Machine Based Surficial Geological Mapping of Arctic Canada at the Geological Survey of Canada

Hazen Russell and Jeff Harris
Collaborators

GSC
• Sharpe, D.

• Broscoe, D. (Algonquin College)
• Lesemann, J. (Ottawa U)
• Lindsay, J. (Guelph U)
• Parkinson, W. (Carleton U)
• Richards, M. (Carleton U)
Program and Project Affiliations

- **GEM**: Geo-mapping for Energy and Minerals
  - Five years (2008-2013)
  - Mapping of geological resources north of 60

- **SMART**: Systematic Mapping of Arctic Canada by Remote Techniques
  - Derivative of RPM: Remote Predictive Mapping
Drift prospecting is required in Canada because most of the bedrock is blanketed by sediment ("drift")
Dispersal trains occur in till, eskers and stream/lake deposits.

These media, and especially till (since it covers so much terrain), are therefore targeted during mineral exploration campaigns.

= dispersal train
The Lac de Gras diamond fields were found based on results from one esker sample.

Compared to till, esker dispersal has received little to no attention.

Canada’s 1st diamond mine was discovered using an esker.
Bedrock RPM

- Remotely sensed
  - Landsat TM
  - Radarsat
  - Aster
  - Hyperspectral

- Geophysics
  - Magnetics
  - Gravity
  - Seismics
  - MT

- DEMS

- Geochemistry

- Field observations

- Landsat TM – 432(RGB)
- Radarsat – RGB
- Landsat TM and Radarsat
- Magnetic – shadow enhanced
- Gamma ray ternary image
- DEM
- Field observations
Outline

- Background
  - Problem / Rational
  - Bedrock
  - Previous work at GSC
- The Challenge
- Case Study
World Around Us

- Public domain
- DEMS
- Sensor proliferation
- Hyperspectral / radar
- Digital Globes
- Computation / visualization
The Problem

- Large landmass
- Shrinking budgets relative to operational costs
- Diminishing research staff numbers
- Limited internal capacity
- Increased demand for products
- High expectations for product delivery
- A philosophical crisis

$ Oil : 1947 - 2009

Z. Deretsky / NSF
Rational for SMART

Surficial Geology Map Coverage
- Mapped Area (500K and Below)
- Airphotos and Field Work - 100K
- EGS Coverage
- Mapped (No Field Work) - 250K
- Reconnaissance - 250K
- Reconnaissance - 1M

GSC Staff

Last Updated October, 2010
Objectives

• Develop a hybrid machine based methodology for the generation of surficial geological maps

• Deconstruct the cognitive process and address as many components as possible through machines based approaches

• Employ public domain data

• Datasets should be national – regional in extent
Software

- Government
- Corporate
  - ArcInfo
  - ENVI
- Open Source Environments
- SAGA
- Whitebox
- R- statistical package
Glacial Flowline Map

Manual Interpretations

SRTM and LANDSAT south 60

LANDSAT north 60

Generic flow

A flowline map of glaciated Canada based on remote sensing data

John Shaw, Davis Sharpe, and Jeff Harris
Subglacial processes of the Scandinavian Ice Sheet in Fennoscandia inferred from flow-parallel features and lithostratigraphy

Mikko Punkari

\[ \text{West Mains Road, Edinburgh EH9 3JW, UK} \]

September 1996

**RECONSTRUCTING THE EVOLUTIONARY DYNAMICS OF FORMER ICE SHEETS USING MULTI-TEMPORAL EVIDENCE, REMOTE SENSING AND GIS**

**CHRIS D. CLARK**

Department of Geography & Sheffield Centre for Earth Observation Science, University of Sheffield, Sheffield, S10 2TN, UK (Email: chris.clark@sheffield.ac.uk)

Quaternary Science Reviews 19 (2000) 1343–1366

Geomorphological reconstruction of the Labrador Sector of the Laurentide Ice Sheet

Chris D. Clark, Jane K. Knight, James T. Gray
Current GSC Benchmark

Grunsky, McMartin and Harris, 2008

Multi-beam RADARSAT I

LANDSAT TM

84.8% accuracy

Classification (Maximum Likelihood)
The Challenge
Challenge: Parsing the Process

A. Photographs = data, e.g. Spectral (LandSat) Radar (RadarSat) Shape files DEM (CDED, SRTM)

B. Stereoscope = DEM derivatives

C. Cogititive Process = Expert knowledge input; legacy data integration; modelling
Sediment – Landform Mapping

- Geological materials do not necessarily have unique spectral signatures
  - Moisture regime
  - Vegetation
    - Reflects:
      - Landscape location
      - Thickness
      - Material provenance

- Hence
  - Conceptual uniqueness does not equal spectral uniqueness
Public Domain Datasets

- Spectral (LandSat, Spot 4; Geobase, Geogratis)
- RadarSat
- DEM

Operational Scale

- 50,000
- 250,000

Operational Scale

- 10 to 30 m

Pixel resolution
Study areas commonly > 40,000 km² (2 NTS sheets)

Landsat Scenes 33,300 km²

Landsat path oblique to NTS tiles
Data Integration challenges

Regional mapping challenges

- Scene leveling
- Seasonality
- Phenology
- Atmospheric (e.g. haze)

LandSat
Good Arctic Coverage
Arctic Island multi channel
leveled mosaic
30 m resolution
Challenge: Parsing the Process

A. Photographs = data, e.g. Spectral (LandSat) Radar (RadarSat) Shape files DEM (CDED, SRTM)

B. Stereoscope = DEM derivatives

C. Cognitive Process = Expert knowledge input; legacy data integration; modelling
Addressing the Problem

Emerging

1. Science Language
   - Parse in to machine Operable Components

Workflow

2. Legacy Data
3a. Morphology
3b. Spectral
3c. Shape
4. Associations
5. Modelling
Expert Knowledge

- Process based models of origin of landscape and sedimentary deposit.
- High level abstract landform definitions
- Analogues from modern glacial environment

**Esker Definition**

A ridge of gravel and sand emplaced during glacial melt by the deposition of sediments from meltwater rivers flowing on the ice (channel fills) or beneath a glacier (tunnel fills)

Rogerson, Canadian Encyclopedia
CDED DEM 0.75 arc secs projected to 12.5 m² pixels

NTDB Eskers
Aylsworth Eskers
l1, l2
r1, r2, r3
esker
ridges
s1, s2
spurs
lakes
1. Science Language / Ontology

- Concept mapping
2. Legacy Data

- Map Data
- Field observations
- Reports
3a. Relief / Morphology: Mapping Eskers
Why start with eskers?
CDED : Data Quality
Process to Quantify eskers

- Multi-step process
  - DEM analysis
  - Spectral analysis
  - Integrated data definition

- Filter DEM - reduce noise due to integer CDED and contour origin.
- Smooth DEM to obtain regional surface
- Subtract local from regional surfaces.
DEM Ridge Extraction
Improving the capture

A. Spectral data – SPOT 4 and 5 Geobase
B. Isocluster unsupervised classification - 100 classes
C. Four highest ranked classes of cluster analysis within esker polygons isolated and merged with polygons
DEM and Spectral Integration
Summarized Volume Estimates

Summarized Esker 'Volumes' (cubic m)
5 Kilometre Intervals along Esker

- 309.915543 - 312233.204842
- 312233.204843 - 619820.209312
- 619820.209313 - 1064500.794411
- 1064500.794412 - 1878406.057358
- 1878406.057359 - 4215809.626579

Aylsworth Eskers Selected for Esker Creation

0 10 20 30 40 50 100 Kilometers

1:1,700,000
3b. Spectral / Radar Signals

Chapter 14
Predictive Mapping of Surficial Materials, Schultz Lake Area (NTS 66A), Nunavut, Canada

ERIC GRUNSKY, JEFF HARRIS, AND ISABELLE Mc MARTIN
Geological Survey of Canada, Ottawa, Ontario, Canada
Spectral / Radar Data Usage

- **Spot Panchromatic**
  - Expert interactive interpretation of landforms

- **LandSat**
  - Unsupervised and supervised classifications of material types
  - Accuracies of 80% with water
  - Accuracies of < 60% excluding water

- **RadarSat**
  - Additional resolution for roughness and moisture refinement
Examples

- Grunsky et al. accuracy results
  - RadarSat = 83%
  - LandSat = 84%
  - Integrated with DEM 93.8% (includes water)
  - No water < 60%
3c. Shape / Spatial Relationships

A. Bedrock
B. Thermokarst
C. Drumlin
D. Complex

Images showing different landforms:
- A: Bedrock
- B: Thermokarst
- C: Drumlin
- D: Complex
Shape Analysis

- Common application to a range of feature analysis
  - Grain shapes
  - Pattern recognition (Micro to Macro)
  - Ecology (Fragstats, Patch)

3c. Shape / Spatial Relationships

- Begin with projected Shapefile of Lakes
  - Compute shape indices
  - Generate statistical surfaces
  - PCA of statistical surfaces
  - Segment into landscape classes
<table>
<thead>
<tr>
<th>Derivative</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Lake Area</td>
<td>Average lake area in moving window</td>
</tr>
<tr>
<td>Lake Density</td>
<td>Lakes per unit area</td>
</tr>
<tr>
<td>Mean Nearest Neighbour Distance</td>
<td>Average distance between lakes in moving window</td>
</tr>
<tr>
<td>Mean Shape Index</td>
<td>Average shape complexity within moving window</td>
</tr>
<tr>
<td>Mean Fractal Dimension</td>
<td>Average shape complexity within moving window</td>
</tr>
<tr>
<td>Mean Elongation</td>
<td>Average elongation within moving window</td>
</tr>
<tr>
<td>Mean Roundness</td>
<td>Average roundness within moving window</td>
</tr>
<tr>
<td>Mean Circularity</td>
<td>Average circularity within moving window</td>
</tr>
<tr>
<td>Mean Related Bounding Figure</td>
<td>Mean compactness within moving window</td>
</tr>
<tr>
<td>Mean nearest neighbour azimuthal difference</td>
<td>A measure of streamlining based on pairwise azimuthal difference of neighbouring lakes</td>
</tr>
</tbody>
</table>
Sample Lake Shape Indices

- Elongation (mean=0.29)
- Relative Bounding Feature (mean = 0.88)
- Fractal Dimension (mean=1.29)
Mean azimuthal difference between nearest neighbour pairs (streamlining)

A measure of streamlining based on pairwise azimuthal difference of neighbouring lakes
5. Modelling - 1

- Robust Classification Method (RCM)

IDL ENVI

monte carlo simulation
random cross-validation to address uncertainty

GEMS
5. Modelling - 2

Random Forest Modeling and Classification of Landsat ETM+ and DEM Data for Surficial Material Mapping

Parkinson W.¹, Richardson M.¹, Russell H.²
Department of Geography and Environmental Studies, Carleton University
² Geologic Survey of Canada
Random Forest - Advantages

- A stochastic implementation of Classification & Regression Trees (CART)
- Powerful against outliers
- Provides variable importance measures
- Superior or equal to others
- Non-parametric
Random Forest

- Maximize available training and validation
  Handle outliers

- Provides measurements of uncertainty
  Absolute statistical
  Spatially across the map
  Per class identified
Random Forest: Classification

Probability Used for Classification

Land Area and Percent Certainty

Probability

Land Area (%)

Percent Certainty (%)

Bedrock  Thick till  Thin till  Organic  Sand & Gravel

70%

60%

GEMS
RF Class Prob 50%

72%
Classification 70%
Case Study
Victoria Island

- 4-6 250,000 scale NTS sheets
- Approximately 80 - 120,000 km²

Ground ice environment
Thick sediment
Raised marine limit
Aeolian
Classified Image

- Bedrock and moisture content map

Minimal vegetation
Poor image
seasonality

- Bedrock (no moisture)
- Low Moisture
- Low Moisture
- High Moisture
- Very High Moisture
Steps to a Surficial Geology Map

- Landform integration
  - Legacy data
  - Automated mapping
- Rule-based material categorization
Isostatic Rebound and Sea Level
Selective Reclassification

- Extraction of glaciomarine/neritic deposits from spectral signature and elevation

Legend

SURFACE MATERIALS

- **Colluvial deposits**
  - Silty to rubbly diamicton resulting from mass wasting of bedrock, till, sand and gravel, and colluvial sediments.

- **Emerged neritic deposits**
  - Stony sandy silt, up to few metres in thickness, composed of reworked till, glaciofluvial sediments, and glaciomarine sediments. May form beaches and strandlines.

- **Sand and gravel**
  - Sandy and gravely deposits of alluvial origin including terrace, fan, deltaic and rare littoral sediments.
  - Sandy and gravely deposits of glaciofluvial origin including kame and eskers, outwash plains and terraces, and outwash deltas.

- **Till veneer**
  - Sandy to bouldery till with high carbonate content and bearing Shield erratics. Occurs commonly downstream of bedrock escarpments and near crests of streamlined landforms. Thickness: 1 m or less.

- **Till blanket**
  - Sandy to bouldery diamicton with high carbonate content and bearing Shield erratics. Occurs within till plains and swarms of streamlined landforms. May occur as boulder log below marine limit.

- **Bedrock**
  - Palaeozoic dolomite, limestone, and shale; subhorizontal. Abraded and fractured near escarpments with thin cover of streamlined till. May contain surface boulders and/or bed sediment free near or below marine limit.

LANDFORMS

- Streamlined landforms
- Eskers
- Meltwater channels
- Raised shorelines
- Hummocky topography
Classification Evaluation

- Verification of classification in targeted areas

---

**Map Legend:**
- **Pink:** Bedrock
- **Green:** Till blanket
- **Light Green:** Till veneer
- **Brown:** Colluvium
- **Yellow:** Sand and gravel

---
Classification Evaluation

- Till veneer
- Till blanket
- Colluvium
Summary

- National integrated datasets required
- Need for geologists to think beyond the stereoscope
- Compilation of machine based methods to replace high level cogitative processes are difficult
- Machine methods can reduce mechanical aspects of geological map production
- Greater quantification is possible using machine based methods
Remote Predictive Mapping Technical Exchange Presentation Series
Winter 2012

Series Coordinators
X. Geng, J. Harris, and H. Russell

A collaborative series of presentations to encourage transfer of technical knowledge on RPM methods. On a weekly basis, a technical exchange presentation will be held at 13:00 on Thursday afternoons. On an alternating-weekly basis, the theme will switch between bedrock and surficial subjects. All bedrock presentations will be at the GSC and surficial presentations will be hosted at both the GSC and Agriculture Canada on the Experimental Farm. This initiative is part of the Geo-mapping for Energy and Minerals Program (GEM).

Presentations from government, academia, and industry are welcome. Presentation may range from techniques development to applications. Please contact Hazen to be added to the schedule.