# DECISION-TREE BASED MAPPING OF FOREST PARAMETERS FROM LANDSAT DATA FOR WILDLIFE HABITAT MODELLING

Dagrun Vikhamar Schuler [a], Jan Paul Bolstad [b], Gunnar Fjone [c], Leif Kastdalen [d] and Mats Finne [e]

[a] Department of Geosciences, University of Oslo,
P.O. Box 1047, Blindern, 0316 Oslo, Norway, email: dagrunv@geo.uio.no
[b] Directorate for Nature Management, 7485 Trondheim, Norway, email: jan-paul.bolstad@dirnat.no
[c] Fjone, 3854 Nissedal, Norway, email: gunnar.fjone@forest.no
[d] Norwegian Space Centre, 0212 Oslo, Norway, email: lkastdal@online.no
[e] Utmarksavdelingen for Akershus og Østfold, P.O. Box 174, 1870 Ørje, Norway,
email: mats.finne@havass.skog.no

## ABSTRACT

Tree species, age groups, wood volume per hectare and productivity classes are mapped using decision trees and Landsat ETM+ data. Reference data consist of forest stand registrations. Accuracy assessment is carried out using national forest inventory data. Furthermore, the mapped parameters are used as input to habitat models, which are based on expert knowledge of associations between habitat variables and vertebrate species.

**Keywords:** Decision trees, Landsat, forest mapping, habitat modelling, nature management.

## 1 INTRODUCTION

Forests cover approximately 37% of the area in Norway, and within these areas the diversity of wildlife and plants depends on the coverage and the composition of the forest. According to the Norwegian Red List (Directorate for Nature Management, 1999) the main threat to biodiversity in Norway is loss of habitats due to forestry, agriculture, road construction, watercourse regulation to produce hydroelectricity and general urbanisation. Productive coniferous forests, wetlands, mires, swamp woodland and certain coastal habitats are all examples of habitats whose acreage has been reduced. The Directorate for Nature Management and the Norwegian Space Centre formed the SatNat-programme in 2001 as an important element of the National Programme for Mapping and Monitoring. The programme aims to produce decision support systems for use in the management of biodiversity.

There are two aims of this work. Firstly, to map forest parameters for large regions using multitemporal Landsat ETM+ images. Secondly, to use the forest maps as input to wildlife habitat models. The habitat models are developed based on expert knowledge of associations between habitat variables and vertebrate species.

## 2 STUDY AREA

Østfold County (4183 $km^2$) is located in southeast Norway, bordering Sweden (Fig.1). The western areas is characterized by a fragmented landscape, alternating between farming, forests and cities with up to 83 000 habitants each. The eastern parts is dominated by more continous boreal forests. The topography is quite flat, ranging from sea level to 336 m. a.s.l. About 2/3 of the County is located below the marine limit. Typically, the farming landscape is located below the marine limit in clay, silt and sandy soils, while hilly forests on subglacial till soils dominate the areas above the marine limit.

# 3 MAPPING FOREST PARAMETERS

## 3.1 DECISION TREE ANALYSIS

The classification scheme is based on decision tree methods, since these techniques can handle large and heterogeneous datasets. The dataset consists of three Landsat ETM+ images acquired within a two years period (spring 2001, summer 2001, autumn 1999), terrain parameters derived from a digital elevation model, map of vegetation zones (Moen, 1999) and a heat index map. For more information about this work, see Vikhamar et al. (2004) and Vikhamar and Fjone (2004).

Decision tree methods strongly depend on large amounts of reference data representing the classes and their spectral variability. Therefore, all available digital maps of forest stands within the test area were used to build and test the decision trees. Reference data for non-forested areas were derived from a database called 3Q. This database includes detailed maps in 1 km² quadrants in a 18 x 18 grid within the farming landscape (crop fields, pastures and impervious areas).

The mapped forest parameters are tree species, age groups, wood volume per hectare and productivity classes (Figs. 2a-d). An accuracy assessment was carried out using data from the national forest inventory (Tables I-IV). Mapped forest parameters were compared to the values registered at sample locations of 250 $m^2$. The accuracy was tested both by using the pixel center value and the 8 neighborhood pixels. Generally, the results show that the accuracy is higher on an area level rather than on pixel level. Software used for the decision tree analysis was See5.

## 3.2 RESULTS

### 3.2.1 Tree species

Four classes were mapped: 1) Spruce; 2) Pine; 3) Deciduous forest; and 4) Mixed forests (Fig. 2a and Table I). Classes were determined based on registered wood volume per hectare per tree species. A threshold of 95% was used to distinguish the classes. This means that other tree species may constitute up to 5%.

### 3.2.2 Age groups

The stage of forest development was mapped using cutting classes, registered in the forest reference data (Fig. 2b and Table II). Three age groups were mapped: 1) Clear cuts and regenerated forests; 2) Young forests; and 3) Old forest.

### 3.2.3 Wood volume per hectare (forest density)

The parameter total wood volume per hectare was used as an indicator for forest density. The wood volume represents the volume of stems of the trees. The stem volume is not necessarily correlated to the crown coverage, which is the actual observation of the satellite sensor. Three classes were distinguished: 1) Low density (1-10 $m^3$/ha); 2) Medium density (10-25 $m^3$/ha); and 3) High density (more than 25 $m^3$/ha) (Fig. 2c and Table III).

### 3.2.4 Productivity class

Productivity class is a classification system for the fertility of the soil. It describes the potential height of trees at the age of 40 years. Productivity class is usually measured from soil samples. The project carried out an experiment to test the possibility to map productivity class using Landsat images. Three classes were mapped: 1) Low productivity (0-11 meter); 2) Medium productivity (12-19 meter); 3) High productivity (over 20 meter) (Fig. 2d and Table IV).

Best results are obtained for low and medium productivity classes. Despite that satellite observations do not provide a direct measure of the productivity, we believe that satellite measurements are a good proxy. This is probably related to the vigorousness of the vegetation.

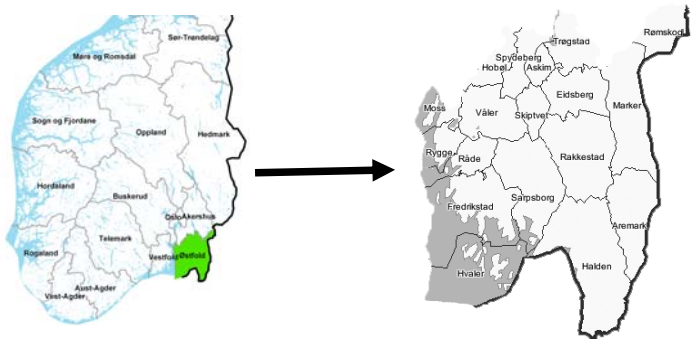



**Figure 1.** Østfold County in southeast Norway is the test area for the forest mapping and the habitat modelling.

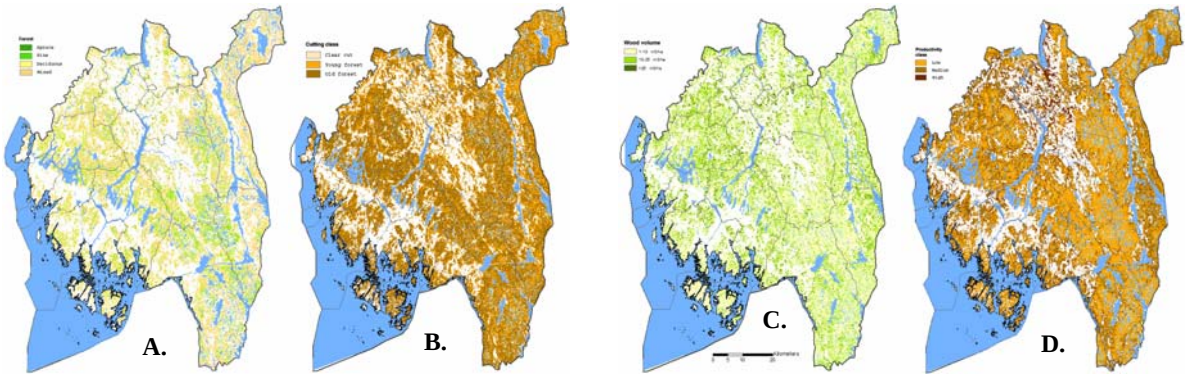



Figure 2. Resulting maps from decision tree analysis: A. Tree species; B. Age groups; C. Wood volume per hectar and D. Productivity class. White represents non-forested areas.

TABLE 1. CONTROL POINT ACCURACY FOR TREE SPECIES

| Class | Center Pixel (%) | 8 Neighboring Pixels (%) |
|---|---|---|
| Spruce[a] | 17 | 56 |
| Pine | 42 | 86 |
| Deciduous[a] | 0 | 4 |
| Mixed forest | 51 | 92 |
| Total | 44 (n=745) | 85 (n=745) |

a. Less than 100 control points

TABLE 2. CONTROL POINT ACCURACY FOR AGE GROUPS

| Class | Center Pixel (%) | 8 Neighboring Pixels (%) |
|---|---|---|
| Clear cuts[a] | 6 | 25 |
| Young forests | 54 | 90 |
| Old forest | 73 | 98 |
| Total | 63 (n=731) | 93 (n=731) |

a. Less than 100 control points

TABLE 3. CONTROL POINT ACCURACY FOR WOOD VOLUME PER HECTAR (FOREST DENSITY)

| Class | Center Pixel (%) | 8 Neighboring Pixels (%) |
|---|---|---|
| Low density | 76 | 99 |
| Medium density | 37 | 84 |
| High density[a] | 21 | 61 |
| Total | 57 (n=745) | 89 (n=745) |

a. Less than 100 control points

TABLE 4. CONTROL POINT ACCURACY FOR PRODUCTIVITY CLASS

| Class | Center Pixel (%) | 8 Neighboring Pixels (%) |
|---|---|---|
| Low productivity | 76 | 94 |
| Medium prod. | 37 | 93 |
| High prod. | 24 | 69 |
| Total | 58 (n=735) | 90 (n=735) |

# 4 PREDICTING BIRD HABITATS

## 4.1 HABITAT MODELLING

To map potential habitats we used in total 21 different land cover variables, among which the mapped forest parameters were included. For non-forest areas we used crop fields, pastures and impervious areas (e.g. roads, parking lots, buildings) with more or less than 50% green areas, provided from another study (Moen, 1999).

Different qualities of the environment tend to attract, or repel, a given wildlife species. We used thematic maps (scale 1:50 000), vegetation zones (Finne, 2005) and a 25 m DEM to discriminate environmental qualities that are used or not used by a given species (e.g. roads, water, built-up areas). Distribution data from the Norwegian Ornithological Association, the Norwegian Zoological Association, and species distribution data from the County Governor of Østfold were used to delimit the species range extent in 10 × 10 km quadrants. The species range extent was used to delimit the habitat maps for each species. Continuous habitats that intersect with the quadrants were included.

The analyses were performed employing a GIS extension in ArcView 3.2, using data from the habitat model and the species range extent. All map layers were transformed to a raster format equivalent to 30m Landsat TM pixels. Habitat variables of land cover, topographic maps, elevation and vegetation zones were related to each species in different map layers. The map layers were reduced to one map layer with the use of overlay and merge operations in the extension. The range extent and majority filter were used to reduce the distribution and to eliminate single pixels.

## 4.2 RESULTS

Predicted habitats for four avian birds are presented (Fig. 3). An accuracy assessment was performed using observational data (Table V). Counts of Black Grouse (Tetrao tetrix) and Capercaillie (Tetrao urogallus) were done by hunters with free running pointing dogs and GPS. Counts and tracks were sampled and used to analyze percent of estimated habitat and percent of counts inside a 60-meter buffer around the tracks. Data from Nightjar (Caprimulgus europaeus) and Woodlark (Lullula arborea) were sampled by drawing circles with a radius of 100 meter on topographic maps in areas were the species had been observed.

We suggested that if there were a low percent of estimated habitats inside the transect area, and the observational points covered a high percent of estimated habitats, we could conclude that the habitat model had a high degree of usability for predicting a distribution. Nightjar and Woodlark (Fig. 3) had a high percent of observational points in the estimated habitat (Tab. V). We conclude that there is a good match between the habitat model and the habitat variables described in section 2. Capercaillie (Fig. 3) had a high percent of observational points in estimated habitat and a high percent of estimated habitat inside the buffer (Tab. V). The result is not unexpected due to the high percent of estimated habitat inside the controlled area (buffer). Black Grouse (Fig. 3) had a low percent of estimated habitats in controlled areas and the observational points covered a low percent of the estimated habitat (Tab. V). The conclusion from the assessment of Capercaillie and Black Grouse is that either the habitat variables from the land cover map are insufficient due to the habitat requirements for these two species or our specification of habitat associations are to coarse. Biased sampling of observational points seems also to be a problem in this assessment.

# 5 CONCLUDING DISCUSSION

Recommendations from the accuracy assessment on four species indicate that the habitat maps are useful as a supplement to wildlife mapping based on field surveys to delimit areas of interest. Especially for species that have a good match between habitat model and land cover variables. Field-tests in some areas with nesting sites and observational points show that the estimated land cover variables forest age (cutting classes) and tree density (wood volume per hectare) are different than the real vegetation. This indicates that it will be essential to classify land cover variables more exactly for use in detailed wildlife mapping. Generally, results from land cover mapping show that the accuracy is higher on a regional level than on a pixel level. This suggests the main potential for this type of wildlife mapping is on a regional level, and as a supplement for wildlife mapping based on field surveys.

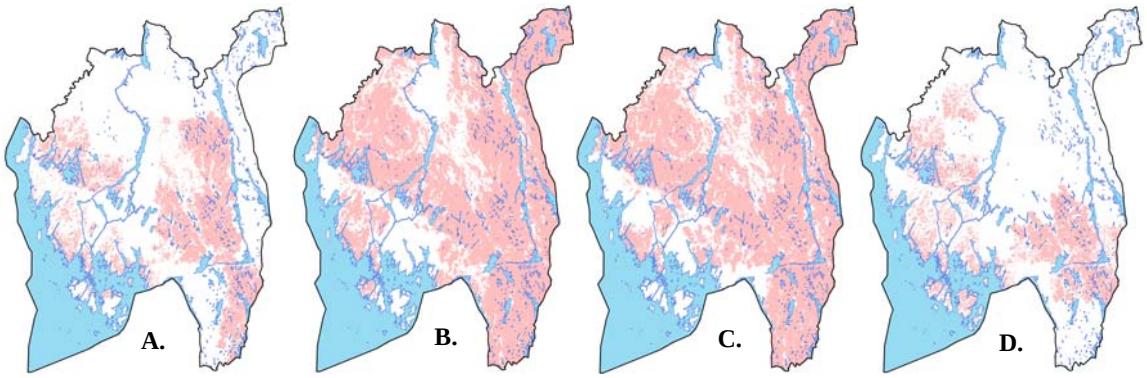



Figure 3 Predicted distributions of: A. Nightjar; B. Woodlark; C. Black Grouse and D. Capercaillie.

TABLE 5. ACCURACY ASSESSMENT BASED ON OBSERVATIONAL DATA AND ESTIMATED HABITATS FOR FOUR WOOD LIVING SPECIES

| Species | A (%) | B (%) |
|---|---|---|
| Nightjar | 39 | 85 |
| Woodlark | 37 | 76 |
| Black Grouse | 43 | 45 |
| Capercaillie | 84 | 92 |

A – Percent cover of estimated habitats in controlled areas
B – Percent cover of observational points in estimated habitats

## ACKNOWLEDGMENTS

The work has been carried out within the national program SatNat, which is funded by the Norwegian Space Centre and the Directorate for Nature Management. Local hunters in Østfold County made observational data of Capercaillie and Black Grouse available. Stein Bukholm and Arnfred Antonsen have been helpful with observational data on leks, Woodlark and Nightjar. Forest stand data were made available by the forest companies Prevista AS and Foran AS. 3Q data and National forest inventory data were provided by the Norwegian Institute of Land Inventory. We gratefully acknowledge the data providers.

## REFERENCES

Directorate for Nature Management 1999. Norwegian Red List 1998. DN-report 1999-3 (in Norwegian with English summary).

Finne, M., 2005. Test av satellittbaserte viltkart for Østfold fylke, Utmarksavdelingen for Akershus og Østfold.

Moen, A., 1999. National atlas of Norway. Vegetation. Norwegian Mapping Authority, Hønefoss, Norway.

St. meld. Nr. 42 (2000-2001). Biologisk mangfold. Sektoransvar og samordning. (Report No. 42 to the Storting (2000-2001), from the Government on Biodiversity - sector responsibility and co-ordination). Ministry of the Environment 2001.

Vikhamar, D. and Fjone, G., 2004. Mapping Forest Parameters Using Decision Trees and Multitemporal Landsat ETM+ Data, IEEE International Geoscience and Remote Sensing Symposium. Proc. of IGARSS 2004, Anchorage, USA.

Vikhamar, D., Fjone, G., Kastdalen, L. and Bolstad, J.P., 2004. Satellittdata til kartlegging av arealdekke. Utprøving av beslutningstremetodikk i Østfold fylke, Direktorat for Naturforvaltning, Trondheim, Norway.