
| C (total); g/kg | VIS-NIR | 400-2498 | PLSR (5) | 76\|32 | 0.65 | 0.91 | Chang and Laird (2002) |
| C (total); g/kg | VIS-NIR | 400-2498 | PCR (7) | 30\|119 | 0.79 | 0.87 | Chang et al. (2001) |
| C; % | UV-VIS-NIR | 250-2450 | PLSR (6) | 59 x-val | 0.06 | | Walvoort and McBratney (2001) |
| C:N ratio | VIS-NIR | 400-2498 | PLSR (6) | 76\|32 | 0.21 | 0.88 | Chang and Laird (2002) |
| CEC; cmol(+)/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.88 | Janik et al. (1998) |
| CEC; cmol(+)/kg | NIR | 1000-2500 | MRA (63 bands) | 35\|56 | | 0.64 | Ben-Dor and Banin (1995) |
| CEC; mmol(+)/kg | NIR | 700-2500 | PCR | 121\|40 | | 0.67 | Islam et al. (2003) |
| CEC; cmol(+)/kg | VIS-NIR | 400-2498 | PCR (8) | 30\|119 | 38.2 | 0.81 | Chang et al. (2001) |
| CEC; cmol(+)/kg | VIS-NIR | 350-2500 | MARS | 493\|247 | 38 | 0.88 | Shepherd and Walsh (2002) |
| CEC; mmol(+)/kg | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.64 | Islam et al. (2003) |
| Σ exch. cations; cmol(+)/kg | MIR | 2500-20,000 | PLSR | | | 0.84 | Janik and Skjemstad (1995) |
| Ca; cmol/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.89 | Janik et al. (1998) |
| Ca; mmol(+)/kg | NIR | 700-2500 | PCR | 121\|40 | | 0.72 | Islam et al. (2003) |
| Ca; g/kg | VIS-NIR | 400-2500 | modified PLSR | 309 | | 0.90 | Cozzolino and Moron (2003) |
| Ca (exch.); cmol(+)/kg | VIS-NIR | 350-2500 | MARS | 493\|247 | 28 | 0.88 | Shepherd and Walsh (2002) |
| Ca (exch.); cmol(+)/kg | VIS-NIR | 400-2498 | PCR (12) | 30\|119 | 40 | 0.75 | Chang et al. (2001) |
| Ca; mmol(+)/kg | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.67 | Islam et al. (2003) |
| Carbonate; % | MIR | 2500-20,000 | PLSR | | | 0.95 | Janik and Skjemstad (1995) |
| Carbonate; % | NIR | 1000-2500 | MRA (25-3113 bands) | 22\|47 | | 0.69 | Ben-Dor and Banin (1995) |
| Cu; mg/kg | VIS-NIR | 400-2500 | modified PLSR | 310 | | 0.82 | Cozzolino and Moron (2003) |
| Cu (Mehlich III); mg/kg | VIS-NIR | 400-2498 | PCR (14) | 30\|119 | 5.37 | 0.25 | Chang et al. (2001) |
| EC; dS/m | MIR | 2500-25,000 | PLSR | 88 | | 0.23 | Janik et al. (1998) |
| EC; μS/cm | VIS-NIR | 400-2400 | SMLR (456, 984, 1014) | 15\|10 | | 0.65 | Shibusawa et al. (2001) |
| EC; mS/cm | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.10 | Islam et al. (2003) |
| Fe (DTPA); mg/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.55 | Janik et al. (1998) |
| Fe (free); % | NIR | 700-2500 | PCR | 121\|40 | | 0.49 | Islam et al. (2003) |
| Fe; mg/kg | VIS-NIR | 400-2500 | modified PLSR | 311 | | 0.90 | Cozzolino and Moron (2003) |
| Fe (Mehlich III); mg/kg | VIS-NIR | 400-2498 | PCR (9) | 30\|119 | 61.4 | 0.64 | Chang et al. (2001) |
| Fe (free); % | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.52 | Islam et al. (2003) |
| K; g/kg | VIS-NIR | 400-2500 | modified PLSR | 317 | | 0.72 | Cozzolino and Moron (2003) |
| K; mmol(+)/kg | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.00 | Islam et al. (2003) |
| K (exch.); mg/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.33 | Janik et al. (1998) |
| K (avail.); mg/kg | VIS-NIR | 400-1100 | NN | 41 | | 0.80 | Daniel et al. (2003) |
| K (exch.); cmol/kg | VIS-NIR | 400-2498 | PCR (13) | 30\|119 | 4.2 | 0.55 | Chang et al. (2001) |
| LR; t/ha | MIR | 2500-25,000 | PLSR | 188 | | 0.86 | Janik et al. (1998) |
| LR; t/ha | NIR | 700-2500 | PLSR | 188 | | 0.73 | Janik et al. (1998) |
| Mg (exch.); cmol/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.76 | Janik et al. (1998) |
| Mg; mmol(+)/kg | NIR | 700-2500 | PCR | 121\|40 | | 0.59 | Islam et al. (2003) |
| Mg; g/kg | VIS-NIR | 400-2500 | modified PLSR | 315 | | 0.90 | Cozzolino and Moron (2003) |
| Mg (exch.); cmol(+)/kg | VIS-NIR | 350-2500 | MARS | 493\|246 | 11 | 0.81 | Shepherd and Walsh (2002) |
| Mg (exch.); mg/kg | VIS-NIR | 400-2498 | PCR (9) | 30\|119 | 12.8 | 0.68 | Chang et al. (2001) |
| Mg; mmol(+)/kg | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.63 | Islam et al. (2003) |
| Mn (DTPA); mg/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.57 | Janik et al. (1998) |
| Mn (exch.); cmol/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.66 | Janik et al. (1998) |
| Mn (Mehlich III); mg/kg | VIS-NIR | 400-2498 | PCR (12) | 30\|119 | 56.4 | 0.70 | Chang et al. (2001) |
| N; % | MIR | 2500-20,000 | PLSR | | | 0.88 | Janik and Skjemstad (1995) |
| N; % | UV-VIS-NIR | 250-2450 | PLSR (11) | 59 x-val | 0.007 | | Walvoort and McBratney (2001) |
| N ($NO_3$); mg/100g | VIS-NIR | 400-2400 | SMLR (589, 1014) | 15\|10 | | 0.54 | Shibusawa et al. (2001) |
| N (miner.); mg/kg | NIR | 1100-2498 | PLSR (1) | 179 x-val | | 0.08 | Reeves et al. (1999) |
| N (pot. min); mg N/kg | VIS-NIR | 400-2498 | PCR (8) | 30\|119 | 26.05 | 0.72 | Chang et al. (2001) |
| N (active); mg/kg | NIR | 1100-2300 | PLSR (8) | 180 x-val | | 0.84 | Reeves and McCarty (2001) |
| N (active); mg/kg | NIR | 1100-2498 | PLSR (9) | 120\|59 | | 0.92 | Reeves et al. (1999) |
| N organic (total); % | MIR | 2500-25,000 | PLSR | 188 | | 0.86 | Janik et al. (1998) |
| N (total); % | NIR | 1100-2500 | MLR (1702, 1870, 2052) | 72\|48 | | 0.92 | Dalal and Henry (1986) |
| N (total); mg/kg | NIR | 1100-2300 | PLSR (10) | 180 x-val | | 0.94 | Reeves and McCarty (2001) |
| N (total); mg/kg | NIR | 1100-2498 | PLSR (8) | 120\|59 | | 0.95 | Reeves et al. (1999) |
| N (total); g/kg | VIS-NIR | 400-2498 | PLSR (7) | 76\|32 | 0.04 | 0.86 | Chang and Laird (2002) |
| N (total); g/kg | VIS-NIR | 400-2498 | PCR (7) | 30\|119 | 0.06 | 0.85 | Chang et al. (2001) |
| Na (exch.); cmol/kg | MIR | 2500-25,000 | PLSR | 183 | | 0.33 | Janik et al. (1998) |
| Na (exch.); cmol/kg | VIS-NIR | 400-2498 | PCR (7) | 30\|119 | 1.3 | 0.09 | Chang et al. (2001) |
| Na; mmol(+)/kg | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.34 | Islam et al. (2003) |
| OC; % | MIR | 2500-20,000 | PLSR | | | 0.92 | Janik and Skjemstad (1995) |
| OC; % | MIR | 2500-25,000 | PLSR | 188 | | 0.93 | Janik et al. (1998) |
| OC; g/kg | MIR | 2500-25,000 | PLSR (17) | 177\|60 | | 0.94 | McCarty et al. (2002) |
| OC; (acidified soil) g/kg | MIR | 2500-25,000 | PLSR (19) | 177\|60 | | 0.97 | McCarty et al. (2002) |
| OC; % | NIR | 1100-2500 | MLR (1744, 1870, 2052) | 72\|48 | | 0.93 | Dalal and Henry (1986) |
| OC; % | NIR | 1100-2500 | RBFN | 140\|60 | 0.32 | 0.96 | Fidêncio et al. (2002) |
| OC; % | NIR | 700-2500 | PCR | 121\|40 | | 0.68 | Islam et al. (2003) |
| OC; g/kg | NIR | 1100-2498 | PLSR (18) | 177\|60 | | 0.82 | McCarty et al. (2002) |
| OC; mg/kg | NIR | 1100-2300 | PLSR (8) | 180 x-val | | 0.94 | Reeves and McCarty (2001) |
| OC (acidified soil); g/kg | NIR | 1100-2498 | PLSR (17) | 177\|60 | | 0.80 | McCarty et al. (2002) |
| OC; g/kg | VIS-NIR | 400-2498 | PLSR (6) | 76\|32 | 0.62 | 0.89 | Chang and Laird (2002) |
| OC; g/kg | VIS-NIR | 350-2500 | MARS | 449\|225 | 0.31 | 0.80 | Shepherd and Walsh (2002) |
| OC; dag/kg | VIS-NIR | 350-1050 | PLSR (5) | 43\|25 | 0.36 | | Viscarra Rossel et al. (2003) |
| OC; % | UV-VIS-NIR | 250-2500 | PCR | 121\|40 | | 0.76 | Islam et al. (2003) |
| OM; % | MIR | 2500-25,000 | PLSR (4) | 31 x-val | 0.72 | 0.98 | Masserschmidt et al. (1999) |
| OM; % | NIR | 1000-2500 | MRA (30 bands) | 39\|52 | | 0.55 | Ben-Dor and Banin (1995) |
| OM; % | VIS-NIR | 400-1100 | NN | 41 | | 0.86 | Daniel et al. (2003) |
| OM; % | VIS-NIR | 400-2400 | SMLR (606, 1311, 1238) | 15\|10 | | 0.65 | Shibusawa et al. (2001) |
| P (avail.); mg/kg | MIR | 2500-25,000 | PLSR | 186 | | 0.07 | Janik et al. (1998) |
| P (avail.); mg/kg | VIS-NIR | 400-1100 | NN | 41 | | 0.81 | Daniel et al. (2003) |
| pH | MIR | 2500-20,000 | PLSR | | | 0.72 | Janik and Skjemstad (1995) |
| pH | NIR | 1100-2300 | PLSR (8) | 180 x-val | | 0.74 | Reeves and McCarty (2001) |
| pH | NIR | 1100-2498 | PLSR (11) | 120\|59 | | 0.73 | Reeves et al. (1999) |
| pH | VIS-NIR | 350-2500 | MARS | 505\|253 | 0.43 | 0.70 | Shepherd and Walsh (2002) |
| $pH_{Ca}$ | MIR | 2500-25,000 | PLSR | 183 | | 0.67 | Janik et al. (1998) |

Another classical paper was published by Viscarra Rossel et al in the year 2006. And they have compiled a table showing the application of this diffuse reflectance spectroscopy in predicting many soil properties, you can see here starting from the exchangeable acidity to from biomass to inorganic carbon to carbon then CN ratio, CEC, then calcium, copper, electrical conductivity, iron, potassium, then magnesium, lime requirement and also magnesium, then nitrogen mineralized nitrogen, then potentially mineralizable nitrogen, active nitrogen, organic nitrogen, total nitrogen, organic carbon and then huge organic matter, phosphorous, pH.

So, those the application of DRS, which has started back in 1990s they have compiled all the literature together for which we tried to predict the soil properties up to 2006 and they have compiled this table and not only they have compiled this table, from this table, we can see which spectral region was utilized because you will see the most of the application focused on VisNIR some applied the NIR, MIR mid infrared region and also the spectral range they have utilized is also mentioned then the multivariate method they have used is also mentioned and then calibration validation R squared values are also given here, RMSE values are given here. So, it shows it gives you a complete picture of the history of diffuse reflectance spectroscopy application in soil from 1992 up to 2006.

(Refer Slide Time: 23:21)



CLASSICAL VISNIR DRS SOIL APPLICATION

| Soil attribute | Spectral region | Spectral range (nm) | Multivariate method[a] | $n_{calib}$ \| $n_{valid}$[b] | RMSE | $R^2$ | Authors |
|---|---|---|---|---|---|---|---|
| $pH_{Ca}$ | VIS–NIR | 400–2498 | PCR (13) | 30\|119 | 0.56 | 0.56 | Chang et al. (2001) |
| $pH_w$ | MIR | 2500–25,000 | PLSR | 183 | | 0.56 | Janik et al. (1998) |
| $pH_w$ | NIR | 700–2500 | PCR | 121\|40 | | 0.70 | Islam et al. (2003) |
| $pH_w$ | VIS–NIR | 400–2400 | SMLR (959, 1214) | 15\|10 | | 0.54 | Shibusawa et al. (2001) |
| $pH_w$ | VIS–NIR | 400–2498 | PCR (13) | 30\|119 | 0.55 | 0.57 | Chang et al. (2001) |
| $pH_w$ | UV–VIS–NIR | 250–2500 | PCR | 121\|40 | | 0.71 | Islam et al. (2003) |
| Clay; % | MIR | 2500–20,000 | PLSR | | | 0.87 | Janik and Skjemstad (1995) |
| Clay; % | MIR | 2500–25,000 | PLSR | 88 | | 0.79 | Janik et al. (1998) |
| Clay; % | NIR | 1000–2500 | MRA (63 bands) | 35\|56 | | 0.56 | Ben-Dor and Banin (1995) |
| Clay; % | NIR | 700–2500 | PCR | 121\|40 | | 0.75 | Islam et al. (2003) |
| Clay; % | VIS–NIR | 400–2498 | PCR (12) | 30\|119 | 4.06 | 0.67 | Chang et al. (2001) |
| Clay; g/kg | VIS–NIR | 350–2500 | MARS | 305\|152 | 7.5 | 0.78 | Shepherd and Walsh (2002) |
| Clay; % | VIS–NIR | 400–2500 | modified PLSR | 321 | | 0.86 | Cozzolino and Moron (2003) |
| Clay; % | UV–VIS–NIR | 250–2450 | PLSR (5) | 59 x-val | 2.9 | | Walvoort and McBratney (2001) |
| Clay; % | UV–VIS–NIR | 250–2500 | PCR | 121\|40 | | 0.72 | Islam et al. (2003) |
| Sand; % | MIR | 2500–25,000 | PLSR | 88 | | 0.94 | Janik et al. (1998) |
| Sand; % | VIS–NIR | 400–2498 | PCR (8) | 30\|119 | 11.93 | 0.82 | Chang et al. (2001) |
| Sand; % | VIS–NIR | 400–2500 | modified PLSR | 319 | | 0.70 | Cozzolino and Moron (2003) |
| Sand; g/kg | VIS–NIR | 350–2500 | MARS | 305\|152 | 10.8 | 0.76 | Shepherd and Walsh (2002) |
| Sand; % | UV–VIS–NIR | 250–2500 | PCR | 121\|40 | | 0.53 | Islam et al. (2003) |
| Silt; % | MIR | 2500–25,000 | PLSR | 88 | | 0.84 | Janik et al. (1998) |
| Silt; % | VIS–NIR | 400–2498 | PCR (8) | 30\|119 | 9.51 | 0.84 | Chang et al. (2001) |
| Silt; % | VIS–NIR | 400–2500 | modified PLSR | 317 | | 0.80 | Cozzolino and Moron (2003) |
| Silt; g/kg | VIS–NIR | 350–2500 | MARS | 305\|152 | 4.9 | 0.67 | Shepherd and Walsh (2002) |
| Silt; % | UV–VIS–NIR | 250–2500 | PCR | 121\|40 | | 0.05 | Islam et al. (2003) |
| Resp. rate; $CO_2$–C/kg/day | VIS–NIR | 400–2498 | PCR (9) | 30\|119 | 205.37 | 0.66 | Chang et al. (2001) |
| Specific surface area; $m^2/g$ | NIR | 1000–2500 | MRA (63 bands) | 35\|56 | | 0.70 | Ben-Dor and Banin (1995) |
| w 10 kPa; % | MIR | 2500–25,000 | PLSR | 23 | | 0.83 | Janik et al. (1998) |
| w 30 kPa; % | MIR | 2500–25,000 | PLSR | 23 | | 0.90 | Janik et al. (1998) |
| w (air dry); % | MIR | 2500–25,000 | PLSR | 303 | | 0.70 | Janik et al. (1998) |
| w (air dry); g/g | NIR | 700–2500 | PCR | 121\|40 | | 0.80 | Islam et al. (2003) |
| w (air dry); % | NIR | 1000–2500 | MRA (25 bands) | 35\|56 | | 0.62 | Ben-Dor and Banin (1995) |
| w (air dry); % | NIR | 1100–2500 | MLR (1926, 1954, 2150) | 72\|48 | | 0.97 | Dalal and Henry (1986) |
| w (oven dry); % | VIS–NIR | 400–2400 | SMLR (606, 1329, 1499) | 15\|10 | | 0.66 | Shibusawa et al. (2001) |
| w; kg/kg | VIS–NIR | 400–2498 | PCR (8) | 30\|119 | 0.005 | 0.84 | Chang et al. (2001) |
| w (air dry); g/g | UV–VIS–NIR | 250–2500 | PCR | 121\|40 | | 0.85 | Islam et al. (2003) |
| Zn (Mehlich III); mg/kg | VIS–NIR | 400–2498 | PCR (13) | 30\|119 | 15.28 | 0.44 | Chang et al. (2001) |

[a] Multivariate techniques include multiple regression analysis (MRA), stepwise multiple linear regression (SMLR), multivariate adaptive regression splines (MARS), radial basis function networks (RBFN), principle components regression (PCR), partial least-squares regression (PLSR). Shown in brackets are the spectral bands used or the number of bands or number of PCR components or number of PLSR factors used in the predictions.

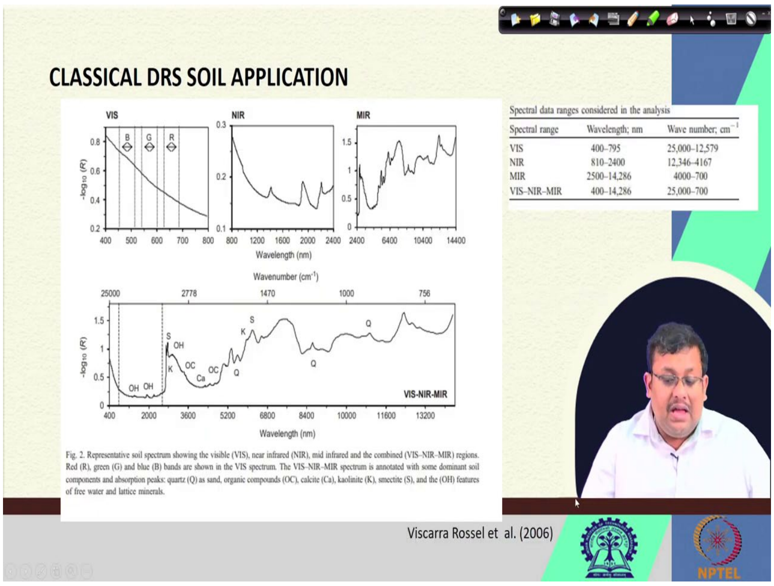
[b] $n_{calib}$\|$n_{valid}$ show the number of samples used in the spectral calibration and the number of factors use in the validation. X-val suggests that the validation was conducted independently using a statistical cross-validation technique.

Viscarra Rossel et al. (2006)

So, it was it is one of the classical paper in the VisNIR DRS soil domain. Also you can see that pH, clay, sand, silt and several other properties, zinc extracted by (())(23:34) extraction. So, all these they have compiled and they have shown the application of DRS using different spectral region like MIR, VisNIR and they have shown the multivariate method also. So, we can have an idea about what are the soil properties which are very much useful, which are very much sensitive to these spectral methods.

And we can see that clay, sand, silt and then organic carbon and many other soil properties are sensitive to the spectral methods and the spectral method has a widespread application for predicting the major important soil fertility and pedological properties for their rapid and cost effective prediction.

(Refer Slide Time: 24:29)



CLASSICAL DRS SOIL APPLICATION

Spectral data ranges considered in the analysis

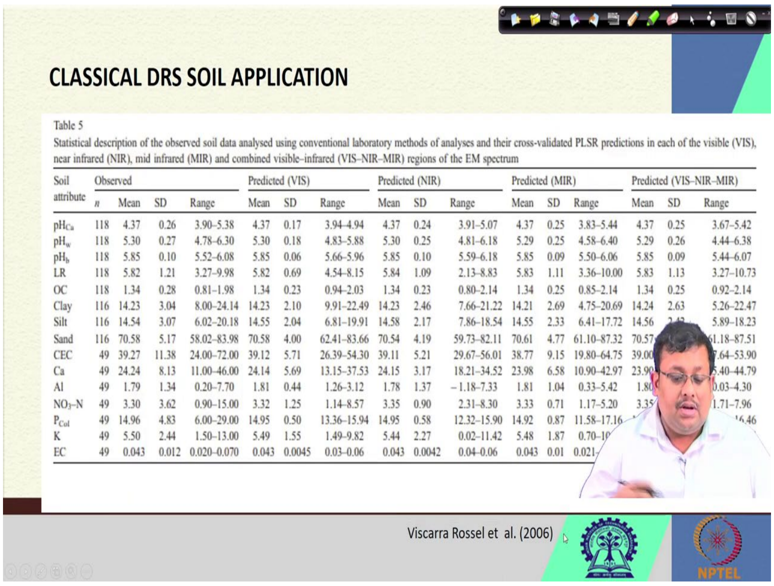
| Spectral range | Wavelength; nm | Wave number; $cm^{-1}$ |
|---|---|---|
| VIS | 400–795 | 25,000–12,579 |
| NIR | 810–2400 | 12,346–4167 |
| MIR | 2500–14,286 | 4000–700 |
| VIS–NIR–MIR | 400–14,286 | 25,000–700 |

Fig. 2. Representative soil spectrum showing the visible (VIS), near infrared (NIR), mid infrared and the combined (VIS–NIR–MIR) regions. Red (R), green (G) and blue (B) bands are shown in the VIS spectrum. The VIS–NIR–MIR spectrum is annotated with some dominant soil components and absorption peaks: quartz (Q) as sand, organic compounds (OC), calcite (Ca), kaolinite (K), smectite (S), and the (OH) features of free water and lattice minerals.

Viscarra Rossel et al. (2006)

Also, they have showed these combined absorbance spectra from this visible zone and the near infrared zone, and mid infrared zone. They have used these visible zone from 400 to 795 nanometer, NIR from 810 to 2400 nanometer, MIR from 2500 to 14,286 nanometer and then the whole VisNIR, MIR region they have used from 400 to 14,286 nanometer and they have used it.

(Refer Slide Time: 25:13)



CLASSICAL DRS SOIL APPLICATION

Table 5

Statistical description of the observed soil data analysed using conventional laboratory methods of analyses and their cross-validated PLSR predictions in each of the visible (VIS), near infrared (NIR), mid infrared (MIR) and combined visible–infrared (VIS–NIR–MIR) regions of the EM spectrum

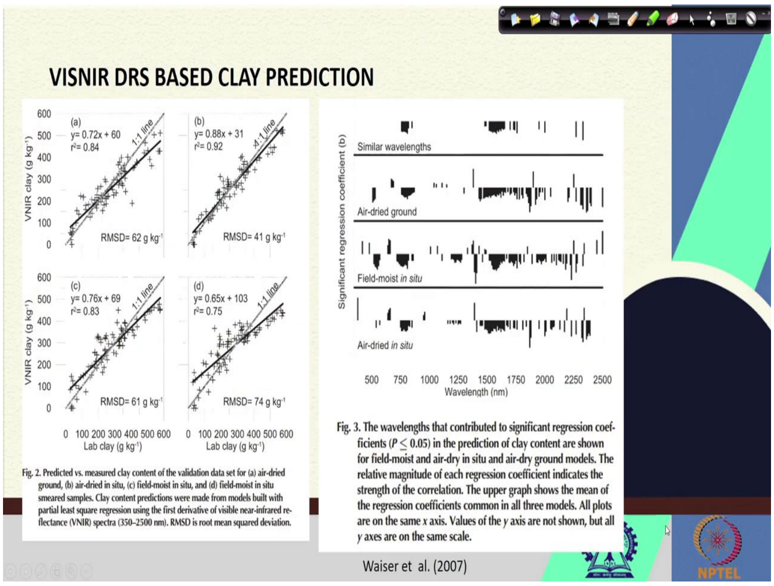
| Soil attribute | Observed | | | | Predicted (VIS) | | | Predicted (NIR) | | | Predicted (MIR) | | | Predicted (VIS–NIR–MIR) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | Mean | SD | Range | Mean | SD | Range | Mean | SD | Range | Mean | SD | Range | Mean | SD | Range |
| $pH_{Ca}$ | 118 | 4.37 | 0.26 | 3.90–5.38 | 4.37 | 0.17 | 3.94–4.94 | 4.37 | 0.24 | 3.91–5.07 | 4.37 | 0.25 | 3.83–5.44 | 4.37 | 0.25 | 3.67–5.42 |
| $pH_w$ | 118 | 5.30 | 0.27 | 4.78–6.30 | 5.30 | 0.18 | 4.83–5.88 | 5.30 | 0.25 | 4.81–6.18 | 5.29 | 0.25 | 4.58–6.40 | 5.29 | 0.26 | 4.44–6.38 |
| $pH_b$ | 118 | 5.85 | 0.10 | 5.52–6.08 | 5.85 | 0.06 | 5.66–5.96 | 5.85 | 0.10 | 5.59–6.18 | 5.85 | 0.09 | 5.50–6.06 | 5.85 | 0.09 | 5.44–6.07 |
| LR | 118 | 5.82 | 1.21 | 3.27–9.98 | 5.82 | 0.69 | 4.54–8.15 | 5.84 | 1.09 | 2.13–8.83 | 5.83 | 1.11 | 3.36–10.00 | 5.83 | 1.13 | 3.27–10.73 |
| OC | 118 | 1.34 | 0.28 | 0.81–1.98 | 1.34 | 0.23 | 0.94–2.03 | 1.34 | 0.23 | 0.80–2.14 | 1.34 | 0.25 | 0.85–2.14 | 1.34 | 0.25 | 0.92–2.14 |
| Clay | 116 | 14.23 | 3.04 | 8.00–24.14 | 14.23 | 2.10 | 9.91–22.49 | 14.23 | 2.46 | 7.66–21.22 | 14.21 | 2.69 | 4.75–20.69 | 14.24 | 2.63 | 5.26–22.47 |
| Silt | 116 | 14.54 | 3.07 | 6.02–20.18 | 14.55 | 2.04 | 6.81–19.91 | 14.58 | 2.17 | 7.86–18.54 | 14.55 | 2.33 | 6.41–17.72 | 14.56 | [illegible] | 5.89–18.23 |
| Sand | 116 | 70.58 | 5.17 | 58.02–83.98 | 70.58 | 4.00 | 62.41–83.66 | 70.54 | 4.19 | 59.73–82.11 | 70.61 | 4.77 | 61.10–87.32 | 70.57 | [illegible] | 61.18–87.51 |
| CEC | 49 | 39.27 | 11.38 | 24.00–72.00 | 39.12 | 5.71 | 26.39–54.30 | 39.11 | 5.21 | 29.67–56.01 | 38.77 | 9.15 | 19.80–64.75 | 39.00 | [illegible] | [illegible] |
| Ca | 49 | 24.24 | 8.13 | 11.00–46.00 | 24.14 | 5.69 | 13.15–37.53 | 24.15 | 3.17 | 18.21–34.52 | 23.98 | 6.58 | 10.90–42.97 | 23.90 | [illegible] | [illegible] |
| Al | 49 | 1.79 | 1.34 | 0.20–7.70 | 1.81 | 0.44 | 1.26–3.12 | 1.78 | 1.37 | −1.18–7.33 | 1.81 | 1.04 | 0.33–5.42 | 1.80 | [illegible] | 0.03–4.30 |
| $NO_3$–N | 49 | 3.30 | 3.62 | 0.90–15.00 | 3.32 | 1.25 | 1.14–8.57 | 3.35 | 0.90 | 2.31–8.30 | 3.33 | 0.71 | 1.17–5.20 | 3.35 | [illegible] | 1.71–7.96 |
| $P_{Col}$ | 49 | 14.96 | 4.83 | 6.00–29.00 | 14.95 | 0.50 | 13.36–15.94 | 14.95 | 0.58 | 12.32–15.90 | 14.92 | 0.87 | 11.58–17.16 | [illegible] | [illegible] | [illegible] |
| K | 49 | 5.50 | 2.44 | 1.50–13.00 | 5.49 | 1.55 | 1.49–9.82 | 5.44 | 2.27 | 0.02–11.42 | 5.48 | 1.87 | [illegible] | [illegible] | [illegible] | [illegible] |
| EC | 49 | 0.043 | 0.012 | 0.020–0.070 | 0.043 | 0.0045 | 0.03–0.06 | 0.043 | 0.0042 | 0.04–0.06 | 0.043 | 0.01 | [illegible] | [illegible] | [illegible] | [illegible] |

Viscarra Rossel et al. (2006)

So, also not only they have used it they have predicted different soil properties using these different zones. They have used both visible zone and also the NIR zone and mid infrared zone and also combined visible near infrared and mid infrared zone and they have shown the, their prediction accuracy also, I mean you can see their range, they have seen they have shown the mean and also standard deviation they are range for all the soil properties.

So, these are the statistical description of the observed soil data. So, these are the soil data they have tried to predict using these individual visible NIR, MIR and combine visible near infrared and mid infrared region and they have predicted and they have compared the prediction accuracy using these 3 to 4 different zones.

(Refer Slide Time: 26:25)



VISNIR DRS BASED CLAY PREDICTION

Fig. 2. Predicted vs. measured clay content of the validation data set for (a) air-dried ground, (b) air-dried in situ, (c) field-moist in situ, and (d) field-moist in situ smeared samples. Clay content predictions were made from models built with partial least square regression using the first derivative of visible near-infrared reflectance (VNIR) spectra (350–2500 nm). RMSD is root mean squared deviation.

Fig. 3. The wavelengths that contributed to significant regression coefficients ($P \leq 0.05$) in the prediction of clay content are shown for field-moist and air-dry in situ and air-dry ground models. The relative magnitude of each regression coefficient indicates the strength of the correlation. The upper graph shows the mean of the regression coefficients common in all three models. All plots are on the same *x* axis. Values of the *y* axis are not shown, but all *y* axes are on the same scale.

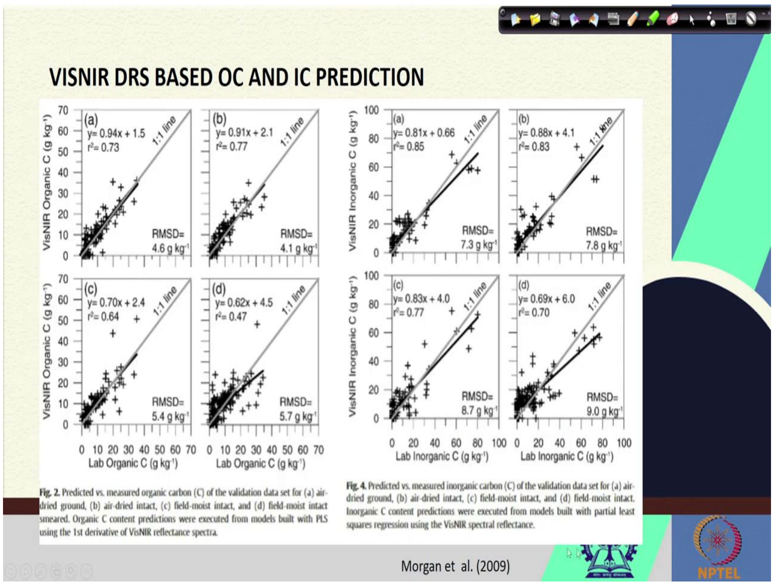
Waiser et al. (2007)

Another great application of VisNIR DRS is for soil was soil clay prediction, which was published by Waiser et al in 2007 in the Soil Science Society of America Journal, you can see here they have predicted the soil clay content in different conditions. For example, here you can see this is they predicted for air dried ground condition and this is for air dried in situ condition and then field moist in situ condition and also field moist smeared condition.

In all these you can see that combining all these you can conclude that for all the applications VisNIR DRS was able to satisfactorily predict the clay content. Of course, the highest prediction accuracy in terms of R square was exhibited in case of these air dried in situ condition, but at the same time, when in field moist condition also they have shown very good results. Not only that, but also they have identified the wavelengths that contributed to significant regression coefficients which are statistically significant.

So, this type of selection of the significant wavelengths are very much helpful for future simplification of the prediction methodology as well as for these will be also helpful if somebody wants to develop a future simplified instrument based on only these selected spectral wavelengths.

(Refer Slide Time: 28:21)



VISNIR DRS BASED OC AND IC PREDICTION

**Fig. 2.** Predicted vs. measured organic carbon (C) of the validation data set for (a) air-dried ground, (b) air-dried intact, (c) field-moist intact, and (d) field-moist intact smeared. Organic C content predictions were executed from models built with PLS using the 1st derivative of VisNIR reflectance spectra.

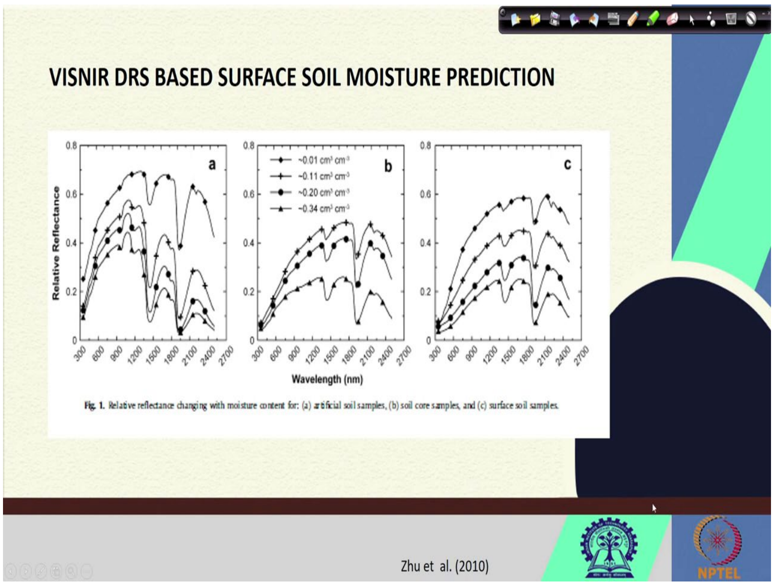
**Fig. 4.** Predicted vs. measured inorganic carbon (C) of the validation data set for (a) air-dried ground, (b) air-dried intact, (c) field-moist intact, and (d) field-moist intact. Inorganic C content predictions were executed from models built with partial least squares regression using the VisNIR spectral reflectance.

Morgan et al. (2009)

So, a Morgan et al also has shown the application of VisNIR DRS for predicting the organic carbon and also the inorganic carbon and you can see their results in different condition like air dried ground condition, air dried intact condition and then field moist intact condition and field moist intact smeared condition and in most of the condition they have seen that VisNIR DRS was able to predict the soil organic carbon satisfactorily.

Similar in case of inorganic carbon also in these 4 condition they have shown the R squared values ranging from 0.70 to 0.85 which shows the adequate accuracy for predicting the inorganic carbon using the diffuse reflectance spectroscopy.

(Refer Slide Time: 29:12)



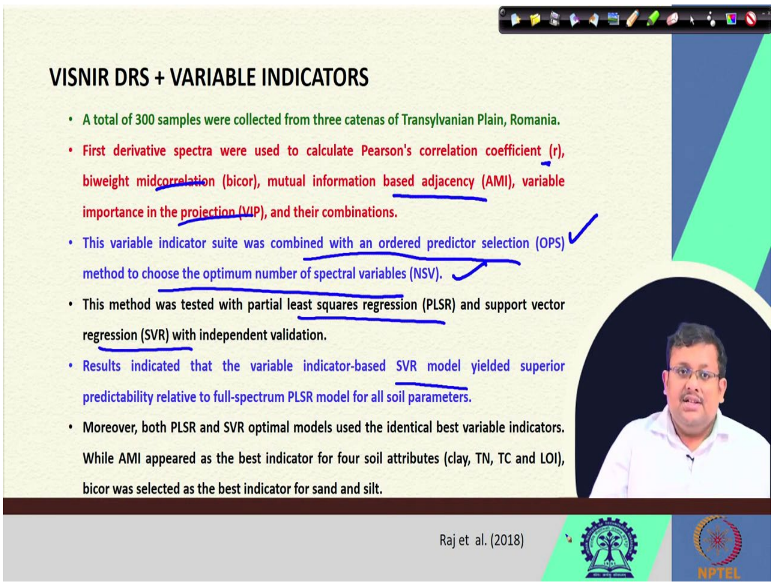
## VISNIR DRS BASED SURFACE SOIL MOISTURE PREDICTION

Fig. 1. Relative reflectance changing with moisture content for: (a) artificial soil samples, (b) soil core samples, and (c) surface soil samples.

Zhu et al. (2010)

Another applications Zhu et al in 2010 which was published in Journal of Hydrology they have shown that VisNIR DRS is able to predict the surface soil moisture condition, surface soil moisture and using different zones like you can see here 1300 for 52, 1450 nanometer and then 1890 to 1990 nanometer and then 2220 to 2280 nanometer they have predicted the moisture content and in all these conditions they have got very good R square values you can see here starting from 0.93, 0.93, 0.93. So, all these conditions they have got a R squared values of 0.93 which is quite good.

(Refer Slide Time: 30:04)



## VISNIR DRS + VARIABLE INDICATORS

- A total of 300 samples were collected from three catenas of Transylvanian Plain, Romania.
- First derivative spectra were used to calculate Pearson's correlation coefficient (r), biweight midcorrelation (bicor), mutual information based adjacency (AMI), variable importance in the projection (VIP), and their combinations.
- This variable indicator suite was combined with an ordered predictor selection (OPS) method to choose the optimum number of spectral variables (NSV).
- This method was tested with partial least squares regression (PLSR) and support vector regression (SVR) with independent validation.
- Results indicated that the variable indicator-based SVR model yielded superior predictability relative to full-spectrum PLSR model for all soil parameters.
- Moreover, both PLSR and SVR optimal models used the identical best variable indicators. While AMI appeared as the best indicator for four soil attributes (clay, TN, TC and LOI), bicor was selected as the best indicator for sand and silt.

Raj et al. (2018)

So, another application which was published by one of my student in 2018. In there, in this paper, we have tried to we have used 300 soil samples, we have collected from the three catenas of Transylvanian Plain of Romania, and then we did the first derivative transformation and also to calculate we have calculated. After the first derivative transformation, we have calculated the Pearson correlation coefficient, then biweight midcorrelation, then mutual information based adjacency and then the variable importance in the projection.

So, all these are and their also combinations. So, these are known as the variable indicators. So, it shows the correlation between different variables. So, these variable indicator suite which is basically R, bicor, AMI, VIP and their combination was combined with an order prediction selection method which is called OPS and then we choose the optimum number of spectral variables. So, the idea behind this research was instead of using all the spectral variables from 350 to 2500 nanometer, can we pinpoint a handful of wavelengths, so, that they can give the comparable prediction accuracy or better prediction accuracy.

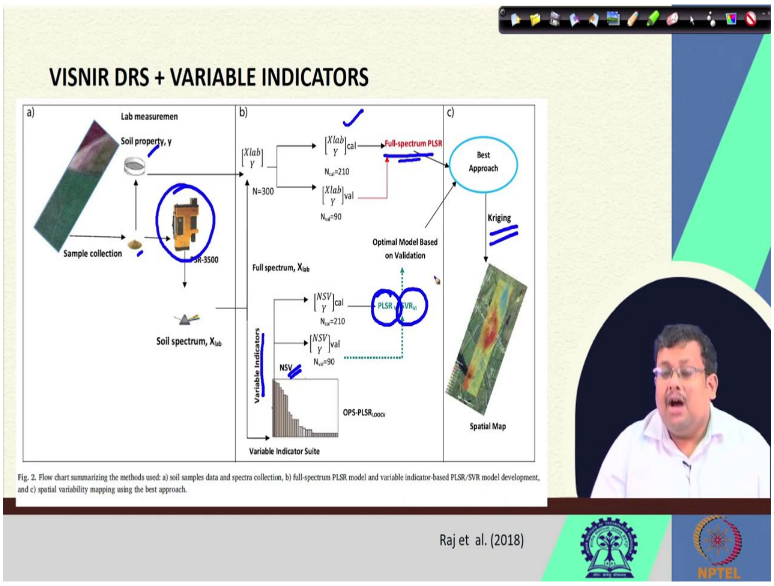
So, we have used this variable indicator suite which is basically composed of these Pearson correlation coefficient, biweight midcorrelation, mutual information based adjacent metric adjacency and also variable importance in the projection and their combinations also and then we have all we have followed the ordered predictor selection approach and then we have selected the optimum number of spectral variables.

So, once we have followed this methodology and we tested with the partial least squares regression and support vector regression, and we compared the accuracy. So, our results indicated that if we combine these variable indicated with more based wavelength selection followed by support vector regression model, it will yield superior prediction relative to the full spectrum PLSR model for all the soil parameters.

So, that shows that using full, using these variable indicator base wavelengths as well and in combination with support vector regression can outperform the partial least squares regression using all the spectral features. Moreover, both PLSR and SVR optimal model used to use the identical based variable indicators. So, while AMI appeared as the best indicator for 4 soil attributes here we have used clay, total nitrogen, total carbon and loss on ignition, organic matter; bicor was selected as the best indicator for sand and silt.

So, this is how we have identified the best processing method I would say for selecting the important spectral variables followed by the best prediction method or chemometric model to give you the optimum result.

(Refer Slide Time: 33:42)



VISNIR DRS + VARIABLE INDICATORS

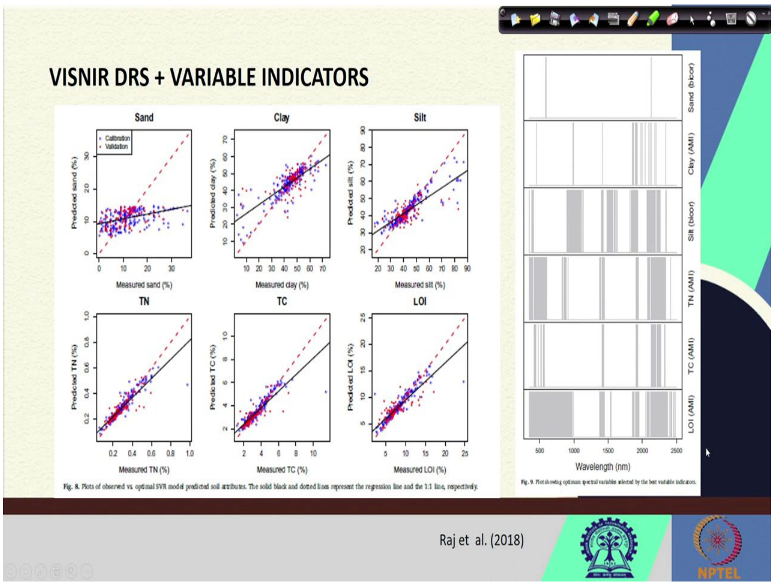
Fig. 2. Flow chart summarizing the methods used: a) soil samples data and spectra collection, b) full-spectrum PLSR model and variable indicator-based PLSR/SVR model development, and c) spatial variability mapping using the best approach.

Raj et al. (2018)

So, this is the total schematic diagram of this total experiment. So, we have collected the samples from the area and then after collection, we have scanned them using the PSR 3500 spectral evolution and then simultaneously we have used the whole spectral variables and did the full spectrum PLSR model which is traditional and then simultaneously we have used the variable indicators and then we have used these OPS ordered predictor selection approach and then after the OPS we have selected the number of important specter variables and then we have used these both partial least square regression as well as support vector regression and then we compare the results.

We have seen support vector results are superior to that partial least square regression. So, ultimately after selecting the best approach, we have utilized the Kriging interpolation to map the soil properties special map, to produce the special map of the soil properties.

(Refer Slide Time: 34:53)



VISNIR DRS + VARIABLE INDICATORS

Fig. 8. Plots of observed vs. optimal SVR model predicted soil attributes. The solid black and dotted lines represent the regression line and the 1:1 line, respectively.

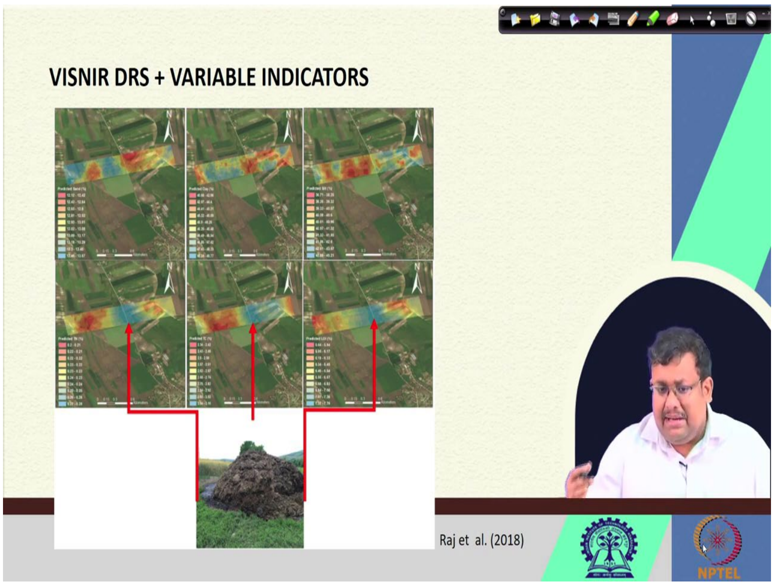
Fig. 9. Plot showing optimum spectral variables selected by the best variable indicators.

Raj et al. (2018)

So, that shows that this method, so, these are the prediction models you can see the predicted soil properties versus measured soil properties this for sand, clay, silt, total nitrogen, total carbon and loss on ignition you can see for total nitrogen and total carbon, loss on ignition, we are getting high accuracy and we have also selected the important spectral variables the optimum spectral variables selected by the based variable indicator.

So, we have compared these based variable indicators for using different types of variable indicators you can see AMI was selected for most of the soil properties, but for silt and sand bicor was selected as the most important variable indicator.

(Refer Slide Time: 35:42)



VISNIR DRS + VARIABLE INDICATORS

Raj et al. (2018)

Not only that also, we have produced the you can see here this special map of soil properties this is for sand, this is sand map, clay map, silt map, total nitrogen, total carbon and loss on ignition and you can see one important thing that these special mapping using the spectral model has identified these zone in the middle as the content of high total nitrogen also total carbon and also loss on ignition, organic matter. And so, we can see that this was actually the location of a heap of or pile of manure.

So, that shows that as expected this area will be dominated by total nitrogen, total carbon and loss on ignition, organic matter or soil organic matter. So, that shows that our approach was useful for predicting the special variability of the soil properties and it is a kind of a ground truth which validated that the accuracy of our approach. So, this is these are a couple of examples of the classical soil science applications of VisNIR DRS.

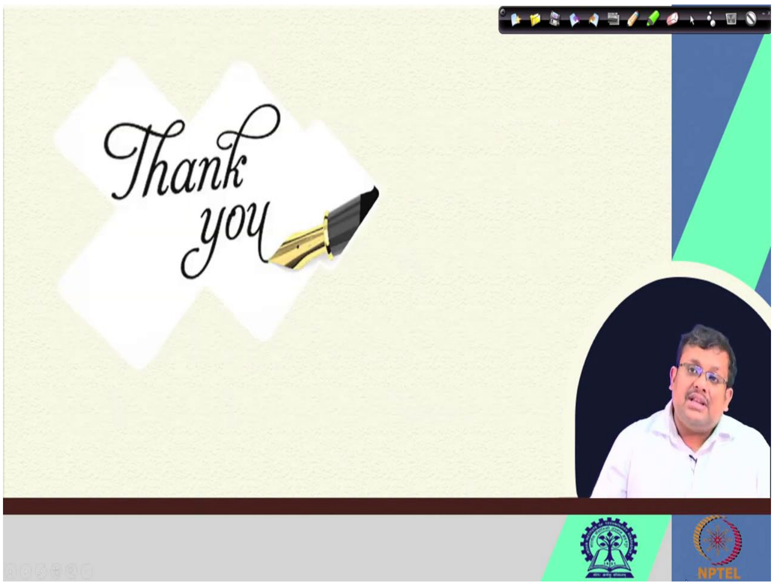
(Refer Slide Time: 37:09)

## REFERENCES

- Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. Geoderma 132 (3–4), 273–290.
- Morgan, C.L.S., Waiser, T.H., Brown, D.J., Hallmark, C.T., 2009. Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. Geoderma 151, 249–256.
- Raj, A., Chakraborty, S., Duda, B.M., Weindorf, D.C., Li, B., Roy, S., et al., 2018. Soil mapping via diffuse reflectance spectroscopy based on variable indicators: an ordered predictor selection approach. Geoderma 314, 146–159.
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131, 59–75.
- Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Science Society of America Journal 71, 389–396.

And hopefully you have learned something new and these are the references which I have used for this lecture. And if you are interested, you should go ahead, reading with this, reading these papers to get more and more information regarding this soil application. And let us wrap up our lecture here.

(Refer Slide Time: 37:34)



And we will start from here. We will see some more soil examples and then we will go with some crop applications also in our next lecture. Thank you.