Analysis of hyperspectral airborne HyMap data for vegetation mapping around Lahnaslampi talc mine, Finland


aGeological Survey of Finland, P.O.Box 77, 96101 Rovaniemi, Finland, tel. +358-20-550 40, fax. +358-20-550 14, e-mail: maarit.middleton@gsf.fi
bGeological Survey of Finland, P.O.Box 96, 02151 Espoo, Finland
cGeological Survey of Finland, P.O.Box 1237, 70211 Kuopio, Finland
dFinnish Forest Research Institute, P.O.Box 16, 96301 Rovaniemi Finland

ABSTRACT

Hyperspectral (450-2500 nm) airborne HyMap imaging spectrometer data (spatial resolution 5 m) were processed for mapping of environmental effects on vegetation due to talc mining activities at Lahnaslampi, Sotkamo, Finland. As a part of the European MINEO project, the HyMap imagery was interpreted by applying vegetation indices, using standard classification and hyperspectral mapping approaches. Classification processes and accuracy assessment were based on forest inventory data, land use data, geochemistry and soil electrical conductivity data, as well as aerial photograph interpretation of tree species composition and recent disturbances due to forest management and mining operations.

First, a general forest stand classification was performed with the maximum likelihood classifier, which resulted in an overall accuracy of 78.2%. Then, vegetation changes, such as decline in foliage, were revealed next to the mine by low NDVI and SAVI index values. Finally, the standard classification and hyperspectral mapping approaches gave more refined results in specifying distribution of stands dominated by mature Norway spruce (Picea abies), and sapling stands of downy birch (Betula pubescens) at contaminated sites (dust and/or seepage water plumes). Image processing results and spectroradiometer measurements of the tree species indicate that the deterioration in plant vitality, especially that of Norway spruce and downy birch, can be detected as slightly increased reflectance in visible (500-700 nm) and clearly decreased reflectance in near-infrared (700-1400 nm) wavelengths. The present study indicated that the impact of the mine on vegetation was minor and the contamination effects were found only in close vicinity of the mine.

1 INTRODUCTION

The study was a part of the MINEO project aimed at developing Earth observation methods for assessing environmental impacts on the European mining sites. The goal was to investigate the feasibility of the HyMap data in vegetation classification including contaminated sites in the Lahnaslampi mine area. The boreal test site for the MINEO project, Lahnaslampi open pit mine is located in the Sotkamo municipality, Finland (Fig. 1). The annual volume of the ore and country rock is two million tons which is processed to 180000 tons of a commercial talc. In these mining operations and enrichment processes, environment has been influenced by dust emissions and seepage of water with increased amounts of elements such as Ni, As, Mg, Na and S. HyMap (Hyvista Corp., Sydney, Australia) airborne hyperspectral scanner data was acquired to assess the extent and quality of the mine’s impact on the surrounding vegetation.

Presented at the 3rd EARSel Workshop on Imaging Spectroscopy, Herrsching, 13-16 May 2003
2 DATASETS

2.1 Environmental ground data

Due to the mining operations, country rock piling and tailings, windblown mineral dust material (talc, silicates, carbonate, sulphides) was spread to the surrounding environment. Tree-ring analysis showed that growth changes in Scots pine (*Pinus silvestris*) trees have occurred from the 1960’s until the early 1980’s. However, reduced growth was observed only in pine trees less than one kilometre from the mine in a down-wind direction (NE). Increased growth was observed in Scots pine trees at 3-6 km distance, but no change was seen in trees at 6-30 km distance. Since the 1980's growth rate has been normal, hence spectral features of Scots pine most likely are not attributed to mining dust in the Lahnaslampi area. The concentrations of chemical elements in moss and humus support the results from the tree-ring analysis, such that anomalous contents of Ni and Mg were revealed less than one km in NE direction from the mine. These features most likely coincide with the dispersal of dust caused by the mining activities.

The seepage plumes were delineated by soil electrical conductivity surveys with electromagnetic induction (EM-31; effective depth 6 m), and galvanic (conductivity fork: effective depth 0.3 m) measurements. The soil electrical conductivity was classified into two intensity groups based on the chemistry of surface, soil and ground water. High conductivity anomalies were located in the immediate vicinity of the country and barren rock piles and road embankments constructed using sulphide bearing rocks. These plumes were high in concentrations of Al, As, Ca, K, Mg, Cd, Mn, Ni, S and Z. The plumes in the second category were located next to the tailing ponds and were rich in Mg, S, Ca, Na and K.

2.2 Acquisition and preprocessing of HyMap

The HyMap data acquisition was conducted in the middle of the growing season on the 28th of July 2000. The flight altitude of 2 km, ground speed of 278 km/h and a 60 degree field of view resulted in data with a spatial resolution of 5 m and a swath width of 2 km. The data has 126 10-nm-wide bands between 450 nm and 2500 nm. False colour aerial photographs were acquired simultaneously with HyMap data and later on digitised to create DEM for geometric correction procedures.

The geometric correction of the data was performed with PARGE software (PARametric GEocoding, University of Zürich, Switzerland) in order to remove distortions due to variations in flight path and altitude. A cross-track
illumination and image balancing (CTIIB) procedure was performed to remove the intensity differences between the left and right sides of the flight lines and match the intensities of adjacent flight lines. After building a mosaic of the flight stripes, a ‘halo’ effect caused by Rayleigh scattering around the mine site was removed with high pass filtering. For atmospheric correction purposes pseudo invariant features 15 m by 15 m in size were prepared, including a black and a white tarp and four natural surfaces, and measured with a spectroradiometer at the time of the flight. The atmospheric correction was performed using an empirical line correction procedure.

3 IMAGE PROCESSING AND RESULTS

Three approaches were used in the image processing of HyMap data. First, vegetation indices were scaled on the basis of soil chemistry and electrical conductivity, indicative of dust and seepage water plumes. Secondly, maximum likelihood method (ML) was used to classify tree species distribution, structure and contamination at forest stand level. Finally, seepage water and dust contaminated vegetation patterns were revealed using hyperspectral mapping procedures. Spectroradiometric measurements of ‘contaminated’ and ‘uncontaminated’ major tree and understorey species were done in laboratory to support the conclusion drawn from image processing.

A mask was created on HyMap data to include only the vegetation pixels in the image processing analysis. Infrastructure, industry, water bodies, bogs and clear-cuts were extracted from a land use database created by the National Land Survey of Finland. The mask was improved to include recent clear-cuts and changes in infrastructure by digitising extra areas based on the interpretation of aerial photographs. Vegetation indices and minimum noise fraction (MNF) transformation for the ML classification were calculated with masked CTIIB and scatter-corrected HyMap radiance data. The hyperspectral mapping was also applied to the same data but with atmospheric correction.

3.1 Vegetation indices

Vegetation indices turned out to be effective in differentiating land use and forest vegetation with varying canopy closure and plant species composition e.g. stands dominated by coniferous or deciduous trees could be separated. The scaled normalized vegetation index (NDVI, Ref. 4) proved to be the best indicator of forest stands with dense foliage. Compared to NDVI, the soil adjusted vegetation index (SAVI, Ref. 2) delineated better the recent forest disturbances (i.e. clear-cuts and site preparations). Forestry land use patterns, e.g. stand structures and ages, were best shown by the scaled SAVI. The red edge position (REP, Ref.1) highlighted contrasts between flat reflectors, such as clear-cuts and bogs, and intensive texture bearing vegetated surfaces, i.e. mature stands dominated by conifer or deciduous trees. Vegetation indices, however, were only useful in identifying stand declines or severe losses of foliage and destruction of understorey vegetation. Hence, the damaged sites next to the country rock piling, discernible also by naked eye, were identified by low NDVI and SAVI index values.

3.2 Vegetation classification using a standard classification approach

The delineation of the training sites was based primarily on a forest inventory database from the Forestry Centre of Finland. The tree species dominance (>80%) was used as criterion to select training sites. These forest compartments were subdivided into the following stand age and structural classes: sapling, young and mature stands (Tab.1). Homogenous sub-areas of the forest polygons were delineated with the aid of both automatic (Oy ArBoreal Ltd., Joensuu, Finland) and visual interpretation of digital false colour imagery and field checking. Non-forest training sites: bogs, recent clear-cuts and sedge (Carex) dominated wetlands, were also included. Forest and non-forest sites were cross-referenced with the geochemistry and geophysical data to identify training sites attributed to mining dust and seepage plumes.
Table 1. Average stand parameters for forest structural classes in the Sotkamo test site.

<table>
<thead>
<tr>
<th>Stand class</th>
<th>Age (years)</th>
<th>Height (m)</th>
<th>Basal area (m²/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sapling stand</td>
<td>19</td>
<td>3.9</td>
<td>4.7</td>
</tr>
<tr>
<td>Young stand</td>
<td>40</td>
<td>10.4</td>
<td>15.3</td>
</tr>
<tr>
<td>Mature stand</td>
<td>81</td>
<td>15.9</td>
<td>20.8</td>
</tr>
</tbody>
</table>

Signature generation and ML classification was performed using 22 selected MNF bands. The separability of signatures was studied i) by calculating the transformed divergence signature separability statistics (Ref. 10, Ref. 5), ii) visually in the MNF feature space and iii) by comparing the mean spectra of each signature. Signatures were merged and split into meaningful information classes, while maintaining the original forest structural classes. This analysis resulted in 22 signatures, out of which 9 were representative of contaminated sites. Classification was carried out using a maximum likelihood classifier (ML) with a 0.5 probability threshold value. A mean filter with a 3 by 3 pixel window was applied to the ML classification map. The classification success was tested by producing an error matrix and calculating overall, producer’s and user’s accuracies (Ref. 7). However, quantitative testing of the contamination could not be done since the pixels showing contamination were used for training purposes to meet the requirement that the number of pixels be at least ten times greater than the number of bands used (Ref. 3).

Mean spectra of each 13 ‘uncontaminated’ signatures were calculated. ‘Birch’ signatures had higher radiance compared to ‘pine’ and ‘spruce’ signatures at wavelengths from 450 nm to 2500 nm (Fig. 2a). ‘Spruce’ signatures were characterized by higher radiance in near-infrared (NIR, 700-1400 nm) range compared to ‘pine’ signatures. The effect of age on radiance is illustrated in figure 2b. As sites became more mature the spectral radiance between 450-2500 nm decreased (Fig. 2b). This phenomenon was similar also at sites dominated by pine and spruce in the Sotkamo test site.

The error matrix for the classification of vegetation is presented in Table 2. An overall classification accuracy of 52.2% (kappa 0.422) was obtained when all classes were input to the accuracy assessment (Tab. 2). The overall accuracy increased to acceptable limits i.e. 78.2% (kappa 0.579) when only the forest stands were considered. The most reliably classified forest type was ‘young/mature pine’ since it occurs only on dry growth sites (Ref. 8), typically covered by understorey with similar edaphic requirements (Ref. 6). These form a biotope with unique spectral characteristics in the Sotkamo test site. The ‘pine sapling’ sites were not stabilized (Ref. 9), but were covered by a variety of understorey species and naturally established deciduous saplings. The poor classification accuracies of ‘pine sapling’ class and confusion with the ‘bog’ sites in the classification confirms that pine had also been cultivated on sites with high soil moisture. Hence the understorey vegetation of many of these ‘pine sapling’ sites is similar to that of bogs, that include moss, shrubs and hay. For the same reason, the ‘bog’ and ‘clear-cut’ classes were in several cases incorrectly classified. Norway spruce and downy birch occur on a variety of growth sites (Ref. 8), hence the spruce and birch biotypes are more diverse in vegetation associations, and results in confusion between these classes.
The 'birch sapling' is overclassified and confused with the 'young/mature pine' class. The 'young/mature spruce' class, on the other hand, is poorly classified and being confused with the 'young birch' class. The deciduous shrubs and saplings growing as understorey in the young coniferous stands have a significant effect on the spectra but were not tall enough (>1.3 m) to be indicated in the forest inventory data. A good classification accuracy was obtained for the 'sedge wetland', which was unique for the study area (see Tab. 2).

Table 2. Error matrix (number of pixels) of forest classification derived from HyMap data over Sotkamo test site.

<table>
<thead>
<tr>
<th>Ground control Classification</th>
<th>Sedge wetland</th>
<th>Young birch</th>
<th>Young/mature pine</th>
<th>Birch sapling</th>
<th>Young/mature spruce</th>
<th>Clear-cut</th>
<th>Bog</th>
<th>Pine sapling</th>
<th>User's accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedge wetland</td>
<td>738</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>95.6</td>
</tr>
<tr>
<td>Young birch</td>
<td>0</td>
<td>95</td>
<td>9</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>75.4</td>
</tr>
<tr>
<td>Young/mature pine</td>
<td>0</td>
<td>13</td>
<td>3767</td>
<td>23</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>236</td>
<td>93.2</td>
</tr>
<tr>
<td>Birch sapling</td>
<td>3</td>
<td>17</td>
<td>305</td>
<td>216</td>
<td>92</td>
<td>2</td>
<td>0</td>
<td>61</td>
<td>31.3</td>
</tr>
<tr>
<td>Young/mature spruce</td>
<td>0</td>
<td>290</td>
<td>46</td>
<td>0</td>
<td>478</td>
<td>0</td>
<td>0</td>
<td>70</td>
<td>54.1</td>
</tr>
<tr>
<td>Clear-cut</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1915</td>
<td>304</td>
<td>0</td>
<td>0</td>
<td>86.1</td>
</tr>
<tr>
<td>Bog</td>
<td>28</td>
<td>1</td>
<td>117</td>
<td>0</td>
<td>0</td>
<td>5118</td>
<td>545</td>
<td>366</td>
<td>8.8</td>
</tr>
<tr>
<td>Pine sapling</td>
<td>1</td>
<td>2</td>
<td>76</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>228</td>
<td>72.6</td>
</tr>
</tbody>
</table>

According to signature separability statistics, signatures representing contaminated sites and having same species dominance were separable from their uncontaminated counterparts outside of the mining area. Examples of spectral signatures are presented in Fig. 2a. Some contaminated signatures, however, were spectrally similar irrespective of the dominant tree species. ‘Contaminated’ signatures were characterised by increase in the visual (450-700 nm) and decrease in NIR spectral ranges. In ML classification, two of these ‘contaminated’ signatures resulted in classes that were confined to the vicinity of the mining area (Fig. 3). The other seven contamination classes did not match with the patterns of dust emissions nor seepage water plumes.

Figure 3. Contaminated 'birch sapling' and 'mature spruce' maximum likelihood classes.
3.3 Hyperspectral mapping approach

Vegetation analysis, performed using the hyperspectral image processing approach of HyMap data, was based on endmember selection and subpixel analysis. First, pixel purity index (PPI) of 17 MNF components was calculated. The ‘pure pixels’ were overlayed on the root-zone (0-30 cm) soil electrical conductivity data. This analysis resulted in 24 contaminated endmember spectra. In addition, ‘dust contaminated’ and ‘uncontaminated’ endmembers representing Norway spruce, Scots pine and downy birch were selected. The endmember spectra were then used as an input into the mixture tuned matched filtering (MTMF) mapping method. As a result, a matched filter score and an infeasibility image for each endmember were obtained. Pixels having a high matched filter score and low infeasibility value were isolated to delineate the distribution of each class. Delineation of the contamination classes in the vicinity of the mine and then matching them with the dust and seepage water plumes was attempted. The results were verified by field investigations, aerial photograph interpretation, forest inventory data, and contamination observed through soil conductivity surveys and geochemical analysis of moss and humus.

![Figure 4. Contaminated 'birch sapling' and 'mature spruce' MTMF classes.](image)

As a result of the MTMF analysis, six endmember spectra representing seepage water and/or dust contaminated sites formed compact classes situated in the vicinity of the mine. One class represented a ‘birch sapling’ site contaminated by seepage water and three classes were combined to represent ‘birch sapling’ site contaminated by dust and seepage water. The ‘mature spruce’ site contaminated with dust and seepage water was also confined to the close surroundings of the mine. The classification result of the ‘sedge-dominated wetland’ also corresponded well with the field checking (Fig. 4).

The MTMF results, based on the dust contaminated ‘mature spruce’ endmember, showed the spruce class to be found only in the mining area. This class overlapped with that of the dust and seepage water contaminated ‘mature spruce’ because both classes had a drop in NIR reflectance and other spectral ranges were unaffected. Dust contaminated ‘pine’ and ‘birch’ classes did not match with the patterns of dust emissions.

The endmember representing ‘uncontaminated mature spruce’ resulted in a class that did not appear in the mining area. In contrast, endmembers representing ‘mature pine’ and ‘mature birch’ resulted in classes that were present also on sites where dust and/or seepage water contamination were confirmed. This suggests that Norway spruce may be more sensitive to changes in environment than downy birch or Scots pine.
4 DISCUSSION AND CONCLUSIONS

The environmental effect of the Lahnaslampi talc mine on the surrounding vegetation was mapped with hyperspectral airborne HyMap imaging spectrometer data. The image processing was based on ground measurement data of dust and seepage water plumes. Tree-ring analysis of Scots pine and geochemical analysis of moss and humus revealed a weak dust emission pattern extending less than 1 km from the mine. The soil electrical conductivity surveys and geochemical sampling of surface and pore water suggested that the extent of the seepage water plumes was limited to the close vicinity of the mine. These results showed that the environmental impact was limited only to the proximity of the mine.

Introduction of chemical elements in seepage plumes have lead to changes in vegetation. Around the tailing ponds, fertilizing effect of nutrient rich water can be seen as dense growth of deciduous species such as willow (Salix), alder (Alnus) and downy birch. In contrast, acidification and introduction of heavy metals in soil was detected visually as changes in tree vitality and decline in understorey around the country rock and sediment pilings. The response of vegetation to contamination remains somewhat unclear at the Lahnaslampi mine. Thus, more detailed studies of plant physiology and species composition are required in the future.

Three processing approaches were taken on HyMap data to reveal vegetation related features. Firstly, scaling of vegetation indices gave information related to land use and forest composition. Changes related to foliage loss and stand decline were highlighted. Secondly, a generic forest stand map was obtained using maximum likelihood classification with an overall accuracy of 78.2%. Thirdly, maximum likelihood and mixture tuned matched filtering classifications delineated two following dust and/or seepage water contaminated forest classes, ‘mature spruce’ and ‘birch sapling’, in the vicinity of the mine.

The spectral characteristics, observed in training and endmember analysis, were firstly related to the dominant tree species and stand age. Young deciduous stands had the highest reflectance in the entire HyMap spectrum. In contrast, the lowest reflectance was recorded for mature coniferous stands. In addition, spruce dominated stands were found to have higher reflectance in NIR than pine dominated. Contaminated tree and understorey species had slightly increased reflectance in visible and clearly decreased reflectance in NIR. However, accuracy assessment revealed confusion between classes that was caused by high spectral variance within a class.

In the hyperspectral mapping approach, the spectral characteristics caused by contamination are well preserved because one pixel represents an endmember. Therefore, selection of an endmember pixel has to be done carefully. In the ML approach, the spectral properties related to contamination might be lost since the training sites contain a greater number of pixels. Hyperspectral data with higher spatial resolution might provide a solution to this dilemma.

Boreal forest stands are sparse, and most of them are dominated by coniferous species. Therefore, they do not form closed canopies. A pixel spectrum representing a 5 by 5 m area is in general a mixture of well-illuminated to shadowed canopy, understorey, plant litter, soil or rock. The spectral response of contamination was relatively weak compared to other stand properties such as species composition, age, structure, genetic differences between plants, forest management operations and other factors related to high spatial variability of growth sites. Therefore, processing of hyperspectral data with 5 m spatial resolution for environmental purposes is challenging in the boreal zone.

5 REFERENCES


