Methods for Retrieval of Forest Parameters from Satellite Remote Sensing

Norsk Regnesentral
ANVENDT DATAFORSKNING

NOTAT/NOTE

SAMBA/26/01

Jostein Amlien
Anne Schistad Solberg

Oslo
December 2001

NR-notat/NR Note

Tittel/Title:

Methods for retrieval of forest parameters from satellite remote sensing

Dato/Date: December
År/Year: 2001
Notat nr/
Note no: SAMBA/26/01

Forfatter/Author:

Jostein Amlien
Anne Schistad Solberg

Sammendrag/Abstract:

One of the key questions in the FOREMMS project, is what kind of forest parameters satellite sensors can provide. In this note, we review previous approaches to estimation of forest parameters from existing satellite optical data, and discuss basic methodological technology and further research needs.
The most fundamental parameter to FOREMMS is the forest cover, which can be derived from a classification of forest type or land cover. This review also includes methods for estimation of forest coverage fraction, and classification of forest type from time series. Biomass and wood volume can be estimated by regression models or by kNN interpolation models. Leaf Area Index (LAI) is a key biophysical variable. Many methods are based on empirical relationship between vegetation indices and LAI. However, the most promising methods seem to be the physically founded ones that have been developed for the production of the standard MODIS LAI product. Detection of forest changes and damages involves time-series of images and separation of effects like seasonal responses of the vegetation, inter-annual variability, and directional change. Deforestation and afforestation can be detected by comparing a classified forest image to the CORINE land cover database. Biodiversity can be mapped from three parameters named 'dominance', 'contagion' and 'fractal dimension'.
We then discuss classification algorithms and algorithms for estimating a continuous parameter. For classification we suggest the Gaussian ML and Haslett's contextual MAP classifier. We also briefly discuss multi-resolution analysis, multi-temporal image classification models and data fusion models. For estimation of a continuous parameter we suggest general regression models that can use local models for spatial context, e.g. by using Markov random field models. This approach would use the same methodological framework as proposed for spatial context and ancillary data in image classification.

Emneord/Keywords: Fjernmåling / remote sensing
Skogparametre / forest parameters

Tilgjengelighet/Availability: Åpen / open

Prosjektnr./Project no.: 900500

Satsningsfelt/Research field: Fjernmåling / remote sensing

Antall sider/No. of pages: 29

Norsk Regnesentral / Norwegian Computing Center
Gaustadalléen 23, Postboks 114 Blindern, 0314 Oslo, Norway
Telefon 22 85 25 00, telefax 22 69 76 60

Contents

Preface

This report is written as a part of the FOREMMS project. Its purpose was initially to give a general documentation and description of the state-of-the-art for forest parameter retrieval from satellite remote sensing. During the process of studying literature and writing this text, the specification and goal of the FOREMMS system has been developing and changing into a system that are more focused, as it addresses s a more limited user group and a smaller parameter set. This document has been developing the same way. The number of forest parameters has been reduced according to the new focus, and the description of them has been developed.

The document has taken advantage to, but also contributed to, other documents in the FOREMMS project, in particular the D2 document (Deliverable 2, System Design) and the reports from the Parameter Working Group (PWG). We would like to thank Rune Strand Ødegård, who has written section 2.7 about biodiversity in this report, and the techniqual managers in FOREMMS, Rune Solberg and Roger Fjørtoft, for their useful comments, questions and suggestions to the text.

1. Introduction

This document aims at giving an overview of state-of-the-art methods for retrieving forest parameters from satellite remote sensing. The document will be the basis for the selection of algorithms for the FOREMMS project. The main goal in the FOREMMS project is to develop and demonstrate an advanced forest environmental and management prototype. The operational prototype shall be able to monitor the whole of Europe giving precise and coherent information on the environmental status and development of European forests.

The interesting units of data that are measured by sensors we call *parameters*. A key question in FOREMMS is what kind of parameters satellite sensors can provide. This is the main topic to be discussed in this report. The answer to this question to a large extent defines the fundament of the FOREMMS system. Some of the parameters can be provided by other sources, these are called *derived parameters*, and some parameters (called *measured parameters*) must be estimated from remote sensing data using existing and new algorithms. In order to compute all the parameters, *ancillary data* are needed for the computational process. These are already existing, and more or less static data, such as land use, previous forest maps, soil type, infrastructure etc. The FOREMMS system can be divided into two parts: a data acquisition system and a data retrieval system. In this document, only algorithms needed for the data acquisition system are discussed.

Multiresolution monitoring

FOREMMS will handle monitoring at three levels. Each level is connected to a geographical area of some typical size:

- **Level-1 areas:** Selected intensive and small-size key-biotype areas of a typical size of 20 km^2 monitored in full detail by automatic field sensors, field studies ("manual sensors") and airborne very-high-resolution remote sensing data (resolution of about 1 m). Automatic field sensors are intended to measure air, precipitation and soil variables. Airborne remote sensing data are typically collected each few years in measurement campaigns and supported by field personnel doing detailed point-location measurements.
- **Level-2 areas:** Areas monitored by fixed and random position high-resolution satellite images (resolution 20-30 m). These satellite data cover areas of typically 3,000-30,000 km^2 and include Level-1 areas. They are acquired at about the same time as or perhaps a bit more often than, the airborne remote sensing data and field campaigns. Together all Level-2 areas cover about 10% of Europe's forests. The basic source of Level-2 data will be Landsat ETM images.
- **Level-3 areas:** Areas for spatial statistical status prediction for Europe's total forested area based on medium resolution satellite data (resolution 250-1000 m), data from Level 2 areas and data from previous monitoring of the same area. The medium-resolution satellite data is acquired frequently through the vegetation season (every few weeks). The basic sources of Level-3 data will be MODIS and AVHRR images.

Organization of the document

In section 2, we review previous approaches to estimation of forest parameters from remote sensing data. We limit the scope to optical data from existing operative satellites. A pre-requisite for extraction of the forest parameters will be that the image data have been properly compensated for geometric and radiometric distortions. FOREMMS will include preprocessing algorithms for such correction, but it is not discussed in this report. Basic methodological technology and areas for further research needed to derive these parameters with sufficient accuracy is discussed in section 3.

2. Extraction of forest parameters

Most forest parameters are related to each other. A large group of parameters is connected to tree size and canopy structure (age, stem volume, crown cover, forest density, tree height, biomass), and another group of parameters is related to the chemical composition of trees and leaves/needles (chlorophyll concentration, water content, pigment concentrations, health condition, defoliation, mineral content, chemical stress). Also soil type, climatic conditions and landscape structure are related to the forest type and forest parameters. In most cases remote sensing instruments do not measure the desired forest parameter directly and parameter retrieval algorithms rely on adequate correlation between the measurable physical properties and the desired forest parameter. The relations between parameters are mostly specific to local species. This fact limits often the applicability of algorithms. These problems can be avoided to some extent by careful sensor/method calibration and validation using detailed and accurate ground truth data. Ancillary data such as digital elevation models and land use maps are often useful in improving the results.

Some forest parameters are directly measurable. Laser altimeters, scanning lidars and ranging scatterometers (radars) measure the tree height directly. Also the crown cover extent can be measured in most cases with these instruments. The backscattering coefficient measured by SAR instruments depends on ground and canopy water content and distribution. At the P-band a major contribution to the backscattering coefficient comes from the double reflection from tree trunks, which contain water. Therefore, a P-band SAR will provide information on the forest density and stem volume. At L- and C-band branch- and leaf-related reflection and scattering become more pronounced. Optical and infrared multichannel instruments will react to the forest structure and spectral properties (determined by chemical composition) of trees and leaves (health condition, content of chlorophyll).

An important factor contributing to structure parameter measurement is tree shadows. The appearance of shadows makes a clear difference between high-resolution and low-resolution images and analysis methods. Shadows add texture to high-resolution images and texture analysis can give information about tree size. On low-resolution images shadows affect only the irradiation level and texture analysis is not feasible. Shadows are most pronounced in infrared images. Another key factor forming a spectral image is the spectral properties and chemical composition of trees, underbrush and ground. Low-resolution hyperspectral instruments provide more information about mainly the chemical composition of an area and high-resolution black and white aerial photographs provide more information on the structure parameters. The instruments in between, as high resolution multispectral and hyperspectral scanners can measure both parameter groups effectively.

2.1.1. Previous projects

Reviews of the feasibility of remote sensing methods for forest parameter extraction have already been done in some European forest remote sensing projects. An extensive feasibility study was part of the FIRS (Forest Information from Remote Sensing) project (Köhl and Päivinen 1996). The main objective of FIRS (http://www.egeo.sai.jrc.it/forecast/firs/) was to contribute to the development of a unified forest information system for the entire European Continent. One of its major aims was to assist in the development and implementation of methodologies for using remotely sensed data to obtain forest information.

Forest mapping methodology was developed in FMERS (Forest Monitoring in Europe with Remote Sensing) project. The objective of FMERS (http://www.vtt.fi/aut/rs/proj/fmers/) was to demonstrate the usefulness of Earth Observation data for forest monitoring at European scale. During the project a methodology was developed to provide standardized geo-referenced and statistical information describing the forest and other wooded land in Europe based on optical and microwave remote sensing data.

The SEMEFOR pilot project (financed by the European Commission as part of the Environment and Climate program, DG XII) is aimed to demonstrate satellite based methods for the assessment of disturbances in European forests, to evaluate the cost-efficiency and to harmonise the used nomenclature.

2.1.2. Outline of this chapter

Before turning to the forest parameters a general discussion of the methodological approaches is given, where we look at classification, continuous parameter estimation, time series analysis and vegetation indices. The forest cover is fundamental to FOREMMS and it can be derived from classification of forest type or land cover. Then we discuss forest parameters like biomass (wood volum), Leaf area index (LAI), forest changes and biodiversity.

2.2. Methodological approaches

Most of the studies described in this document use general methods for the parameter estimation. The studies have tried e.g. to establish a relationship between remote sensing data and a given forest parameter using e.g. regression analysis or K-nearest neighbour interpolation. Thus, there are no detailed algorithm descriptions given, but the method class has been identified. Before turning to the parameter list, we give an introductory overview over these general methodological approaches.

The forest parameters discussed here are of two main types: categorical variables and continuous variables. Categorical variables can be obtained through a classification process, in which each pixel is attributed to one of a set of categories or classes. In this project we mainly consider supervised classification, in which case the properties of each class are known or can be deduced from ground truth. Estimators of continuous parameters are established by identifying a function that describes the relationship between the observed pixel values and the desired forest parameter.

Parameter estimation and classification can be performed directly on the observed data, but in many cases it is advantageous to compute certain features from the original image data, and then effectuate the parameter estimation on the feature vectors. The features may e.g. describe the texture of a set of neighbouring pixels. The features should be selected so as to simplify the discrimination between the different classes, and the choice of well-suited features is a key problem.

2.2.1. Image classification

The general image classification problem consists in classifying the image X into a set of K classes, $c_1 \ldots c_K$. X can be multivariate and the elements of the feature vector corresponding to each pixel may stem from several spectral bands, multitemporal acquisitions and/or multiple sensors. The task is to estimate the most probable class label image C given the observed image data X, which in a Bayesian framework means maximising the posterior probability $P(C \mid X) = P(X \mid C) P(C) / P(X)$. $P(X|C)$ is often called the image data model and represents the radiometric properties of each class. $P(C)$ is the prior model for the class labels. Markov random fields are frequently used to impose spatial regularity constraints through a prior model.

2.2.2. Estimation of a continuous parameter

Finding an estimator of a continuous forest parameter typically consists of establishing a relationship between the image data and measurements of this forest parameter. This may include physical models or statistical models. Statistical modelling relies on estimation of the relationship between the image data and the parameter using a combination of training data and prior knowledge regarding the type of relationship, e.g. linear or non-linear models with regularisation to ensure reasonably smooth functions. The prior knowledge mentioned here is generally based on knowledge about physical processes. A frequently used class of methods is regression analysis. There is a large family of well-established models for multivariate data in the class of generalised linear models in literature (see e.g. Venables and Ripley 1999). Interpolation based on the kNN algorithm is also often used for the estimation of forest parameters. Yet another alternative is to use multilayer perceptron neural networks, but this approach is essentially similar to general regression.

Holmgren *et al.* (2000) describe the weighted kNN method as a way of combining satellite image data and plot measurements of forest parameters. In the kNN method each target plot is assigned to a distance weighted average of the attribute data from the k closest reference plots in the spectral feature space. The distances between a target and a reference plot are expressed as different transforms of spectral values and/or ancillary data. A disadvantage of the method is that the lowest values are overestimated and the highest values underestimated.

2.2.3. Multitemporal and time series analysis

In some approaches it is insufficient to consider remote sensing data from one acquisition only, as the vegetation phenology is important for estimating the forest parameter. By selecting an appropriate combination of satellite images according to the phenological development, it will be easier to include the phenology into the analysis. Various methods for analysis of multitemporal data are therefore reviewed. Detection of changes will also involve analysis of time series or other ways of analysing multitemporal data.

2.2.4. Vegetation indices

One important feature in vegetation and forestry studies is the concept of vegetation index (VI). The most common index is the Normalized Difference Vegetation Index (NDVI), but other indices also exist. Vegetation and forest studies of time series will often involve the NDVI.

NDVI is derived from red (R) light in the visual part of the spectrum (600-700 nm) and near-infrared (NIR) radiation (700-1100 nm) The main idea is that green vegetation will increase the reflection in NIR, but reduce it in R. The physical rationale is that the photosynthesis absorbs red light, and that the intercellular tissue in the leaves will reflect NIR.

The NDVI is defined as NDVI = (NIR–R) / (NIR+R). The nominator (NIR–R) represents the vegetation, as it will vary according to variations in the vegetation. But it will also depend on how much solar energy that the ground receives. One important factor for variations in the solar irradiation is spatial variations in the solar incidence angle caused by the topography. The denominator (NIR+R) will be more invariant to the vegetation as NIR and R reacts the opposite way. But since it is strongly correlated to the irradiation, this term is used for normalizing the expression.

Another VI is the simple ratio, SR = NIR/R. This VI will also vary according to variations in the vegetation and be invariant to the irradiation. Generally, both indices have shown to be linearly correlated to biophysical properties, like (green) biomass. An important difference is that NDVI have an upper bound of one, while SR is unlimited. According to their definitions, we can derive NDVI from SR and opposite: NDVI = (SR–1)/(SR+1) = 1 – 2/(SR+1), and SR = 2/(1–NDVI) – 1.

One problem of NDVI and SR is the fact that they depend on the soil background if the vegetation is not dense enough to cover the soil. In order to compensate for the soil effect, other vegetation indices have been suggested. The soil adjusted VI (SAVI) is a modification of NDVI, and it is defined as: SAVI = (1 + L) (NIR - R) / (NIR +R + L), where the adjustment factor L accounts for the spectrally different extinction of the radiation from the soil through the canopy (Huete 1988). Also note that R and NIR represent red and near-infrared reflectance, not absolute radiance reflected.

Variations in the VI may also be induced by the atmosphere, but can be minimized in the soil and atmospherically resistant VI (SARVI) by considering the difference between blue and red reflectance values: SARVI2 = 2.5 (NIR - R) / (1 + NIR + 6R - 7.5B) (Huete *et al.* 1997).

Simulations for different soils backgrounds (sand, silt, and clay) have shown that NDVI values increase with moist soil backgrounds, while Tasselled Cap greenness index (GVI) values are less influenced by soil background variations (Todd and Hoffer 1998). For all soil backgrounds the Tasselled Cap wetness index (WI) increased as green vegetation cover increased.

2.3. Forest cover and type

2.3.1. Introduction

Land cover classification schemes involve identification of forest cover as well as different forest types. Thus, forest cover and forest type can be derived from a land cover classification. Alternatively, algorithms for classification of forest type can be obtained by simplifying an existing land cover classification method. This section will therefore describe algorithms for derivation both of forest type, forest cover and land cover.

Forest cover

The most important parameter in the FOREMMS system is the forest cover, in the sense that FOREMMS will mainly consider forested areas only. In order to obtain efficient processing, it is important to avoid unnecessary processing of unforested areas, especially when the algorithms include heavy computer processing. For unforested areas, some parameters will be meaningless, like tree height. Though many parameters may be said to have a meaning outside the forests, like biomass, the algorithms will not be calibrated for such regions. Only a few parameters need to consider unforested areas, for instance reforestation. We will therefore include a fixed forest mask in the system, and we need to define this mask.

In order to monitor changes in forest cover, it may be necessary to apply Level 3 data, and derive the fraction of forest cover and compare to former values of forest coverage or a coverage derived from the finer forest mask. Therefore some algorithms for forest coverage will be reviewed.

Forest type and tree species

The main factor in separating different forest types is the tree species composition. Therefore, tree species classification or identification will be relevant for classification of forest type.

In some cases, it will be of interest to use the composition of the species in the forest. It may therefore be a need to conduct a sub-pixel analysis of the pixels in order to estimate the fractional coverage of each species or groups of species.

Multi-temporal approaches

It is often the case that a given set of tree species cannot be separated at a given time (or based on a single image), but that a multi-temporal data set describing how the spectral signature of the tree species varies according to its phenology. Vegetation phenology means the development of the vegetation during the growing season. The phenology is closely linked to the tree species and forest type, and the best time to identify a certain tree species depends on the actual forest type involved, the species, and also the geographical location.

The temporal data can be merged at a data-fusion level or a decision-fusion level. When the data is fused at decision level, selected features that are relevant for forest type is selected according to some carefully selected phenological model, (see e.g. Reed *et al.* 1994). Wolter *et al.* (1995) noted that a common multitemporal approach for forest type classification of large areas is difficult because the phenological models will depend on the tree species, forest types, and geographical areas involved.

Concatenating the multi-temporal images into one image is a more simple approach to multi-date forest classification; (see e.g. Mickelson *et al.* 1998). The multitemporal information is treated just like different spectral bands in a multivariate classification algorithm. This approach does not require a full phenological model, but it assumes that the actual images are taken at appropriate times. Therefore, phenological knowledge must be used in the selecting the appropriate images given the forest type and the geographical location.

2.3.2. Methodology

Forest cover

The CORINE landcover database (CEC 1993) covers large areas of Europe, and it classifies the area into several land cover classes, among them several forest classes. CORINE is a high-resolution database based on SPOT data (Champeaux *et al.* 2000). The ESA forest map (ESA 1992) may be used when CORINE land cover doesn't exit. If existing forest masks are not available or reliable, the forest mask can be derived from a forest type map. Forest/non-forest discrimination can also be derived by thresholding a biomass map (Häme *et al.* 1997).

In a vegetation mapping over Western Europe it was observed that clustering on NDVI time series was not able to separate forests from grasslands (Champeaux *et al.* 2000). The method was improved by focusing on the low reflectance in the visible part of the spectrum. This effect was most prominent in summertime for northern Europe and in the spring and autumn in southern Europe. The forests were identified by thresholding the visible reflectances in monthly composites over three years, where the thresholds were determined from reference maps.

Forest cover fraction

The percentage of forest cover is often estimated by means of the NDVI vegetation index derived from NOAA AVHRR. In order to improve the estimation, other spectral bands were tested and compared to NDVI in a study of temperate coniferous forests in the Cascade Range of Oregon (Boyd and Ripple 1997). The MIR and TIR spectral bands of AVHRR were individually not strongly related to percentage forest cover, but when they were included in vegetation indices, an improvement in the derivation of forest cover was obtained. Indices including visible, NIR and TIR were found to be the most suitable ones. In general channel 3 (MIR) did not cause further improvements of the indices. The exception was the complex division index, C_3 / (C_1 C_2 C_4 C_5), which was able to separate among four different forest success ional stages present in addition to the improvement of forest cover derivation.

The forest density for the conterminous USA, defined as the percentage of forest covers for larger areas, was mapped from AVHRR data by means of a mixed pixel model and linear regression between the AVHRR data and classified Landsat TM data covering limited areas (Zhu and Evans 1994; Zhu 1994). The TM data were classified into forest and non-forest, aggregated into a forest density value for each AVHRR cell and then regressed to the spectral data of the AVHRR. It was necessary to develop separate models for each of the 15 physiogeographic region defined in the study. Forest information derived from sub-pixel measurements was also found to aid forest type classification.

Global continuous fields of the fraction of woody vegetation were derived from 8 km AVHRR Pathfinder Land data by means of a linear mixing model (DeFries *et al.* 2000). The model input were linear discriminants derived from 30 different metrics of the annual phenological cycle, using training data derived from a global network of Landsat scenes. In general the results from different years were consistent, except in high latitudes where variations in snow cover occurred, and there were also apparent problems with artefacts in the multi-year data set. The agreement between the derived data and the reference data was highest when the data was averaged over many years. Change detection will require more refinement and improved inclusion of end-members.

Forest type

Two previous studies at NR have investigated classification of tree types and cutting classes based on multi-spectral aerial images with a spatial resolution of 0.5 m (Aas *et al.* 1996; Aas *et al.* 1997). Feature extraction methods were used to extract spectral and textural information from local windows in the scene. The classification of tree species and cutting classes was much more robust when features computed from windows of size 25x25 m were used compared to classifying single pixels.

In another study at NR the performance of a multi-parameter SAR sensor was compared to an airborne spectrometer (Volden *et al.* 1998). The hyper-spectral data combined with a model for spatial context gave good results for tree-species classification (95% correct). A simulated comparison between SPOT, Landsat TM and photographic film showed that hyper-spectral images give much better results. These two studies also propose a method for incorporating previous forest maps or stand borders in the classification. This improved the classification performance significantly.

Multi-temporal approaches

An image consisting of 18 channels was composed from the six reflective bands of three Landsat TM images during spring, summer and autumn, and the forest types classified with an overall accuracy of 79% (Mickelson *et al.* 1998). The classification result was a genus level forest classification with 20 classes characterizing dominant canopy species and 13 subclasses characterizing the under-story vegetation.

Vegetation phenology is important for vegetation monitoring (Reed *et al.* 1994). NDVI time series were derived from AVHRR, and 12 different metrics of phenological variability and key events computed from the series. The metrics included the onset of greenness, time of peak NDVI, maximum NDVI, rate of greenup, rate of senescence, and integrated NDVI. There was a strong coincidence between the metrics and the predicted phenological characteristics. The metrics established the phenological consistency of deciduous and coniferous forests.

The technique of applying classification trees on multitemporal data is a good alternative to more traditional methods for land cover mapping, since such trees are hierarchical and nonlinear they will be suited to handling non-parametric training data as well as categorical or missing data (Hansen *et al.* 1996). They compared the performance of a tree to a ML classifier using a global data set, and found that the accuracy of the two methods was comparable (82%). Finally, they pointed out that a tree also could be used to reduce the dimensionality of the data sets and to find those metrics that are most useful for discriminating among cover types. In a later work Hansen *et al.* (2000) applied a hierarchical global land cover classification scheme, which uses a set of 41 multitemporal metrics that is created from a set of monthly AVHRR composites. The metrics will typically be maximum, mean or minimum of a spectral channel or an index during the four warmest months ("summer") or the 8 warmest months ("annual"). The forests are identified at the third level in the decision tree. The metrics being applied to identify the forest cover include NDVI annual peak, NDVI summer mean, red and NIR reflectance annual minimum, MIR radiance summer mean, red reflectance annual mean. The various forest types are included at next levels in the decision tree. Mixed forests are identified by annual mean of channel 4 (TIR) and maximum NDVI. Needle-leaf forests are identified by maximum NDVI in the temperate zone. Temperate broad-leaved forests are identified by the minimum of MIR and the amplitude of NDVI.

Clustering methods on NDVI time series was developed for mapping the land cover of the USA (Loveland *et al.* 1991). Each NDVI set is composed from AVHRR data covering a 10-day or one month period. The selection of the composition procedure will influence the accuracies when the NDVI composites are applied of for forest/non-forest classification (Roy 1997), and the recommended procedure is to use the maximum value of NDVI in the composition period for each pixel (Holben 1986).

A hierarchical unsupervised approach was applied on Iberia (Lobo *et al.* 1997), because of its high environmental diversity and very strong environmental gradients. A time series of 12 monthly NDVI values was derived from AVHRR, and the first two principal components were determined. Unsupervised classification by means of agglomerative clustering of 3000 randomly sampled pixels yielded a dendrogram, from which three classification levels was identified. For each level, the classified samples were used as training data for determining linear canonical discriminant functions and corresponding class signatures in terms of centroids and covariance matrices. All pixels were first transformed from temporal space to canonical space, and then for each three levels assigned to a class using the maximum likelihood method with a 95% probability threshold. Finally the temporal NDVI spectrum for each class was calculated as the average of the assigned pixels. The highest classification level identified the distinction between Atlantic and Sub-Mediterranean vegetation, being characterised by a summer and a spring peak of NDVI, respectively. The lower levels in the hierarchical classification gave maps with a high degree of spatial continuity, and it was verified that the result was bio-climatically coherent. It was demonstrated that the NDVI time series is an accurate signal of vegetative phenology, which is not obtained from global maps of land cover. The ecological information at finer scales were found to be relevant and with detailed legends, and therefore the method could be suitable for regional scale applications.

2.3.3. Discussion and conclusion

Forest cover

Forest cover is related to forest type/tree species as it can be derived from a classification of an image into different forest types or tree species (if the classification includes all forest types/tree species found in the area).

Reasonably good estimates of forest cover are possible to find from remote sensing data. The studies discussed above derived forest cover from vegetation indices, but it should also be possible to derive from a direct classification of the multi-spectral image data. It is likely that multi-temporal data that includes signature development during the season is needed to achieve acceptable performance. The multi-temporal data can be incorporated by using two different approaches, either to develop feature extraction methods that derive the essential information from the time series, or to input the time series model to the classification module and develop means of fusing this information in the decision process.

Forest cover fraction

The studies reviewed above are related to AVHRR images and use subpixel methods to estimate the forest density for AVHRR image pixels. These studies used spectral mixing models, which is one class of methods that can be applied to subpixel analysis. Other approaches may include statistical mixed pixel models (Solberg *et al.* 1995).

For medium resolution data, sub-pixel methods might not be required if one is interested in forest density estimates on a scale coarser than the image resolution.

Forest type

It should be possible to derive tree species (and thus derive information about forest type) from satellite images with accuracy high enough to make it interesting for the FOREMMS application. Classification methods should be used for this. It might be necessary to apply feature extraction to compute textural features from high and medium resolution images, and also to use models for spatial context in the classification process.

Different tree species develop at different times during the season. As discussed for forest cover it is probably an advantage to develop classification models which can utilize multitemporal data, e.g. in terms of monthly image data to include phenology in the model. Information about the phenology can be derived from e.g. time series of NDVI or other vegetation indices during the season.

2.4. Biomass

2.4.1. Introduction

Biomass is defined as the total amount of dry organic matter (t/ha). It will include also biomass below ground, but sometimes only biomass above the ground is considered. In this report, we define total biomass as the total biomass above and below ground (roots), while total biomass above ground will be specified when used.

The total biomass above ground in a forest is composed of tree biomass above ground and understorey biomass. Tree biomass above ground can be calculated as the sum of stem wood, stem bark, living branches and needles biomass (Fazakas *et al.* 1999). The organic matter of a tree including bark, branches and roots can be computed from existing models based on singletree measurements of diameter (dbh) and height (Häme *et al.* 1997).

Wood volume (stem volume or tree volume) is one of the most important variables in traditional forestry resource assessments (Fazakas *et al.* 1999). Since it is a volume measure (m^3/ha), not a mass measure, it is not a measure of biomass. However, it is very closely related to biomass, and connections may be established as linear relationships. Stem volume refers to the volume of the stems only, and since it is measured on ground in practical forestry (Häme *et al.* 1997), we must assume that the reference data in FOREMMS will be given in terms of stem volume.

The primary production is related to the amount of green biomass, which is a small fraction of the total biomass of a tree. Green biomass is often estimated from remote sensing data by means of some vegetation index, for instance the NDVI. Green biomass is also related to the leaf area, which is measured by the leaf area index (LAI), which is discussed in another section in this report.

Strong relationships between the total biomass and the green biomass are often the case, but it is dependent on factors like tree species and age. NDVI has shown to be effective in estimating green biomass and LAI of non-wooden canopies, but it has been found to be a very poor indicator of forest biomass (Häme *et al.* 1997), since the majority of the forest biomass is non-green. The primary controls on the reflectance during the development and lifetime of the trees are other factors than the tree biomass because of saturation at an early development stage.

The most common methods found in the literature for estimation of biomass from remote sensing data are based on regression functions, but the 'k-nearest-neighbour' (kNN) approach is also used. The following section gives a review of selected papers describing methods for estimation of biomass.

2.4.2. Methodology

The relationships between forest parameters and TM data were examined for a lodgepole pine forest in Yellowstone National Park (Jakubauskas and Price 1997). The analysis was carreid out by means of stepwise multiple regression analysis, which results in an optimal set of variables that explain the dependent variable and where the remaining (excluded) variables are not able to improve the explanatory power significantly. When only the original spectral bands were examined, TM7 were found to explain 58% of the variance in total biomass. No significant improvement was obtained by introducing any other spectral bands. When transformed TM data were included in the analysis, NDVI were found as the best variable, but only a very small improvement was obtained as NDVI explained 59% of the total biomass.

Large areas of conifer-dominated Boreal forests have limited ground reference data. In a semi-physical two-step approach developed for boreal forest biomass estimation from AVHRR data (Häme *et al.* 1997), TM data was applied as a link between the ground data and AVHRR data. In the first step tree volumes were modelled by linear regression models of Landsat TM spectral data (R and NIR). Broadleaved trees have a 1.5-2 times higher NIR reflectance compared to conifers, and the volumes of coniferous and broadleaved trees were therefore modelled separately. The proportion of the total tree volume related to broadleaved trees was also estimated directly by linear regression models, partly because this proportion was considered as a fundamental variable of Boreal forests itself, and partly as an alternative for estimating the biomass of broadleaved trees. The next main step in the model estimation was the transformation of the corresponding spectral data in AVHRR into this spectral model. Two different methods were given for the estimation: a regression model and a mean/variance based approach. Finally, in order to estimate total biomass instead of stem volume, the relationship between these two closely related parameters was determined from a regression between outputs from known estimation models where plot measurement data is input. The results showed that the red channel was the best single channel, and that NDVI was not appropriate for biomass estimation. The method applies only for large areas and works best on boreal forest zone.

The 'k-nearest-neighbour' (kNN) method was applied to combine reflective TM data with NFI (National Forest Inventory) plot data to estimate tree biomass and wood volume for a forest estate (60°N, 17°E) in central Sweden (Fazakas *et al.* 1999). The biomass for each pixel is predicted as weighted averages of the biomass values of the spectrally closest reference plots. The chosen weights were proportional to the inverse squared distances in the spectral domain. They considered Mahalanobis distance as more appropriate to use than the Euclidean distance, but did not find any significant difference. The number of spectral neighbours is recommended not be higher that 5. The results were evaluated both by means of cross-validation, where each plot in turn is tested against all the other plots, and by means of plots from an intensively sampled validation area. The results showed that though the estimates on a plot level had a large variation, the variation for the entire estate (510 ha) was much lower (RMSE was 8.7% for biomass and 4.6% for wood volume). The accuracy increased for aggregations of cells, and the method gave satisfactory estimates (RMSE <10%) for mean wood volume for areas > 1 km^2. The RMSE for mean wood volume and mean biomass decreases as the unit area increases. As the relationship between wood volume and digital numbers is non-linear, the method generates biased estimates if the spectral distances between reference plots are too long, but the method is asymptotically unbiased as the number of available reference plots increases. However, when historical stand characteristics were used as co-variates in an experiment where timber volumes were estimated in radiata pine plantations in New Zealand, the results indicated that the estimates were unbiased both for plots and stands, though with high RMSEs (Tomppo *et al.* 1999). In an experiment utilizing Swedish National Forest (NFI) inventory plots, the site index, forest age, and mean tree height from forest stand data were introduced as features in the 'spectral' domain (Holmgren *et al.* 2000), and the standard error of stem volume was reduced from 36% to 17%.

2.4.3. Discussion and conclusion

It is desirable in FOREMMS to get estimates of biomass or stem volume at a continuous scale. Two categories of methods for estimation of biomass are found in the literature: regression-based methods like linear regression, and the kNN interpolation method. It is likely that more general regression models can perform better than linear regression. If continuous scale cannot be done robustly, it might be possible to classify the scene into areas with biomass levels in certain categories.

Several of the studies have found that more reliable estimates have been found for entire forest stands than for single pixels. Models that can utilize ancillary information in terms of stand border maps and forest characteristics for the stands are thus desirable. It has also been found that the estimates will be more reliable for larger cells that for smaller ones.

The reported studies have been performed on limited test areas, and how the performance of a regression model for regions outside the area the model was derived for is an open question, which has to be addressed for operational applications. One solution will be to determine separate regression models for each forest type.

This review shows that both regression models and kNN interpolation models may be utilized in the estimation of biomass and tree volume. A decision of which of the two methods that we will select for biomass estimation in FOREMMS cannot be deduced from the review. The final choice of method for the baseline of FOREMMS will depend on a synoptic view of all parameters that will derive from remote sensing in FOREMMS.

2.5. Leaf area index and FPAR

2.5.1. Introduction

Leaf Area Index (LAI) is defined as the one sided green leaf area per unit ground area in broadleaf canopies, or as the projected needleleaf area per unit ground area in needle canopies. One related parameter is FPAR, which means the Fraction of Photosynthetically Active Radiation absorbed by vegetation canopies (see the Boston University website http://cybele.bu.edu/modismisr/laifpar/laifpar.html)

LAI is a key biophysical variable influencing land surface photosynthesis, transpiration, and energy balance. Strong relationships between LAI and NDVI or SRVI have been found in coniferous forests in the western USA over large LAI ranges (Bonan 1995). Both LAI and FPAR are key variables in most ecosystem productivity models, and in global models of climate, hydrology, biogeochemistry and ecology. Global LAI surfaces were an early product of the MODIS Land Science Team (Turner *et al.* 1999), and methods based on radiative transfer models have been developed for that purpose.

2.5.2. Methodology

A simple way of estimating LAI is by means of regression of spectral data or some derived vegetation indices (VIs). Nilson *et al.* (1999) studied coniferous forest in central Sweden and found decreasing non-linear trends between several different TM bands and LAI, with relative errors about 30-40%. Studies of individual stands or well-defined forest types have revealed better correlation results. Baynes and Dunn (1997) studied 8-year-old pine stands and found that LAI could be modelled by strong inverse non-linear correlation functions of TM band 5

(r^2 = 0.91) and of band 3 (r^2= 0.71). LAI were also modelled by positive linear correlation with NDVI (r^2= 0.81) and SR (r^2= 0.78).

Jakubauskas and Price (1997) analysed the relationships between LAI and TM data in a lodgepole pine forest in Yellowstone, USA. Both spectral data and derived VIs were analysed by means of stepwise multiple regression analysis. When considering the spectral bands only, TM5 were found to explain 62% of the variance in LAI, and with no significant improvement from the other spectral bands. When both types of TM data were considered, LAI was best explained (R^2=66%) by a combination of NDVI and the greenness-component of a scene-specific Tasselled Cap transformation (see Crist and Cicone 1984).

Different VIs (NDVI, SR, SAVI) were derived from TM images and compared with LAI measured in field at three sites within the temperate zone (Turner *et al.* 1999). The general finding was a strong general relationship for low values of LAI, as the VIs were increasing up to LAI values of 3-5. For LAI values above 5 the sensitivity of the VIs was low, and they even decreased at the highest LAI values in the coniferous forests. This decrease is explained by the reduced NIR reflectance in the complex canopies in mature forest stands. Atmospheric correction improved the relationship between LAI and VI, but topographic corrections had little or no effect. A serious problem is that forest type and other vegetation properties that are independent of LAI, will have significant effects on the VIs. None of the investigated VIs were found to be generally better than the others, but it was suggested to select between the various VIs according to forest type and successional stage.

LAI is commonly estimated using NDVI, and other VIs based on VNIR radiation. However, it has been demonstrated that LAI is more closely related to middle-infrared (MIR) radiation than to visible light, and that MIR can be applied to improve the correspondence between the VI and LAI in boreal forests. Brown *et al.* (2000) suggested a MIR based modification of the simple ratio, the reduced simple ratio, RSR = (NIR/R) ($MIR-MIR_{min}$) / ($MIR_{max}-MIR_{min}$). In a study of boreal forests of Canada they found that RSR showed increased sensitivity to LAI and was reducing background effects in conifer canopies. Furthermore, RSR had the potential to unify deciduous and conifer species in LAI retrieval, which has impact when information about forest type is missing or where the forest types are mixed within a pixel. These results are verified by the findings of Boyd *et al.* (2000), who compared NDVI to a vegetation index based on both NIR and MIR radiation, VI3, for a boreal forest canopy. The relationship with LAI was stronger than for NDVI, and the variation in LAI was better explained. However, in a study of LAI in a larger landscape scale in Glacier National Park, USA, NDVI was found to be the best index for estimating LAI, though the accuracy decreased with coarser pixels (White *et al.* 1997). The MIR corrected SR was found to overestimate the LAI because of difficulties in deriving the appropriate reflectance scale of the MIR correction, but it was a good indicator of understory canopy cover.

The MODIS instrument is designed to provide global imagery at one single viewing angle and seven short-wave spectral bands (visible, NIR and MIR) , while the MISR (multi-angle imaging spectroradiometer) instrument is designed to provide global imagery at nine discrete viewing angles and four VNIR spectral hands. Algorithms for retrieval of LAI and FPAR fields both from MISR alone (Knyazikhin *et al.* 1998a) and from a combination of MISR and MODIS (Knyazikhin *et al.* 1998b) are applied in the production of the Terra standard product of LAI and FPAR. Algorithms for deriving LAI / FPAR from MODIS reflectance data are found at http://modarch.gsfc.nasa.gov/MODIS/ATBD/atbd_mod15.pdf (see also Knyazikhin *et al.* 1999). The algorithms are independent of any particular canopy reflectance model, and utilise information of the canopy spectral properties and structural attributes. They depend on a structural land cover classification into one of six global biomes, where broadleaf forest and needle forest are the two relevant for FOREMMS.

The algorithm for producing global LAI and FPAR fields from the combination of MODIS and MISR canopy reflectance data (Knyazikhin *et al.* 1998b) is a synergistic algorithm based on a 3-D formulation of the radiative transfer process in vegetation canopies. The radiative transfer in plant canopies have some specific features, which have made it possible to split the complicated 3-D radiative transfer problem into two independent and simpler sub-problems, and store their solutions in a look-up-table LUT. A land cover classification map is required in order to identify what biome the ground cover belongs to, and misclassification between distinct biomes can fatally impact the quality of the retrieval, but the impact of misclassification between spectrally similar biomes is negligible (Tian *et al*. 2000). The algorithm is dependent on the spatial resolution of the data, and therefore the algorithm must be adjusted for data resolution in order to be utilized with data from other sensors.

The algorithm for retrieval of LAI and FPAR from MISR data only (Knyazikhin *et al*. 1998a), is a two-step process that utilizes all the information provided by atmospherically corrected MISR data. The fist step is based on a set of models, which represent the biome type, canopy structure, and soil/understory reflectances. Candidate models are identified according to their correspondence with the retrieved spectral hemispherically integrated reflectance. The second step identifies the best one of the candidates. Each of the candidate models are evaluated by means of their correspondence to the retrieved spectral directional reflectances at each of observed the MISR angles. The most probable values of LAI and FPAR are specified by means of measure theory using the set of all acceptable solutions. The algorithm was tested with bidirectional reflectance data over Africa from the POLDER instrument (Zhang *et al*. 2000). The test results demonstrated some advantages of using multi-angle data, as they can:

1) decrease the dispersion and saturation of LAI,
   and increase the quality of retrieved LAI and FPAR,
2) improve the accuracy of LAI retrievals in geometrically complex canopies such as shrubs,
3) help determine biome (land cover) types correctly
   by using the minimum value of LAI dispersion.

The algorithms for deriving LAI from MODIS reflectance data (Knyazikhin *et al*. 1999) also have a backup solution based on NDVI. For each of the six biomes, LAI and FPAR have been modelled as non-linear functions of NDVI. The functions are given in terms of LUTs and are listed in table 2.3 in http://modarch.gsfc.nasa.gov/MODIS/ATBD/atbd_mod15.pdf.

2.5.3. Discussion and conclusion

In FOREMMS, large areas covering a wide variety of forest types are going to be mapped. Solutions based on regression with single bands or with VIs, will require large amounts of training data in order to be representative for the whole area. However, VI based methods that include MIR may be possible solutions.

The full algorithms for MODIS will be very complicated to implement, and the results can also be obtained as standard products from MODIS. There is therefore no reason to implement those algorithms into FOREMMS unless we want to improve them or adapt them to other sensors. Instead FOREMMS could include the LAI/FPAR product from MODIS.

For FOREMMS it therefore seems that the best solution for a robust baseline algorithm will be to implement the MODIS backup solution algorithm, and if possible refine the LUTs that are being used. An advantage of this solution is that NDVI can be derived from any multispectral satellite image.

2.6. Forest changes

2.6.1. Introduction

Multi-temporal images are essential for detection of forest changes. In addition to change detection, such images are also applied for deriving information related to phenological phenomena, and for instance estimation of growth during a season.

For forest change detection from time series we must distinguish between seasonal responses of the vegetation, inter-annual variability, and directional change (Coppin and Bauer 1994). Directional change may be caused by intrinsic vegetation processes, land use or other human-induced processes (e.g. pollution stress), and alterations in global climatic patterns. For a given forest monitoring system, its ability to detect changes will depend on its ability to account for variability at different scales. Coppin (1994) discusses four different types of signature changes: no change (closed canopy remained closed), canopy depletion (net overall canopy loss), recent storm damage (structural/textural canopy changes), and canopy increment (process of canopy closure).

A simple approach to change detection is to compare two classified images, for instance two forest cover maps for detection of deforestation and afforestation. Other approaches may include thresholds, linear transformations, classifications, regression models and change vector analysis (Häme *et al.* 1998).

The two main variables for forest growth are Gross primary production and aboveground primary production (NPAA). Some important aspects of NDVI concerning growth were reviewed by Bonan (1995). The NDVI is related to FPAR, which is important for the photo-synthesis rate. The seasonal cycle of NDVI is correlated to the seasonal uptake and release of CO_2. The annual integrated NDVI is related to annual net primary production (NPP).

2.6.2. Methodology

The most common approach to change detection is to study changes of individual pixels. Varjo (1996) studies changes on the stand level based on the assumption that at least part of the changes follow stand borders, especially man-made changes. They study changes as standwise differences of the three first central moments for individual Landsat TM channels using a nonparametric kernel method.

Time series can be applied to detect e.g. defoliation by using well-established change detection techniques for optical sensors, like vegetation indexes or principal component analysis (Collins and Woodcock 1996). Changes are detected using a linear model relating the spectral signature at two different dates. The most common linear change detection approach is multidate principal component analysis. These methods are easier to use if the images are from the same sensor, and if they are well calibrated. Comparisons of linear change detection techniques for forest applications are given by Collins and Woodcock (1996) and Muchoney and Haach (1994). Nielsen *et al* (2001) presented a method for change detection based on multivariate alteration detection and maximum autocorrelation factors.

Häme *et al.* (1998) presented an unsupervised change detection method that is based on clustering of two images. The detection and identification of the changes are performed in two separate clustering procedures. In the first phase the two images are clustered separately. In the next phase, the clusters in the latter image are reclustered into finer classes, which are compared to the first image. A detailed description of the method is given by Häme *et al.* (1998). When the method was tested in southern Finnish Boreal forest using TM data, it could reliably detect and identify clearcuts. The method also provided information on forest damage.

Cohen *et al* (1998) compared two unsupervised classification methods for mapping forest harvest activity from a temporal dataset consisting of five sequential date-pairs of difference images derived from Landsat. The merged image differencing method was merging the classification results from the five separate difference image pairs into a single map of forest harvest activity. The simultaneous image differencing method applied one single unsuper-vised classification of the full sequential difference image data set. Both methods mapped the clearcuts with more than 90% accuracy.

Sohn *et al* (1999) applied a new spectral pattern matching approach that utilizes the spectral angle concept for mapping deforestation and successional forest regrowth. Forest regrowth stages were mapped by assigning spectral clusters to reference classes, based on the minimum spectral angle rule. By applying the spectral pattern matching approach, spectral clusters can be assigned into information classes in an objective way. The conceptual difference between the spectral distance and spectral angle in feature space is also reviewed.

Varjo and Folving (1997) monitored rapid forest changes, like cuttings, in large areas of boreal forests by utilizing an unsupervised clustering method and multitemporal radio-metrically calibrated Landsat TM data. In order to obtain a change detection of acceptable accuracy, the classification units must be the forest stands instead of the pixels.

Sachs *et al* (1998) used Landsat TM and MSS imagery to map forest cover and detect major disturbances between 1975 and 1992 for a large area of interior British Columbia. Changes in landscape pattern were examined by describing both conifer and harvested patches in each biogeoclimatic zone by means of calculated indices like interior area, perimeter/area ratio, and patch-shape complexity.

Multi-temporal TM data showed that the ratios of band 5 to band 4 increased as longleaf pine were subjected to a one year prolonged drought (Pinder and McLeod 1999). Mean ratios increased significantly (from 0.42 to 0.55) during the drought, but the amount of the increase varied among the forest stands. Ratios returned to pre-drought levels once the drought was broken.

Royle and Lathrop (1997) studied the defoliation of an eastern hemlock forest in the New Jersey Highlands during a 10-year period by comparing the simple ratio VI from Landsat TM data. A regression model relating estimates of canopy condition to the temporal difference in the VI was developed to predict hemlock condition across the study area. The VI difference was highly correlated to hemlock damage as measured on the ground ($r^2 = 0.73$).

Forest growth is closely related to the carbon exchange. Bonan (1995) discussed remote sensing applications for modelling of the seasonal and annual carbon balance in terrestrial ecosystems. The applications will take advantage of models that combine the biophysical and biogeochemical controls of CO_2 exchange. The annual production of biomass and the seasonal cycle of CO_2 exchange in boreal forests have been well approximated by such a model, where the required input is the beginning and end of the growing season, absorbed photosynthetically active radiation, foliage nitrogen concentration, and vegetation type

In order to estimate Gross primary production (P-G) and aboveground net primary production (NPPA), Coops (1999) applied the 3PGS model using NDVI from AVHRR and Landsat MSS. The method is based on a monthly time-step model (Physiological Principles Predicting Growth using Satellite data (3PGS)), which requires monthly weather data, soil texture, and rooting depth. It will also require the fraction of photosynthetically active radiation absorbed by the forest canopies (fPAR), which is estimated from a satellite-derived NDVI. AVHRR and Landsat MSS data were used by the model, yielding 3PGS predictions at a more refined land-scape scale. Field tests resulted in a linear relation between predicted and measured wood production ($r^2 = 0.4$).

2.6.3. Discussion and conclusion

The time issue concerning detection of harvesting / deforestation and regrowth / afforstation is very different. The former is happening fast and need to be detected fast, while the latter is slow processes, and can therefore be detected by more laborious and not so frequent methods.

Forest harvesting and deforestation seems to be detectable by clustering methods within the forest mask. Afforestation and regrowth will occur outside the forest mask and should be detected by a forest classification of the whole FOREMMS area.

Forest health should be monitored by studying time series of NDVI or other forest parameters. Serious changes in forest health may also be detected by a change in forest class, or by the same clustering methods that will be used for detecting harvesting and deforestation.

As the primary production is dependent on the incoming solar energy, FPAR seems to be a relevant parameter for determining the primary production. The total amount of absorbed PAR (photosyntetically active radiation) is given by FPAR and the incoming PAR, which can be determined from meteorological models and observations.

2.7. Biodiversity

2.7.1. Introduction

The ecological interpretation of landscape patterns is one of the major objectives of landscape ecology. These landscape patterns needs to be quantified to establish relationships to ecological processes and to detect changes.

Wickham *et al* (1995) estimated 3 diversity estimates in the USA from AVHRR satellite data: land cover richness, vegetation richness and vegetation clustering.

The definition of a system of nomenclature for mapping European Forest, (Köhl and Päivinen 1996) suggested 3 parameters called 'dominance', 'contagion' and 'fractal dimension' (Dale *et al.* 1995; O'Neill *et al.* 1988), which is related to the spatial arrangement of patches. In the suggested nomenclature the spatial arrangement of patches is suggested as indicator attributes to describe the potential of forested areas to fulfil the functions described by the attributes 'Threats to Species Diversity', 'Wildlife Habitats', 'Scenic Beauty' and 'Environmental Impact' (FIRS project - key parameter study).

2.7.2. Patchiness / patch attributes

Dominance is the complement to evenness, provides a measure of how common one land cover is over the landscape. Its value indicates the degree of which species dependant on a single habitat can pervade in an area.

Contagion measures the degree to which land cover units are clumped or aggregated. Contagion is a useful metric for those species that require a large contagious area of particular forest type or land cover.

Fractal dimension is an indicator for the complexity of spatial patterns. This indicator can be used to select areas that are suitable for species that inhabit at edges or require multiple habitats. Can also be used to select areas that are suitable for species that inhabitat large contagious areas.

The GOFS (Global observation of forest cover) project has also identified a product related to forest change called Forest Fragmentation Product. This type of parameter could be implemented in FOREMMS on different levels.

A whole range of indices has been suggested to quantify the spatial arrangement of patches. In FOREMMS the most logical approach is to calculate indices based on the forest mask or available land cover databases (Level 2 data) with a large filtering window (e.g. 5km x 5km).

The FRAGSTATS program (http://www.fsl.orst.edu/lter/data/software/fragstat.htm) was developed to quantify landscape structure. FRAGSTATS offers a comprehensive choice of landscape metrics and was designed to be as versatile as possible. Two separate versions of FRAGSTATS exist: one for vector images and one for raster images. Both versions of FRAGSTATS generate the same array of metrics, including a variety of area metrics, patch density, size and variability metrics, edge metrics, shape metrics, core area metrics, diversity metrics, and contagion and interspersion metrics. The raster version also computes several nearest neighbour metrics.

3. Discussion of algorithms

As previously mentioned, the literature for extraction of forest parameters reviewed is mainly based on using general classes of method to estimate the parameter, and not specially dedicated algorithms. Two main classes of algorithms are needed: classification algorithms and algorithms for estimating a general parameter.

3.1. Methods for image classification

As the default method for image classification, we propose to use the ML classifier with a multivariate Gaussian density function. For optical images, the Gaussian data model is generally very well suited. Besides the elementary pixelwise ML classifier, a contextual version (e.g. Haslett's algorithm) will be considered. Texture parameters computed on a neighbourhood could be included in the feature vector. One alternative to ML classification is to introduce Markov random fields or Markov chains as a prior model and perform maximum *a posteriori* (MAP) or maximum posterior marginals (MPM) classification. However, the computational cost is considerable and it is difficult to estimate the spatial regularity parameters of such prior models from a ground truth of limited size. Another alternative is to use non-parametric methods such as neural networks or clustering techniques, e.g. the k-nearest neighbours (kNN) algorithm. In this case, no precise data model is needed. This is a considerable advantage in many applications, but in our case the Gaussian model is well justified and multivariate ML classification seems to be the best methodological starting point.

3.1.1. Multi-resolution analysis

The multilevel analysis in FOREMMS will consist of estimating a parameter (e.g. a forest parameter) at Level 1 based on data from one sensor (or more), estimation of the same parameter at Level 2 based on data from a second sensor, and possibly at Level 3 based on data from a third sensor. Furthermore, a model describing the relationship between the parameter at Level 1, Level 2 and Level 3 must be established.

Only a limited set of studies have used multiresolution data to predict forest parameters. The classification accuracy of low-resolution images can be estimated by using high-resolution images as ground reference information. Kloditz *et al* (1998) carried out a statistical analysis on simulated low-resolution data derived from TM data, which resulted in two models for combining the information from the aggregated TM data in order to predict NDVI from NOAA data.

Multi-resolution pyramids for segmentation of a single image

For image segmentation purposes, a popular approach is to create a pyramid representation of the original, single-sensor image by sampling it at different spatial resolutions. In some cases, this can speed up and/or simplify the segmentation problem. An example is given by Wang and Liu (1999). Multi-resolution Markov random fields (MRMRF) are often used to model the observed data at the different spatial resolutions and how this is related to the scene labels which one seeks to estimate. The use of a MRMRF to model data at different spatial scales could be relevant to FOREMMS.

Multi-resolution pyramids for estimation of statistical parameter processes

Image pyramid models or quad-tree representations are also useful to establish relationships when the original image data are measured at different resolutions. One example of such a model is given by Luettgen and Willsky (1995). A pyramid model describes the observed data or the underlying stochastic process that one seeks to estimate at different spatial resolutions, and how the different resolutions are related.

Multiple scale modelling in terms of Kalman filtering is discussed for soil moisture prediction by Kumar (1999). Here, scale replaces the role time normally has in Kalman filtering. Instead of predicting new observations at a new time, observations at a finer scale are predicted. We will study such models in FOREMMS.

3.1.2. Multitemporal image classification models

We will also need classification models that incorporate multitemporal information from time series of images, e.g. to be able to use phenology. A principal decision in multi-temporal image analysis is whether the images are to be combined on the data-fusion level or the decision-fusion level (Jeon and Landgrebe 1999). Data-level fusion consists or combining the multi-temporal images into a joint data set and performing the classification based on all data at the same time. In decision-level fusion, a classification is first performed for each time, and then the individual decisions are combined to reach a consensus decision. If no spectral-signature changes (affecting the information classes) have occurred between the image acquisitions, this is very similar to classifier combination (Benediktsson and Kanellopoulos 1999). A few studies have considered decision-level fusion in a model allowing class changes (Jeon and Landgrebe 1999; Raviv 1967; Solberg *et al.* 1996; Swain 1978).

One possibility might be to start by using the methodological framework presented by Solberg *et al.* (1996). In this model, Markov chains are used to model the time dimension, and a model for legal transitions between development states can be built. As an example, consider a model of NDVI during the growing season (first leaf, max leaf etc.). Different tree species or forest types will be modelled using different transition times and states.

3.1.3. Models for fusion of ancillary and remote sensing data

Most of the studies for forest parameter extraction discussed in section 3 use only remote sensing images. The performance of the algorithms is not perfect, there is a need of methodological improvements to achieve better accuracy. One possible source for algorithm improvement is to combine the image data with ancillary data in term of e.g. general land cover maps or other map information from CORINNE or other sources. There are generally no standard methods for combining map information with image information, but in previous studies we have tried to establish a methodological framework for this purpose (Solberg *et al.* 1996). A brief review and discussion of multisource image classification is given below.
A prerequisite for multisource fusion is that the image data are co-registered. Image fusion can either be applied to visualize the combined image data set, or to use the combined image data set to classify or segment the scene into ground-cover classes. We will only consider the latter approach.

The main approaches to multi-sensor data fusion found in the remote sensing literature are statistical methods (Benediktsson and Swain 1989; Solberg *et al.* 1996), Dempster-Shafer evidence theory (Lee *et al.* 1987; LeHegaratMascle *et al.* 1997) and neural networks (Benediktsson *et al.* 1996). The most common approach to multi-sensor classification is to concatenate the data into one vector and treat it as if it were a set of measurements from a single sensor. Such a classifier is difficult to use when the data cannot be modelled with a common probability density function, or when the data set includes ancillary data, e.g., from a GIS system.

Many multi-sensor studies have used neural nets because no specific assumptions about the underlying probability densities are needed. A drawback of neural nets in this respect is that they act like a black box in that the user cannot control how the different data sources are used. Another drawback is that specifying a neural network architecture involves specifying a large number of parameters.

Hybrid approaches combining statistical methods and neural networks for data fusion have also been proposed. Benediktsson and Sveinsson (1997) apply a statistical model to each individual source, and use neural nets to reach a consensus decision. Most applications involving a neural net use a multi-layer perception, but other neural network architectures can be used.

The fusion of the images can take place at different levels, the pixel, feature, or decision level (Solberg *et al.* 1996). Pixel-level fusion consists of merging information from different images on a pixel-by-pixel basis to create a new multisensor image, which is then input to a segmentation/ interpretation process. In feature-level fusion, a set of features from each image is merged and used for further interpretation. In decision-level fusion, the single-sensor images are input to single-sensor segmentation / classification algorithms, and then the computed probabilities for each ground cover type are combined to reach a consensus classification of the scene. The latter approach might be most useful in FOREMMS, where the sensors used gives images with very different spectral and spatial characteristics. We would also recommend to use the mathematical framework presented in (Solberg *et al.* 1996) as the basis for further improvements regarding the use of multisource data in FOREMMS. This framework includes means of combining image data with existing map information and very simple models for class changes or development.

3.2. Estimating continuous parameters

Two main approaches are suggested in the literature: kNN and linear regression. Unless the amount of training data is huge, kNN is likely to be very little robust. Regression-based approaches are much more robust. The performance of regression-based approached can probably be improved by using more general regression models than linear models. This is a well-established field of work in the statistics literature, and other classes of regression models should be tried, e.g. generalized linear models (Venables and Ripley 1999). Further research in FOREMMS should compare the performance of both the simple regression models, the kNN approach, and more general regression models for a given data set and forest parameter.

These parameter estimation methods are designed for single locations, and they do not consider the spatial neighbourhood of a location. There are many models for spatial statistics in the statistics literature, (see e.g. Høst *et al.* 2001). However, operating on a global level, many of them will be computationally intense, compared to operating on a local neighbourhood. With computational aspects in mind, we suggest that FOREMMS can use local models for spatial context, e.g. by using Markov random field models. This would make us able to use the same methodological framework as proposed for spatial context and ancillary data in image classification.

3.3. Baseline methodology for FOREMMS

A limited set of baseline algorithms have been implemented in the FOREMMS prototype:

- Pixelvise classification has been implemented as maximum likelihood (ML) classification using Gaussian distribution models.
- Contextual classification has been implemented by Haslett's maximum *a posteriori* (MAP) algorithm. The algorithm considers the contextual contribution to the

posterior probability in addition to the *a priori* probability and the likelihood. The context is derived from a local neighbourhood around each pixel.

- Estimation of continuous variables by means of non-linear regression has been implemented by means of neural network (NN) algorithms.
- The kNN interpolation method will not be imple mented in the prototype.
- Estimation of the patchiness attributes has been implemented.
- Extraction of extract forest parameters by means of time series analysis of NDVI values is not implemented in the prototype.
- Extraction of NDVI data for storage in the database is not yet implemented.

3.3.1. Suggested improvements

After the implementation of baseline algorithms into the FOREMMS prototype, we suggest the following options for further development after the prototype version, according to the discussions given in this note.

Classific ation of forest cover is most relevant for change detection compared to existing databses like CORINE land cover. It must be undertaken on Level 3 data in order to ensure results within a reasonable time range. Therefore we suggest to devellop a forest cover fraction algorithm for Level 3 data.

Derivation of forest type from from Level 3 data should take advantange of the frequent acquistions of those data, and utilize the phenological information in such data. Such algorithms often utilize NDVI data, and therefore derivation and storage of NDVI should be implemented in order to be able use these data later. We have suggested to apply a Markov chain based algorithm.

Extraction of biomass is implemented by non-linear regression in a neural network (NN) algorithm. We suggest to evlauate this solution carefully before turning to the other option, implementing a kNN interpolation algorithm. Our suggestion for improvement and delvelopment of continuous paramteres in general, is to use local models of spatial context and implement Markov random field models.

The baseline algorithms do not include LAI/FPAR, but we suggest that the Modis backup algorithm is implemented, as it is given as a simple LUT function of the NDVI and main biome (forest type). In addition, FOREMMS should include the LAI/FPAR PRODUCT from MODIS.

4. References

Baynes, J. and Dunn, G.M. (1997). “Estimating foliage surface area index in 8-year-old stands of Pinus elliottii var. elliottii x Pinus caribaea var. hondurensis of variable quality” *Canadian journal of forest research - Revue canadienne de recherche forestier*, **27** (9): 1367-1375.

Benediktsson, J.A. and Swain, P. (1989). "A method of statistical multisource classification with a mechanism to weight the influence of the data sources." *International Geoscience and Remote Sensing Symposium 1989, IGARSS'89, July 1989, Vancouver, Canada*, pp. 517-520.

Benediktsson, J.A., Sveinsson, J.R. and Ersoy, O.K. (1996). "Optimized combination of neural networks" *IEEE International Symposium on Circuits and Systems (ISCAS'96), May 1996, Atlanta, Georgia*, pp. 535-538.

Benediktsson, J.A. and Sveinsson, J.R. (1997). “Feature extraction for multisource data classification with artificial neural networks” *International journal of remote sensing*, **18** (4): 727-740.

Benediktsson, J.A. and Kanellopoulos, I. (1999). “Classification of multisource and hyperspectral data based on decision fusion” *IEEE Transactions on Geoscience and Remote Sensing*, **37** (3): 1367-1377.

Bonan, G.B. (1995). “Land atmosphere interactions for climate system models - Coupling biophysical, biogeochemical, and ecosystem dynamical processes” *Remote sensing of environment*, **51** (1): 57-73.

Boyd, D.S. and Ripple, W.J. (1997). “Potential vegetation indices for determining global forest cover” *International journal of remote sensing*, **18** (6): 1395-1401.

Boyd, D.S., Wicks, T.E. and Curran, P.J. (2000). “Use of middle infrared radiation to estimate the leaf area index of a boreal forest” *Tree physiology*, **20** (11): 755-760.

Brown, L., Chen, J.M., Leblanc, S.G. and Cihlar, J. (2000). “A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis” *Remote sensing of environment*, **71** (1): 16-25.

CEC (1993) "CORINE land cover, technical guide" *Office for official publications of the European Communities*, Report no. EUR 1285 EN.

Champeaux, J.L., Arcos, D., Bazile, E., Giard, D., Goutorbe, J.P., Habets, F., Noilhan, J. and Roujean, J.L. (2000). “AVHRR-derived vegetation mapping over Western Europe for use in numerical weather prediction models” *International journal of remote sensing*, **21** (6-7): 1183-1199.

Cohen, W.B., Fiorella, M., Gray, J., Helmer, E. and Anderson, K. (1998). “An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery” *Photogrammetric engineering and remote sensing*, **64** (4): 293-300.

Collins, J.B. and Woodcock, C.E. (1996). “An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data” *Remote sensing of environment*, **56** (1): 66-77.

Coops, N. (1999). “Improvement in predicting stand growth of Pinus radiata (D. Don) across landscapes using NOAA AVHRR and Landsat MSS imagery combined with a forest growth process model (3-PGS)” *Photogrammetric engineering and remote sensing*, **65** (10): 1149-1156.

Coppin, P.R. and Bauer, M.E. (1994). “Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features” *IEEE Transactions on Geoscience and Remote Sensing*, **32** (4): 918-927.

Crist, E.P. and Cicone, R.C. (1984). “A physically based transformation on Thematic Mapper data - The TM tasseled Cap” *IEEE Transactions on Geoscience and Remote Sensing*, **GE-22** (3): 256-263.

Dale, V.H., Offerman, H., Frohn, R. and Gardner, R.H. (1995). "Landscape characterization and biodiversity research" *Measuring and montoring biodiversity in tropical and temperate forests, IUFRO symposium, Chiang Mai, Thailand*, pp. 47-63.

DeFries, R.S., Hansen, M.C. and Townshend, J.R.G. (2000). “Global continuous fields of vegetation characteristics: a linear mixture model applied to multi-year 8 km AVHRR data” *International journal of remote sensing*, **21** (6-7): 1389-1414.

ESA (1992). Remote sensing forest map of Europe, ESA/ESTEC, ISY Office, Noordvijk, the Nederlands.

Fazakas, Z., Nilsson, M. and Olsson, H. (1999). “Regional forest biomass and wood volume estimation using satellite data and ancillary data” *Agricultural and forest meteorology*, **98-99**: 417-425.

Hansen, M., Dubayah, R. and DeFries, R. (1996). “Classification trees: An alternative to traditional land cover classifiers” *International journal of remote sensing*, **17** (5): 1075-1081.

Hansen, M.C., Defries, R.S., Townshend, J.R.G. and Sohlberg, R. (2000). “Global land cover classification at 1km spatial resolution using a classification tree approach” *International journal of remote sensing*, **21** (6-7): 1331-1364.

Holben, B.N. (1986). “Characteristics of maximum-value composite images from temporal AVHRR data.” *International journal of remote sensing*, **7** (11): 1417-1434.

Holmgren, J., Joyce, S., Nilsson, M. and Olsson, H. (2000). “Estimating stem volume and basal area in forest compartments by combining satellite image data with field data” *Scandinavian journal of forest research*, **15** (1): 103-111.

Huete, A.R. (1988). “A soil-adjusted vegetation index (SAVI)” *Remote sensing of environment*, **25** (3): 295-309.

Huete, A.R., Liu, H.Q., Batchily, K. and vanLeeuwen, W. (1997). “A comparison of vegetation indices global set of TM images for EOS-MODIS” *Remote sensing of environment*, **59** (3): 440-451.

Häme, T., Salli, A., Andersson, K. and Lohi, A. (1997). “A new methodology for the estimation of biomass of conifer-dominated boreal forest using NOAA AVHRR data” *International journal of remote sensing*, **18** (15): 3211-3243.

Häme, T., Heiler, I. and SanMiguelAyanz, J. (1998). “An unsupervised change detection and recognition system for forestry” *International journal of remote sensing*, **19** (6): 1079-1099.

Høst, G., Steinbakk, G.H., Skare, Ø., Tvete, I.F. and Varin, C. (2001) "FOREMMS WP2-Task 9 Scientific Report: Statistical data analysis; version 3-1." *Norwegian Computing Center*, Report no.

Jakubauskas, M.E. and Price, K.P. (1997). “Empirical relationships between structural and spectral factors of Yellowstone lodgepole pine forests” *Photogrammetric engineering and remote sensing*, **63** (12): 1375-1381.

Jeon, B. and Landgrebe, D.A. (1999). "Decision fusion approach for multitemporal classification" *IEEE Transactions on Geoscience and Remote Sensing*, **37** (3): 1227-1233.

Klöditz, C., vanBoxtel, A., Carfagna, E. and vanDeursen, W. (1998). "Estimating the accuracy of coarse scale classification using high scale information" *Photogrammetric engineering and remote sensing*, **64** (2): 127-133.

Knyazikhin, Y., Martonchik, J.V., Diner, D.J., Myneni, R.B., Verstraete, M., Pinty, B. and Gobron, N. (1998a). "Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data" *Journal of Geophysical Research-Atmospheres*, **103** (D24): 32239-32256.

Knyazikhin, Y., Martonchik, J.V., Myneni, R.B., Diner, D.J. and Running, S.W. (1998b). "Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data" *Journal of Geophysical Research-Atmospheres*, **103** (D24): 32257-32275.

Knyazikhin, Y., Glassy, J., Privette, J.L., Yian, Y., Lotcsh, A., Zhang, Y., Wang, Y., Morisette, J.T., Votava, P., Myneni, R.B., Nemani, R.R. and Running, S.W. (1999). MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation absorbed by vegetation (FPAR) product (MOD15) algorithm theoretical basis document. http://eospso.gsfc.nasa.gov/atbd/modistables.html

Kumar, P. (1999). "A multiple scale state-space model for characterizing subgrid scale variability of near-surface soil moisture" *IEEE Transactions on Geoscience and Remote Sensing*, **37** (1): 182-197.

Köhl, M. and Päivinen, R. (1996) "Definition of a system of nomenclature for mapping European forests and for compiling a pan-European forest information system" *Office for Official Publications of the European Commission*, Report no. CL-NA-16-416-EN-C.

Lee, T., Richards, J.A. and Swain, P.H. (1987). "Probabilistic and evidential approaches for multisource data-analysis" *IEEE Transactions on Geoscience and Remote Sensing*, **25** (3): 283-293.

LeHegaratMascle, S., Bloch, I. and VidalMadjar, D. (1997). "Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing" *IEEE Transactions on Geoscience and Remote Sensing*, **35** (4): 1018-1031.

Lobo, A., Marti, J.J.I. and GimenezCassina, C.C. (1997). "Regional scale hierarchical classification of temporal series of AVHRR vegetation index" *International journal of remote sensing*, **18** (15): 3167-3193.

Loveland, T.R., Merchant, J.W., Ohlen, D.O. and Brown, J.F. (1991). "Development of a land-cover characteristics database for the conterminous United-States" *Photogrammetric engineering and remote sensing*, **57** (11): 1453-1463.

Luettgen, M.R. and Willsky, A.S. (1995). "Likelihood calculation for a class of multiscale stochastic-models, with application to texture-discrimination" *IEEE transactions on image processing*, **4** (2): 194-207.

Mickelson, J.G., Civco, D.L. and Silander, J.A. (1998). "Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery" *Photogrammetric engineering and remote sensing*, **64** (9): 891-904.

Muchoney, D.M. and Haack, B.N. (1994). "Change detection for monitoring forest defoliation" *Photogrammetric engineering and remote sensing*, **60** (10): 1243-1251.

Nielsen, A.A., Conradsen, K. and Andersen, O.B. (2001). "Change detection in the 1996-1997 AVHRR oceans pathfinder sea surface temperature data." *The 12th*

Scandinavian conference on image analysis, 11-14 June 2001, Bergen. Norwegian Society for Image Processing and Pattern Recognition, NOBIM, pp. 725-732.

Nilson, T., Anniste, J., Lang, M. and Praks, J. (1999). “Determination of needle area indices of coniferous forest canopies in the NOPEX region by ground-based optical measurements and satellite images” *Agricultural and forest meteorology*, **98-9**: 449-462.

O'Neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H. and Graham, R.L. (1988). “Indices of landscape pattern.” *Landscape ecology*, **1** (3): 153-162.

Pinder, J.E. and McLeod, K.W. (1999). “Indications of relative drought stress in longleaf pine from Thematic Mapper data” *Photogrammetric engineering and remote sensing*, **65** (4): 495-501.

Raviv, J. (1967). “Decision making in markov chains applied to the problem of pattern recognition” *IEEE transactions on information theory*, **13**: 536-559.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W. and Ohlen, D.O. (1994). “Measuring phenological variability from satellite imagery” *Journal of vegetation science*, **5** (5): 703-714.

Roy, D.P. (1997). “Investigation of the maximum normalized difference vegetation index (NDVI) and the maximum surface temperature (T-s) AVHRR compositing procedures for the extraction of NDVI and T-s over forest” *International journal of remote sensing*, **18** (11): 2383-2401.

Royle, D.D. and Lathrop, R.G. (1997). “Monitoring hemlock forest health in New Jersey using Landsat TM data and change detection techniques” *Forest science*, **42** (3): 327-335.

Sachs, D.L., Sollins, P. and Cohen, W.B. (1998). “Detecting landscape changes in the interior of British Columbia from 1975 to 1992 using satellite imagery” *Canadian journal of forest research - Revue canadienne de recherche forestier*, **28** (1): 23-36.

Sohn, Y.S., Moran, E. and Gurri, F. (1999). “Deforestation in north-central Yucatan (1985-1995): Mapping secondary succession of forest and agricultural land use in Sotuta using the cosine of the angle concept” *Photogrammetric engineering and remote sensing*, **65** (8): 947-958.

Solberg, A.H.S., Bosnes, V. and Storvik, G. (1995). "Tissue classification in MR images based on a mixed pixel model" *9th Scandinavian Conference on Image Analysis., June 6-9, 1995, Uppsala, Sverige*.

Solberg, A.H.S., Taxt, T. and Jain, A.K. (1996). “A Markov random field model for classification of multisource satellite imagery” *IEEE Transactions on Geoscience and Remote Sensing*, **34** (1): 100-113.

Swain, P.H. (1978). “Bayesian classification in a time-varying environment” *IEEE Transactions on systems man and cybernetics*, **8**: 879-883.

Tian, Y.H., Zhang, Y., Knyazikhin, Y., Myneni, R.B., Glassy, J.M., Dedieu, G. and Running, S.W. (2000). “Prototyping of MODIS LAI and FPAR algorithm with LASUR and LANDSAT data” *IEEE Transactions on Geoscience and Remote Sensing*, **38** (5): 2387-2401.

Todd, S.W. and Hoffer, R.M. (1998). “Responses of spectral indices to variations in vegetation cover and soil background” *Photogrammetric engineering and remote sensing*, **64** (9): 915-921.

Tomppo, E., Goulding, C. and Katila, M. (1999). “Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory” *Scandinavian journal of forest research*, **14** (2): 182-192.

Turner, D.P., Cohen, W.B., Kennedy, R.E., Fassnacht, K.S. and Briggs, J.M. (1999). “Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites” *Remote sensing of environment*, **70** (1): 52-68.

Varjo, J. (1996). “Controlling continuously updated forest data by satellite remote sensing” *International journal of remote sensing*, **17** (1): 43-67.

Varjo, J. and Folving, S. (1997). “Monitoring of forest changes using unsupervised methods: A case study from boreal forest on mineral soils” *Scandinavian journal of forest research*, **12** (4): 362-369.

Venables, W.N. and Ripley, B.D. (1999). *Modern applied statistics with S-PLUS*. Springer. 501 pp.

Volden, E., Solberg, A.S., Huseby, R.B. and Solberg, R. (1998) "Forest mapping by airborne remote sensing techniques" *Norwegian Computing Center, NR*, Report no. 928.

Wang, L. and Liu, J. (1999). “Texture classification using multiresolution Markov random field models” *Pattern recognition letters*, **20** (2): 171-182.

White, J.D., Running, S.W., Nemani, R., Keane, R.E. and Ryan, K.C. (1997). “Measurement and remote sensing of LAI in Rocky Mountain montane ecosystems” *Canadian journal of forest research - Revue canadienne de recherche forestier*, **27** (11): 1714-1727.

Wickham, J.D., Wade, T.G., Jones, K.B., Riitters, K.H. and Oneill, R.V. (1995). “Diversity of ecological communities of the United States” *Vegetatio*, **119** (2): 91-100.

Wolter, P.T., Mladenoff, D.J., Host, G.E. and Crow, T.R. (1995). “Improved forest classification in the Northern Lake-states using multitemporal Landsat imagery” *Photogrammetric engineering and remote sensing*, **61** (9): 1129-1143.

Zhang, Y., Tian, Y., Knyazikhin, Y., Martonchik, J.V., Diner, D.J., Leroy, M. and Myneni, R.B. (2000). “Prototyping of MISR LAI and FPAR algorithm with POLDER data over Africa” *IEEE Transactions on Geoscience and Remote Sensing*, **38** (5): 2402-2418.

Zhu, Z.I. and Evans, D.L. (1994). “United-States forest types and predicted percent forest cover from AVHRR data” *Photogrammetric engineering and remote sensing*, **60** (5): 525-531.

Zhu, Z.L. (1994). “Forest density mapping in the lower 48 states - A regression procedure” *USDA Forest Service Southern Forest Experiment Station research paper* (280): 1-12.

Aas, K., Solberg, A.S., Koren, H. and Solberg, R. (1996) "Semi-automatic revision of forest maps combining spectral, texture, laser altimeter and GIS data." *Norwegian Computing Center, NR*, Report no. 909.

Aas, K., Solberg, A.H.S., Koren, H. and Solberg, R. (1997). "Semi-automatic revision of forest maps combining spectral, textural, laser altimeter and GIS data" *Third International Airborne Remote Sensing Conference and Exhibition, København*, pp. 405-411.