# Visible near-infrared (VNIR) and mid-infrared (MIR) Lab spectroscopy for soil texture classification: Analysis of machine learning and spectral bands selection techniques

Presented by- Ternikar Chirag Rajendra

PhD student under guidance of Prof. D Nagesh Kumar

Department of Civil Engineering, IISc, Bangalore

Date of Joining – 31/07/2018

INDIAN INSTITUTE OF SCIENCE

**International Workshop On Land Cover/Land Use Changes, Forestry, and Agriculture in South/Southeast Asia**

# Highlights

- VNIR, MIR and VNIR+MIR lab spectroscopy and machine learning was used to classify soils as per USDA texture triangle.
- Models were compared for classification with reduced number of spectral bands as opposed to all the bands.
- Physical interpretation of the important bands selected was tabulated.
- Inaccuracies in classification of individual texture classes were compared in terms of neighbour and far classes accuracy.

# Contents

- Importance of Soil as function and it’s properties
- How have researchers classified soil texture
- Proposed Objectives and Methodology
- Physiochemical and Spectroscopy dataset
- Evaluation strategy for classification
- What did we find? Is it useful?
- What should be done next?

# Introduction-Importance of Soil

- Soil is a vital component of Earth surface providing ecosystem services, filters water, supplies nutrients to plants, provides us with food, fibre and energy, stores carbon and regulates the emissions of greenhouse gases and it affects our climate
- Unprecedented pressures on soil from degradation and urbanisation, threatening above functions, agro-ecological balances and food security. Sustainable soil management is important as per Sustainable Development Goals (Goal 2 – Zero Hunger)
- India Total Land Area – 328 M Ha
- Soil Health card Scheme for India was conceptualized in 2015:
  - The government is planning to cover as many as all farmers under the scheme
  - The scheme will cover all the parts of the country
  - In the form of soil card, the farmers will get a report and this report will contain all the details about their particular farm
  - A farm will get the soil card once in every 3 years
- It will contain status of the soil with respect to 12 parameters: (N, P, K, S, Zn, Fe, Cu, Mn, B, pH, EC, OC)

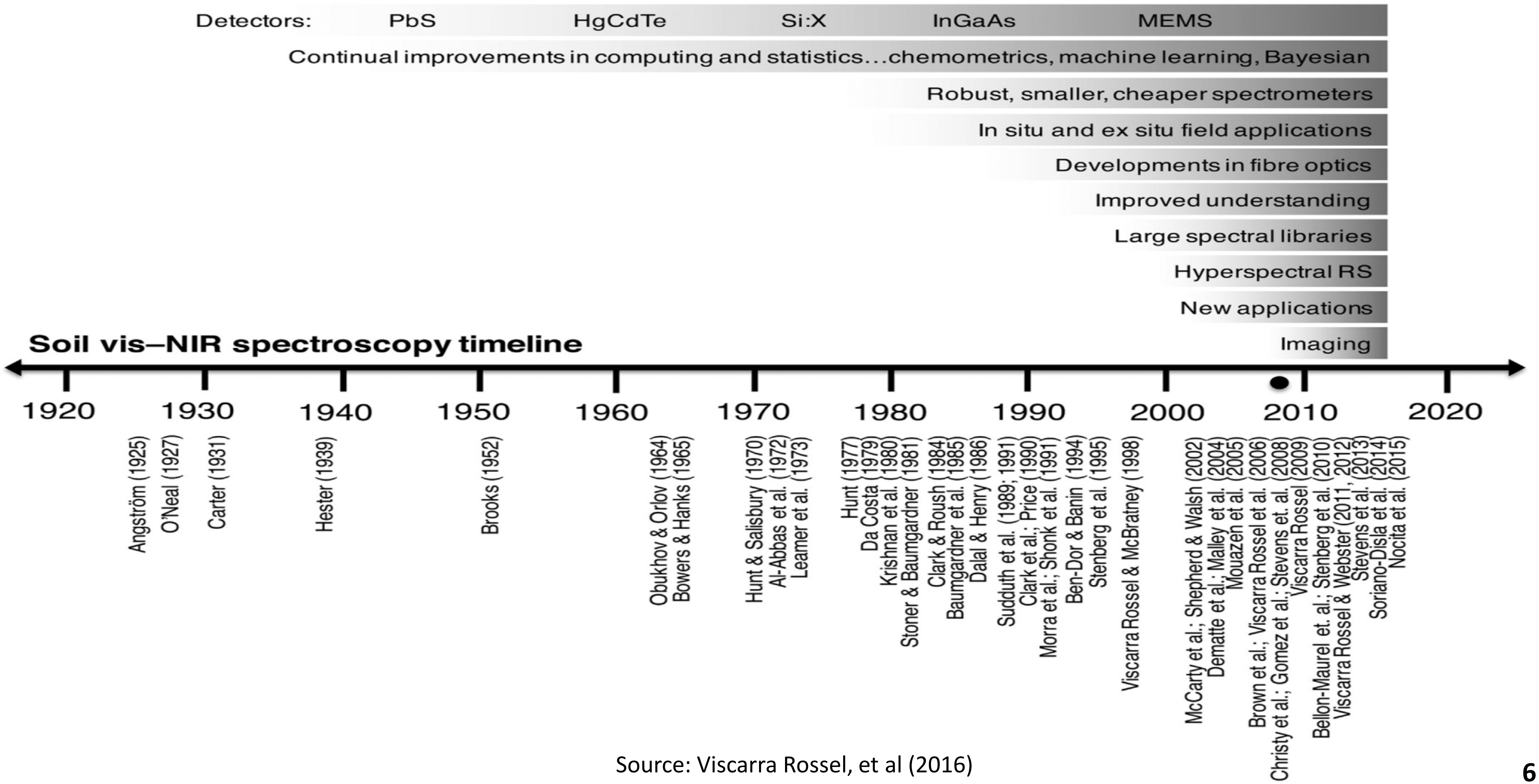
Source: https://sustainabledevelopment.un.org & https://www.soilhealth.dac.gov.in

# Soil Mapping in India

| Organization | Type of Survey | Scale | Area/Districts covered (M ha) |
|---|---|---|---|
| NBSS&LUP (National Bureau of Soil Survey and Land Use Planning) | Small scale soil mapping | 1:2,50,000 | 300.5 |
| | Soil resource mapping | 1:50,000 | 198.4 |
| | Detailed soil survey | 1:4,000/15,000 | 8.48 |
| | Detailed soil survey (Sujala III project) | Cadastral | 11 districts |
| | Detailed soil survey (LRI flagship programme) | Cadastral | 115 blocks (22 states) |
| SLUSI (State Land Use Survey of India) | Rapid reconnaissance survey for watershed prioritization | 1:50,000 | 200 |
| | Land degradation mapping | 1:50,000 | 65 districts |
| | Detailed soil survey | 1:4,000/15,000 | 13.5 |
| | Soil resource mapping under NRIS (DOS) | 1:50,000 | 89 districts |
| NRSC (National Remote Sensing Centre) | Waste land mapping | 1:50,000 | India |
| | Soil resource mapping under NRIS (DOS) | 1:50,000 | 200 |

Source: Dharumarajan et. al (2019)

# Literature Review




Source: Viscarra Rossel, et al (2016)

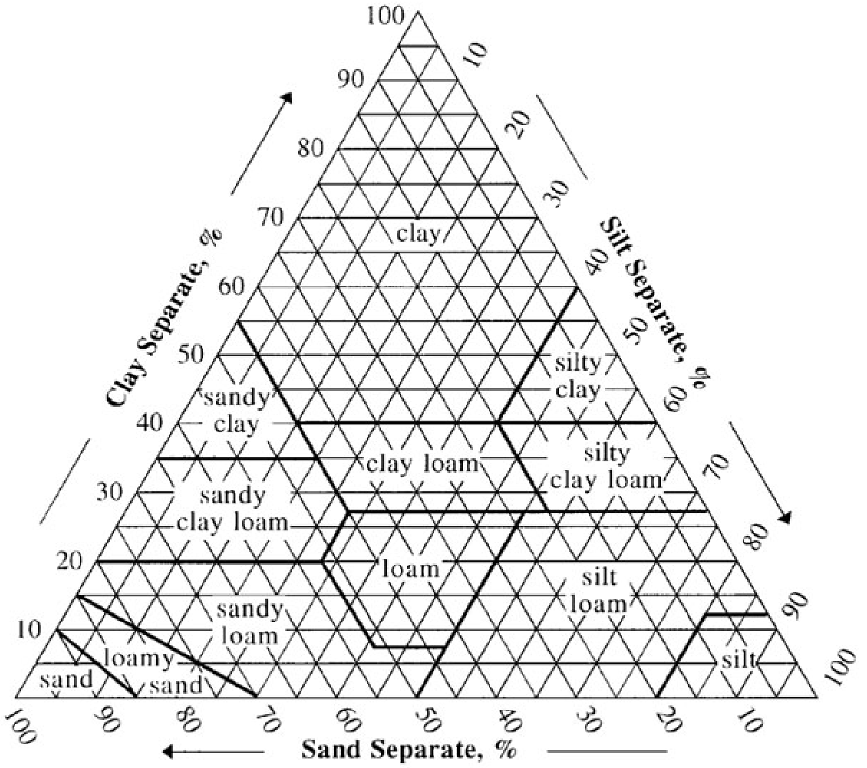
# Literature Review

- Currently, new technologies are used to produce miniaturised, rugged and economical hand-held instruments
- Adoption of VNIR and MIR in laboratories is taking traction
- The black disc in 2008 represents the conception of the global soil spectroscopy project
- Characterize soil and its variability
- Deriving a spectral classification to describe the associations between spectra, soil, land cover and geography
- Usefulness of the global database for predicting soil attributes, such as soil organic and inorganic carbon, clay, silt, sand and iron contents, cation exchange capacity, pH and many other properties.
- DSM (Digital Soil Mapping) efforts on large scale:
  - OzDSM, Australia– 2008
  - Global Soil Map, DSM, UN – 2007
  - DSM, Europe – 2010
  - ISRIC, World Soil Information – 2005
  - India, NBSS&LUP – 2018

Source: Viscarra Rossel, et al (2016)

# Soil Texture

- Soil texture data are utilised in following studies (Agriculture, Water Resources, Landscape Management):
  - Crop suitability
  - Crop yield and growth pattern
  - Precision Agriculture
  - Surface runoff modelling
  - Soil erosion modelling
  - Soil moisture patterns
  - Slope-stability analysis
  - Disaster mitigation and management
  - Landscape management
  - Belowground C, N storages

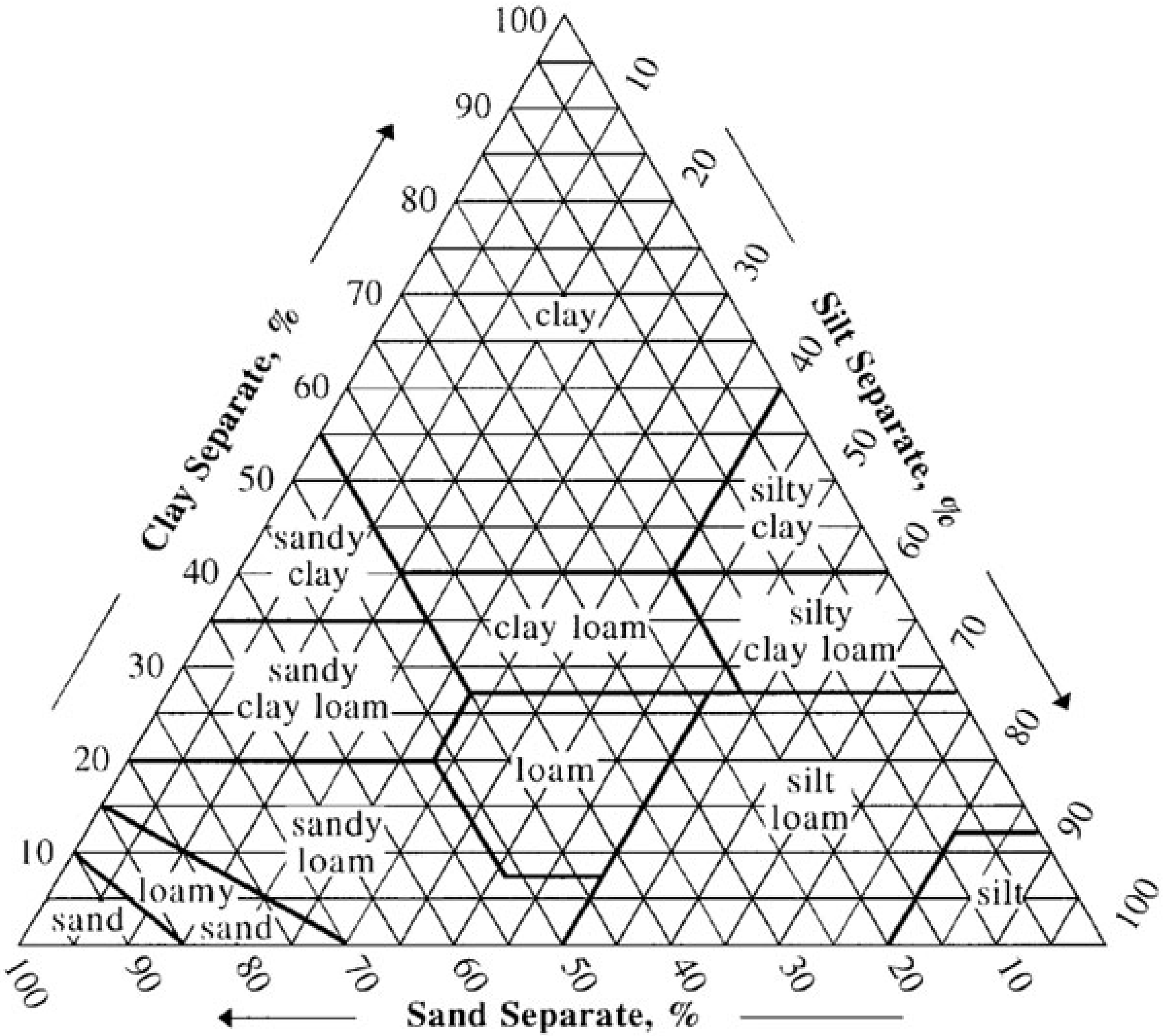



- DSM (Digital Soil Mapping) provides topsoil properties. No depth information/profile.

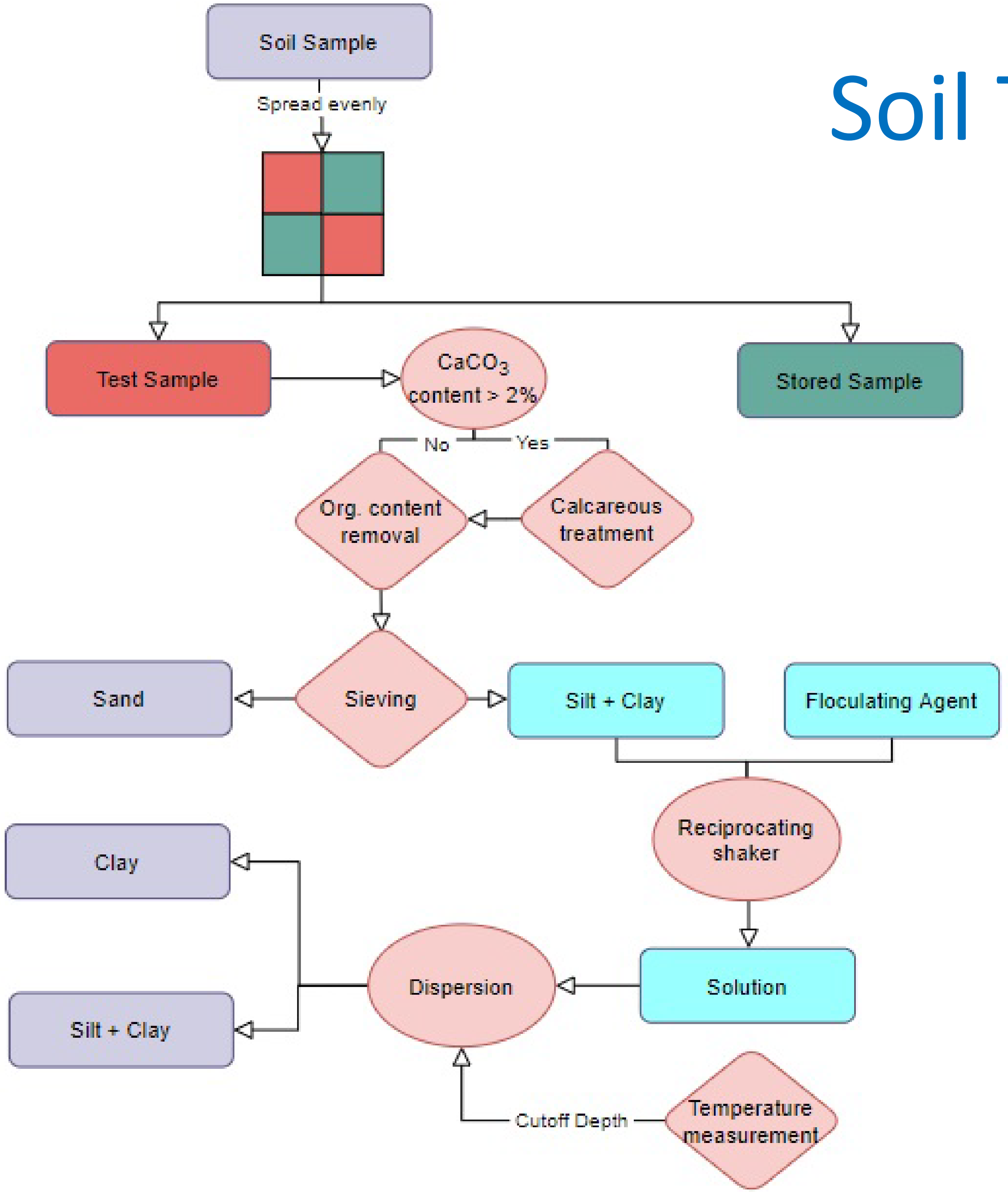
# Soil Texture Classification

| Name of soil separate | USDA Diameter Limits (mm) | WRB Diameter Limits (mm) |
|---|---|---|
| Clay | < 0.002 | < 0.002 |
| Silt | 0.002 – 0.05 | 0.002 – 0.063 |
| Very fine sand | 0.05 – 0.10 | 0.063 – 0.125 |
| Fine sand | 0.10 – 0.25 | 0.125 – 0.20 |
| Medium sand | 0.25 – 0.50 | 0.20 – 0.63 |
| Coarse sand | 0.50 – 1.00 | 0.63 – 1.25 |
| Very coarse sand | 1.00 – 2.00 | 1.25 – 2.00 |
| Coarse | > 2.00 | > 2.00 |

USDA – United States Department of Agriculture
WRB – World Reference Base for Soil




Source: http://www.fao.org

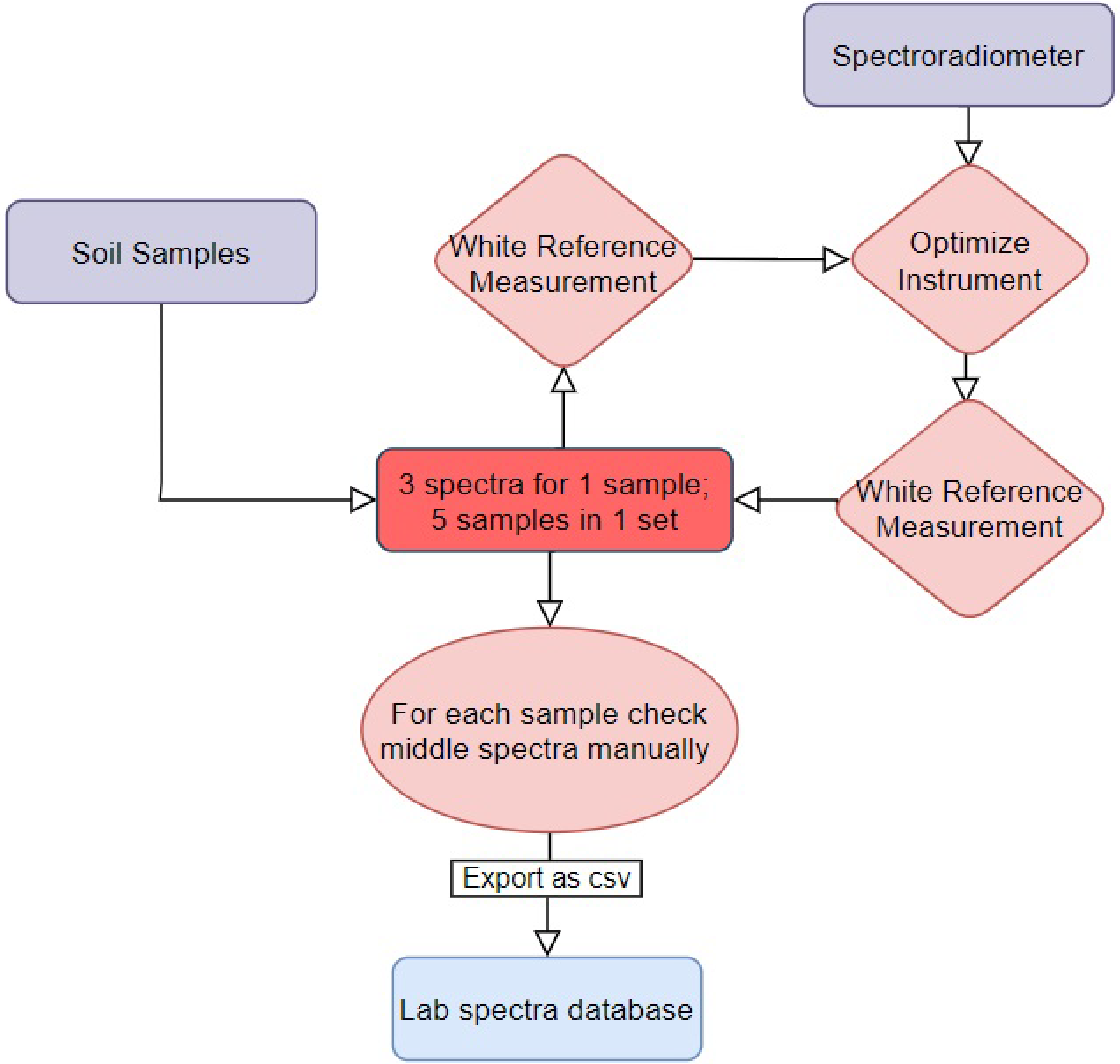



# Soil Texture Measurement

- Particle size distribution using International Pipette method / Robinson's pipette method
- Test sample is dried, tested for calcium carbonate and treated for calcareous content and organic content removal
- Sand is separated using sieving
- Clay is separated using dispersion.
- Silt content is calculated as:

  Silt % = 100 – (Sand % + Clay %)
- Time taken: ~3-5 days

# Lab Spectra Measurement



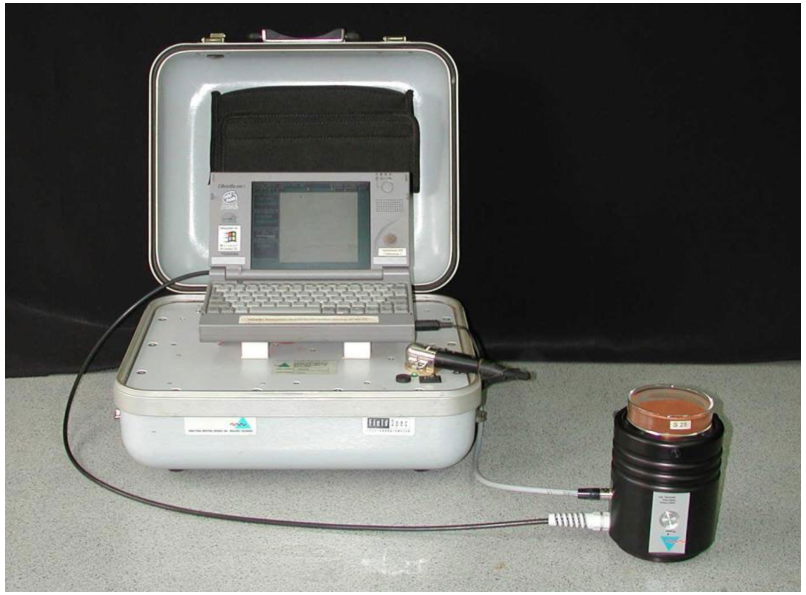

**Materials:**

- 1 Spectroradiometer
- 1 White reference panel, 5x5'
- 2-4 Tungsten Halogen Lamp

**Set up:**

- Instrument warm up time – 60 mins
- Lamps warm up time – 20 mins
- No. of spectra per target – 30
- Dark current average of scans – 50
- White reference number of scans – 50

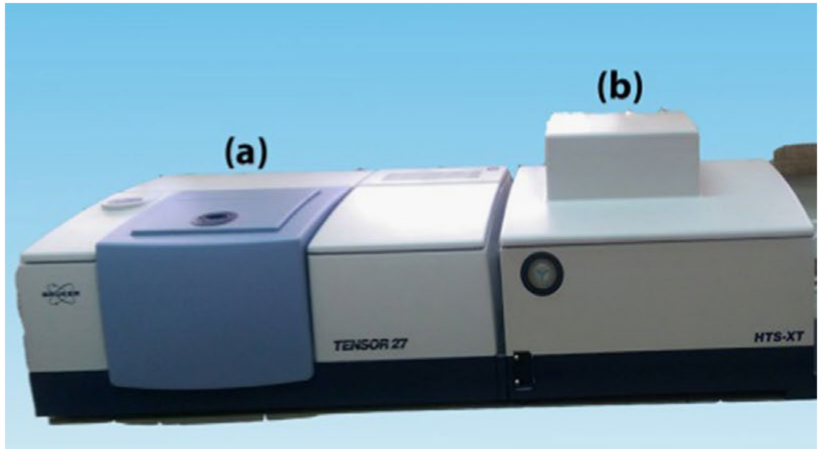
# Lab Spectra Measurement



ASD Field Spec VNIR with optical setup

**VNIR** setup: ASD Spectroradiometer and optical accessory for diffuse reflectance measurement

Source: ICRAF-ISRIC, 2019, Vagen et al., 2020




Bruker MIR FT-IR (a) with HTS-XT setup (b)


Aluminium plate filled with soil samples

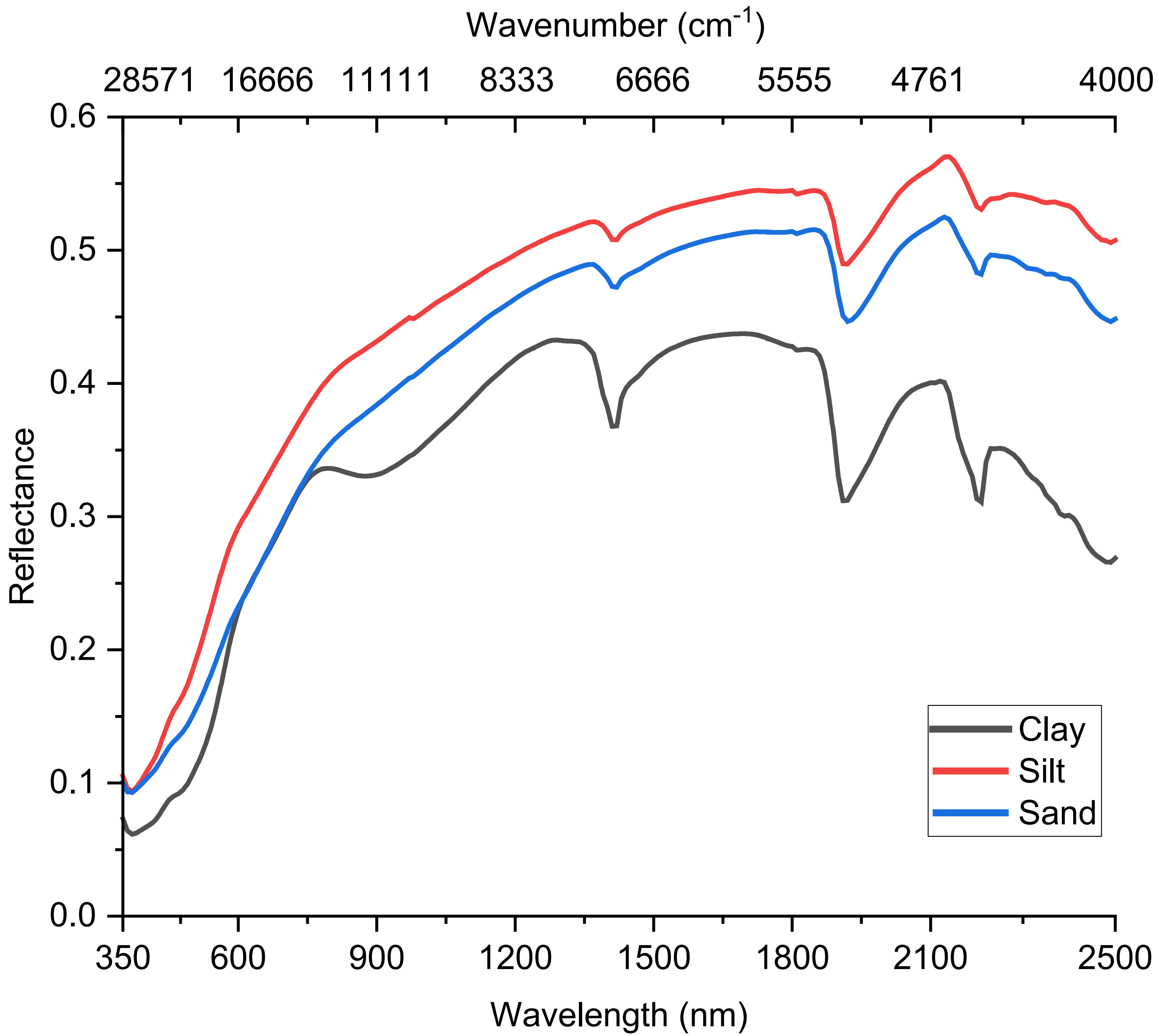
**MIR** setup: Bruker MIR FT-IR with HTS-XT for diffuse reflectance measurement

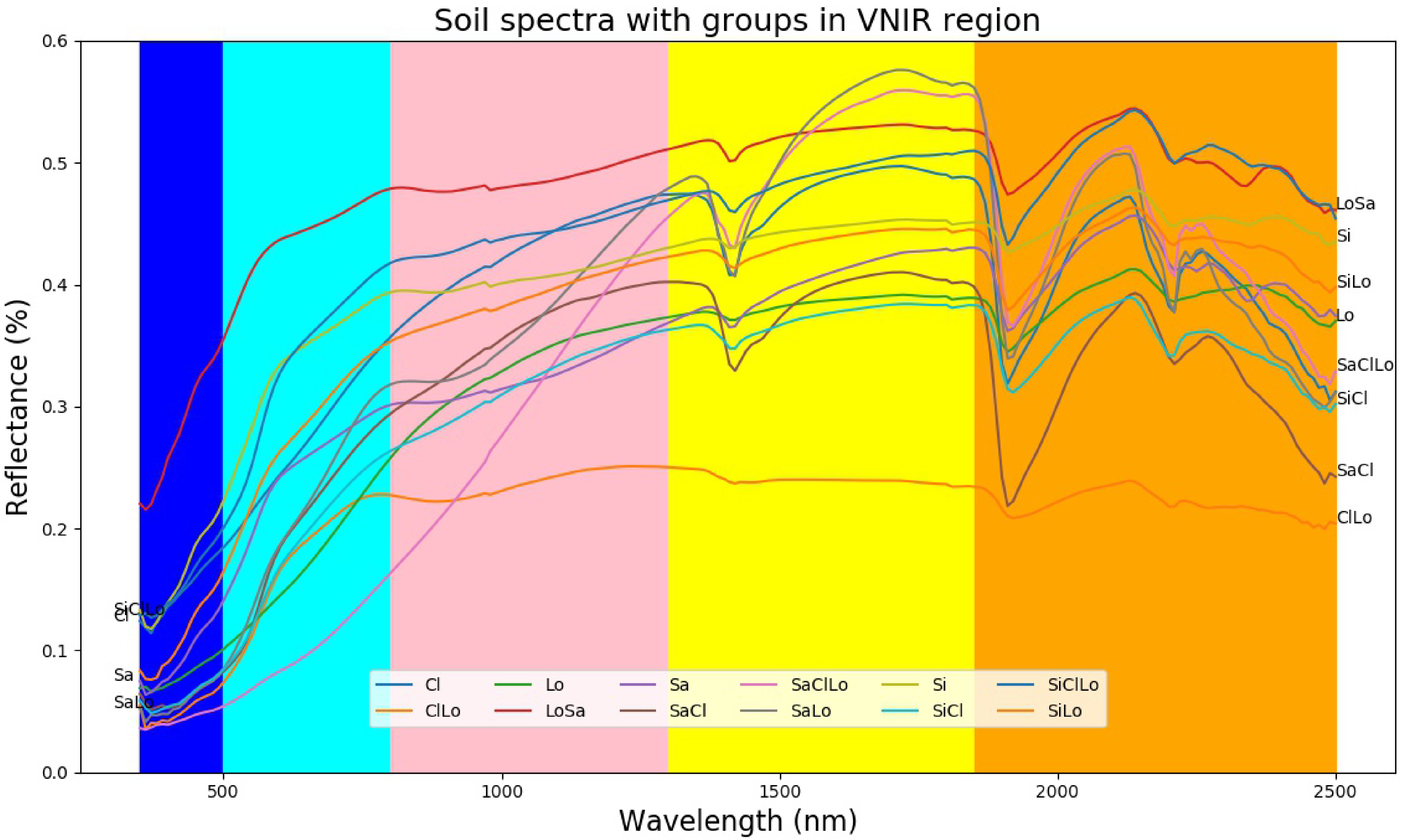
# Physiochemical and Spectroscopy dataset

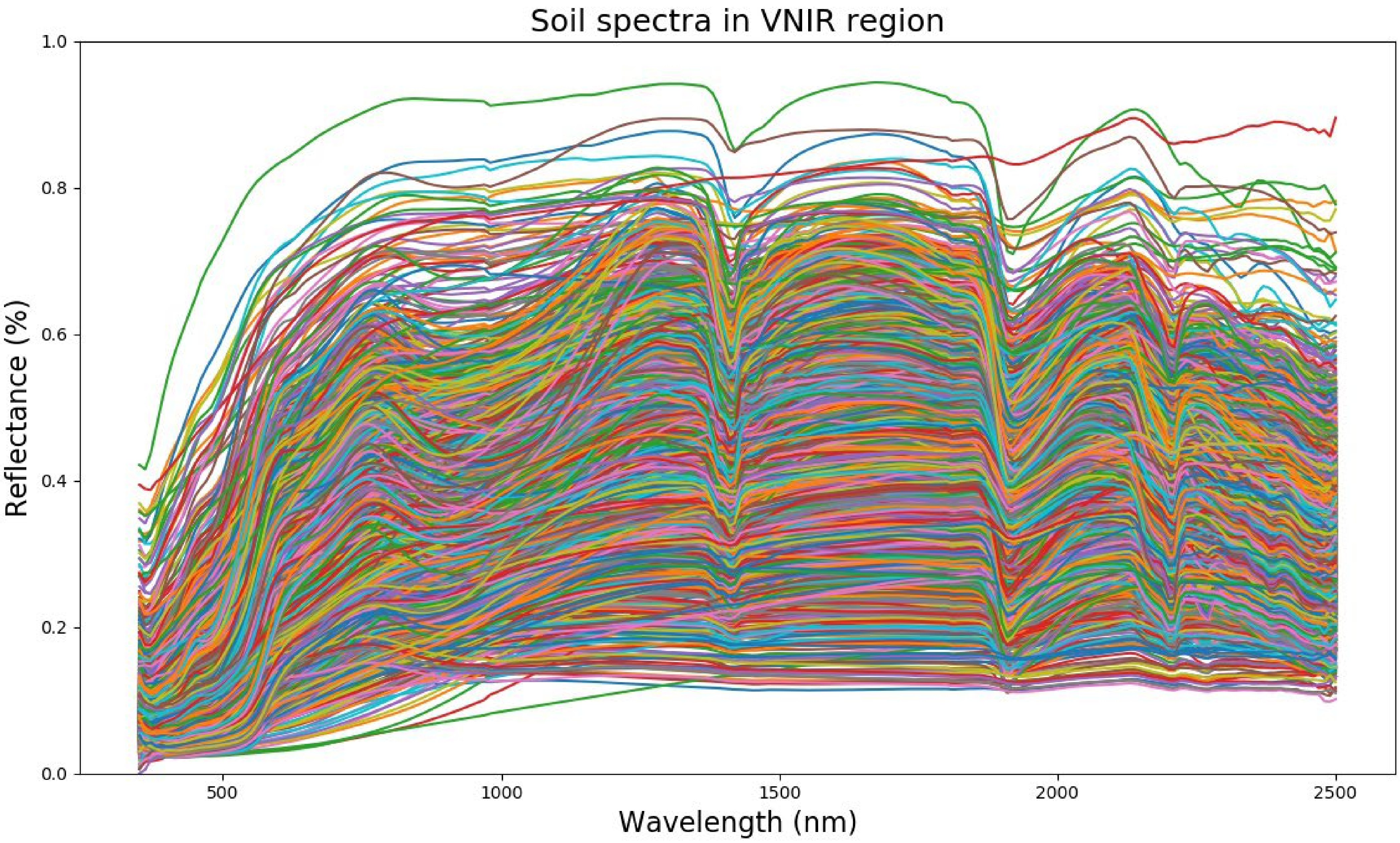
- ICRAF-ISRIC Soil Spectral Library
  - 4438 soil samples from 754 profiles totally in 58 countries
  - Soil texture by International pipette method in different labs and time
  - After preprocessing 3643 samples were used
- VNIR spectra
  - 204 bands (@10 nm) in 410-2440 nm
  - ASD field spec at ICRAF plant lab
- MIR spectra
  - 1762 bands (@ 4 $cm^{-1}$) in 2441-14286 nm
  - Bruker FTIR at ICRAF plant lab
- VNIR+MIR spectra
  - 1966 bands in 410-14286 nm

Source: Garrity & Bindraban, 2004; Shepherd & Walsh, 2007

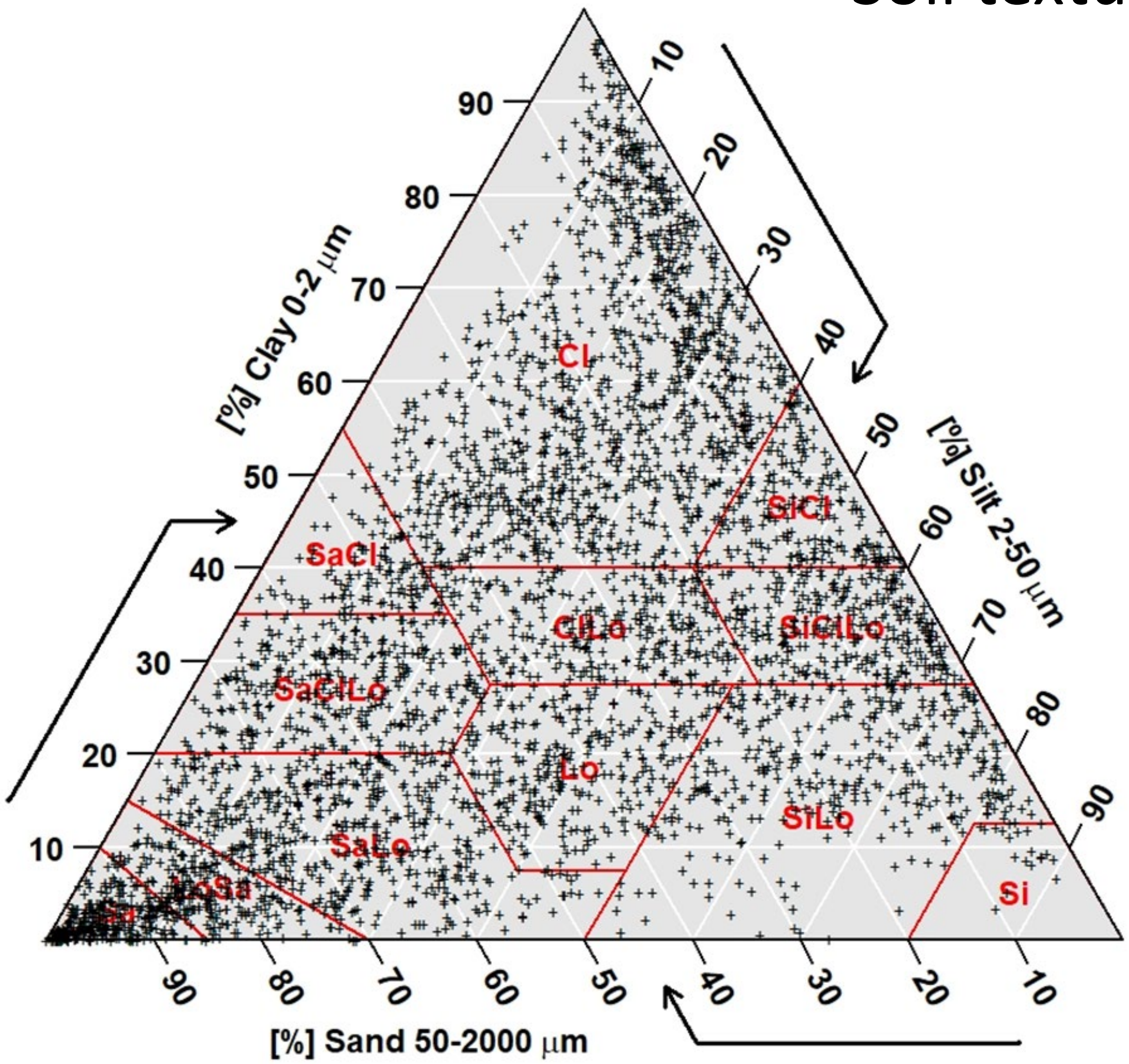
# Soil spectra in VNIR region

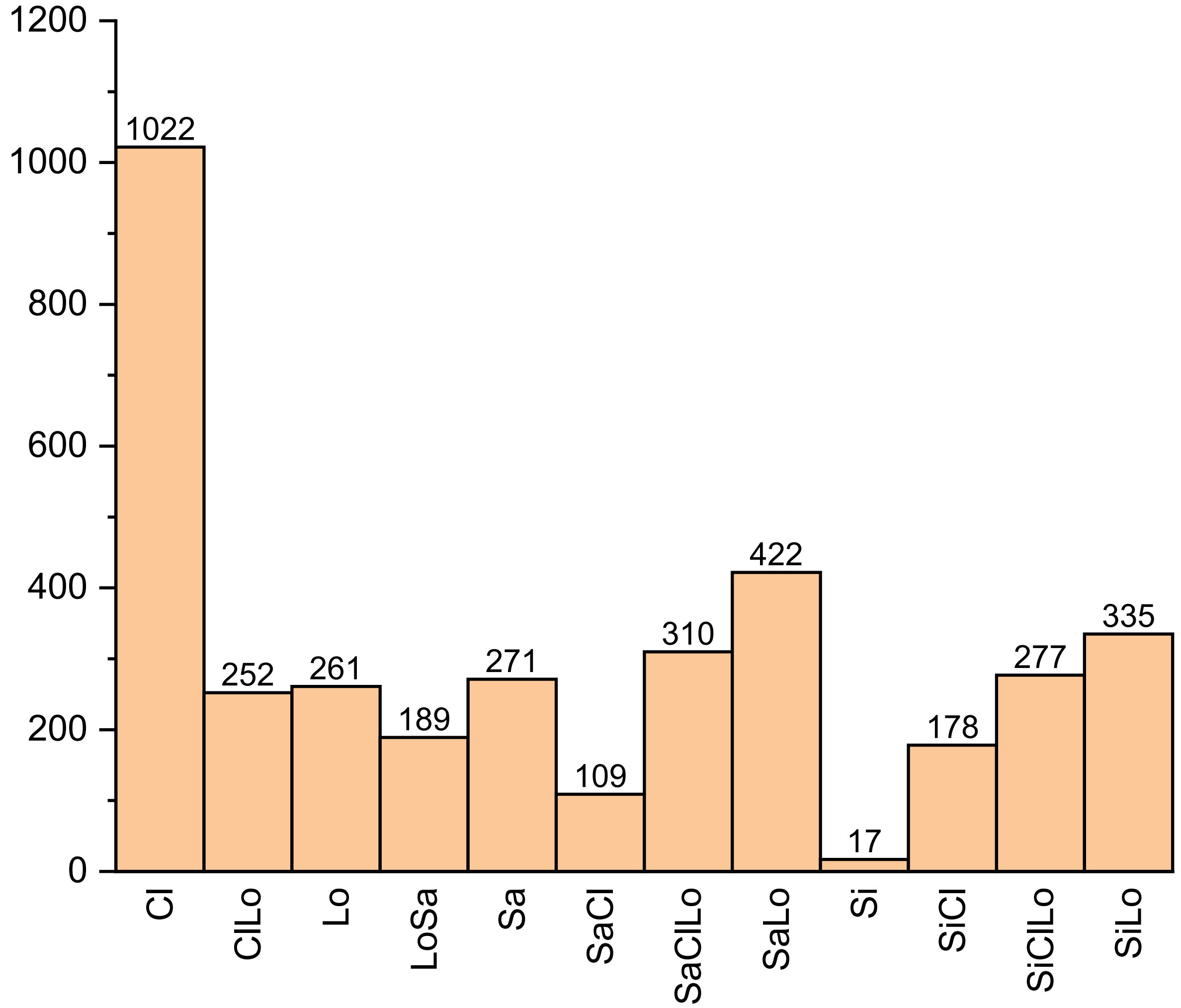



Soil spectra with groups in VNIR region
Reflectance (%)
Wavelength (nm)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
500
1000
1500
2000
2500
LoSa
Si
SiLo
Lo
SaClLo
SiCl
SaCl
ClLo
SiClLo
Cl
Sa
SaLo
Cl
ClLo
Lo
LoSa
Sa
SaCl
SaClLo
SaLo
Si
SiCl
SiClLo
SiLo

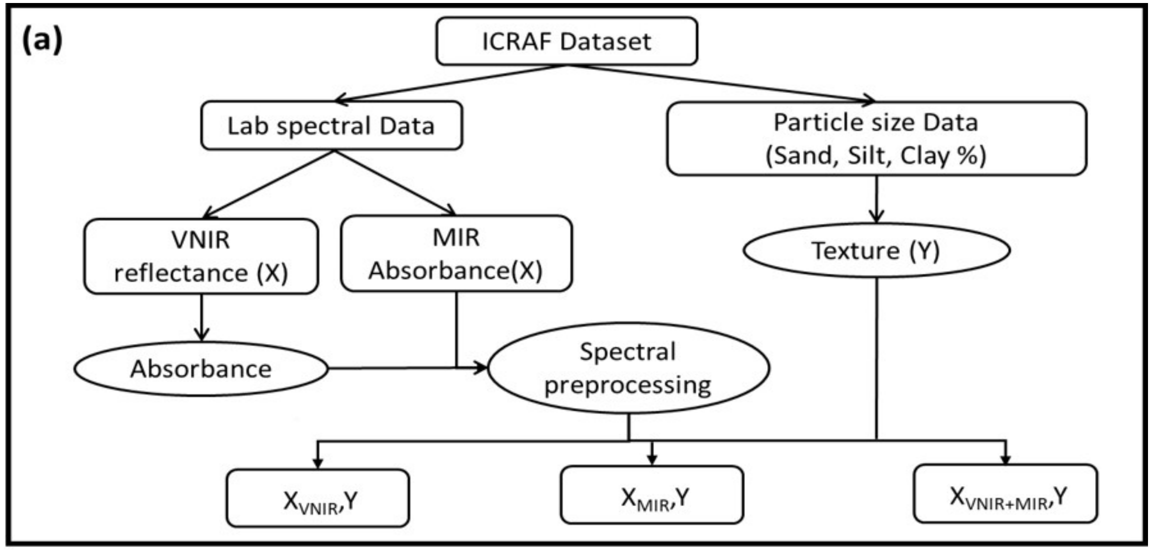

Soil spectra in VNIR region
Reflectance (%)
Wavelength (nm)
0.0
0.2
0.4
0.6
0.8
1.0
500
1000
1500
2000
2500

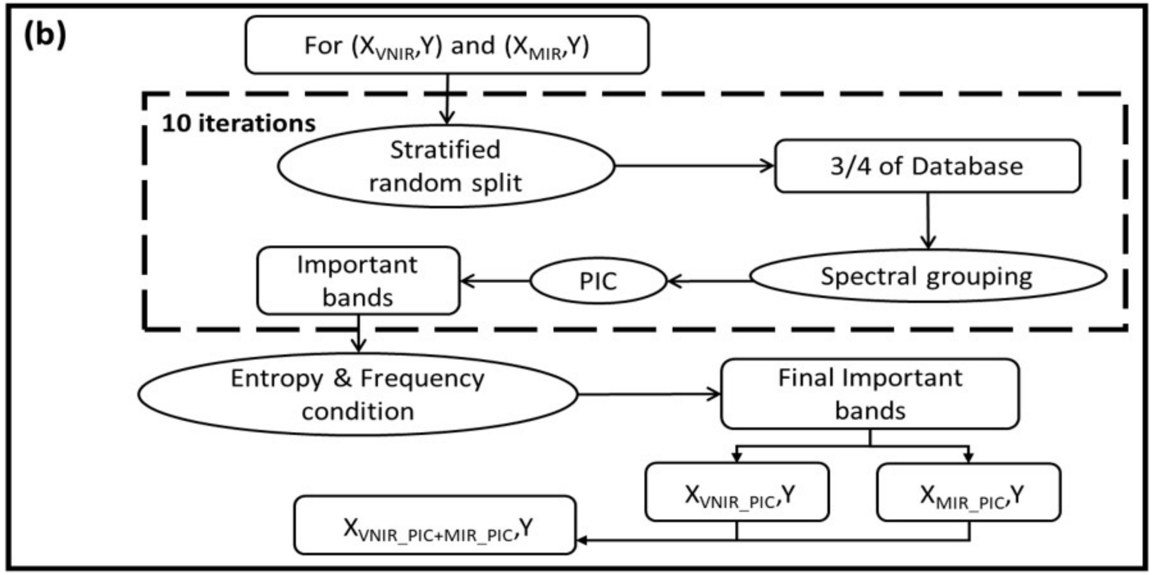
# Soil texture distribution

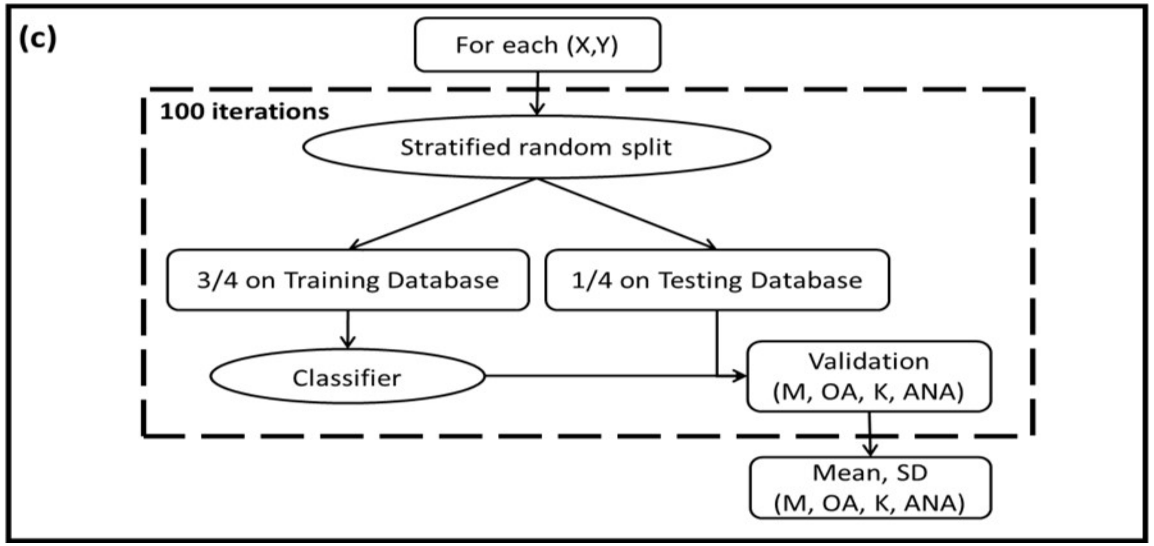


# Methodology




a) Dataset preprocessing
b) PIC based band selection
c) Classification routine

Classifiers - MNLR, SVM

Evaluation - Confusion Matrices,
Overall Accuracy,
Kappa,
Added Neighbourhood Accuracy

# Evaluation (OA,K)

| VNIR (MNLR) | | Measured Classes (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predicted Classes (%) | **Texture** | **Sa** | **LoSa** | **SaLo** | **SaClLo** | **SaCl** | **Cl** | **ClLo** | **Lo** | **SiCl** | **SiClLo** | **SiLo** | **Si** |
| | **Sa** | 50 | 23 | 14 | 0 | 0 | 0 | 3 | 12 | 5 | 1 | 7 | 0 |
| | **LoSa** | 4 | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **SaLo** | 18 | 38 | 37 | 12 | 0 | 2 | 13 | 20 | 7 | 10 | 8 | 0 |
| | **SaClLo** | 3 | 6 | 9 | 17 | 7 | 1 | 3 | 0 | 0 | 0 | 0 | 0 |
| | **SaCl** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **Cl** | 7 | 4 | 14 | 68 | 93 | 94 | 57 | 22 | 61 | 24 | 11 | 0 |
| | **ClLo** | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | **Lo** | 1 | 4 | 3 | 0 | 0 | 0 | 2 | 5 | 0 | 1 | 2 | 0 |
| | **SiCl** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **SiClLo** | 1 | 0 | 2 | 1 | 0 | 2 | 6 | 8 | 7 | 19 | 5 | 0 |
| | **SiLo** | 15 | 17 | 19 | 3 | 0 | 1 | 14 | 34 | 20 | 44 | 67 | 100 |
| | **Si** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$$OA = \frac{\sum_{i=1}^{Nc} M_{i,i}}{\sum_{i,j=1}^{Nc} M_{i,j}} = \frac{l}{s}$$

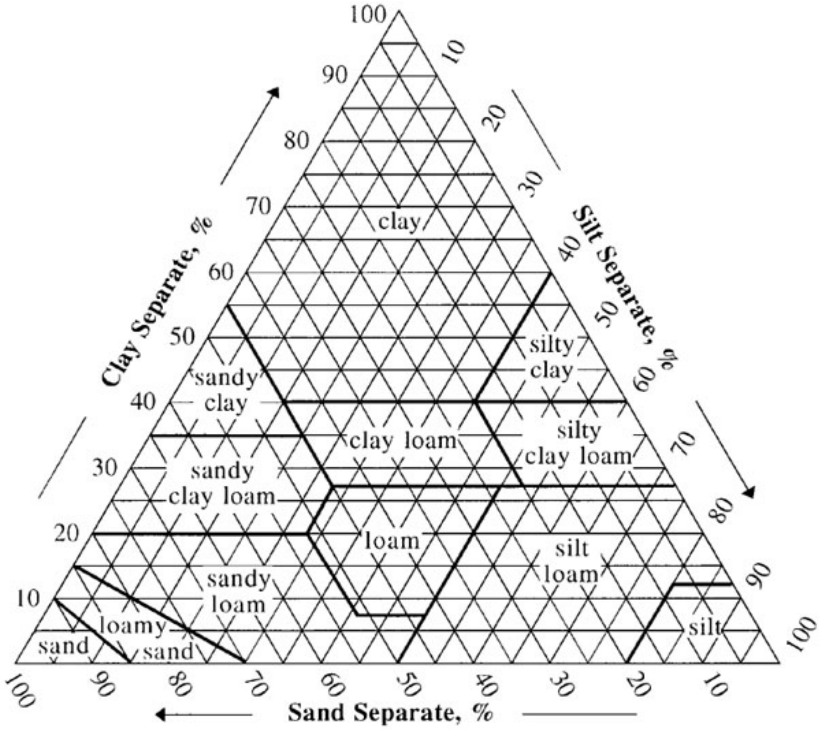
$$K = \frac{l * s - \sum_{j=1}^{Nc} \left( \sum_{i=1}^{Nc} M_{i,j} * \sum_{i=1}^{Nc} M_{j,i} \right)}{s^2 - \sum_{j=1}^{Nc} \left( \sum_{i=1}^{Nc} M_{i,j} * \sum_{i=1}^{Nc} M_{j,i} \right)}$$

Source: Cohen, (1960); Congalton, (2001)

# Evaluation (NA, ANA)

Table of USDA texture classes and their corresponding neighbours derived from USDA texture triangle

| Texture Class | % Area covered in Texture triangle | Texture Class abbreviation | No. of Neighbours | Neighbours |
|---|---|---|---|---|
| Sand | 1.50 | Sa | 1 | LoSa |
| Loamy Sand | 3.00 | LoSa | 2 | Sa, SaLo |
| Sandy Loam | 11.45 | SaLo | 4 | LoSa, SaClLo, Lo, SiLo |
| Sandy Clay Loam | 7.65 | SaClLo | 4 | SaLo, SaCl, ClLo, Cl |
| Sandy Clay | 4.00 | SaCl | 3 | SaClLo, ClLo, Cl |
| Clay | 29.75 | Cl | 4 | SaCl, ClLo, SiClLo, SiCl |
| Clay Loam | 6.25 | ClLo | 7 | Lo, SaClLo, SaCl, Cl, SiCl, SiClLo, SiLo |
| Loam | 7.45 | Lo | 4 | SaLo, SaClLo, ClLo, SiLo |
| Silty Clay | 4.00 | SiCl | 3 | Cl, ClLo, SiClLo |
| Silty Clay Loam | 5.00 | SiClLo | 4 | SiCl, Cl, ClLo, SiLo |
| Silt Loam | 16.50 | SiLo | 5 | Si, SaLo, Lo, ClLo, SiClLo |
| Silt | 3.45 | Si | 1 | SiLo |

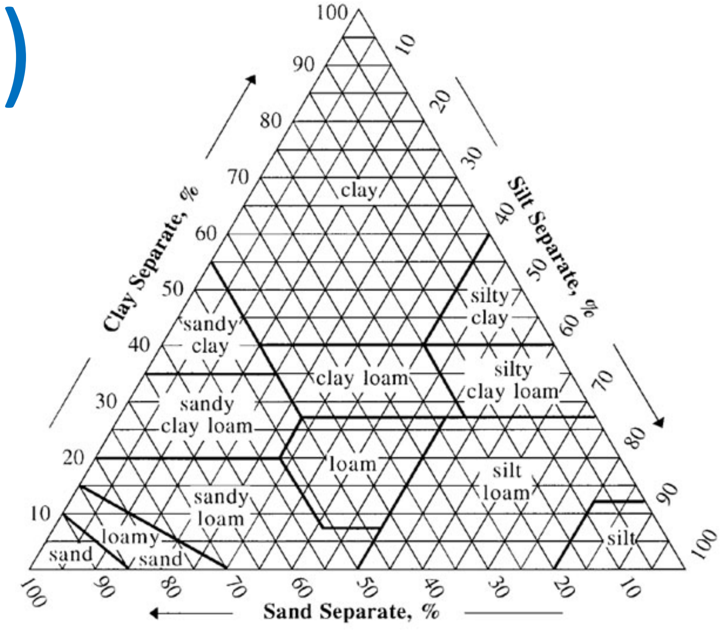


# Evaluation (NA,ANA)

Neighbour Classes Matrix (*N*) derived from the USDA texture triangle

N =

| Sa | LoSa | SaLo | SaClLo | SaCl | Cl | ClLo | Lo | SiCl | SiClLo | SiLo | Si | Classes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Sa |
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | LoSa |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | SaLo |
| 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | SaClLo |
| 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | SaCl |
| 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | Cl |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | ClLo |
| 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | Lo |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | SiCl |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | SiClLo |
| 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | SiLo |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | Si |




| Texture Class | % Area covered in Texture triangle | Texture Class abbreviation | No. of Neighbours | Neighbours |
|---|---|---|---|---|
| Sand | 1.50 | Sa | 1 | LoSa |
| Loamy Sand | 3.00 | LoSa | 2 | Sa, SaLo |
| Sandy Loam | 11.45 | SaLo | 4 | LoSa, SaClLo, Lo, SiLo |
| Sandy Clay Loam | 7.65 | SaClLo | 4 | SaLo, SaCl, ClLo, Cl |
| Sandy Clay | 4.00 | SaCl | 3 | SaClLo, ClLo, Cl |
| Clay | 29.75 | Cl | 4 | SaCl, ClLo, SiClLo, SiCl |
| Clay Loam | 6.25 | ClLo | 7 | Lo, SaClLo, SaCl, Cl, SiCl, SiClLo, SiLo |
| Loam | 7.45 | Lo | 4 | SaLo, SaClLo, ClLo, SiLo |
| Silty Clay | 4.00 | SiCl | 3 | Cl, ClLo, SiClLo |
| Silty Clay Loam | 5.00 | SiClLo | 4 | SiCl, Cl, ClLo, SiLo |
| Silt Loam | 16.50 | SiLo | 5 | Si, SaLo, Lo, ClLo, SiClLo |
| Silt | 3.45 | Si | 1 | SiLo |

$$NA = \frac{\sum_{i,j=1}^{Nc}(M * N)_{i,j}}{\sum_{i,j=1}^{Nc} M_{i,j}} \qquad ANA = OA + NA = \frac{\sum_{i=1}^{Nc} M_{i,i}}{\sum_{i,j=1}^{Nc} M_{i,j}} + \frac{\sum_{i,j=1}^{Nc}(M * N)_{i,j}}{\sum_{i,j=1}^{Nc} M_{i,j}}$$

Adapted from Brown et al., (2006)

# Evaluation (NA, ANA)

| All Bands | | | | |
|---|---|---|---|---|
| **Region (Best Classifier)** | **Texture class** | **% in Correct Classes** | **% in Neighbour Classes** | **% in Far Classes** |
| **VNIR** (MNLR) | **Sa** | 50 | 4 | 46 |
| | **LoSa** | 8 | 61 | 31 |
| | **SaLo** | 37 | 32 | 31 |
| | **SaClLo** | 17 | 12 | 71 |
| | **SaCl** | 0 | 100 | 0 |
| | **Cl** | 94 | 2 | 4 |
| | **ClLo** | 2 | 82 | 16 |
| | **Lo** | 5 | 54 | 41 |
| | **SiCl** | 0 | 68 | 32 |
| | **SiClLo** | 19 | 68 | 13 |
| | **SiLo** | 67 | 15 | 18 |
| | **Si** | 0 | 100 | 0 |

Percentage Distribution of the classifications for a given texture class into correct class, neighbouring classes and far classes in the testing database using all bands in VNIR region

| VNIR (MNLR) | | Measured Classes (%) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Predicted Classes (%) | Texture | Sa | LoSa | SaLo | SaClLo | SaCl | Cl | ClLo | Lo | SiCl | SiClLo | SiLo | Si |
| | Sa | 50 | 23 | 14 | 0 | 0 | 0 | 3 | 12 | 5 | 1 | 7 | 0 |
| | LoSa | 4 | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | SaLo | 18 | 38 | 37 | 12 | 0 | 2 | 13 | 20 | 7 | 10 | 8 | 0 |
| | SaClLo | 3 | 6 | 9 | 17 | 7 | 1 | 3 | 0 | 0 | 0 | 0 | 0 |
| | SaCl | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Cl | 7 | 4 | 14 | 68 | 93 | 94 | 57 | 22 | 61 | 24 | 11 | 0 |
| | ClLo | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| | Lo | 1 | 4 | 3 | 0 | 0 | 0 | 2 | 5 | 0 | 1 | 2 | 0 |
| | SiCl | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | SiClLo | 1 | 0 | 2 | 1 | 0 | 2 | 6 | 8 | 7 | 19 | 5 | 0 |
| | SiLo | 15 | 17 | 19 | 3 | 0 | 1 | 14 | 34 | 20 | 44 | 67 | 100 |
| | Si | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Confusion matrix

| | Sa | LoSa | SaLo | SaClLo | SaCl | Cl | ClLo | Lo | SiCl | SiClLo | SiLo | Si | Classes |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N = | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Sa |
| | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | LoSa |
| | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | SaLo |
| | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | SaClLo |
| | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | SaCl |
| | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | Cl |
| | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | ClLo |
| | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | Lo |
| | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | SiCl |
| | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | SiClLo |
| | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | SiLo |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | Si |

Neighbouring class matrix

# Results – All bands

- Which region is better? VNIR, MIR or VNIR+MIR
- Which classifier is better? MNLR or SVM

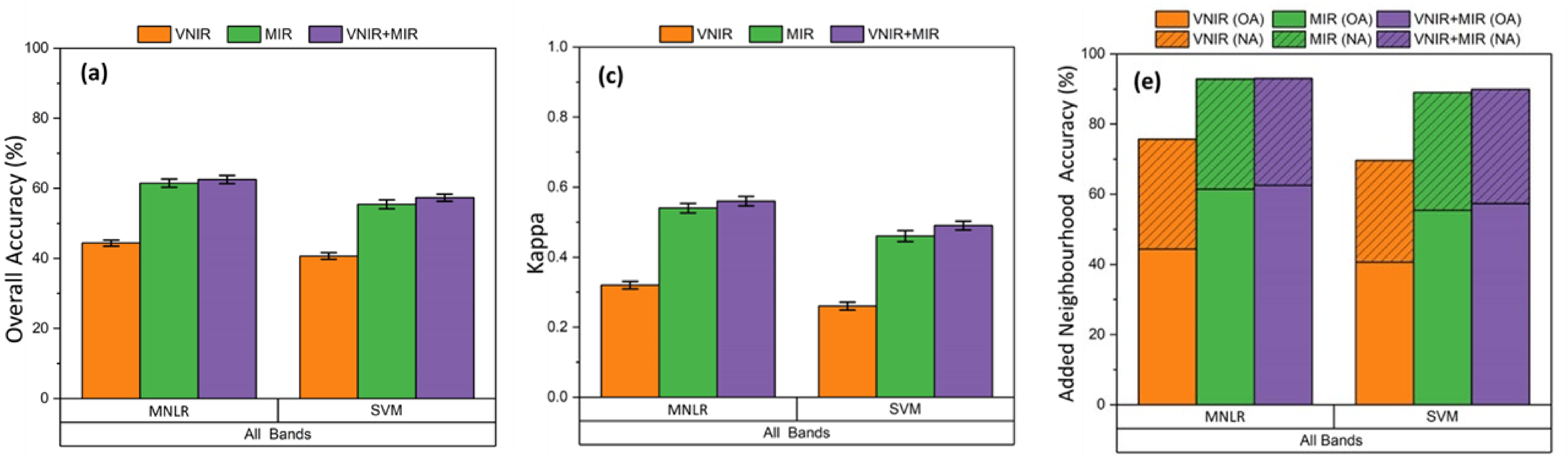
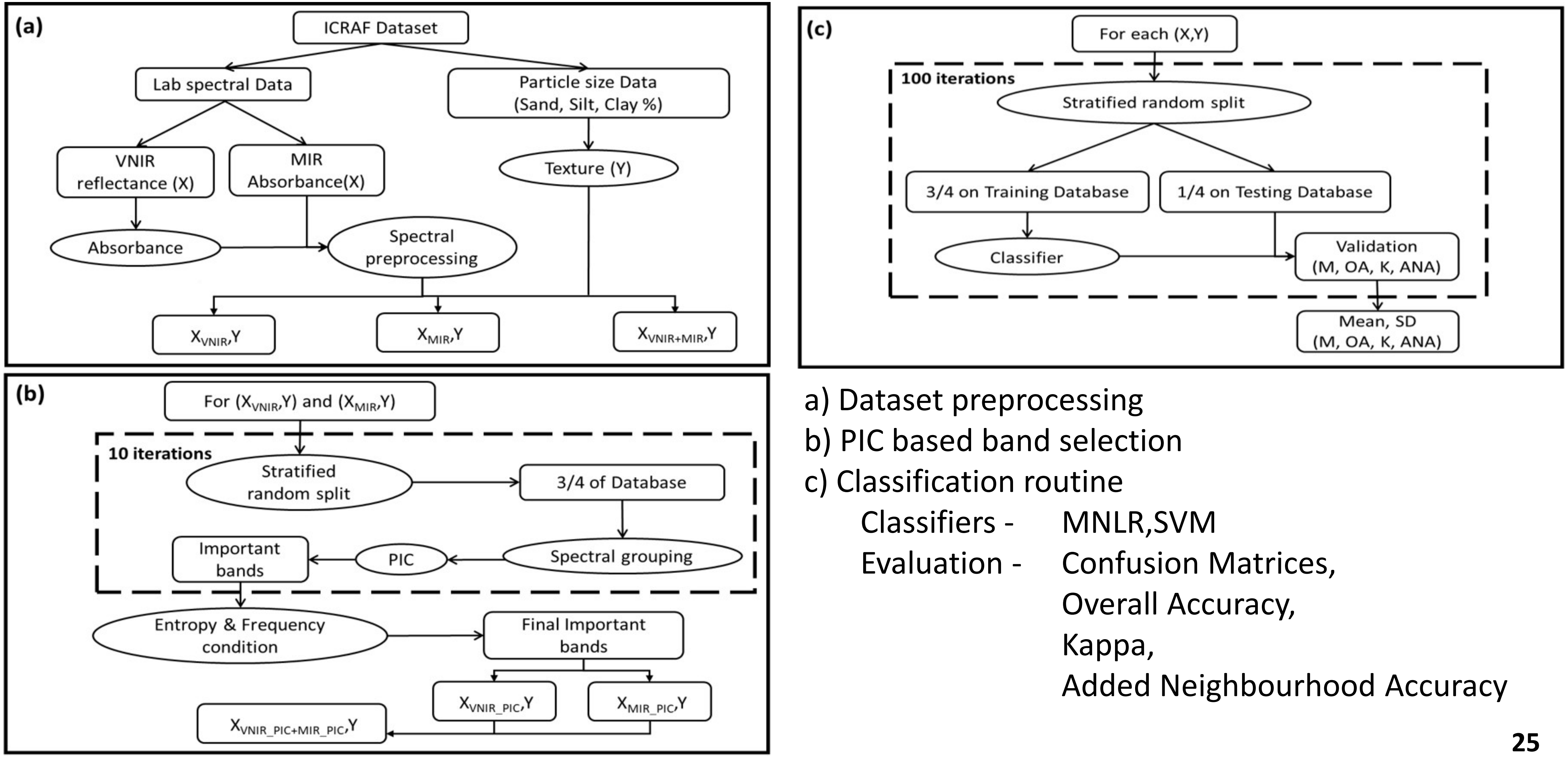



Figure The texture classification on the testing database for MNLR and SVM classifiers indicating : (a) Overall Accuracy (%), (c) Kappa and (e) Added Neighbourhood Accuracy (%) using all bands

# Results – Key points

- VNIR+MIR regions classified with mean OA 62.53 %, mean Kappa of 0.56 and mean ANA of 93.05 %
- Top 3 easily classified textures – Clay, Silt Loam, and Sand
- Top 3 difficult to classify textures – Silt, Sandy Clay, and Silty Clay
- No spectral region or classifier classified – Silt
- Difficult to classify textures – Silt, Sandy Clay, and Silty Clay; were majorly misclassified into neighbour classes
- Extensive far class misclassification (≥ 30%) in 6 texture classes in VNIR and none in MIR, VNIR+MIR regions
- MNLR outperformed SVM in all regions

# Methodology




a) Dataset preprocessing
b) PIC based band selection
c) Classification routine

Classifiers - MNLR,SVM
Evaluation - Confusion Matrices, Overall Accuracy, Kappa, Added Neighbourhood Accuracy

# Band Selection

- Partial Information (PI) measures the partial dependence and selects predictor variables depending on the response variable.
- PI can identify the predictor variables without making any assumptions about its form or model representation.
- A sample estimate of PI(R, P|Z) (i.e., partial dependence of response variable R with a potential predictor P conditional to the preselected predictor set Z) is calculated as:

$$\widehat{PI}(R, P|Z) = \frac{1}{n}\sum_{i=1}^{i=n} \log\left[\frac{f_{R|Z,P|Z}(r_i, p_i|Z_i)}{f_{R|Z}(r_i|Z_i) * f_{P|Z}(p_i|Z_i)}\right]$$

$f_{R|Z}(r_i|Z_i)$ ,$f_{P|Z}(p_i|Z_i)$ marginal probability density function of R and P conditional on Z respectively

$r_i, p_i, Z_i; i = 1, \dots \dots, n,$ are sample observations of R,P and Z respectively

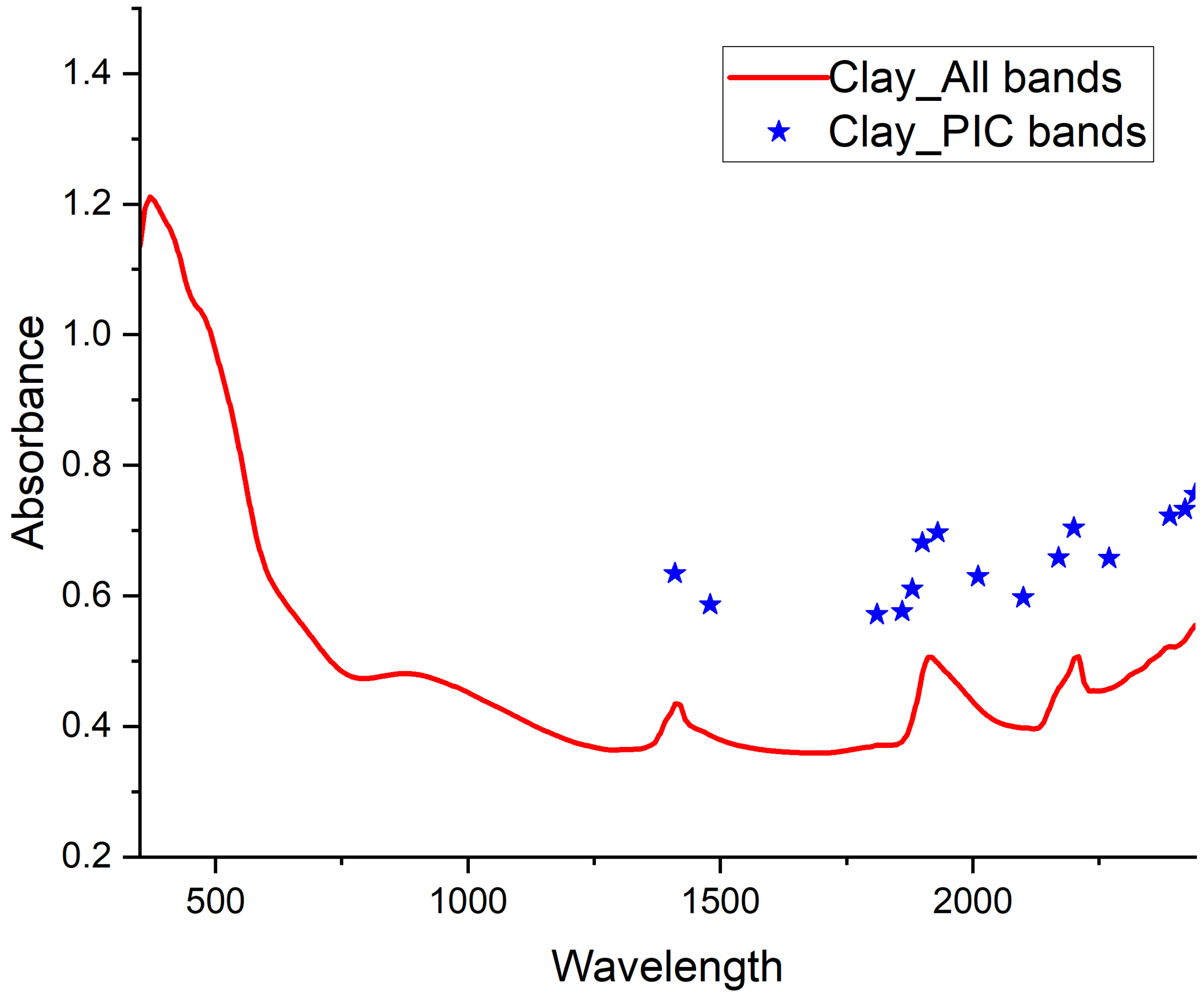
- The PIC is derived from PI by scaling it to a (0,1) as:

$$\widehat{PIC} = \sqrt{1 - \exp\left(-2\widehat{PI}\right)}$$

- 20 spectral groups were used with entropy calculation by equal width discretization

Source: Paul S. & Nagesh Kumar, (2019)

# Band Selection




Clay absorbance spectra with all bands and PIC selected bands in VNIR region (separated by factor of 0.2 for illustration)

# Results – PIC selected bands

- Which region is better? VNIR, MIR or VNIR+MIR
- Which classifier is better? MNLR or SVM
- Does reduced band help in classification?

| No. of bands | VNIR | MIR | VNIR+MIR |
|---|---|---|---|
| All bands | 204 | 1762 | 1966 |
| PIC selected bands | 15 | 29 | 44 |

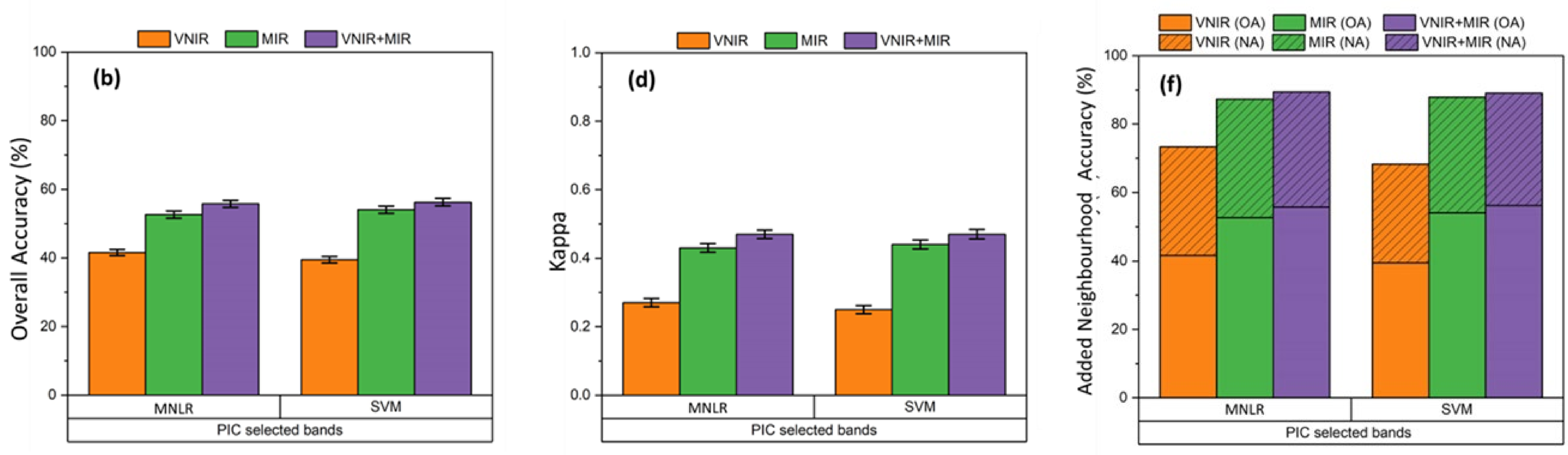
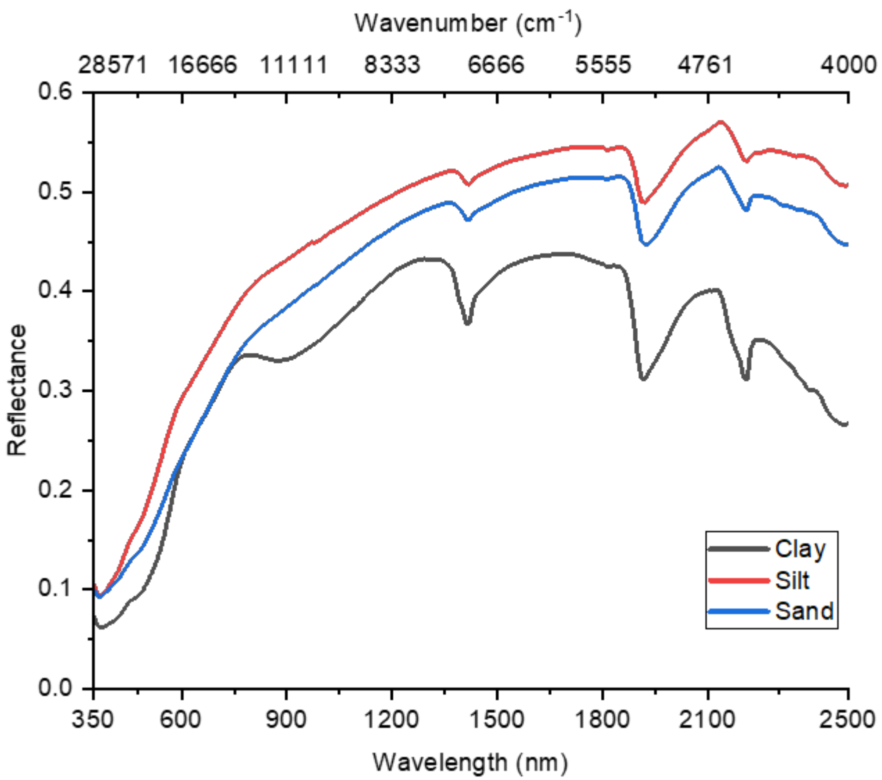
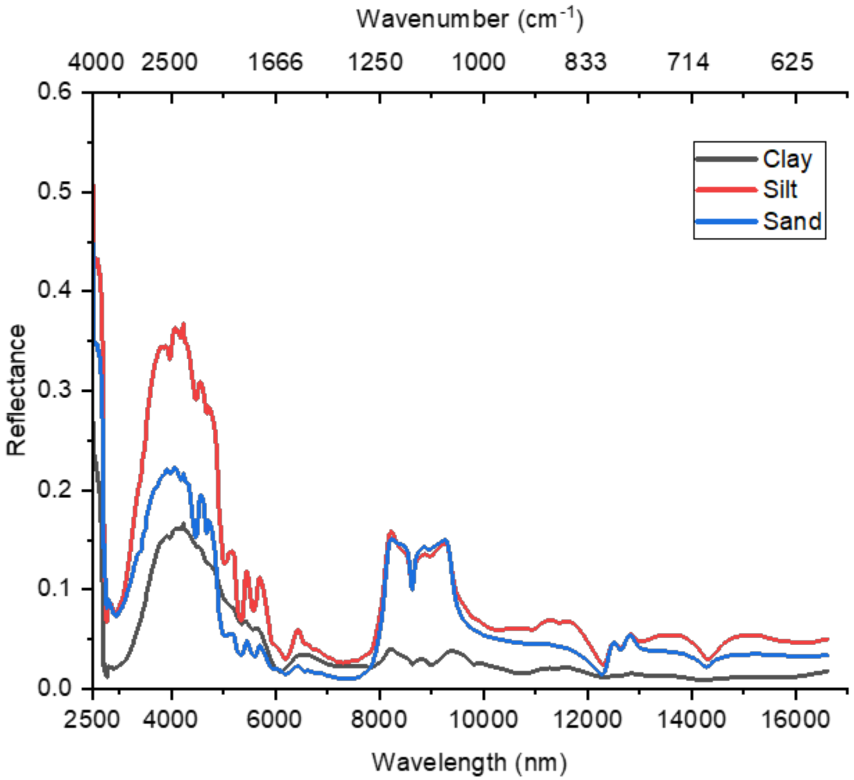



Figure The texture classification on the testing database for MNLR and SVM classifiers indicating :
(b) Overall Accuracy (%), (d) Kappa and (f) Added Neighbourhood Accuracy (%) using PIC selected bands

# Results – Important spectral features of soil



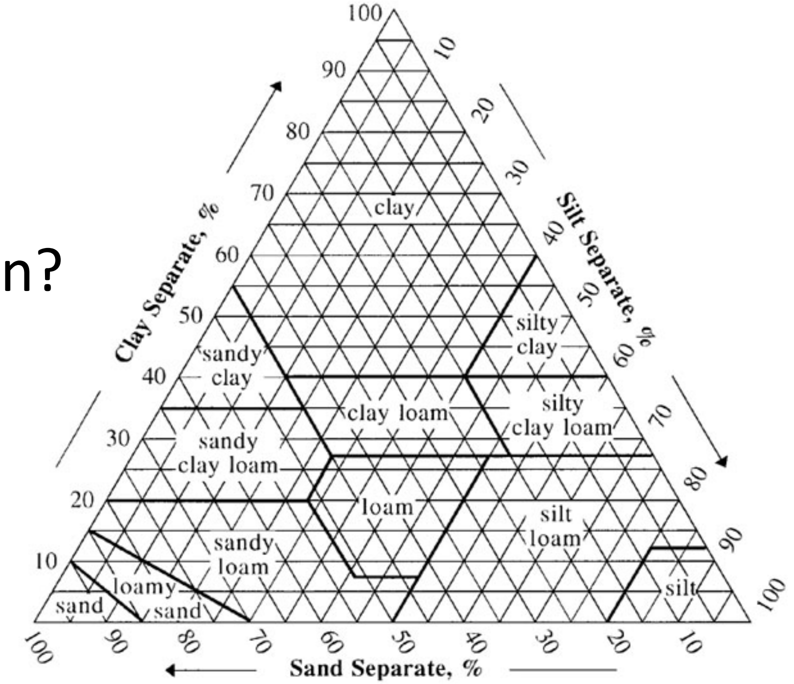




- Single bond stretching of O-H from free water and clay lattice in 400 – 2440 nm
- Single bond stretching mainly from C-H, O-H of water and clay lattice in 2441 – 4000 nm
- Triple bond stretching of C≡C, C≡N in 4001 – 5000 nm
- Double bond stretching of C=O and features of quartz in 5001 – 6666 nm
- Fingerprint region with feature of silicates and quartz in 6667 – 14286 nm

# Results – Key points

- VNIR+MIR_PIC classified with mean OA 56.23 %, mean Kappa of 0.47 and mean ANA of 89.08 %
- Top 3 easily classified textures – Clay, Sand and Sandy Loam
- Top 3 difficult to classify textures – Silt, Sandy Clay, and Silty Clay
- No spectral region or classifier classified – Silt, Sandy Clay
- Difficult to classify textures – Silt, Sandy Clay, and Silty Clay; were majorly misclassified into neighbour classes
- Extensive far class misclassification (≥ 30%) in 4 texture classes in VNIR and in 2 texture classes in MIR, VNIR+MIR regions
- SVM outperformed MNLR in MIR and VNIR+MIR regions

# Discussions




- Which region is better? Do reduced bands help classification?
  - MIR (more fundamentals); Yes (suitable chromophores)
- Which soil texture class is suitably classified?
  - Clay, Sand (Clay minerals, spectrally active)
- Is misclassification more in neighbouring classes than far classes?
  - Yes (~30%)
- Why certain texture classes perform poor in classification?
  - Silt (Measurement error, methodology, areal representation, decision rules, low no. in training)

# Comparison with literature

| Reference | Data Used\|Method | Total (Train\|Test) | No. of Classes | Class type | Testing Accuracy ± Std. Dev (%) |
|---|---|---|---|---|---|
| **Barnes, 2000** | Landsat 5\|ISODATA | 303(NA) | 3 | Sandy Loam, Sandy Clay Loam, Clay Loam | 51 |
| **Zhai, 2006** | Landsat 5\|NN | 443(354\|89) | 3 | Loam, Clay Loam, Clay | 65.7 ± 1.8 |
| **DeMatte, 2016** | Landsat 5\|GMLC | 504(300\|204) | 4 | Sand, Sandy Loam, Clay Loam, Clay | 63.8 |
| **Gomez, 2019** | Sentinel 2\|Lin-SVM | 130(91\|39) | 4 | Sandy Loam, Sandy Clay Loam, Sandy Clay, Clay | 50 |
| **Mouazen, 2005** | Lab Spectra\|FDA | 365(244\|121) | 3† | Sand, Loam, Clay | 85.1 |
| **Jia, 2019** | Lab Spectra\|RBF-SVM | 198(132\|66) | 4‡ | Clay, Clay Loam, Loam, Sand | 78.8 |
| **Gouda,2021** | Lab Spectra LUCAS\|LightGBM | 14454(12087\|2367) | 3* | Fine, Medium, Coarse | 75 |
| **Gouda,2021** | Lab Spectra ICRAF\|LightGBM | 2416(2021\|395) | 3* | Fine, Medium, Coarse | 75 |
| **This paper** | Lab Spectra ICRAF\|MNLR | 3643 (2737\|906) | 12 | All 12 Classes | 62.5 ± 1.2 |

*Canadian soil texture classification; † Belgium soil texture classification; ‡ International Soil Society classification;

# Conclusions

- Best classification performance using MNLR in combined VNIR+MIR region with OA 62.53%, K of 0.56 and ANA of 93.05%
- PIC bands provide slightly lower classification but huge reduction in the number of bands (>93% reduction in number of bands)
- MIR compared to VNIR region provides around 11 to 17% higher accuracy
- VNIR+MIR provide only slight improvement in accuracy over the MIR region
- A texture class is more misclassified into its Neighbouring classes than in far class. Allowing this misclassification, the overall accuracy increases by around 30%
- Remote textural classes i.e. Clay, Silt Loam, and Sand texture, have good classification performance as compared to intermediate textural classes i.e. Silt, Sandy Clay, and Silty Clay

# Applications

- Quick qualitative inference on texture (~few hours)
- Improved quantitative predictions using qualitative predictions
  - Clay content, OM, OC, N, moisture content, hydraulic conductivity, erosion modelling, etc
- Reduced model complexities
- Quantifying uncertainties in neighbouring classes
  - Definition of neighbour in distance term needs to be studied
- Classification performances from the regression techniques for sand, silt, and clay fractions
- Simulation of lab spectra to upcoming satellite mission for evaluating soil texture classification either in quantitative or qualitative fashion

# Working details

- Data available at (www.isric.org)
- Codes available at (https://github.com/ternikarcr/Texture_Classification.git)
- Codes in Python – sklearn package, R – NPRED package
- Graphs in Origin Pro

# Memories

- SARI meeting in Philippines, 2018
- GEE, HPC – Global Landsat Mosaic
  - Don’t be intimidated, Subject matter experts never go out of job
- Prof. Chirag Jain – Why the similarities and Why the differences?
  - All my discussion sections are structured in this manner
- Dr. Thuy Le Toan – Young researchers need to open source the data
  - No data over India

# Acknowledgement

**Ternikar Chirag Rajendra**
Research Scholar WREE, Civil Engg. at **IISc**, Bangalore (India)

**Dr. Cécile Gomez**
Research Scientist at IRD (France), **IFCWS at IISc**, Bangalore (India)

**Dr. D Nagesh Kumar**
Professor in Civil Engg., CEaS, DCCC at **IISc**, Bangalore (India)

Thanks to **Dr. Laurent Ruiz** (Indo-French Cell of Water Science) and **Prof. Sekhar Muddu** (Civil Engineering Department) of IISc for their thoughts and discussion through which the core idea of this paper was conceptualized.

# Agrocares Scanner

- Temperature Sensor NTC resistance probe
- Near Infra Red Spectrometer (1) MEMS technology Wavelength Range (1300-2550 nm)
- EC Probes (6) electrical conductivity probes with alternate 1 kHz bi-polar measurement
- One of a kind soil nutrient device based on spectroscopy
- Measure pH, Organic Carbon, Total P, N, K exchange, Cation exchange, Soil Temperature; Suggests suitable crop types
- Cost - 3000 Euros ~ 3 lakh Rs. (Wageningen, Netherlands)



Source: https://www.agrocares.com

# THANK YOU