# Optimization of Sampling Designs for Validating Digital Soil Maps

Yakun Zhang (M.Sc.)
Supervisor: Prof. Asim Biswas
Department of Natural Resource Sciences



## Outline

- Introduction and objectives
- Methods
- Results
  - 3D digital soil mapping of SOC content (theory and results)
  - Comparison of sampling designs



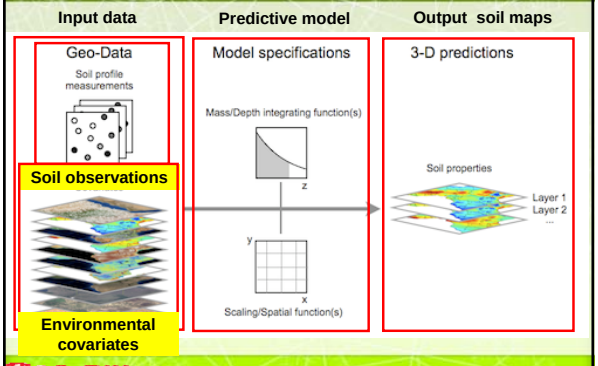
## Digital soil mapping

- Spatial variability of soil properties

  Jenny's theory: *S = f (cl, o, r, p, t, …) (Jenny 1941)*
- Traditional soil maps
  - Delineation of soil type or properties
- Digital soil maps
  - Georeferenced soil database
  - High resolution
  - Quantitative relationship between soil properties and environmental covariates



## Digital soil mapping- Workflow




http://africasoils.net/services/data/digital-soil-mapping/



## Sampling design

- Sample size
  - Trade-off between budget and accuracy
- Sample locations
  - determined by sampling design
  - Sampling design is not the actual pattern of soil locations, but the procedure used to select it.

**Provide reliable input for predictive model**

http://africasoils.net/services/data/digital-soil-mapping/



## Challenges and opportunities

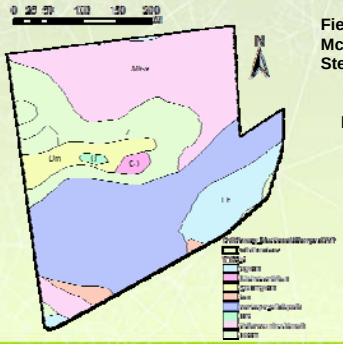
- Due to the technical constraints, most studies focus on top soil. However, the vertical variability, and horizontal variability of deep soil need more work.
- Development of proximal soil sensing allows us to accurately obtain the deep soil properties, which needs more detection and application.

## Objectives

- Produce 3D digital soil maps to quantify horizontal and vertical variability of soil properties
- Compare different sampling designs for digital soil maps in multiple layers (3D)



## Study area

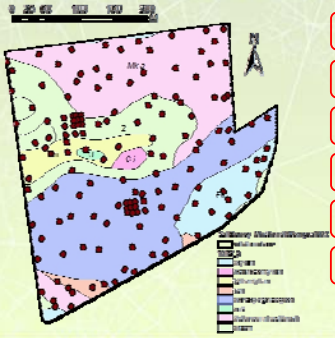


**Field 26 of Macdonald farm**
**McGill University**
**Ste-Anne-de-Bellevue**

Field size-11 ha



## Sampling designs



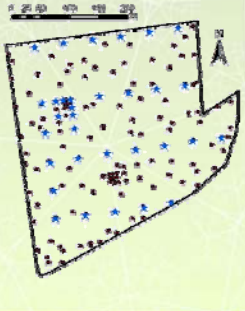
- Grid sampling 56
- Grid random sampling 45
- Transect sampling 37
- Random sampling 44
- Stratified random sampling 42
- All designs together 148

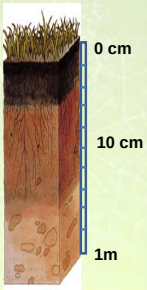


## Sample collection

**Input data**

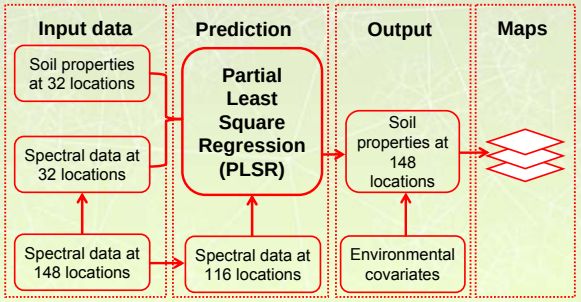
- **148 vis-NIR spectral data**
- **32 soil samples**





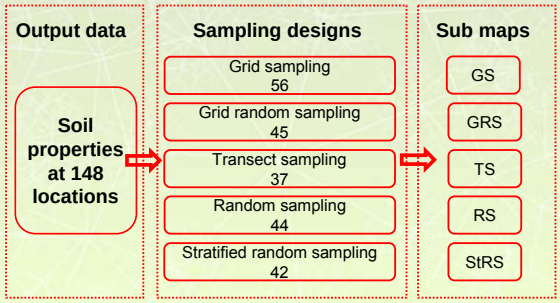



## Digital soil mapping- Procedure






## Sampling designs- Procedure



## Soil properties and environmental covariates

- **SOC content**

  Laboratory: loss-on-ignition

  **Normalization and transformation:** $y = \log(\frac{\frac{x}{100}}{1-\frac{x}{100}})$

- **Environmental covariates**
  - Elevation(RTK)
  - Gamma-ray radiation- Total count, K, U, Th, Cs
  - Apparent electrical conductivity (DUALEM)- 1m-HCP, 1m-PRP, 2m-HCP, 2m-PRP



## Results- 1

**3D digital soil mapping of SOC content by using regression-kriging**



## 3D mapping procedure

- **Regression Kriging (RK)**

Prediction = Trend Prediction (Regression) + Residual Prediction (Kriging)

- **3D- Regression Kriging (RK)**

$$\hat{z}(s_0, d_0) = \sum_{j=0}^{p} \hat{\beta}_j \cdot X_j(s_0, d_0) + \hat{g}(d_0) + \sum_{i=1}^{n} \lambda_i \cdot e(s_i, d_i)$$

Regression; Depth function; Residual kriging



## 3D mapping procedure

- **3D variogram**

Exponential model, consisting of three standard parameters ($c_0$, $c_1$, r), and 2 to 3 geometric anisotropy parameters ($a_x$, $a_y$, $a_d$).

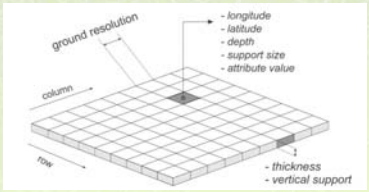
$$\gamma(\mathbf{h}) = \begin{cases} 0 & \text{if } |\mathbf{h}|=0 \\ c_0 + c_1 \cdot [1 - e^{-(\frac{h}{r})}] & \text{if } |\mathbf{h}|>0 \end{cases} \qquad \mathbf{h}=[h_x, h_y, h_d]$$

The isotropic lag ($h$) is calculated by scaling the horizontal and vertical distances by using anisotropy parameters:

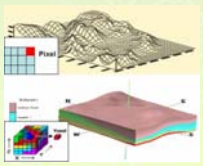
$$h = \sqrt{(\frac{h_x}{a_x})^2 + (\frac{h_y}{a_y})^2 + (\frac{h_d}{a_d})^2}$$



## 3D mapping procedure




Voxels (X,Y,d … ΔX, ΔX, Δd)






## Spectral model- PLSR

- **Cross-validation results**

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \qquad RPD = \frac{\text{Standard Deviation}}{RMSE}$$

## Original SOC maps- Google earth

- Soil SOC maps
- Sampling variance (standard error)
- 90% Confidence Interval (Lower boundary and upper boundary)

General linear model (GLM) | Regression tree (RT) | Random forest (RF)

- Depth intervals: 0-5cm, 5-15cm, 15-30cm, 30-60cm, 60-100cm.



## Results- 2

**Comparison of different sampling designs for calibrating digital soil maps**



## Comparison of sampling designs

GS | GRS | RS | StRS | TS

Ori

Spatial coverage | Distribution of values | Accuracy var+RMSE