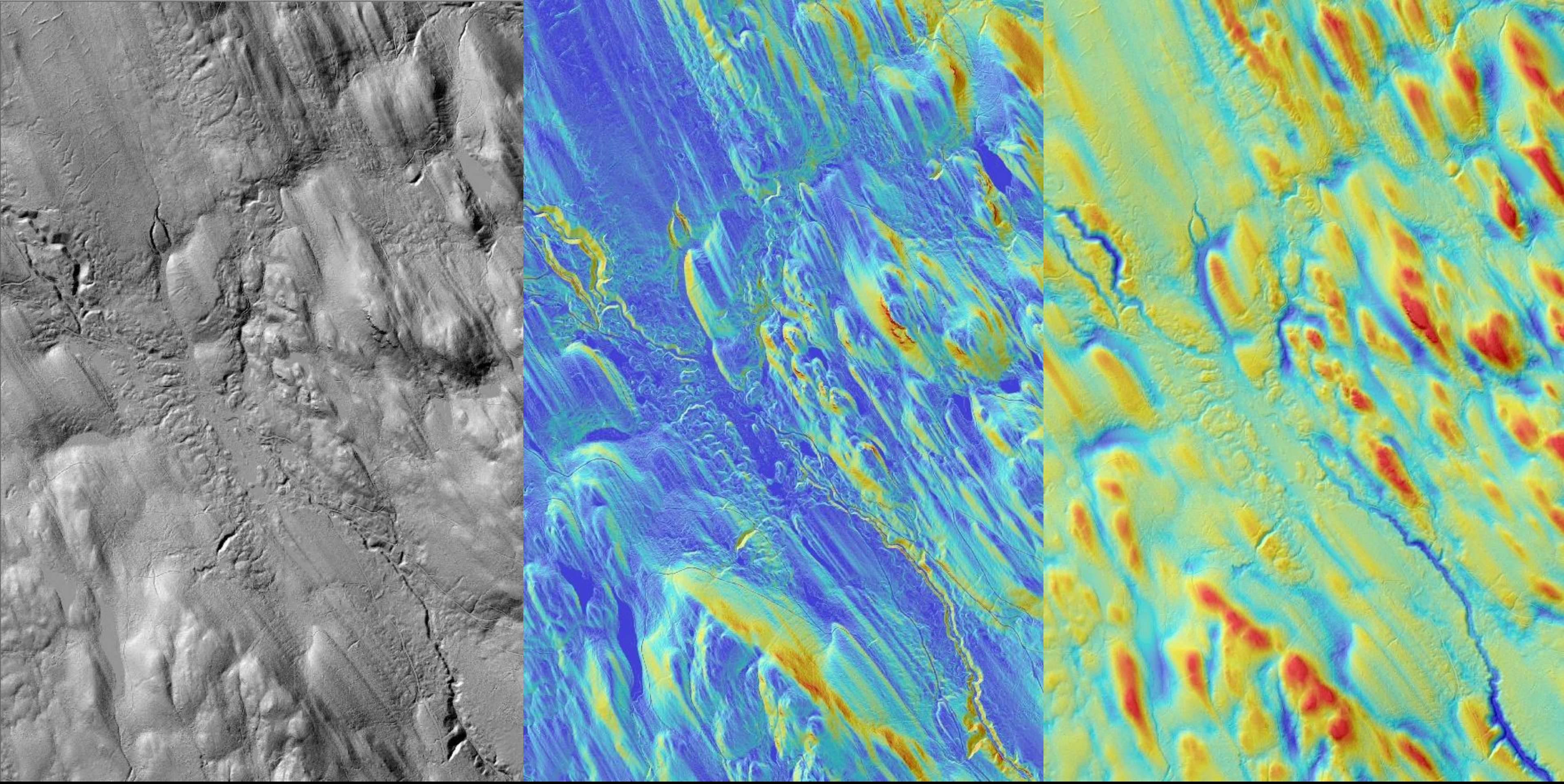


3. Environmental Covariates for Digital Soil Mapping

Canadian Digital Soil Mapping Workshop, 2022

CDSM WG
CANADIAN SOCIETY OF SOIL SCIENCE
SOCIÉTÉ CANADIENNE DE LA SCIENCE DU SOL
Ontario
UNIVERSITY of GUELPH
DALHOUSIE UNIVERSITY
FACULTY OF AGRICULTURE

About Us

Brandon Heung

Dalhousie University

Brandon.Heung@dal.ca

Daniel Saurette

Ontario Ministry of Agri-Food and Rural Affairs

University of Guelph

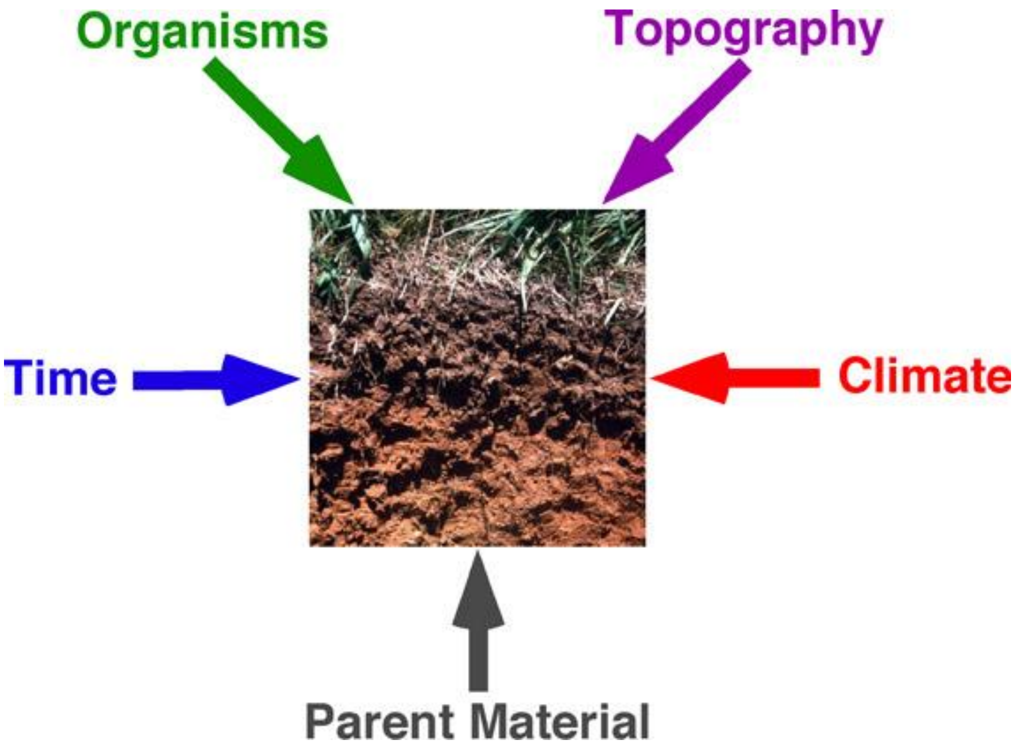
Daniel.Saurette@ontario.ca

Outline

1. **Overview of Soil-Environmental Variables**

2. Dealing with multi-collinearity

Environmental Covariates



c = climate (temp, rainfall)
o = organisms (vegetation/land cover)
r = relief (topography, aspect)
p = parent material (bedrock and surficial geology)

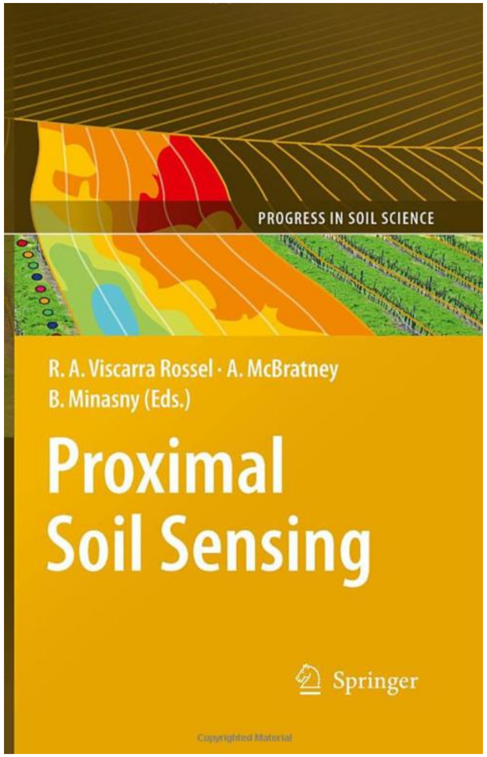
Examples of Data Sources

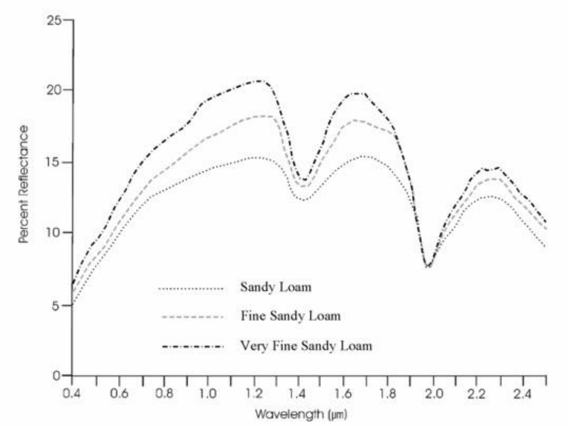
Proximal Soil Sensing	Remote Sensing	Other Geospatial Data
Electrical Conductivity Surveys	Satellite Imagery	AAFC Crop Inventory
Gamma Radiometric Surveys	Gamma Radiometric Surveys	Surficial Geology
Soil Resistivity	Digital Elevation Models	Bedrock Geology
Ground Penetrating Radar Surveys	LiDAR Surveys	Climate Model Data
Hyperspectral Surveys	Hyperspectral Surveys	Soil Surveys

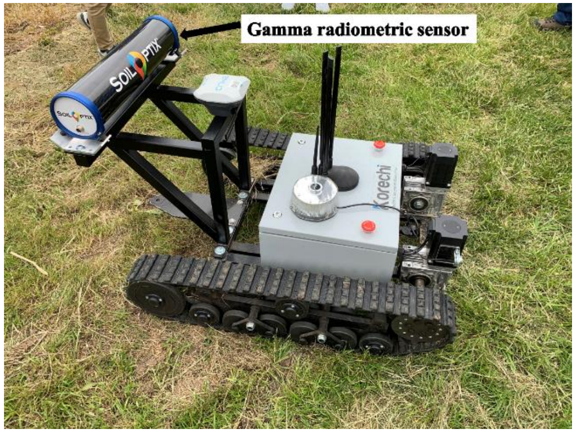
Soil (*s*)

- **Remote Sensing Data**

- **Proximal Soil Sensing Data**







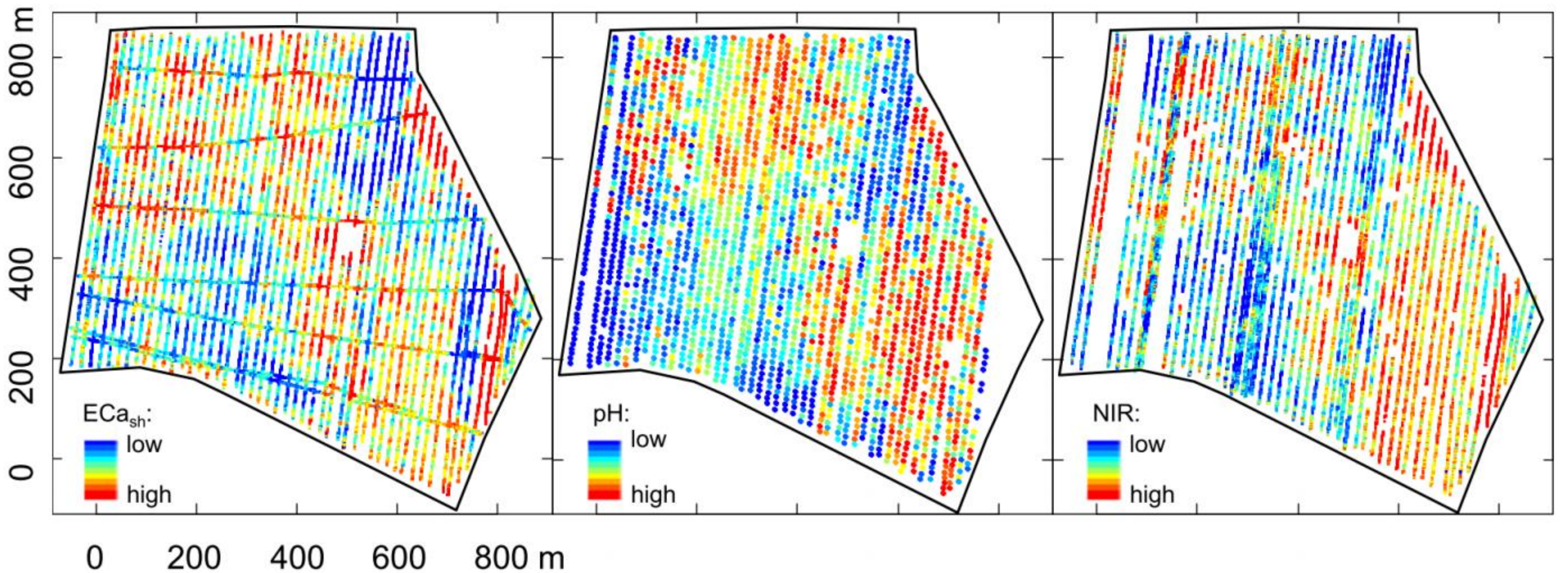
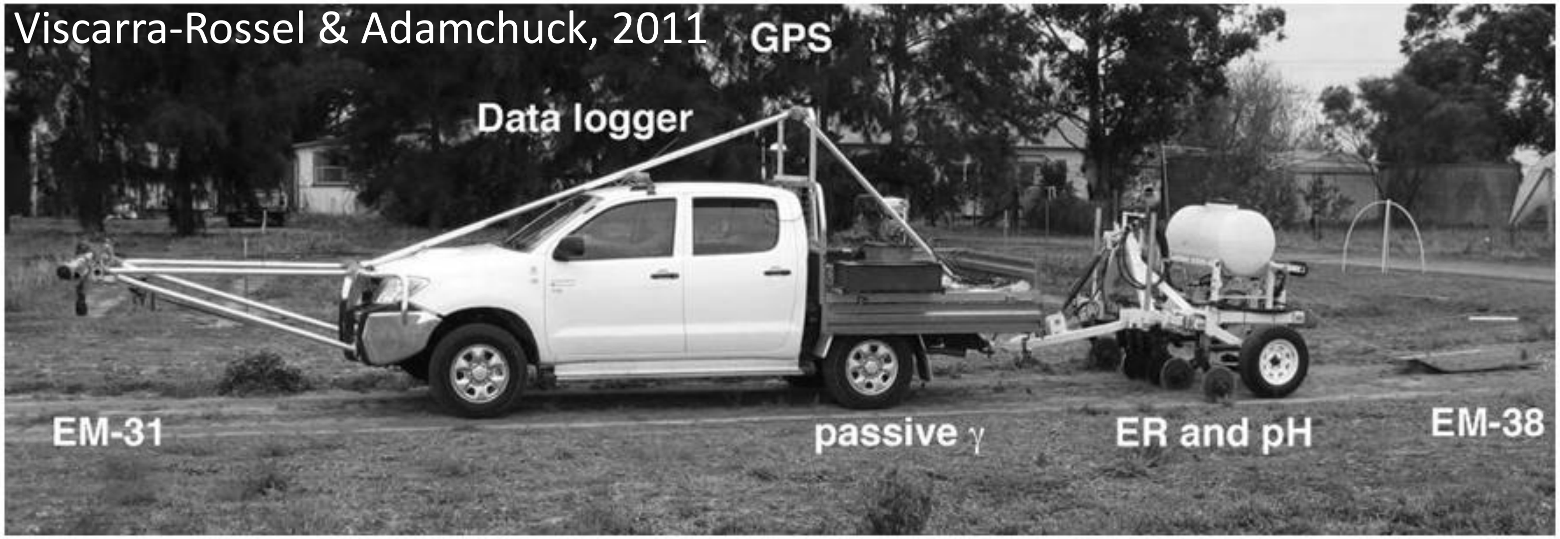


Fig 3. Proximal sensor measurements. Spatial locations of the sensor readings by the three-sensor platform. Apparent electrical conductivity (ECa_{sh}), pH, and light absorbance (NIR_{avg}) in the topsoil are depicted from left to right. Quantile classification scheme with 10 levels used.

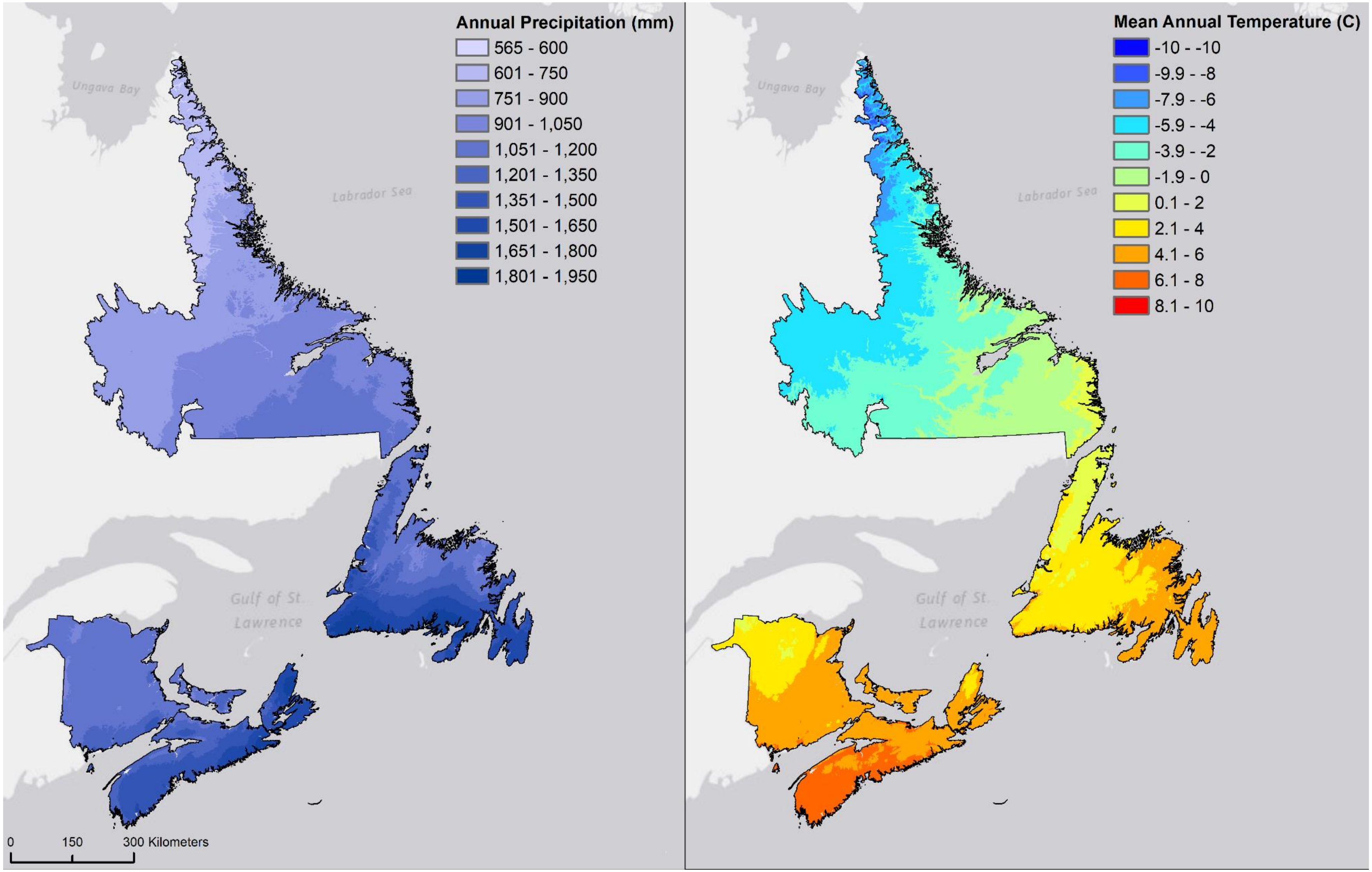
Schirrmann et al., 2016

Viscarra-Rossel & Adamchuck, 2011



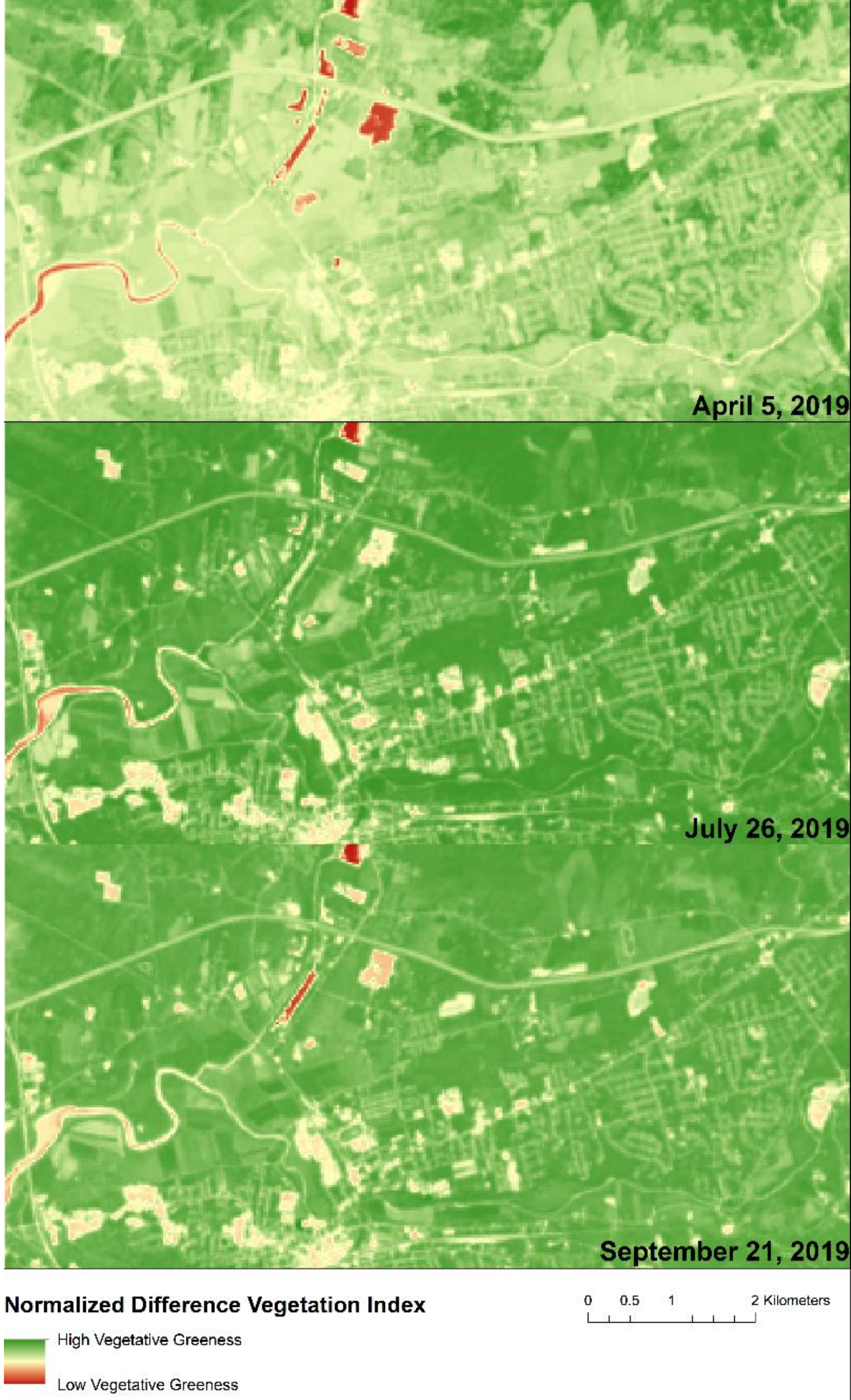
Multi-Platform Soil Sensor

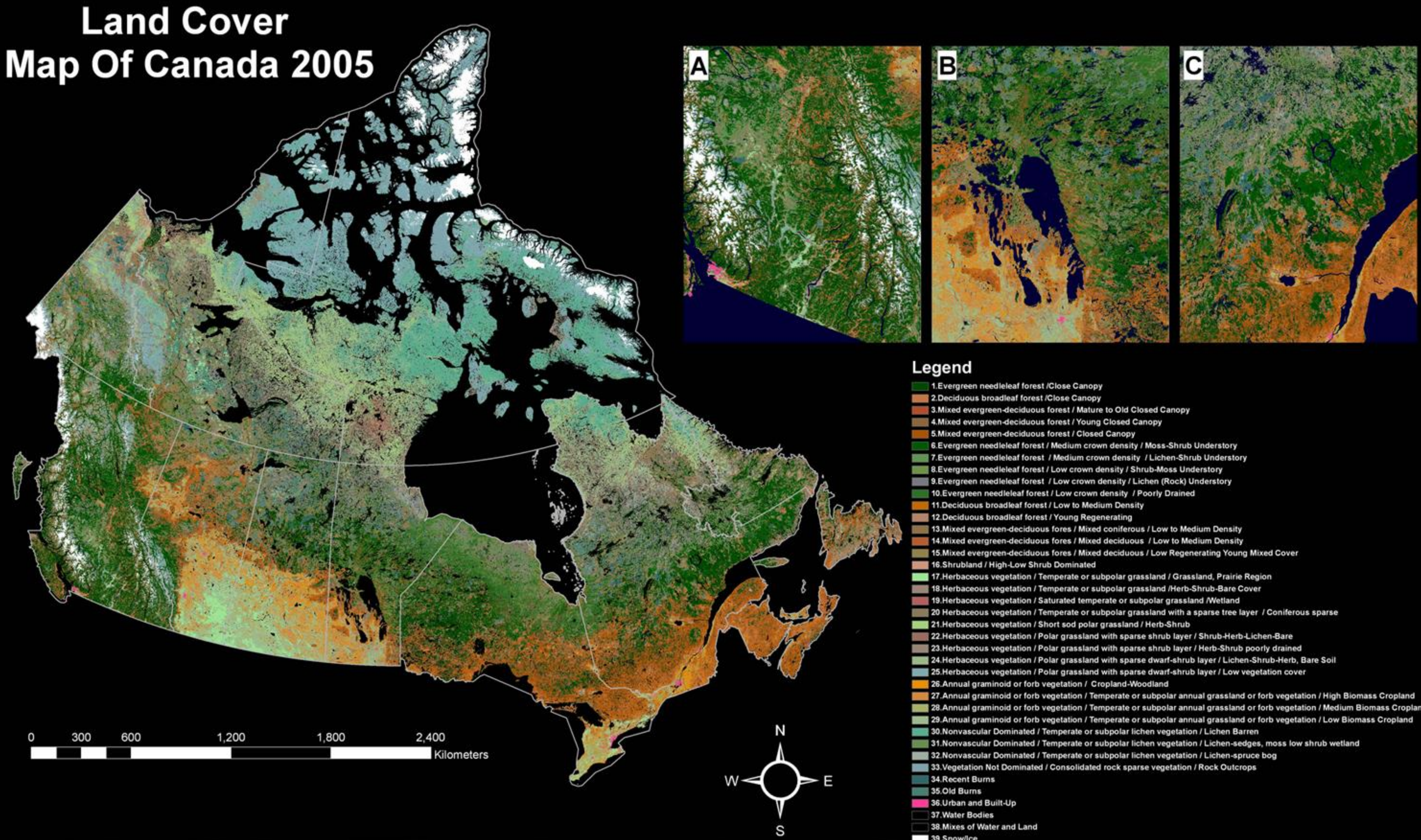
Climate (*c*)


Annual Precipitation (mm)
565 - 600
601 - 750
751 - 900
901 - 1,050
1,051 - 1,200
1,201 - 1,350
1,351 - 1,500
1,501 - 1,650
1,651 - 1,800
1,801 - 1,950
Ungava Bay
Labrador Sea
Gulf of St. Lawrence
0
150
300 Kilometers
Mean Annual Temperature (C)
-10 - -10
-9.9 - -8
-7.9 - -6
-5.9 - -4
-3.9 - -2
-1.9 - 0
0.1 - 2
2.1 - 4
4.1 - 6
6.1 - 8
8.1 - 10
Ungava Bay
Labrador Sea
Gulf of St. Lawrence

Organisms (*o*)

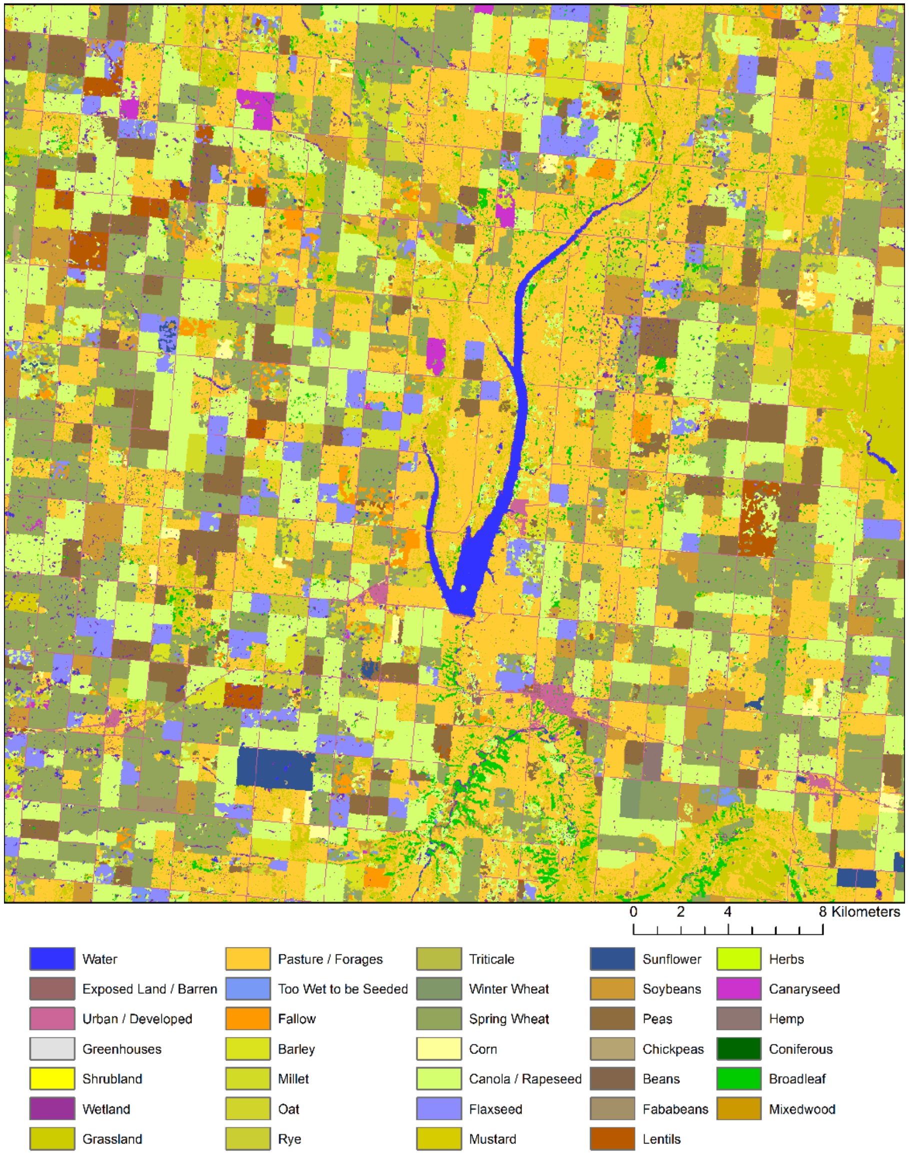
Mainly derived from remote sensing data and are focused on vegetation





Canadian Land Cover Classification Map (circa 2000)

Annual Crop Inventory (AAFC)



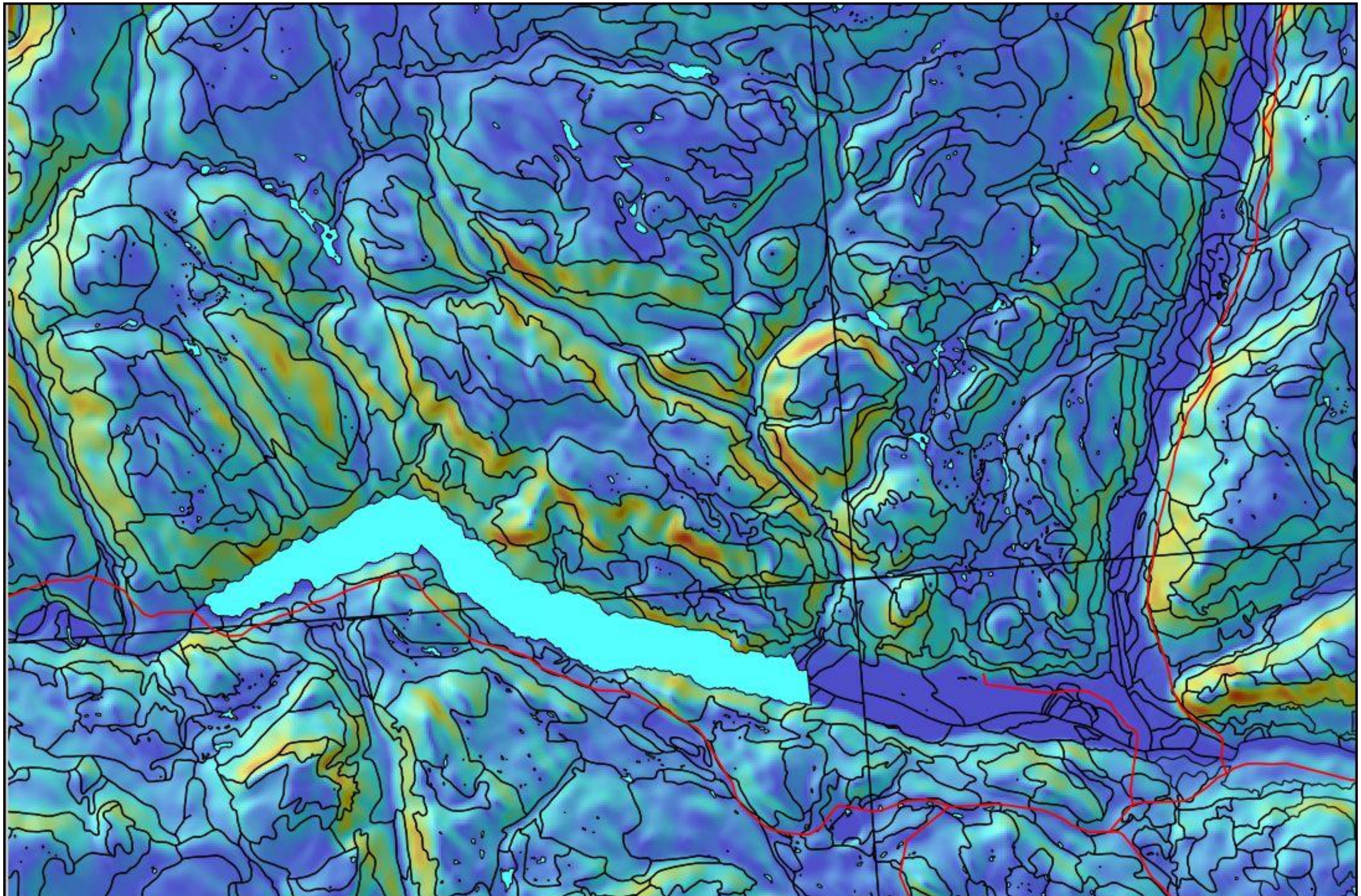
1 m spatial resolution | 15 m spatial resolution | 25 m spatial resolution

Relief (*r*)

Digital Elevation Model (DEM)

- Provides elevation data for each pixel with respect to coordinate system
- Readily available with extensive coverage
- Used as a base layer for nearly all digital soil mapping studies

Relief (*r*)

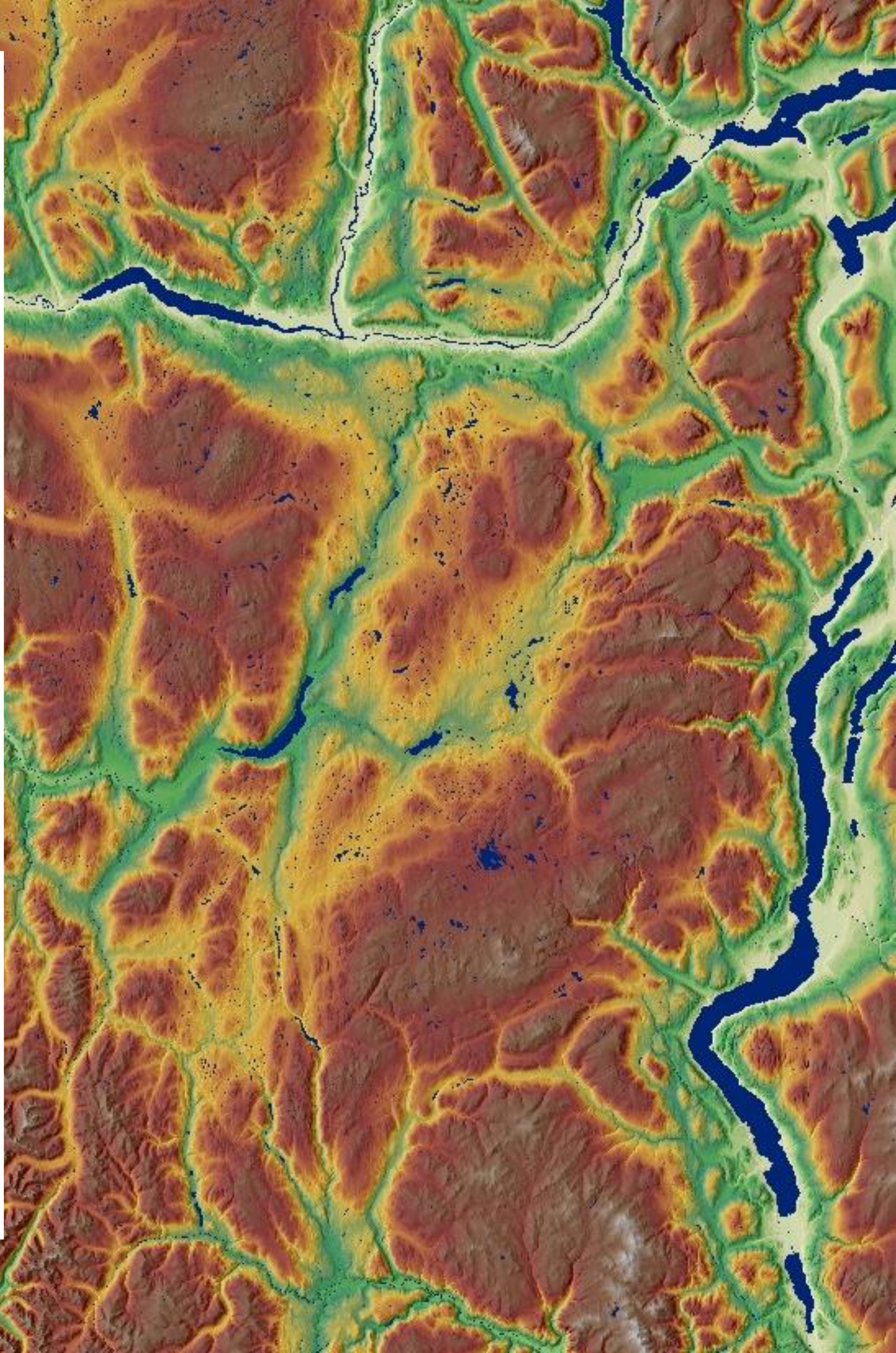


Digital Elevation Model (DEM)

Most commonly used dataset because it is widely available and cover large spatial extents

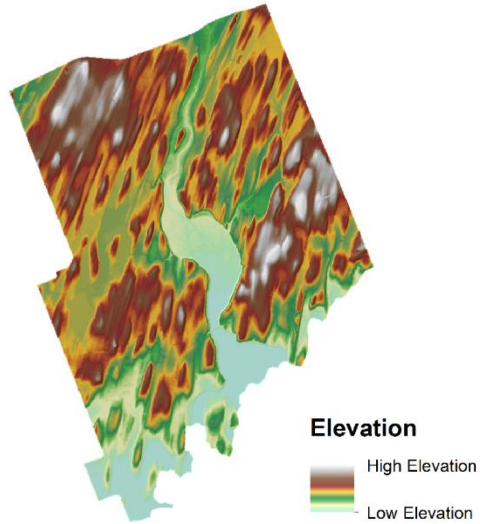
Used to derive many topographic derivatives:

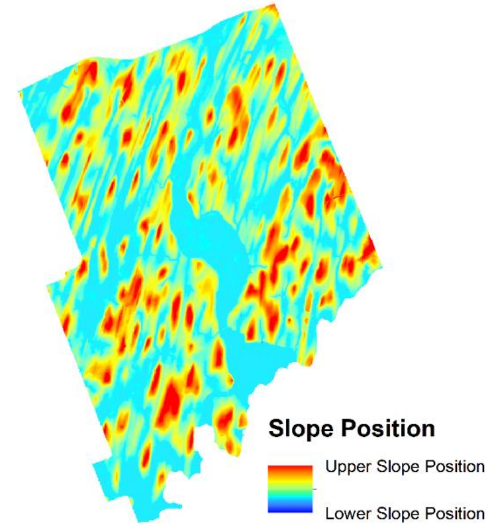
- Elevation, Slope, Aspect, Curvatures
- Solar Insolation, Skyview Factor
- Topographic Wetness Index, LS-Factor
- Landform Classifications
- Relative Slope Position
- Multi-Resolution Derivatives

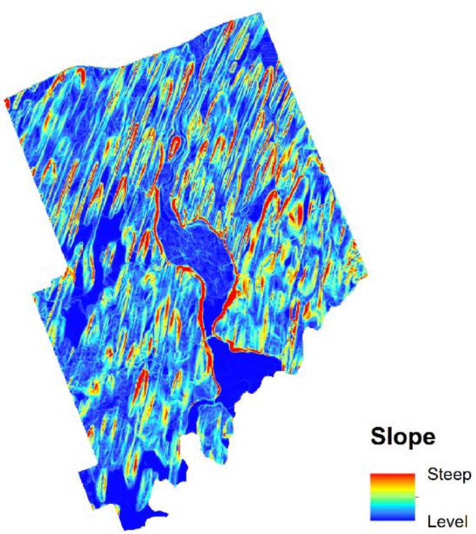


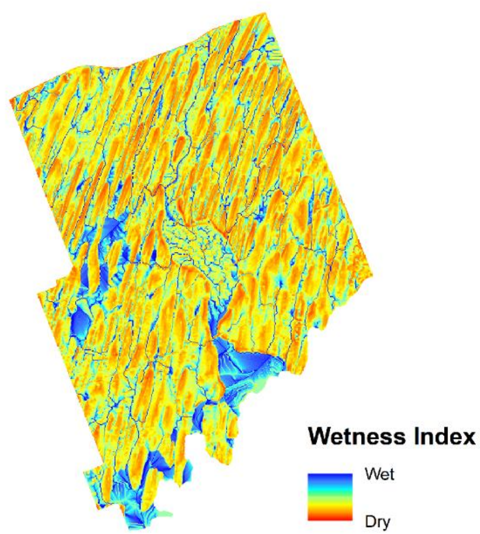
Basic Terrain Derivatives

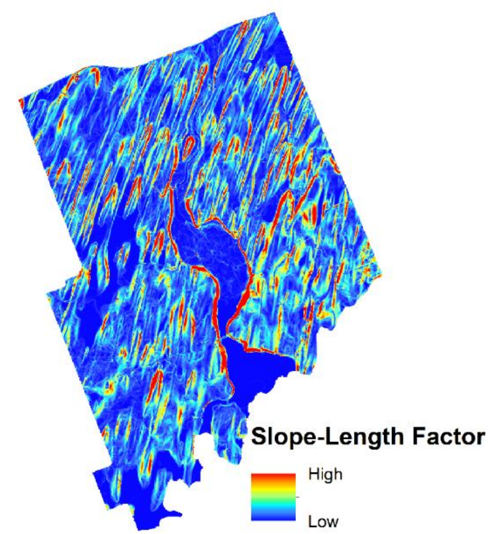
20 m Digital Elevation Model





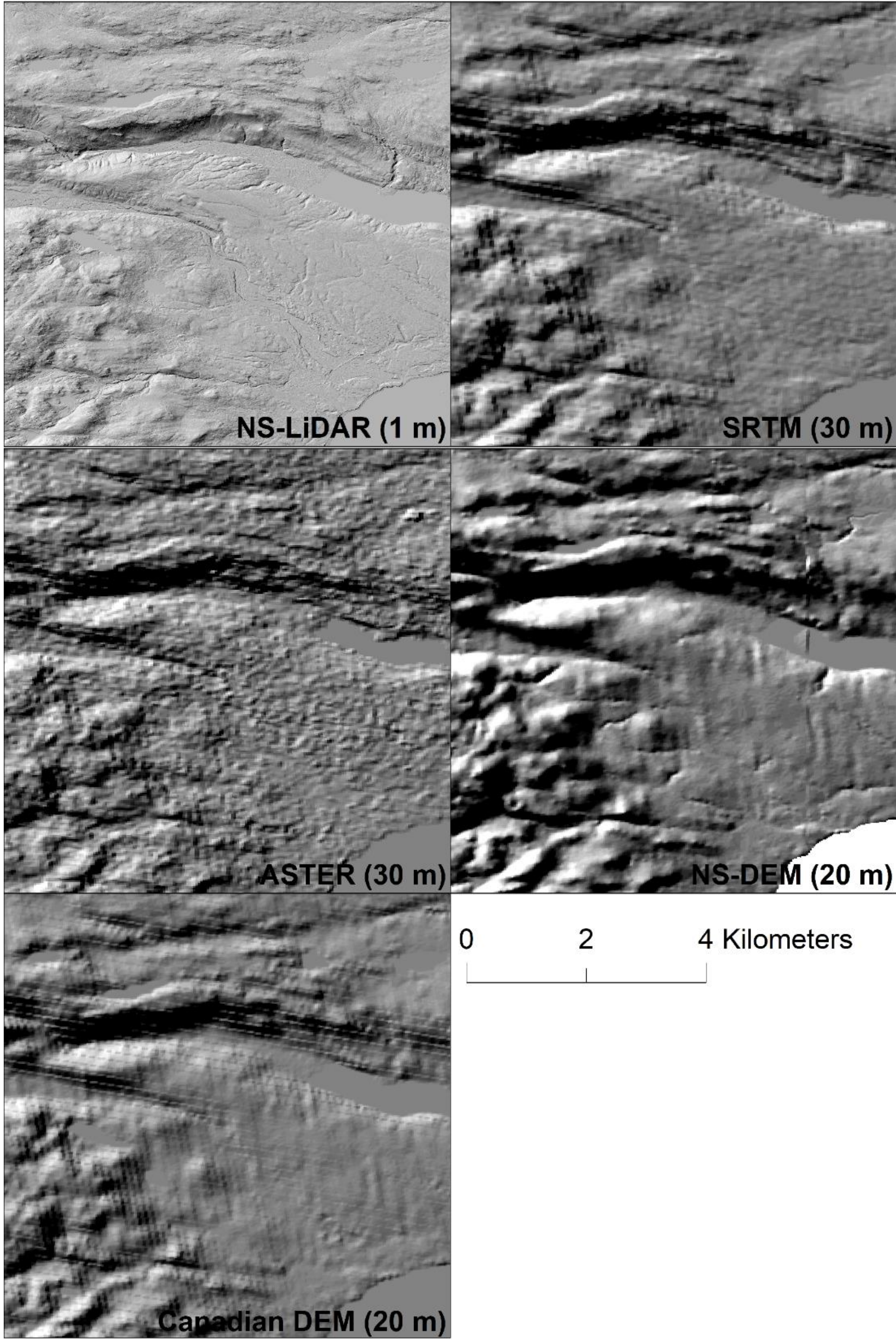


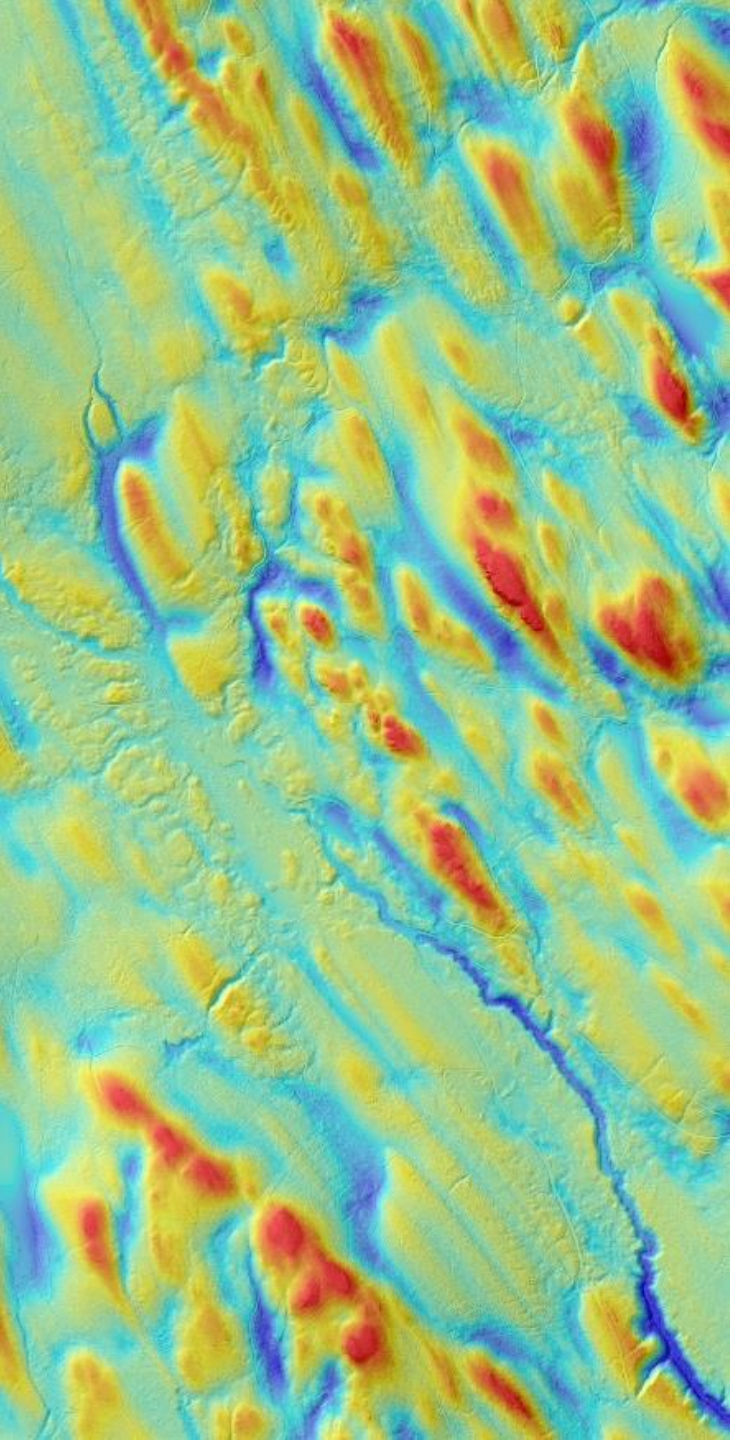




Some Problems

- Data anomalies – especially from satellite derived products.
- Increase in Resolution = Increase in Noise
- Influences of anthropogenic features
- Data storage
    - Float vs. Integer





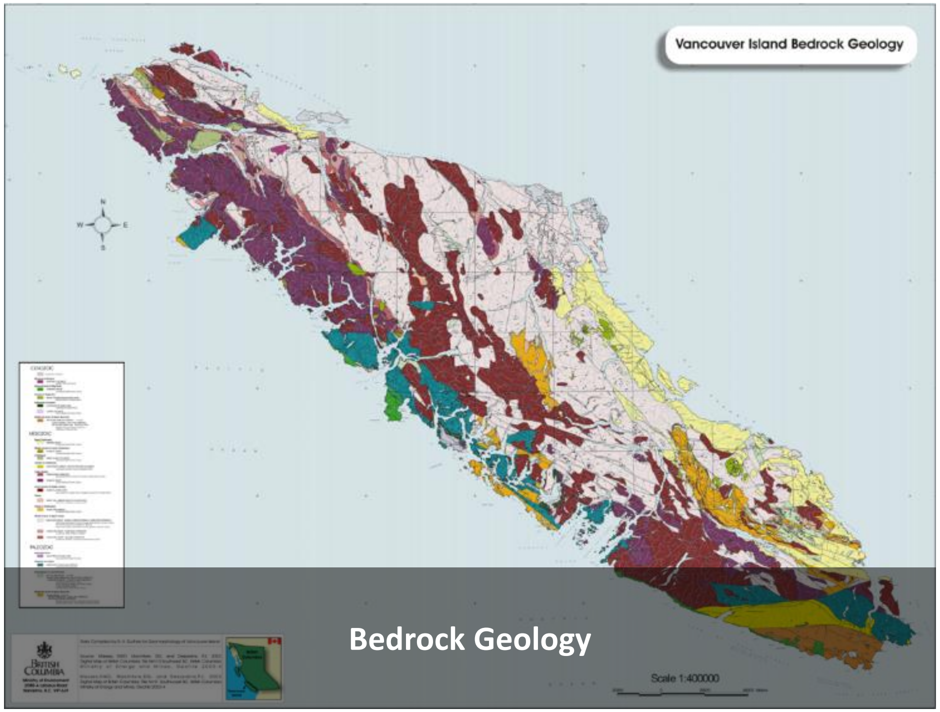
Some Tips

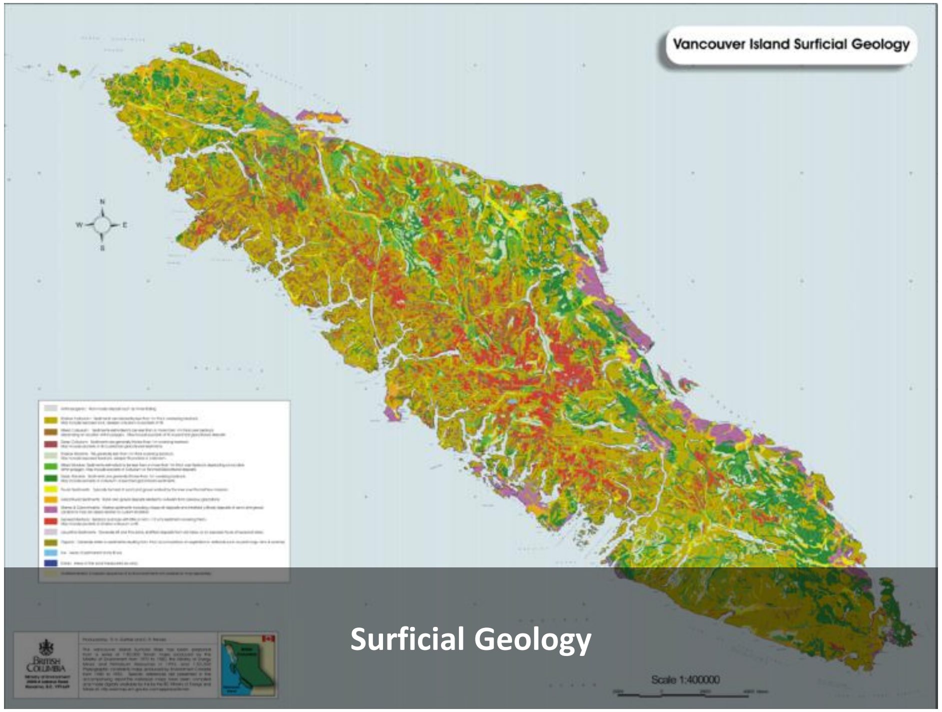
1. Always look at a hillshade model of the raw DEM.

2. Identify any anomalies or artifacts. These are usually the result of interpolation and data acquisition methods used for producing the DEM.

3. Apply pit removal algorithm and/or filtering.

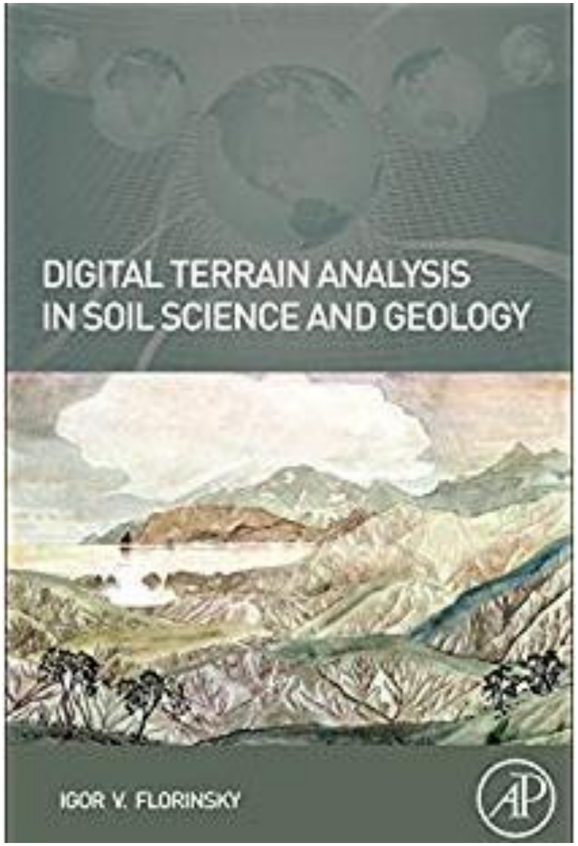
4. Hydrological correction by reinforcing stream network.

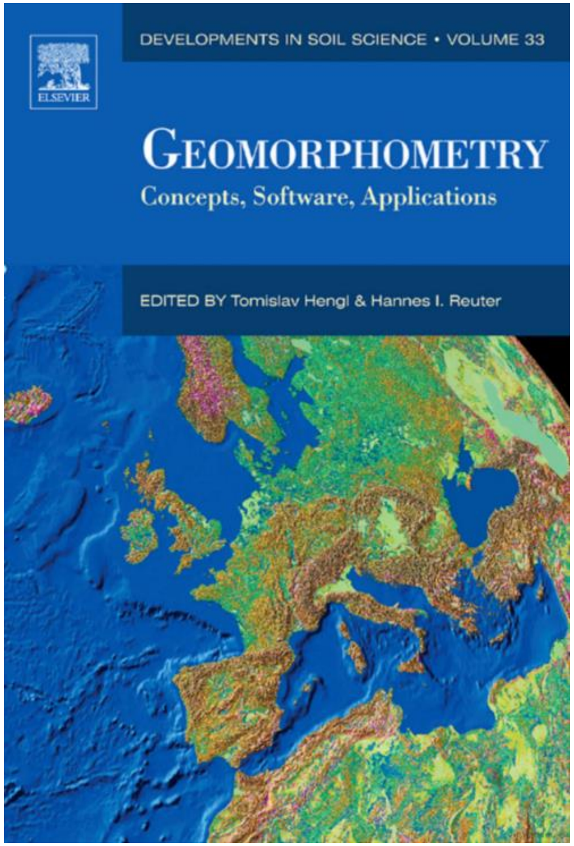




Parent Material (*p*)

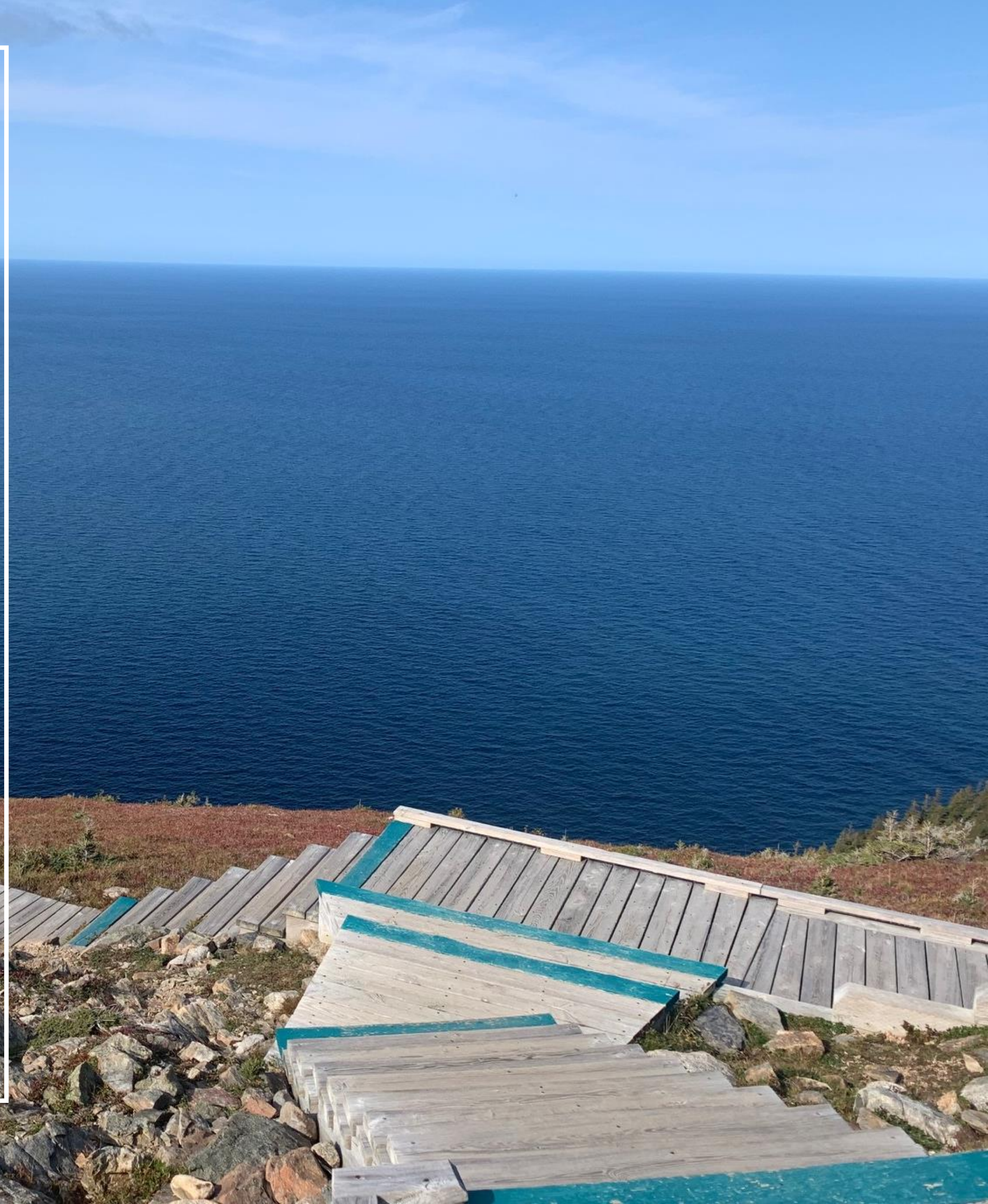
Useful Resources





Module 3.1: Creating Environmental Covariates

- Navigating the RSAGA package
- Geoprocessing with RSAGA
- Geoprocessing with Whitebox
- **Case Study:** A digital elevation model for Keene, ON will be used to generate a suite of topographic variables



Outline

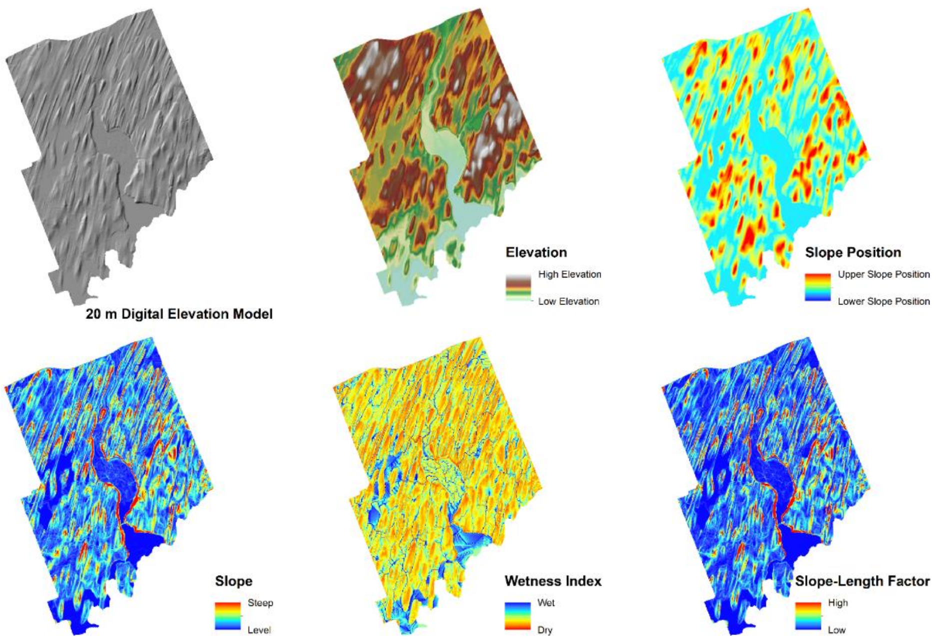
1. Overview of Soil-Environmental Variables

2. **Dealing with Multi-Collinearity**

On Multi-Collinearity—The Curse of Dimensionality

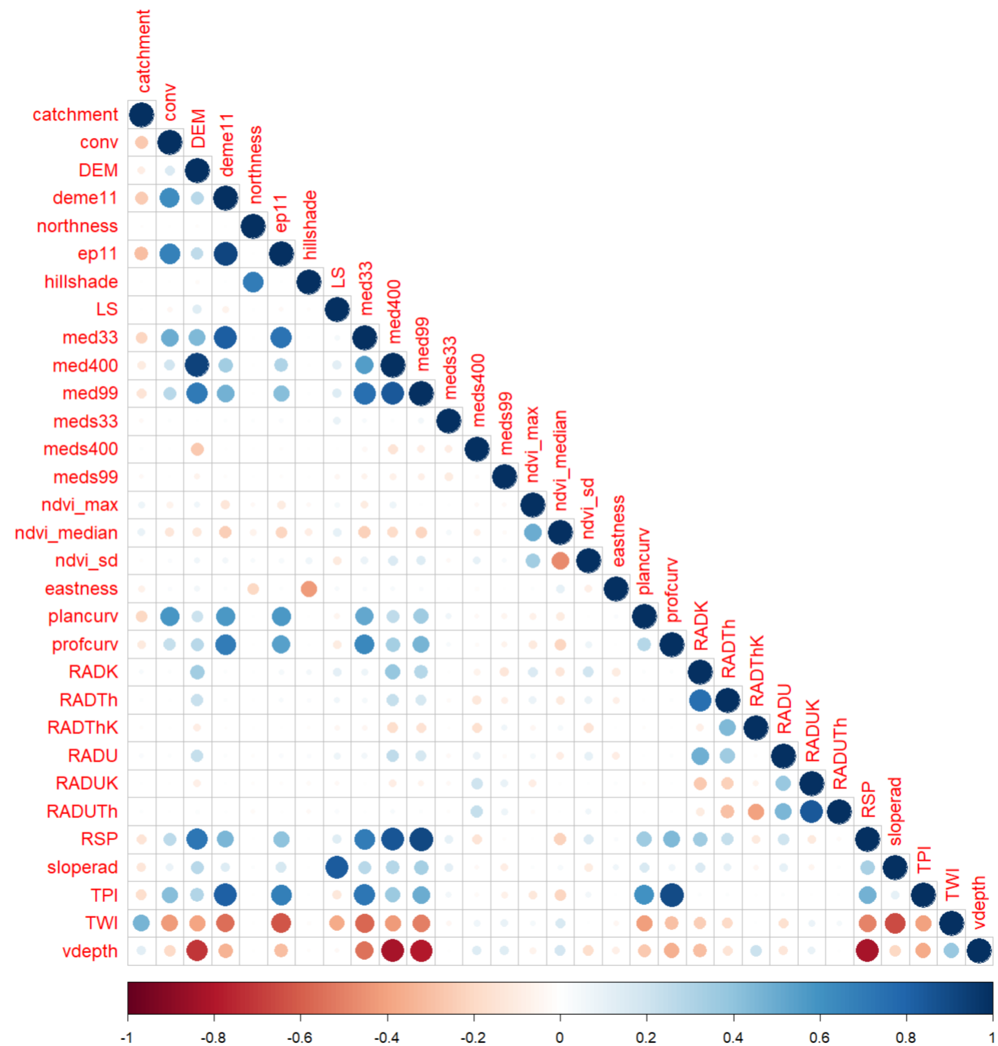
- Environmental covariates are often correlated.
  - Derivatives produced from the same DEM
  - Derivatives produced from multi-temporal remote sensing data

- Some models are sensitive to multi-collinearity.
- Difficult to interpret soil-environmental relationships.
- Reduces computational efficiency
- High-model complexity



On Multi-Collinearity—The Curse of Dimensionality

- Environmental covariates are often correlated.
  - Derivatives produced from the same DEM
  - Derivatives produced from multi-temporal remote sensing data
- Some models are sensitive to multi-collinearity.
- Difficult to interpret soil-environmental relationships.
- Reduces computational efficiency



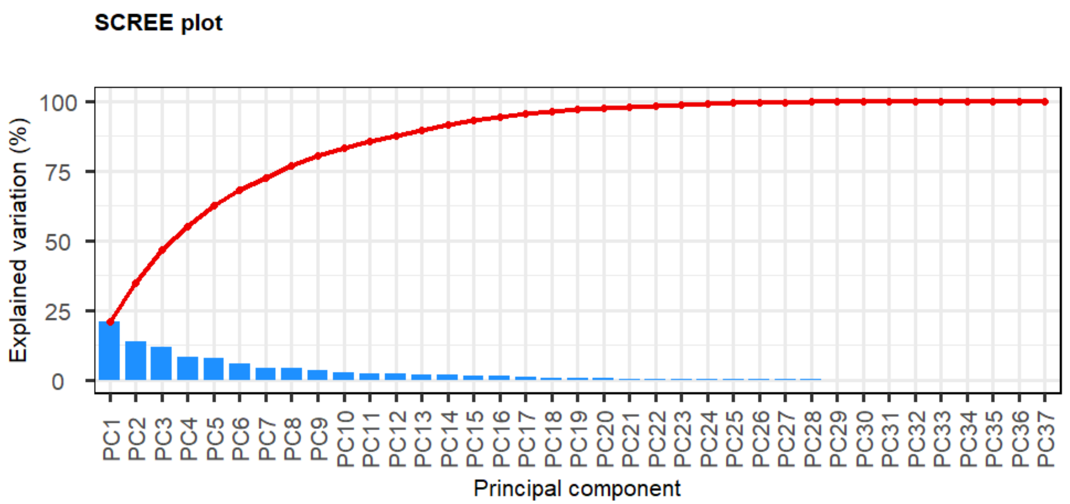
Solution 1: Principal Component Analysis

- Linear transformation of predictors into a series of non-correlated principal components.

- Each principal component accounts for some portion of the total variance.

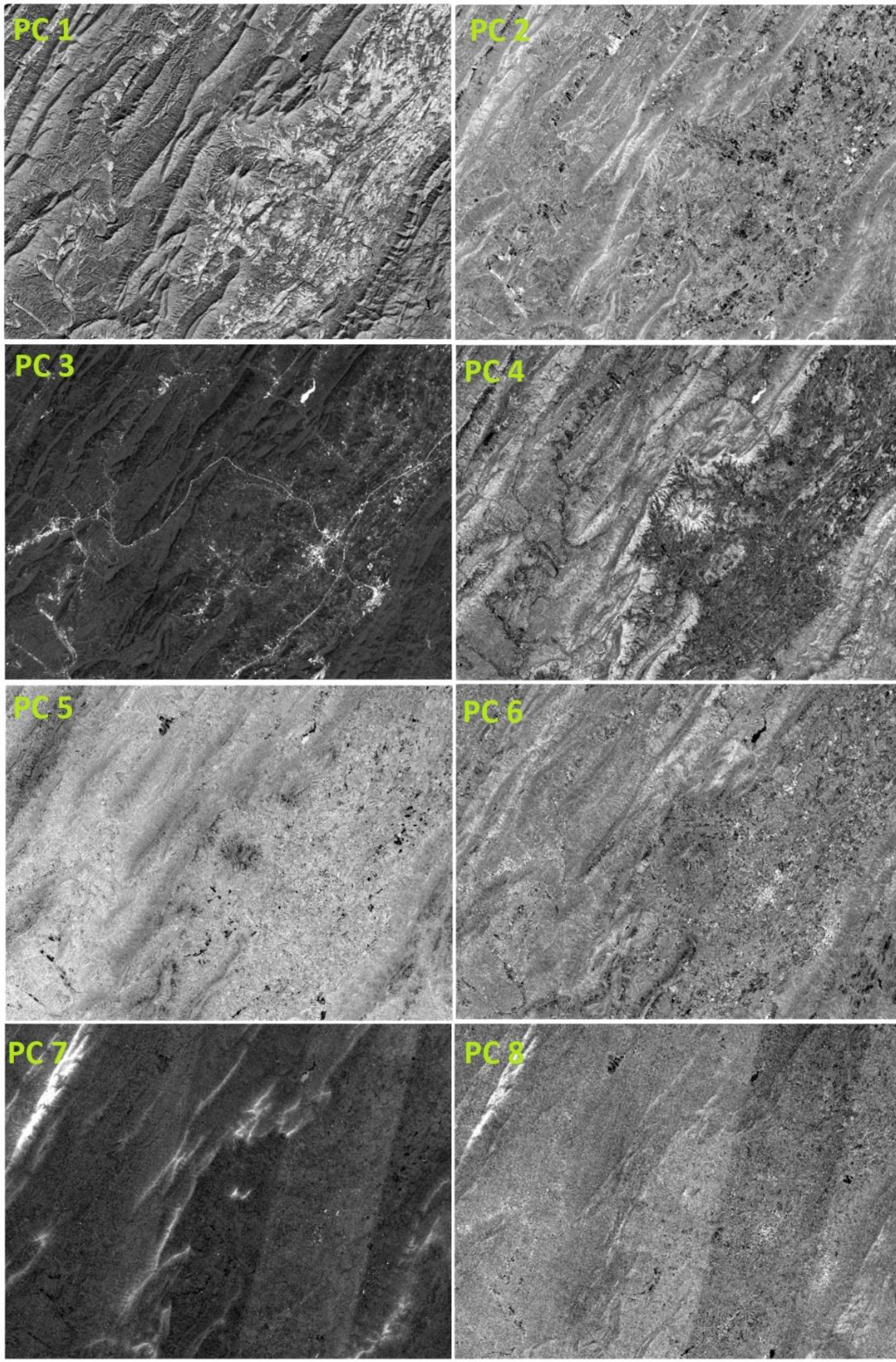
- The principal components that account for 90-95% of the total variance are retained as predictors for modelling

- Higher PCs are noise



Solution 1: Principal Component Analysis - Issues

- Difficult to interpret the soil-environmental relationships because all covariates are used to calculate each PC.
- Original covariates are lost.
- Could be computationally demanding for large datasets.



Solution 2: Variable Inflation Factor Analysis

- **Step 1:** Run ordinary least squares regression whereby one predictor is fitted to all other predictors. Calculate R^2.
- **Step 2:** Calculate variable inflation factor for each variable.
- **Step 3:** Remove variable with highest VIF.
- **Step 4:** Repeat Steps 1-4 until all variables have a VIF of less than a threshold (e.g., VIF<5 or VIF<10)

$$X_1 = \alpha_0 + \alpha_2 X_2 + \alpha_3 X_3 + \cdots + \alpha_k X_k + e$$

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2}$$

Solution 2: Variable Inflation Factor Analysis

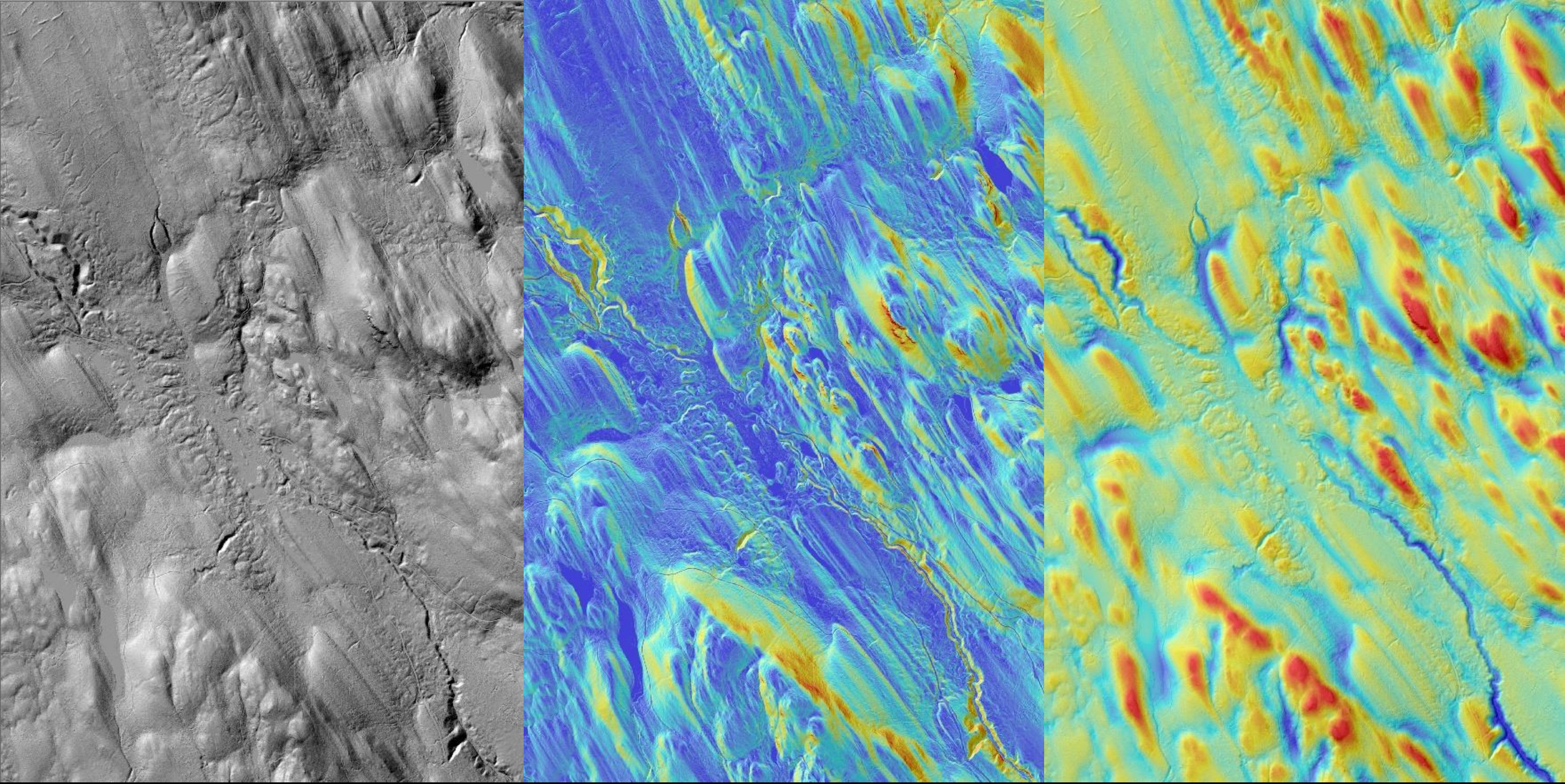
- The original covariates are retained.
- Easier to interpret soil-environmental relationships because the covariates are not transformed (in comparison to PCA).

$$X_1 = \alpha_0 + \alpha_2 X_2 + \alpha_3 X_3 + \cdots + \alpha_k X_k + e$$

$$\mathrm{VIF}_i = \frac{1}{1 - R_i^2}$$

Module 3.2: Covariate Reduction

- **Case Study:** Variable inflation factor analysis will be carried out using environmental predictors for Keene, ON.



3. Environmental Covariates for Digital Soil Mapping

Canadian Digital Soil Mapping Workshop, 2022

CANADIAN SOCIETY OF SOIL SCIENCE
SOCIÉTÉ CANADIENNE DE LA SCIENCE DU SOL