Topographic normalization, satellite data source and classification method for alpine vegetation mapping in Sweden

Heather Reese and Håkan Olsson
Section of Forest Remote Sensing
Department of Forest Resource Management
SLU, Umeå
E-mail: heather.reese@slu.se
Objectives

To map alpine vegetation types and to test:

- Different satellite data sources
- Effect of including elevation derivatives
- Different quantity and configurations of training data
- Two supervised classification algorithms

To develop a reliable procedure for performing c-correction
Objectives

Evaluated satellite data sources:

- SPOT (10 m)
- Landsat (25 m)
- IRS AWiFS (60 m)
Objectives

Effect of including elevation derivatives,

Using a 50 m DEM:

- Elevation
- Slope
- Aspect
- Topographic wetness index
Objectives

Test of different quantity and configurations of training data

- 2300 air photo interpreted plots, from systematic sample
- 200 air photo interpreted polygons, subjectively chosen
- 100 field plots, from the national inventory of landscapes (NILS)
Test of two supervised classification algorithms

- Quadratic Discriminant Analysis
- Random Forest
Vegetation types

Classes are defined by vegetation composition, height, and density, and moisture regime

15 Classes
Bare rock
Extremely Dry Heath
Dry Heath
Mesic Heath
Wet Heath
Grass Heath
Short alpine meadow
Tall alpine meadow
Snow bed vegetation
Willow
Mire
Mountain Birch (shrub layer)
Mountain Birch (forb layer)
Snow/Ice
Water
Study area

Vindelfjällen Nature Reserve

Is an area with both northern and southern alpine vegetation types

Vegetation season
mid-June to mid-August
Satellite data

Used the overlap area between sources

SPOT 5 HRG 10 m
Landsat TM 25 m
AWiFS 60 m

Two scene mosaic
Satellite data

- Used single date of imagery (due to cloud cover)

- Where possible, two dates of imagery were used to classify vegetation typified by its quickly changing nature, namely snow bed vegetation and wetlands

- Indices included: NDVI, NDII and SAVI
Topographic Normalization

Topographic normalization with C-correction

Before

After
Topographic normalization with C-correction

After applying a topographic normalization using "c" calculated from a large random sample, the classification accuracy became lower than that for an non-normalized image.

Something was wrong!

C-correction (Teillet et al., 1982)

Corrects for shadowing in an image based on
- sun angle
- surface slope
- plus the empirical parameter "c", calculated from a sample of the satellite data and the corresponding angle of illumination, $i$

\[
\hat{\rho}_{\lambda h} = \hat{\rho}_{\lambda t} \frac{\cos z + c_{\lambda}}{\cos i + c_{\lambda}}
\]

Where \( \hat{\rho}_{\lambda t} \) is the reflectance for band $\lambda$ still influenced by topography ($i$), $z$ is the solar zenith angle, and $i$ is the solar incidence angle upon the surface.
To calculate $c$, a sample of values of cosine of $i$ and the corresponding spectral data are needed.

Very few papers give any information about the sample used to derive the empirical $c$-parameter (except Civco 1989). They often only say they’ve used a ”large” sample.

In the literature, two ways to sample emerge:

1) subjective selection of small number of observations ($n<100$) for a single vegetation type for a range of topographic conditions

2) random sample over an entire image or subset of image
Methods

1) “random sample”
   a random sample with a “large” number of observations

2) “aspect sample”
   a stratified random sample where observations were stratified by
   aspect (north- and south-facing slopes)

3) “cosi sample”
   a stratified random sample where observations were stratified by cosine of $i$

All samples were taken on alpine vegetation only, excluding rock, snow, water, mires, and forest
Results

Large random sample (n=16,500)

Most studies use Smaller random sample (n=1,600)

Aspect stratified sample (n=5,000)

Best results Cosine of $i$ stratified sample with optimal (power) allocation (n=5,000)
Conclusions from topographic normalization with c-correction

Calculation of \( c \) was more precise and accurate using a sample stratified on cosine of \( i \)

The sampling method does influence the outcome of \( c \), and therefore the topographic normalization

Different vegetation types, most likely with influence of vegetation cover and background soil reflectance, have a different value of \( c \) for optimal topographic normalization.

\( \text{NIR reflectance} \)

Background – sample based national inventories

**NFI**
- field visits >10,000 plots per year
- much data on forest, ground cover...
- plots revisited on a 5 to 10 year cycle

**NILS**
- 631 5×5 km squares
- 12 field plots in inner 1×1 km square
- photo-interpretation of 1×1 km
- all land cover types

**THUF/MOTH**
MOTH project:
- 200 points photo-interpreted in 2×5 km in NILS squares
- field visit to plots of interest
Background – National Inventory of Landscapes in Sweden (NILS)

NILS
631 5x5 km tracts
C-IR 1:30,000 aerial photos

Inner 1x1 km has 12 field-inventoried plots

EU MOTH project photo-interprets 200 plots in each NILS tract
Doubled the NILS 5 x 5 tracts
Doubled the NILS 5 x 5 tracts

110 plots
Systematic

20 polygons

NILS field plots

Interpretation of all vegetation types present (%) at center point, 5, 10, 20 and 30 m radii
Training data

2 x NILS (N=2331)

2 x NILS 50% (N=1165)

1 x NILS (N=1368)

1 x NILS 50% (N=684)

<table>
<thead>
<tr>
<th>Training data</th>
<th>Nr plots after quality control for alpine only</th>
</tr>
</thead>
<tbody>
<tr>
<td>2xNILS</td>
<td>803</td>
</tr>
<tr>
<td>2xNILS50</td>
<td>436</td>
</tr>
<tr>
<td>1xNILS</td>
<td>532</td>
</tr>
<tr>
<td>1xNILS50</td>
<td>268</td>
</tr>
<tr>
<td>Subjective</td>
<td>200</td>
</tr>
<tr>
<td>NILS Field</td>
<td>89</td>
</tr>
</tbody>
</table>
Evaluation data

- Seven 1 x 10 km areas, 17 non-overlapping photos per area

- Multi-scale photo-interpretation of high resolution (1:2000) CIR photos.

- Systematic sample of 5 m, 10 m, and 30 m plots (n = 666)
Results

Highest accuracy using SPOT + Elevation with RF and all training data

Overall accuracy (pixel-level):
- 88% (alpine + subalpine)
- 73% (alpine classes)

SPOT 5 (10 m)

- 63% (alpine classes)
- Landsat 5 TM (25 m)

AWiFS (60 m)

- 47% (alpine classes)
## Results

Overall accuracy (%) for the alpine vegetation classification, for the Random Forest (RF) with and without DEM derivatives and Quadratic Discriminant Analysis (QDA) without DEM derivatives

<table>
<thead>
<tr>
<th></th>
<th>SPOT</th>
<th>Landsat</th>
<th>AWIFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RF w/DEM RF QDA</td>
<td>RF w/DEM RF QDA</td>
<td>RF w/DEM RF QDA</td>
</tr>
<tr>
<td>2xNILS</td>
<td>72.9 64.0 60.7</td>
<td>62.7 58.5 61.5</td>
<td>46.5 36.4 38.9</td>
</tr>
<tr>
<td>2xNILS50</td>
<td>67.0 -- 60.3</td>
<td>61.1 -- 61.1</td>
<td>--2 --2 --2</td>
</tr>
<tr>
<td>1xNILS</td>
<td>64.9 -- 56.5</td>
<td>59.4 -- 60.3</td>
<td>--2 --2 --2</td>
</tr>
<tr>
<td>1xNILS50</td>
<td>63.6 -- 54.4</td>
<td>59.4 -- 59.2</td>
<td>--2 --2 --2</td>
</tr>
<tr>
<td>Field data</td>
<td>--1 --1 --1</td>
<td>--1 --1 --1</td>
<td>51.2 ... 22.3</td>
</tr>
<tr>
<td>Subjective</td>
<td>69.9 62.3 60.3</td>
<td>56.5 53.2 60.3</td>
<td>45.9 33.4 33.9</td>
</tr>
</tbody>
</table>
Results

Elevation data increased accuracy of classification by ~9%

In particular:

- elevation to help separate grass heath and dry heath
- wetness index to identify willow
## Results

### Accuracy for individual classes (SPOT, RF)

<table>
<thead>
<tr>
<th>Class</th>
<th>Accuracy (pixel level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare rock</td>
<td>76%</td>
</tr>
<tr>
<td>Dry heath (torr rished)</td>
<td>73%</td>
</tr>
<tr>
<td>Mesic heath (frisk rished)</td>
<td>80%</td>
</tr>
<tr>
<td>Grass heath (gräshed)</td>
<td>93%</td>
</tr>
<tr>
<td>Willow (vide)</td>
<td>31%</td>
</tr>
<tr>
<td>Alpine meadow (örtäng)</td>
<td>50%</td>
</tr>
<tr>
<td>Mountain birch</td>
<td>88%</td>
</tr>
</tbody>
</table>

Willow classification most difficult due to:
- spectral similarity with other types
- definition of willow type (> 50% willow)
Full coverage vegetation maps with 10 m pixel size

Air photo interpreted Mountain Vegetation Map

Classification from SPOT data
Now being tested as a method for mapping mountain area
Conclusions

• SPOT data (10m pixel) was the best choice for detailed alpine vegetation classification

• Elevation derivatives increased map accuracy

• Random Forests classification gave better results than QDA

• Larger training data sets gave higher accuracy in general.

• NILS field data were too few for use with single SPOT or Landsat images … but for larger scenes areas like AWiFS or Sentinel-2 (300 km wide) they may be useful

• Role of subjectively collected data needs to be considered (faster, cheaper)

• Ballpark suggestion of ~ 40 to 60 samples per class + extra 30% to account for clouds, data quality problems
Future Directions

Sentinel-2 + Laser DEM + Laser DSM

Elevation, slope, etc

Vegetation height and density
Thank you!

Questions?