

## GENERALIZATION OF HETEROGENEOUS ALPINE VEGETATION IN AIR PHOTO-BASED IMAGE CLASSIFICATION, LATNJAURE CATCHMENT, NORTHERN SWEDEN\*

LINDBLAD, K. E. M.<sup>1\*\*</sup>, NYBERG, R.<sup>2</sup> & MOLAU, U.<sup>1</sup>

<sup>1</sup> Department of Plant and Environmental Sciences, Göteborg University, Box 461, SE 405 30 Göteborg, Sweden

<sup>2</sup> Department of Physical Geography, Karlstad University, SE 651 88 Karlstad, Sweden  
[rolf.nyberg@kau.se](mailto:rolf.nyberg@kau.se)

\*\* Author for correspondence Tel.+46 (0)31 773 4804; e-mail [karin.lindblad@botany.gu.se](mailto:karin.lindblad@botany.gu.se)

*ABSTRACT.*— Mapping alpine vegetation at a meso-scale (catchment level) using remote sensing presents difficulties due to a patchy distribution and heterogeneous spectral appearance of the plant cover. We discuss issues of generalization and accuracy assessment in this case study when using a digital CIR air photo for an automatic classification of the dominant plant communities. Spectral information from an aerial photograph was supplemented by classified plant communities in field and by topographical information derived from a DEM. 150 control points were tracked in the field using a GPS. The outcome from three alternative classifications was analysed by Kappa statistics, user's and producer's accuracy. Overall accuracy did not differ between the classifications although producer's and user's accuracy for separate classes differed together with total surface (ha) and distribution. Manual accuracy assessment when recording the occurrence of the correct class within a radius of 5 meters from the control points generated an improvement of 16 % of the total accuracy. About 10 plant communities could be classified with acceptable accuracy where the chosen classification scheme determined the final outcome. If a high resolution pixel mosaic is generalized to units that match the positional accuracy of simple GPS this generalization may also influence the information content of the image.

**Key words:** Tundra, plant community, GIS, CIR-air photo, digital classification.

---

\* Recibido 2-VIII-2006. Aceptado 1-X-2006.

**RÉSUMÉ.**— *La cartographie de la végétation à moyenne échelle (bassin versant) par télédétection n'est pas aisée d'une part du fait de la structure en mosaïque complexe du tapis végétal et d'autre part l'hétérogénéité des réponses spectrales de la végétation. Cet article s'interroge sur le degré de précision et la généralisation possible à partir de photographies aériennes numérisées de classifications automatiques des communautés végétales dominantes. L'information spectrale obtenue par photographie aérienne a été contrôlée par une classification des communautés végétales sur le terrain et par l'information topographique d'un MDT. De plus, 150 points de contrôle ont été définis sur le terrain à l'aide d'un GPS. Les résultats obtenus par trois types de classification ont été analysés par la méthode statistique de Kappa, précision de l'utilisateur et du producteur. D'une manière générale, le degré de précision n'est pas différent pour les trois classifications, bien que des nuances apparaissent selon les classes, la surface totale et la distribution. La présence sur le terrain d'une classe correctement identifiée à moins de 5 mètres d'un point de contrôle augmente la précision de 16%. Une dizaine environ de communautés végétales peuvent être identifiées avec une précision acceptable en fin de classification. Si une mosaïque de pixels à très haute résolution est généralisée à des unités dont la précision est comparable à celle d'un simple GPS, cette généralisation influence la quantité d'information de l'image.*

**Mots clé:** Tundra, communautés végétales, GIS, photos aériennes CIR, classification digitale.

**RESUMEN.**— *Hemos llevado a cabo la cartografía de la vegetación alpina a escala media (nivel de cuenca experimental) mediante interpretación remota. Esta metodología plantea dificultades debido a la distribución en mosaico de la vegetación y a la heterogeneidad del espectro obtenido. Se discuten las posibilidades de generalización de los resultados y el grado de precisión alcanzado en este caso experimental mediante fotografía aérea digital CIR aplicada a una clasificación automática de las comunidades vegetales dominantes. La información espectral obtenida por foto aérea se complementó con la clasificación de las comunidades vegetales in situ y la información topográfica derivada de un Modelo Digital de Terreno. Además se marcaron 150 puntos de control en el campo por medio de GPS. Los resultados de tres clasificaciones alternativas se analizaron mediante el estadístico Kappa y la precisión del usuario y del productor. El grado de precisión obtenido apenas difirió entre clasificaciones, a pesar de que sí había diferencias significativas entre la precisión del usuario o del productor para las diferentes clases, así como para la superficie total y la distribución. La presencia sobre el terreno de una clase correctamente identificada a menos de 5 m de un punto de control, aumentó la precisión en un 16 %. Unas 10 comunidades vegetales pueden ser identificadas con un grado de precisión aceptable al terminar la clasificación. Si un mosaico de píxeles de alta resolución se generaliza a unidades cuya precisión es comparable a la de un simple GPS, tal generalización puede también influir en la cantidad de información de la imagen.*

**Palabras clave:** Tundra, comunidades vegetales, GIS, fotos aéreas, CIR, clasificación digital.

**Abbreviations:** CIR air photo (Colour Infrared air photo), DEM (Digital Elevation Model), GIS (Geographical Information System), GPS (Global Positioning System), TWI (Topographic wetness index).

## 1. Introduction

Vegetation maps are important in terrestrial ecology as soon as an area with more than one plant community is described or analyzed. Scales vary from meter to kilometre squares or larger, depending on purpose, requested resolution and available source materials (field mapping, remote sensing). In landscape ecology, the basic unit for study is often a catchment (watershed) and in alpine regions catchments are normally well-defined due to the dominant topographical relief. These catchments are also adequate units in terms of functional ecology (e. g., home ranges of predators; MOLAU *et al.*, 2003). However, mapping alpine vegetation at the intermediary scale (mesoscale), e. g. corresponding to the catchment scale, imposes some particular difficulties, including a continuous snow-melt during the growing season, temperature/elevational gradients, slopes of various inclination, and aspect (direct interception of solar radiation). Detailed studies of arctic and alpine plant communities show diverse and heterogeneous floristic patterns due to soil moisture as well as substrate, geology, hydrology and microclimate which all have a great impact on vegetation zonation and plant composition (EVANS *et al.*, 1989; BARRIO *et al.*, 1997; GRABHERR, 1997; KÖRNER, 1999). Therefore, the normally patchy and heterogeneous appearance of the alpine and arctic vegetation resulting from the harsh environment in combination with the rugged terrain, complicates mapping of plant communities by image classification and processing of spectral data (FRANK, 1988; SKIDMORE, 1989; TREITZ *et al.*, 1992; FRANKLIN & WOODCOCK, 1997). Here we present a case study from an alpine landscape in northern Swedish Lapland that addresses some of these issues.

Digital vegetation classifications using remote sensing and GIS techniques are common tools in ecological research and mapping of today (FRANK, 1988; BROWN, 1994; MOSBECH & HANSEN, 1994; GOODCHILD, 1994; GOWARD *et al.*, 1994; NILSEN *et al.*, 1999; DIRNBÖCK *et al.*, 2003).

In Sweden the use of air photos for vegetation mapping in the mountains has earlier been evaluated by IHSE (1975) and ALLARD *et al.* (1998), with the main interest focussed upon manual interpretation. Air photos were applied in the production of vegetation maps at the scale of 1:100 000 covering the entire Swedish mountain area (e. g. ANDERSSON, 1981). These maps provi-

de valuable generalized information on the vegetation, but do not show sufficient detail for studies at an intermediary scale.

At the Latnjajaure Field Station (LFS) in the Abisko mountains of northern Sweden, extensive research at a smaller plot scale on species reproduction, species diversity, plant community interactions and potential effects from Global change on the alpine flora has been conducted since 1992 (STENSTRÖM & MOLAU, 1992; STENSTRÖM & JÓNSDÓTTIR, 1997; ALATALO, 1998; MOLAU & ALATALO, 1998; STENSTRÖM, 1998; STENSTRÖM, 1999, 2000; MOLAU & LARSSON, 2000; MOLAU, 2001; LARSSON, 2002; JÄGERBRAND, 2005; JÄGERBRAND *et al.*, in press). However, at the catchment level less is known about plant community patterns, with no existing detailed vegetation map. Consequently, there is a need for an extended classification of the vegetation and distributions of various plant communities at a resolution between the small-scale plot level and the more generalized landscape level.

In July 2000, mesoscale CIR air photos were acquired over the area to allow further studies of the alpine flora in an overview perspective. In the present study, the spectral information from a digitized air photo was supplemented by topographical information derived from a DEM and a digital vegetation classification was carried out (FRANKLIN, 1995 for a review; HOERSCH *et al.*, 2002; DIRNBÖCK *et al.*, 2003; PFEFFER *et al.*, 2003).

The study highlights the significance of using alternative classification schemes and degrees of generalization based on ground reference data when dealing with heterogeneous landscapes. Furthermore we discuss some issues e. g. generalization of classes when producing signature files with the goal of creating ecologically meaningful mapping results, and the problems of dealing with positional uncertainty in image and field data when assessing the accuracy of classifications.

## 2. Site description

The study was conducted at the Latnjajaure Field Station (LFS; 68°21'N, 18°30'E, 1000 m a. s. l.) in northern Swedish Lapland (Figure 1). The lake's catchment area, Latnjavagge, approximately 12 km<sup>2</sup>, is phytogeographically regarded as subarctic-alpine tundra, but has a typically arctic climate with an annual mean temperature of -2.1 ranging from -2.9 to -1.2 °C (1993-2004) and mean annual total precipitation of 844 mm ranging from 605-1091 mm (1990-2004). July is the warmest month with a mean temperature ranging from +5.4 to +11.6 °C (1990-2005).

The valley is snow-covered most of the year. Approximately 2 km<sup>2</sup> of the area are permanent snowfields and 1 km<sup>2</sup> lakes. Of the remaining 9 km<sup>2</sup> 60 % is situated in the mid-alpine zone and 40 % in the high-alpine zone. Hence, snowbeds are a major component of the landscape, covering vast areas, and comprising unique organism communities (MOLAU *et al.*, 2003). The bedrock in Latnjavagge belongs to the upper Caledonian nappe called Köli nappe composed of mica-garnet schist and inclusions of marble on the west facing slopes. Intrusions of acidic granites can be found in the northern part of Latnjavagge. The valley is a well developed glacial trough valley. The lower part of the valley floor is dominated by Lake Latnjajaure and a series of transverse moraine ridges, 10 to 20 m high, with material of granitic origin. The crests are snowfree during the winter and many boulders are wind polished on the north-facing sides. Steep scree slopes below rockwalls are frequent along the valley sides (BEYLICH, 2003, 2004). Gelifluction characterizes the more gentle slopes, expressed as slowly moving lobes and sheets. Sorted circles are common on the valley floor (KLING, 1996).

### 3. Material and Methods

The procedure employed in this study comprised classification of vegetation in the field, rectification and classification of a CIR air photo, post processing and labelling of regions, and accuracy assessment as an intrinsic tool to adjust and evaluate the image classifications. The primary data were 1) an IR air photo in scale 1:30 000 which was scanned to a resolution of 1 meter in pixel size. The image was taken in July 2000 under favourable weather conditions and contains three bands of information: visible green (500-600 nm), visible red (600-700 nm), and near infrared (750-1000 nm); 2) a DEM from the Swedish National database with a resolution of 50 meters; 3) field data on existing plant communities. For GIS-processing and image analysis ERDAS IMAGINE v. 8.5 and v. 8.7 and Arc View v. 3.3 were used. The CIR air photo was rectified using orthophotos and the national DEM as reference.

#### *Vegetation data and field classification*

A grid-net encompassing 118 staked grid points at every 50 meters (400 x 1000 meters) was established in Latnjavagge in 1998 (MOLAU *et al.*, 2003). Five additional smaller grids (25 x 25 meters grid squares) were staked at altitudes ranging from 980 to 1400 meters altitude, in total 104 grid points (Figure 1). Dominating vascular plant species, bryophytes, lichens and approximate

cover of stones were recorded in every 10 × 10 meter square within the grids and around each stake. Relative time of snowmelt and soil moisture were estimated and categorized into three groups respectively: early, late, very late snowmelt and dry, mesic-moist and wet soil. The vegetation within the grids was classified in 10 meter (main grid) and 5 meter (smaller grids) squares according to the Scandinavian classification system by Nordiska Ministerrådet (PÅHLSSON, 1998) and used as a base for identification of training fields (Appendices 1 and 2). Nomenclature follows for vascular plants NILSSON (1991), liverworts DAMSHOLT (2002), bryophytes SÖDERSTRÖM & HEDENÅS (1998), and lichens MOBERG & HOLMÅSEN (2000).

#### *Classification procedures*

Training sets of pixels selected to represent vegetation classes in a supervised classification were identified in the CIR image with the use of ground truth data (Figure 1). The training data was checked for normality by histogram visualisation and signature files were evaluated using contingency tests with maximum likelihood and minimum distance algorithms (e. g. LILLESAND & KIEFER, 2001). Two versions, A and B, of supervised classifications were generated, including a test of fuzzy classification technique. An unsupervised classification (version C) was performed using the ISODATA-clustering algorithm (ERDAS IMAGINE Field Guide 1999). Iterations were computed with the maximum number of clusters set to 100 with 0.95 convergence threshold. New means for the clusters were calculated after each iteration, until most pixels (95 %) were permanently assigned to particular clusters during the procedure.

To allow a sufficiently detailed recovery of training areas in the image, and to minimize mixed pixel-effects in classifications, the original 1 m resolution image was used. For accuracy assessment, a generalization was made to spatial units approximately corresponding to the smallest area that could be tracked in field using a hand-held GPS. A digital elevation model of 10 m cell size was derived from the Swedish national elevation database. These elevation data were originally created from photogrammetric profile measurements or 5 m contour maps and have an accuracy of about 2.5 m in the vertical dimension. The topographic descriptors elevation, slope and index for soil water content (TWI) were thereafter computed (JENSON & DOMINIQUE, 1988; MOORE *et al.*, 1991; MOORE *et al.*, 1992; GUISAN & ZIMMERMAN, 2000).

We used classification rules according to altitudinal limits and value for TWI based on previous knowledge of vegetation distribution and on the field sampling of control points. Mountain summits somewhat resemble heath

communities at lower altitude in spectral appearance and plant composition (Lindblad, unpublished). But the high percentage of stones and boulders, thin soils, extreme wind exposure and lower temperature influence the summits to be impoverished in vascular plants but rich in lichens and bryophytes. Hence *Fellfield* (FF) was chosen as a more accurate classification for altitudes greater than 1300 m.a.s.l.

Above 1100 meters *Mesic meadow* (MM) was reclassified to *Grass heath* (HG) consisting of a thick vegetation cover mainly dominated by *Racomitrium lanuginosum*, *Festuca vivipara* and *Juncus trifidus* (Appendices 1 and 2).

Slope gradient was used to alleviate problems with confusion between shadow and water pixels on sloping ground. As shadow effects were common predominantly on steep, rocky slopes, reclassification was made to rock and stone. TWI was used to reclassify a few cases of vegetation erroneously classified as wet varieties in dry, usually high, locations. To avoid image effects of sun glint and shallow water, pixels that should be classified as lakes were added to the image using overlay from a vector layer created by on screen digitizing from the air photo. Aspect computed from the DEM was tested as complementary information to model spatial variation in altitudinal zonation but was dropped from the classifications as the number of control points within affected areas was too low to allow an evaluation of its significance.

#### *Field control data and verification*

The collection of field data for verification and improvement of the preliminary image classification was conducted in July-August 2003, at a time which corresponds well with the date of the CIR air photo in terms of vegetation development. Initially, 400 points were randomly chosen by stratified random sampling (CONGALTON, 2001) and plotted over the area. About 150 points could be tracked and reached in field using a GPS (Garmin 12). Vascular plants and bryophytes together with characteristic features were noted within a 10 m square and the vegetation around each control point was assigned to the dominant plant community. Later revision of image classes was done to capture all plant communities found in field. The grid points from the staked areas in field were used as complementary field control points (Figure 1).

Accuracy assessment was performed by retrieving dominant class within 11 x 11 pixel windows centred upon each control point, as well as directly against classifications down sampled to 10 m pixels. A further accuracy assessment was conducted by recording the occurrence of the correct class

within a radius of 5 meters from the control points and from the staked field grid points.

Overall accuracy was calculated together with omission error (indicating the probability of a reference pixel being correctly classified), commission error (indicating the probability that a pixel classified on the map actually represents the category on the ground) and Cohen's Kappa for the classifications (COHEN, 1960; ROSENFELD & FITZPATRICK-LINS, 1986; CONGALTON, 1991). Similarity assessment between the supervised classifications was conducted where version B, containing 15 classes, was defined as a reference classification zone layer and version A as the class layer for comparison (Table 3).

#### 4. Results

##### *Signature files and Overall accuracy*

In total 22 plant communities were identified in field within the plots of grid points. In the supervised classification process the 22 communities were grouped into 11 spectrally separable vegetation classes and 4 non-vegetation classes. After evaluation by contingency tests, selected signatures were merged to two alternative signature files, A) 14 classes and B) 15 classes (Table 1, Figure 4). The principal problem in the classification was how to deal with highly heterogeneous plant communities.

Plant communities found on ground with active (or relatively recently active) frost processes forming polygons were omitted. The mosaic pattern within 1-5 m<sup>2</sup> made the definition of homogeneous training fields difficult.

The snowbed plant communities generated low producer's accuracy in training data classifications of the separate classes due to spectral similarity to classes dominated by stones and boulders or by water (for wetter snowbeds) causing extensive problems of separating snowbeds from wet fen areas. In version A, the final signature file included just one snowbed/fen class incorporating waterlogged, overflowed snowbeds and wet areas (producer's accuracy 92%). For signature file B) the fen class was kept separate from one generalized snowbed class resulting in a producer's accuracy of 58% for the fen class and 92 % for the snowbed class.

The *rich patch* class (RP) is an amalgamation of diverse rich areas with high herbs. Flat areas dominated by grasses and sedges are spectrally very similar to the high herbs and difficult to separate why some grassy areas are incorporated although lacking the characteristics of the high herb plant communi-

ties (Appendices 1 and 2).

Overall accuracy was generally rather low for all three classifications with a Kappa statistic of slightly over 0.6. User and producer's accuracy for the three most frequently occurring vegetation map classes Dry heath (HD), Mesic heath (HM) and Mesic meadow (MM) (Figure 2) are presented in Table 2. Overall, producer's accuracy was on average 3-5% higher than user's accuracy and both generally higher in version A.

The manual accuracy assessment when recording the occurrence of the correct class within a radius of 5 meters from the control points generated an improvement of 16 % of the total accuracy in version A and B respectively. No improvement of total accuracy was found comparing grid points and image classes.

Figure 3 shows the relation between the number of correctly classified reference points and number of map classes found within an 11 x 11 pixel window. With higher heterogeneity (more than three classes within the window), the risk of misclassification of the control point increases.

When applying a fuzzy classification procedure on signature file version B, one alternative class assignment per pixel was used, i. e. a layer with the second best class assignment was created (ERDAS IMAGINE Field Guide 1999). Including the alternative class assignment when assessing the accuracy of the classification before topographic descriptors were applied, the overall accuracy increased about 15 %. Some of these improvements could be accomplished also by using topographic variables.

Applying a combination of spectral and topographic information (elevation, TWI and slope), primarily to adjust misclassifications due to shadow effects and terrain location resulted in an e 1. increased overall accuracy (Table 2). Further generalization of the number of classes was tested and found to produce slightly higher values for accuracy, but also ecologically less meaningful results. Such classifications are therefore not discussed here.

#### *Differences between classifications*

To highlight the significance of using alternative classification schemes and degrees of generalization, the three versions of the map classification were compared. The classified images differed in number of classes, and total accuracy of the signature files but not in overall accuracy and kappa statistics (Table 2). Version B with separate snowbed and fen classes gave a similar result as version A despite a lower overall accuracy for the signature file.

Figure 2 shows the area (ha) of each vegetation class within the three classifications, which differ substantially in several classes. The dominating vegetation classes in all three classifications are HD, HM, FF and MM. Classes HD and MM cover a 13 and 5 % larger area respectively in version A and HM has a 17 % larger area in version B. When comparing the unsupervised classification result with the supervised classifications the area of HD and HM is 39 and 12 % larger in Version A and 29 and 28% larger in version B. The unsupervised classification (version C) only comprises 11 synthesised classes analogous to HD, FF, HM, HG, MM/MR, MG, fen/water, HB, outcrop/scree slope, snow and lake (Table 1) and one manually digitized class for Patterned heath (HP). Important vegetation classes such as HSS, MR, RP are lacking. Cohen's kappa, overall accuracy and presence accuracy presented in Table 2 show that this classification only marginally differed from the two supervised classifications.

The percent agreement (similarity) and diversity of classes between the two supervised classifications are shown in Table 3. A high number for diversity means that the class in question was represented by many other classes in the second classification.

Agreement of the whole area is 83.9 percent where similarity ranges from 31-100 %. The mean number of diversity for classes is  $6 \pm 3.2$  which together with % agreement implies a considerable difference in total area and distribution among classes.

## 5. Discussion

### *Spectral similarities*

The heterogeneous characteristics of the vegetation in Latnjavagge complicates the classification of several crucial but scattered plant communities which are ecologically meaningful at a high resolution and important for determining major ecological processes at landscape level. In our case we used essentially floristic based vegetation classes following the Nordic classification system by Nordiska Ministerrådet (PÅHLSSON, 1998).

The polygon ground represents a floristically and physiologically complicated area with patterned ground resulting in a very heterogeneous spectral appearance. These areas are easy to identify using manual interpretation due to the geometric appearance but difficult to accurately capture using digital classification. Still these seasonal frost or permafrost zones are important when discussing landscape dynamics in a changing climate. Possibly texture

analysis could be applied to delimit such areas in a classification.

The snowbed communities are complicated to classify as these areas are heterogeneous in terms of time of snowmelt, stone frequency, and bryophyte and vascular plant cover. Spectrally, the communities resemble stony, water saturated, rich fens or areas overflowed by water. The attempt to classify the snowbed plant communities separately did not generate acceptable results due to the diverse and heterogeneous ground cover. In areas with scattered vegetation cover or a high percentage of stones, the soil component will be the main determinant for the spectral appearance resulting in confounding with rocky and stony areas (NILSEN *et al.*, 1999). In the heath snowbed communities the cover of mosses (mainly *Dicranum* sp. and *Kiaeria starkeii*) varies between 40-80 % and vascular plants between 16-41% depending on the day of final snowmelt (pers. comm. R. G. Björk, 2005) which may cause varying spectral appearances within the same plant community. MOSBECH & HANSEN (1994) commented that wet areas dominated by green mosses have similar reflectance curves as green luxuriant areas such as *Eriophorum* sp. so that an extensive cover of mosses will just emphasize the luxuriant fen class during vegetation monitoring.

Using a fuzzy classification procedure, it is possible to identify problematic pixels during the class assignment process, that is, pixels which spectrally are the most different from their class mean value. The next most likely class can then be included as a possible alternative when mapping those areas. We obtained some improvement in accuracy using this procedure, but confusion between several classes still remained, and the whole process of classification and vegetation mapping becomes more complex.

#### *Use of topographic descriptors*

The topographic complexity of the study site will affect the level of uncertainty in the derived topographic variables. In particular, sources of error in the primary data of DEMs may be propagated to the new data layer, hence the quality of the DEM is crucial (GUISAN & ZIMMERMAN, 2000; SCHMIDT & PERSSON, 2003; VAN NIEL *et al.*, 2004). When setting criteria for post-classification using derived variables, some field knowledge of likely zonation and altitudinal restrictions of plant communities is important as even small changes in criteria may cause substantial differences in the derived vegetation map. Recognizing that we used a DEM of rather low resolution, we nevertheless found an improvement of approximately 8-10 % in clas-

sification accuracy from the use of topographic descriptors.

*Classification accuracy assessment*

Sample size is dictated by the need of expressing accuracy in an error matrix. Previous studies have recommended a minimum of 30 samples per map class to create an adequate error matrix (STORY & CONGALTON, 1986; CONGALTON 1991). In our study the sample size did not reach this level due to a very sparse and scattered appearance of some map classes in field combined with a great loss of sampled control points due to inaccessible topography. The presented statistics of the chosen classifications must be considered as indicative rather than the "absolute truth".

LANDIS & KOCH (1977) characterised the ranges for KAPPA values into six groupings where values > 0.80 represents strong agreement, 0.61- 0.80 substantial, 0.41-0.60 moderate, 0.21-0.40 fair and values < 0.20 poor agreement. From this point of view, our kappa value of 0.62 (62 %) may be considered to show a substantial agreement with the actual pattern of the ground cover, but agreement is weaker when considering the vegetation classes alone.

However, the accuracy assessment seemed to give a conservative bias of accuracy (cf. VERBYLA & HAMMOND, 1995). Uncertainty in the precise location of control points in the classified image in combination with a high heterogeneity of the ground cover were complicating factors when comparing field data and classified data, and this may account for part of the low classification accuracy (cf. SMITH *et al.*, 2003). Imprecise matching of field data and image data can lead to poor results in accuracy assessments, even though the labelling of pixels by the classification procedure could be correct (FISHER, 1990). The pixel size usually cannot be used as the sampling unit because of our inability to accurately locate it on the ground (FISCHER, 1990; CONGALTON, 2001). However, with heterogeneous vegetation, pixel groupings present difficulties in accuracy assessment if spatial majority filters are used to capture dominant class. In our case (Figure 3) more than three classes within an 11 x 11 pixel window, i. e., a higher heterogeneity, was associated with numerous misclassifications even though the correct class was often found within the window. The accuracy increased substantially by looking at occurrence of a class within the pixel window.

*Differences between classifications*

Several studies have reported on differences in overall accuracies dependent on used techniques and the importance of reporting individual accura-

cies when evaluating the final result of an image classification (ROSENFELD & FITZPATRICK-LINS, 1986; STORY & CONGALTON, 1986; CONGALTON, 2001).

For the two supervised classifications there are only small differences in Kappa statistics suggesting that both vegetation models are equally valuable representations of reality. On the contrary, when comparing the individual producer's and user's accuracies for the three most abundant map classes (HD, HM and MM) there are differences of as much as 15% between the supervised versions and as much as 29% compared with the unsupervised version. Furthermore, the total area and distribution of the separate classes differed substantially between the two supervised classifications. When it comes to the separate classes, the results imply that, despite an overall accuracy of substantial agreement, there is a fairly large difference in the local distribution within the valley of the dominating map classes dependent on the chosen classification scheme.

The unsupervised classification generated a Kappa statistic on level with the results from the supervised classifications. One limitation worth noting when comparing it with the other versions is the reduced number of map classes. This most likely results in a somewhat false improvement of overall accuracy and Cohen's kappa compared with version A and B. Despite this, the result implies that the unsupervised method might generate a result of equal value to the more time consuming supervised work.

One particular problem when comparing areas between vegetation classes is the loss of surface in the images from steep slopes and rock-ledges in the area of interest. Preliminary calculations from the LFS show that approximately 30-40% of the actual area is lost in steeper parts. This is a problem when considering the distribution and % cover of map classes for calculations of for example potential diversity at landscape level. Using slope information from the DEM, a correction term could be applied in such calculations to obtain more true areas.

## 6. Conclusions

Using remote sensing for vegetation mapping at the community level presents a difficulty in distinguishing ecologically meaningful categories at a high resolution, which are important in determining major ecological processes at landscape level. The result of classifications of inherently heterogeneous vegetation may be highly dependent upon the classification scheme followed. To create meaningful classes from the image classification one must

decide for what purpose the final product is intended. The result may differ depending on how generalization was carried out through refinement steps, signature fusions, and use of topographical descriptors. The classified images of the study area, while being far more detailed than the existing vegetation map to the scale 1:100 000, should be considered as of a 'fuzzy' nature regarding borders between different plant communities.

When downsampling is applied to a heterogeneous data set, a coarse, generalized cellular raster structure cuts across natural vegetation boundaries, resulting in a blend of spectral information which is difficult to classify correctly. On the other hand, if a high resolution pixel mosaic must be generalized to units that match the positional accuracy of simple GPS-based control point sampling, this generalization may also to some extent influence the information content of the image.

It was possible to classify about 10 plant communities with acceptable accuracy. A single image as used in the present study does not take advantage of possible time-dependent spectral variations among the plant communities. To obtain a higher reliability in the classification of heterogeneous mountain vegetation, other modelling techniques than simple rule-based modelling could be applied, or spectrally more varied as well as multitemporal imagery and more comprehensive field data put to use.

### Acknowledgements

The authors thank the Abisko Scientific Research station and its staff for help and hospitality. This research has been founded by the following grants: The foundation of Helge Ax:son Johnson, Stockholm (K. L.), the Royal Swedish Academy of Sciences (Abisko scholarship K. L.), the Swedish Natural Science Research Council (NFR, to U. M.), grant no. B-AA/BU 08424-330 for photographing of the high-resolution infrared photo, the Royal Society of Arts and Sciences, Göteborg, and the Kempe Foundations, Örnsköldsvik, for management and administration of the Latnjajaure field station.

### References

- ALATALO, J. M. (1998). Response to simulated climatic change in an alpine and subarctic pollen-risk strategist, *Silene acaulis*. *Global Change Biology*, 3 (Suppl. 1): 74-79.
- ALLARD, A.; IHSE, M. & NORDBERG, M.-J. (1998). Vegetationsförändringar i

- fjällen-metodstudier i norra fjällen med hjälp av IRF-flygbilder och satellitbilder. Research report No. 109. University of Stockholm, Department of Physical Geography, Sweden. 47 pp. (in Swedish).
- ANDERSSON, L. (1981). *Vegetationskarta över de svenska fjällen*. Kartblad nr 2 Abisko.
- BARRIO, G. del; ALVERA, B.; PUIGDEFÀBREGAS, J. & DÍEZ, C. (1997). Response of high mountain landscape to topographic variables: Central Pyrenees. *Landscape Ecol.*, 12: 95-115.
- BEYLICH, A. A. (2003). Present morphoclimates and morphodynamics in Latnjavagge, the northern Swedish Lapland and Austadalur, east Iceland. *Jökull.*, 52: 33-54.
- BEYLICH, A. A. (2004). Intensity and spatio-temporal variability of chemical denudation in an arctic-oceanic periglacial drainage basin in northernmost Swedish Lapland. *Nord. Hydrol.*, 36 (1): 21-36.
- BROWN, D. G. (1994). Predicting vegetation types at treeline using topography and biophysical disturbance variables. *J. Veg. Sci.*, 5: 641-656.
- COHEN, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Measurement.*, 20: 37-46.
- CONGALTON, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.*, 37: 35-46.
- CONGALTON, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. *Int. J. Wildland Fire*, 10: 321-328.
- DAMSHOLT, K. (2002). *Illustrated flora of Nordic Liverworts and Hornworts*. Nordic Bryological Society, 837 pp., Lund.
- DIRNBÖCK, T.; DULLINGER, S.; GOTTFRIED, M.; GINZLER, C. & GRABHERR, G. (2003). Mapping alpine vegetation based on image analysis, topographic variables and Canonical Correspondence Analysis. *Appl. Veg. Sci.*, 6: 85-96.
- EVANS, B. M.; WALKER, D. A.; BENSON, C. S.; NORDSTRAND, E. A. & PETERSEN, G. W. (1989). Spatial interrelationships between terrain, snow distribution and vegetation patterns at an arctic foothills site in Alaska. *Holarctic Ecol.*, 12: 270-278.
- ERDAS IMAGINE Field Guide (1999). Fifth edition. ERDAS®, Inc., Atlanta, Georgia.
- FISHER, H. S. (1990). Simulating the distribution of plant communities in an alpine area. *Coenoses*, 5 (1): 37-43.
- FRANK, D. (1988). Mapping dominant vegetation communities in the Colorado Rocky Mountain Front Range with Landsat Thematic Mapper and digital terrain data. *Photogramm. Engin. Remote Sens.*, 54: 1727-1734.
- FRANKLIN, J. (1995). Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progr. Phys.*

- Geogr.*, 19: 447-499.
- FRANKLIN, J. & WOODCOCK, C. E. (1997). Multiscale vegetation data for the mountains of southern California: Spatial and categorical resolution. In: QUATTROCHI, D. A. & GOODCHILD, M. F. (eds.) *Scale in remote sensing and GIS*, pp. 141-171. CRC-Lewis Publishers, Boca Raton, FL., USA.
- GOODCHILD, M. F. (1994). Integrating GIS and remote sensing for vegetation analysis and modelling: methodological issues. In: WALSH, S. J.; DAVIS, F. W. & PEET, R. K. (eds.) *Special features in Vegetation Science 6. J. Veg. Sci.*, 5: 615-626.
- GOWARD, S. N.; HUENNRICH, K. F. & WARING, R. H. (1994). Visible-Near Infrared spectral reflectance of landscape components in Western Oregon. *Remote Sens. Environ.*, 47: 190-203.
- GRABHERR, G. (1997). Polar and alpine tundra. In: WIEGOLASKI, F. E. (ed.) *Ecosystems of the world 3*, pp. 97-121, Elsevier, Amsterdam.
- GUISAN, A. & ZIMMERMAN, N. (2000). Predictive habitat distribution models in ecology. *Ecol. Model.* 135: 147-186.
- HOERSCH, B.; BRAUN, G. & SCHMIDT, U. (2002). Relation between landform and vegetation in alpine regions of Wallis, Switzerland. A multiscale remote sensing and GIS approach. *Computers, Environment and Urban Systems*, 26: 113-139.
- IHSE, M. (1975). *Flygbildstolkning av fjällvegetation-en metodstudie för översiktlig kartering*. SNV PM 596, Sweden. 134 pp. (In Swedish).
- JENSON, S. K. & DOMINIQUE, J. O. (1988). Extracting topographic structure from digital elevation data for geomorphic information system analysis. *Photogramm. Eng. Rem. Sens.*, 54: 1593-1600.
- JÄGERBRAND, A. K. (2005). *Subarctic bryophyte ecology: phenotypic variation and responses to simulated environmental change*. Ph. D. Thesis, Botanical Institute, University of Göteborg. Sweden, 34 pp.
- JÄGERBRAND, A.; LINDBLAD, K. E. M.; BJÖRK, R. G.; ALATALO, J. M. & MOLAU, U. (2005). Bryophyte and lichen diversity under simulated environmental change compared with observed variation in unmanipulated alpine tundra. *Biodiversity and Conservation* (in press).
- KLING, J. (1996). *Sorted circles and polygons in northern Sweden*. Ph. D. thesis, University of Göteborg, Botanical Institute, Sweden, 28 pp.
- KÖRNER, C. (1999). *Alpine plant life: functional plant ecology of high mountain ecosystems*. Springer-Verlag, Berlin. 343 pp.
- LANDIS, J. R. & KOCH, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33: 159-174.
- LARSSON, E. (2002). *Seed banks and dispersal in subarctic and arctic environments*. Ph. D. Thesis, University of Göteborg, Botanical Institute, Sweden. 41 pp.
- LILLESAND, T. M. & KIEFER, R. W. (2001). *Remote sensing and image interpre-*

- tation. 4th ed. John Wiley & Sons, Inc. 704 pp. USA.
- MOBERG, R. & HOLMÅSEN, I. (2000). *Lavar: En fälthandbok*. Interpublishing, Stockholm. (In Swedish).
- MOLAU, U. (2001). Tundra plant responses to experimental and natural temperature changes. *Mem. Nat. Inst. Polar Research*, Special Issue, 54: 445-466.
- MOLAU, U. & LARSSON, E.-L. (2000). Seed rain and seed bank along an alpine altitudinal gradient. *Can. J. Botany*, 78: 728-747.
- MOLAU, U. & ALATALO, J. M. (1998). Responses of subarctic-alpine plant communities to simulated environmental change: Biodiversity of bryophytes, lichens, and vascular plants. *Ambio*, 27: 322-329.
- MOLAU, U.; KLING, J.; LINDBLAD, K.; BJÖRK, R.; DÄNHARDT, J. & LIESS, A. (2003). A GIS Assessment of Alpine Biodiversity at a Range of Scales. In: NAGY, L., GRABHERR, G., KÖRNER, C. & THOMPSON, D. B. A. (eds.) *Alpine Biodiversity in Europe. Ecological Studies* Vol. 167, pp. 221-229. Springer-Verlag, Berlin, Heidelberg.
- MOORE, I. D.; GRAYSON, R. B. & LADSON, A. R. (1991). Digital terrain modelling: a review of hydrological, geomorphological and biological applications. *Hydrol. Process*, 5: 3-30.
- MOORE, I. D.; GESSLER, P. E.; NIELSEN, G. A. & PETERSON, G. A. (1992). Soil attribute prediction, using terrain analysis. *Soil Society of American Journal*, 57: 443-452.
- MOSBECH, A. & HANSEN, B. U. (1994). Comparison of satellite imagery and infrared aerial photography as vegetation mapping methods in an arctic study area; Jameson Land, East Greenland. *Polar Res.*, 13: 139-152.
- NILSEN, L.; BROSSARD, T. & JOLY, D. (1999). Mapping plant communities in a local arctic landscape applying a scanned aerial photograph in a geographical information system. *Int. J. Remote Sensing*, 20: 463-480.
- NILSSON, Ö. (1991). *Nordisk fjällflora*. Bonniers fakta bokförlag, Stockholm, Sweden. 272 pp. (In Swedish).
- PÅHLSSON, L. (1998). *Vegetationstyper i Norden*. TemaNord 1998:510. Nordiska Ministerrådet, Köpenhamn. European vegetation types: the Nordic countries. -CD-Rom, DIVS 1998:801.
- PFEFFER, K., PEBESMA, B. J. & BURROUGH, P. A. (2003). Mapping alpine vegetation using vegetation observations and topographic attributes. *Landscape Ecol.*, 18: 759-776.
- ROSENFELD, H. G. & FITZPATRICK-LINS, K. (1986). A coefficient of agreement as a measure of thematic classification accuracy. *Photogramm. Engin. Remote Sens.*, 52: 223-227.
- SCHMIDT, F. & PERSSON, A. (2003). Comparison of DEM data capture and Topographic Wetness Indices. *Precision Agriculture*, 4: 179-192.

- SKIDMORE, K. (1989). An expert system classifies Eucalypt forest types using Thematic Mapper Data and digital terrain model. *Photogramm. Engin. Remote Sens.*, 55: 1449-1464.
- SMITH, J. H.; STEHMAN, S. V.; WICKHAM, J. D. & YANG, L. (2003). Effects of landscape characteristics on land-cover class accuracy. *Remote Sensing of Environment*, 84: 342-349.
- SÖDERSTRÖM, L. & HEDENÄS, L. (1998). Checklista över Sveriges mossor. (in swedish). *Myrinia*, 8: 58-90.
- STENSTRÖM, M. (1998). *Controls on reproductive effort and success in Saxifraga oppositifolia, an arctic-alpine perennial*. Ph. D. Thesis, University of Göteborg, Botanical Institute, Sweden.
- STENSTRÖM, A. (1999). Sexual reproductive ecology of *Carex bigelowii*, an arctic-alpine sedge. *Ecography*, 22: 305-313.
- STENSTRÖM, A. (2000). *From pollination to variation reproduction in arctic clonal plants and the effects of simulated climate change*. Ph. D. Thesis, University of Göteborg, Botanical Institute, Sweden.
- STENSTRÖM, A. & JÓNSDÓTTIR I. S. (1997). Responses of the clonal sedge, *Carex bigelowii*, to two seasons of simulated climate change. *Global Change Biol.*, 3 (Suppl. 1): 89-96.
- STENSTRÖM, M. & MOLAU, U. (1992). Reproductive biology in *Saxifraga oppositifolia*: phenology, mating systems, and reproductive success. *Arct. Alp. Res.*; 24: 337-343.
- STORY, M. & CONGALTON, R. G. (1986). Accuracy assessment: A User's perspective. *Photogramm. Engin. Remote Sens.*, 52: 397-399.
- TREITZ, P. M., HOWARTH, P. J. & SUFFLING, R. C. (1992). Application of detailed ground information to vegetation mapping with high spatial resolution digital imagery. *Remote Sens. Environ.*, 42: 65-82.
- VAN NIEL, K. P.; LAFFAN, S. W. & LEES, B. G. (2004). Effect of error in the DEM on environmental variables for predictive vegetation modelling. *J. Veg. Sci.*, 15: 747-756.
- VERBYLA, D. L. & HAMMOND, T. O. (1995). Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *Int. J. Remote Sensing*, 16: 581-587.

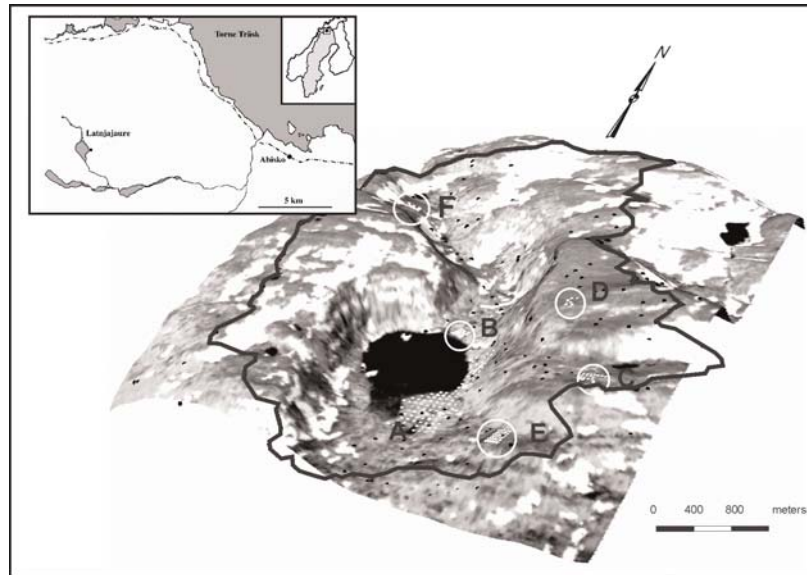


Figure 1. Site map of Latnjavagge. White dots, circles and letters A-F show permanent reference grids. Black dots represent field control points. Black line delimits the Latnjajaure catchment area.

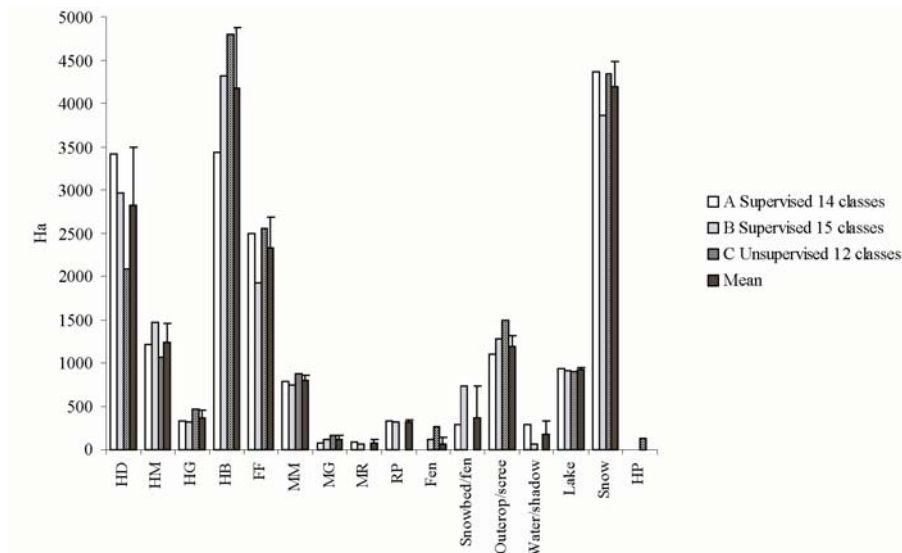


Figure 2. Cover in hectare of each map class in the three classifications (A-C) together with mean  $\pm$  SD. For abbreviations see Table 1.

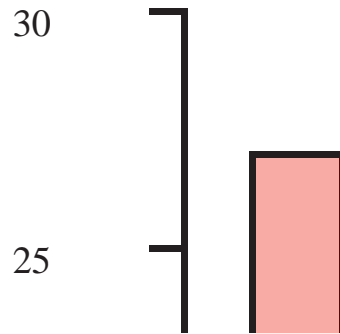


Figure 3. Number of reference points (y-axis) in relation to number of map classes within an 11 x 11 pixel window (x-axis). Filled bars represent number of correctly classified reference points. Whitebars represent number of incorrectly classified reference points.

Table 1. Classified map classes and abbreviations are presented together with producer’s accuracy for the signature file of the supervised classification A and B. For plant communities and dominant species see Appendix 1 and 2.

Class	Code	A	B
Dry heath	HD	95	81
Fellfield	FF	(Created from HD)	(Created from HD)
Mesic heath	HM	92	81
Grass heath	HG	87	84
Snowbed	HSS	-	58
Mesic meadow	MM	58	58
Grassy meadow	MG	78	62
Moist meadow	MR	88	86
Fen/HSS	Fen	99	92
Rich patch	RP	100	100
Boulder field	HB	99	86
Rock/Stone	R/S	100	99
Outcrop	O	100	100
Snow	Snow	100	100
Number of classes		3	14

Table 2. Summary information for the three classifications including post classification with the topographic descriptors elevation, slope and TWI. Producer's accuracy and user's accuracy for the three most frequently occurring map classes Dry heath (HD), Mesic heath (HM) and Mesic meadow (MM) are presented in column 5-7. Occurrence accuracy represents the percent agreement between the field classification of control points or grid points and map classes using an 11 x 11 pixel window. Last column presents the Cohen's kappa statistics.

	No of Total accuracy classes (%)		Producer's/User's accuracy						Occurrence accuracy (%)		Kappa	
	Signature file	Image file	HD	HM	MM	Control points	Grid points					
Supervised A	14	89	66	85	60	53	53	68	81	82	67	0.623
Supervised B	15	83	67	75	67	47	44	52	78	83	72	0.624
Unsupervised C	12		69	73	73	50	47	68	85	74		0.655

Table 3. Supervised classification version A is compared with version B, by using version B as reference classification zone layer and version A as the class layer for comparison. Similarity means percent agreement between pixels in version A and B for each map class. Diversity means number of map classes in the class layer for comparison falling within each map class in the reference classification zone layer. The fen class in version B was set to correspond with the snowbed/fen class in version A. For abbreviations see Table 1.

Map class	Similarity (%)	Diversity
HD	89	6
HM	77	10
HG	99	6
HB	71	8
FF	93	6
MM	92	8
MG	64	6
MR	100	2
RP	89	12
(Fen)	-	-
Snowbed/fen	31	9
Outcrop/screeslop	77	4
Water/shadow	59	4
Lake	100	1
Snow	99	2
Mean $\pm$ SD	81.5 $\pm$ 19.8	6 $\pm$ 3.2
Total agreement	83.9	

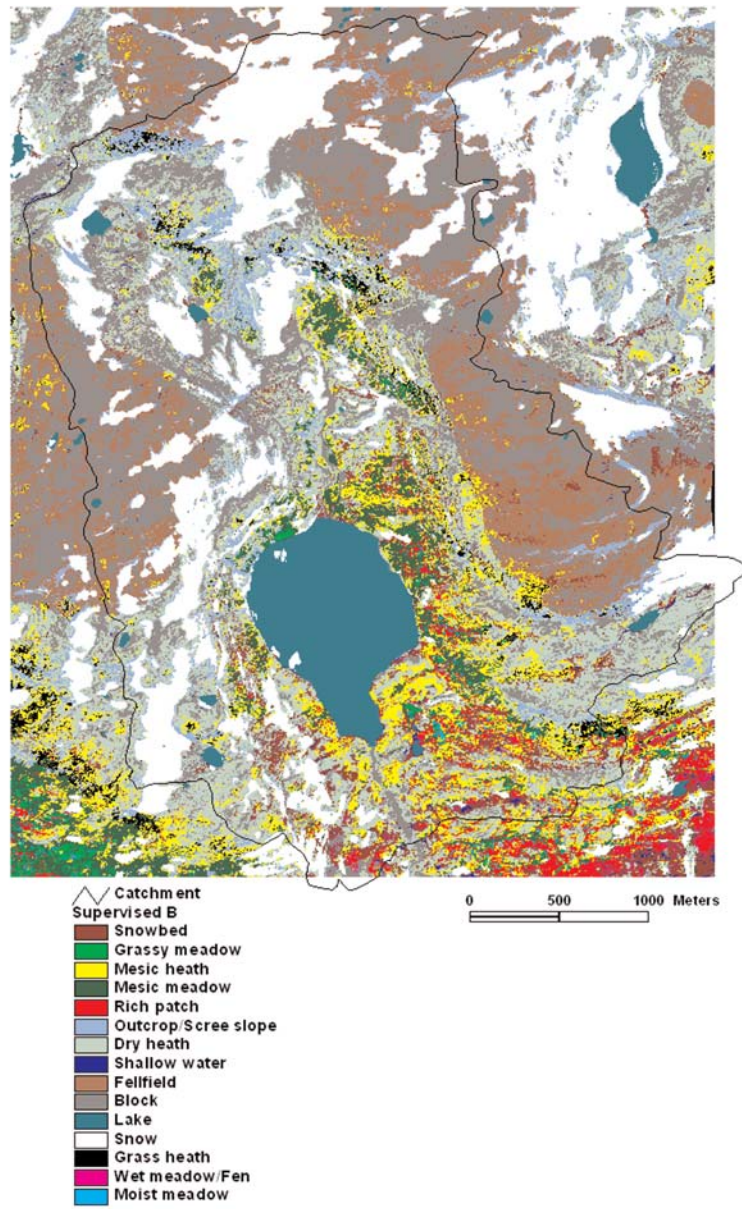


Figure 4. Vegetation map Supervised B. Generalized with the Arc View functions: Four orthogonal majority filter and Remove noise with smallest region of 5 cells to create more homogeneous areas for each map class.

## APPENDICE 1

CLASSIFIED PLANT COMMUNITIES IN FIELD, ABIOTIC CRITERIA FOR CLASSIFICATION AND ABBREVIATIONS FOR THE PLANT COMMUNITIES.

Numbers in column 1 represents the code for plant community classification given in PÅHLSSON (1998). Nutrient status: Rich and Poor (acidic). Soil moisture: Dry, mesic-most and wet. Snow protection: Low, moderate, high and approximate time for melt out. For dominant species see Appendix 1. Producer's accuracy for the signature files of the supervised classification A and B are presented under the abbreviations used for the final map classes.

Plant community		Code	A	B
1.1.1.1	Poor, acidic.	Dry	HD	HD
Loiseleuria procumbens- Arctostaphylos alpinus-Empetrum hermaphroditum type.	Dry moraine ridges Low snow protection. Wind blown.	heath	95	81
1.1.2.1	Poor.	Fellfield	FF	FF
Festuca ovina-Juncus trifidus-Cladonia spp.-type (Northern form with Deschampsia lapponica)	Dry Bare ground and blocks dominating Bryophytes >60% cover, vascular plants<10%. Low snow protection. Wind-blown.		(Created from HD)	(Create from HD)
1.2.1.5.	Poor, acidic	Mesic	HM	HM
Juncus trifidus-Salix herbacea-type	Mesic Moderate snow protection.	heath	92	81
1.1.1.2	Medium	Grass	HG	HG
Empetrum hermaphroditum- Racomitrium lanuginosum -type	Dry-mesic soil. Low snow protection. Altitude >1200 m	heath	87	84
1.3.1.1.	Poor, acidic	Heath		HSS
Cassiope hypnoides- Salix herbacea-type.	Dry, well drained. Melt out mid July.	snowbed		58
1.3.1.1.a	Poor.	Stony		
Polytrichum sexangulare-type	Wet, continuous water supply. High snow protection. Melt out late July-early August.	snowbed		
1.3.2.1.a.	Rich, calcareous.	Rich		
Distichium	Wet, continuous water	stony		

Plant community		Code	A	B
capillaceum-type.	supply. High snow protection. Melt out late July- early August.	snowbed d		
1.3.2.1. Salix polaris-type	Rich, calcareous. Moist, relatively well drained. Overflowed in springtime. Melt out late July	Meadow snowbed		
1.2.2.1. Cassiope tetragona- type.	Rich, calcareous. High snow protection. Melt out early June.	Mesic meadow	Mesic meadow MM 58	MM 58
1.2.4.1. Potentilla crantzii- Bistorta vivipara- type.	Rich, calcareous. Mesic-moist. Moderate snow protection. Melt out late June.	Mesic- moist meadow		
1.1.3.2. Dryas octopetala- type	Rich, calcareous. Dry Low snow protection. Melt out early-mid June.	Dryas meadow		
1.1.3.1. Kobresia mysurooides-Dryas octopetala -type.	Rich, calcareous scree slopes. Dry Low snow protection Very early melt out, latest mid June.	Dry scree meadow		
Grassy meadow Not dealt with in Påhlsson	Rich, continuous nutrient supply Fluvial, sandy deposits Moderate snow protection	Grassy meadow	MG 78	MG 62
3.5.2.3. Ranunculus nivalis- Paludella-type	Rich Moist-wet Good snow protection. Melt out mid-late June.	Moist meadow	MR 88	MR 86
3.3.2.5. Carex aquatilis- Drepanocladus spp.- type	Rich Mobile surface water.	Sedge fen	HSS/fen	Fen
3.3.3.2/3.3.2.4. Eriophorum scheuchzeri-Carex lachenalii- Drepanocladus spp.- type	Rich- extremely rich Mobile surface water. Moderate snow protection. Melt out late June.	Medium rich-rich fen	99	92
1.2.6.3. Trollius europaeus- type	Rich, calcareous Mobile surface water. Early melt out.	Rich patch	RP 100	RP 100

<b>Plant community</b>		<b>Code</b>	<b>A</b>	<b>B</b>
Block field	Poor, acidic		HB	HB
	Dry-mesic heath.		99	86
	Vegetation <10% cover.		HB	
Not dealt with in Pahlsson	Poor, mosaic vegetation	Patterned	HP	
	Active frost processes.	heath	-	-
	Poor fen in wet depressions and dry heath on polygon crests.			
	Moderate snow protection. Melt out mid June			
Not dealt with in Pahlsson	Poor, mosaic vegetation	Wet	HPW	
	Active frost processes	patterned	-	-
	Poor fen in wet depressions	heath		
	Tussock tundra on polygon summits. Rich in stones and boulders. Moderate snow protection. Melt out mid June.			
Not dealt with in Pahlsson	Poor	Tussock	HT	
	Mesic-moist	tundra	-	-
	Moderate snow protection. Melt out mid June.			
Rock/Shadow		R/S	100	99
Outcrop/scree slope		O/S	100	100
Snow		Snow	100	100
Lake		Lake		
Number of classes		23	13	14









