IIASA

International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-98-085 / October 1998

Classification and Estimation of Forest and Vegetation Variables in Optical High Resolution Satellites: A Review of Methodologies

Gidske L. Andersen (gidske@nrsc.no)

Approved by
Sten Nilsson (nilsson@iiasa.ac.at)
Leader, *Forest Resources Project*

Foreword

IIASA, the Russian Academy of Sciences and several Russian governmental agencies, signed agreements in 1992 and 1994 to carry out a large-scale study on the Russian forest sector. The overall objective of the study is to focus on policy options that would encourage sustainable development of the sector.

The first phase of the study concentrated on the generation of extensive and consistent databases for the total forest sector in Russia. In its second phase, the study encompassed assessment studies of greenhouse gas balances, forest resources and forest utilization, biodiversity and landscapes, non-wood products and functions, environmental status, transportation infrastructure, forest industry and markets, and socioeconomics.

The remote sensing activities within this project aims at the following three main objectives:

- to produce an up-to-date forest information database of the Russian forestsector;
- to develop and test methods to produce an up-to-date land use and land cover database for Russia; and
- to develop and test operative forest information and decision support system, with monitoring and revision capabilities, in a GIS environment.

The current study was done during the 1998 Young Scientists Summer Program (YSSP).

About the Author

Gidske L. Andersen is currently working on her master of science thesis in Botany at the Nansen Environmental and Remote Sensing Center in Bergen, and the Botanical Institute, University of Bergen, Norway. In her thesis, she applies remote sensing to study vegetation changes in the Eastern desert of Egypt.

Acknowledgements

Firstly, I would like to express my gratitude to the Norwegian Research Council for funding my stay at IIASA, and to Ola M. Johannesen who introduced me to IIASA's Young Sscientists Summer Program.

Secondly, my thanks to Alf Öskog for his contribution to and supervision of my work during the summer months. In addition, a special mention of thanks is extended to Professor Sten Nilsson and all the other colleagues, not listed here, in the Forest Resources Project for their interest in and support of my research. It was much appreciated.

Finally, a big thank you to all the 1998 YSSPers for making the summer an unforgettable one!

Contents

Classification and Estimation of Forest and Vegetation Variables in Optical High Resolution Satellites: A Review of Methodologies

Gidske L. Andersen (gidske@nrsc.no)

1 Introduction

Several international organizations have highlighted the importance of developing sustainable forest management (i.e., the Intergovernmental Panel on Forests, World Commission on Forests and Sustainable Development), both in terms of long-term management, and biodiversity conservation and mitigation of climate change impacts.

Russia has vast forest resources, comprising 23% of the world's forest areas (forest area as defined by FAO, 1995), and in world-wide growing stock terms this amounts to 21–22%. In addition, 11% of living forest biomass and 55% of the world's growing stock of conifers of the world's growing stock is located in Russia. The majority of Russia'sn forest resources are located in Siberia and the boreal forest of Central and Western Siberia is the largest unbroken tract of forest on earth. Accordingly, Russian and Siberian forests are of global importance, not only ecologically but also economically.

Russian forests received increased attention after the political and economic transformation of the former Soviet Union. The debate is now about the future development of forest resources. Any new policies, resulting from the debate, will inevitably have an impact on both the management and exploitation of forest resources and significant changes are considered to be likely.

The current forest inventory system is central in the debate as several weaknesses are recognized [18, 23, 24]. Weaknesses in the current system include:

- no forest fire monitoring system;
- no inventories of disturbances due to pests, disease and other biotic factors;
- no complete survey of the extent and intensity of abiotic impacts;
- the production and utilization of so-called non-wood products and functions of the forest are not efficiently monitored;

- inventories for individual regions have been carried out infrequently; existing forest maps are of poor quality (1:2500000), and the last inventory of approximately 90 million hectares was carried out during the period 1948–1956; and
- the current inventory system is unable to capture many of the important dynamics of the forest.

Hence, it is recognized that there is a need for a new land information system which should be policy and market oriented and in which remote sensing plays a crucial role [18]. Both the economic conditions and the vast forest areas require a more intensified use of remote sensing.

Some of the variables of interest for a future inventory system using satellite data are:

1. Factors reflecting disturbances of areas; forest fires, insect damage, harvests and air pollution.
2. Primary and secondary forests, and harvested areas.
3. Species composition regarding coniferous and deciduous species.
4. Natural regeneration status regarding coniferous and deciduous species.
5. Rough distribution of volume classes, primarily three to four classes.
6. Transportation infrastructure including road network.
7. Indicators of surface parameters such as humidity and the extent of permafrost which can be used as variables to forecast the risk of large scale disturbances (e.g., forest fire).

Combining the existing forest inventory information with high resolution satellite data (SPOT) will produce new and detailed information, and offer not only the possibilitiy to update faster but also the possibility for historical studies. The classification of satellite data offers the possibility of recognizing variables of interest.

This study reviews and describes methods for the classification and estimation of forest and vegetation variables in optical high resolution satellite data.

2 Classification of Boreal Forests

Forestry planning is, to a certain degree, assisted by remote sensing, mainly in North America and Europe. The goal of applying remote sensing to forestry is to add value (utility) in the planning process [11], which should mean one of the following:

- the information retrieved from satellite remote sensing should be less expensive than collection by other means;

- it should contain better information at the same cost; or
- it should be both cheaper and better in some way.

The potential of applying remote sensing to forestry has been recognized by several authors [27, 31].

The use of ground based survey methods to obtain detailed inventory data is currently both too expensive and time consuming [12, 22, 27]. If detailed data were available, it would be possible to optimize a response to the changing market demands. The need for a less expensive and more comprehensive approach has prompted much research into extracting forest information from aerial and satellite remote sensing [27].

Woodcock, *et al.* [31] mentions the costs which can be saved, both in the long term and per unit area. The conversion of satellite image format to GIS is easy, making map production possible within a short time and also allows the fast update of computer-based maps following fires, floods, and other catastrophic events. This is particularly useful for areas where resource information is limited and fast changes make the information collected by conventional methods outdated [19]. Another advantage is the high degree of consistency between locations that is offered by remote sensing [31].

In the years since the launch of the first optical satellite, Landsat MSS in 1972, technology has developed and improved so much that multiple sensors are available today. In particular, development has been toward a better spectral, spatial and radiometrical resolution. The main satellites are described in *Table 1.* The bands which are usually available are: green, red, and near infrared; however, with newer sensors, these are extended to blue, and two mid-infrared bands.

Table 1: Some characteristics for commonly used optical satellites.

Satellite Sensor	**Ground Resolution**	**No. of Visible Bands**	**No. of Infrared Bands**
Landsat MSS	80[a]	3	1
Landsat TM	30	3	3
SPOT P	10	1	0
SPOT XS	20	2	1–2[b]
RESURS-O1 MSU-E	45	2	1

[a] Varies slightly for the different Landsat generations. [b] SPOT4 has a mid-infrared band.

2.1 Methods of classification and segmentation

A raster image is basically a grid of pixels, i.e., squares with digital numbers, distributed in a spatial framework of rows and columns. The human brain recognizes

patterns in the image and the interpretation gives us information of the spatial distribution of features described by the sum of pixels. Classification consists of recognizing these features and patterns and their assignment to specified land cover classes. Both spectral and spatial patterns are of interest and the approach can be either manual or digital.

The visual/manual technique uses the skills of the human brain to recognize factors such as shape, size, pattern, shadow, tone/color, texture and site [15]. Some of these factors can be limited by the ground resolution of the satellite. Manual interpretation as described by Borry, *et al.* (1990) (referred to in [11]) deems to be more reliable than digital classification of forest type. Roy, *et al.* [19][1] emphasize that visual interpretation is a subjective method and hence mapping borders will vary between interpreters. However, this is of limited importance for broad scale studies. For detailed mapping computer-aided techniques are suggested. If the area studied is over a certain size a visual approach is, however, very cumbersome, time, and labour consuming. Visual interpretation is also applied as part of the digital approach in the segmentation and ensuing evaluation of the defined segments [3]. Woodcock, *et al.* [31] uses manual techniques to label clusters derived from a complex classification.

Digital classification methods normally use stastical or spectral decision rules to assign the unit considered to different classes. The unit considered can be either the pixel or a region. As opposed to the per-pixel approach a region-based classification first requires segmentation.

2.1.1 Segmentation

Image segmentation is a method of defining discrete objects or classes of objects. The method was presented in the 70's by Ketting and Landgrebe (referred to in [11]) but is still not commonly applied as a satellite image interpretation tool for forestry purposes [3, 11]. However, the advantages for forestry are recognized by several authors [2, 3, 10, 11, 20] which is due to the fact that forestry planning and management still uses forest stand as the basic unit. Ideally, the use of segmentation will define forest stands. A next step is the classification of regions and normally the more conventional methods are applied. The classification of regions avoids the problems of single mixed pixels but also provides more parameters to classify than the brightness normally used. These parameters can be texture, perimeter and shape. In particular, texture measures are very important features for discriminating individual stands [2, 3]. Skewness and kurtosis are other parameters which can be put into the classification of the segments. However, a reliable calculation of these require the regions to be above a minimum size [2].

Several methods of segmentation exist, four of these are: pixel-based, edge-based, object-based/region growing, and hybrid [2]. Burger and Steinwendner [2] select the object-based approach due to its simplicity, efficacy, generality and its capability of being combined with edge-based methods. These object/region based methods can be further divided into split and merge-, region-growing-, and pyramid-linking-segmentation techniques. The region-growing algorithms are most often applied

[1] This is not a study of boreal forests but the principles and conclusion are of general interest.

to forestry [2, 3, 10, 20, 30, 31]. There are four different types of region-growing algorithms:

- spatial clustering;
- single linkage region-growing;
- hybrid linkage region-growing; and
- centroid linkage region-growing.

In principle the segmentation algorithms consist of several steps, however the specific approach depends on the algorithm. The following describes the general approach of a region-growing algorithm.

The first step is the combination of pixels into regions. A neighborhood function, commonly consisting of either four or all eight neighbor pixels, is applied and these pixels are included in the region if the spectral difference between the center- and neighbor-pixels are below a certain threshold. This threshold can be a user-specified Euclidean distance in a n-dimensional space [20] or an adaptive threshold [3], t_{adapt},

$$t_{adapt} = (1 - min(0.8, \frac{standarddeviation}{mean}))t \quad (1)$$

where t is a given threshold. In the latter case, the user specifies the final highest allowed threshold for including a pixel. Below this value, the threshold can change as the region grows; it decreases as the standard deviation of the growing region increases. Thus, this adaptiveness of the threshold is a simple measure of the texture and, as such, an alternative way to include texture measures in segmentation.

The addition of single texture bands were tested by Ryherd and Woodcock [20] and Woodcock and Harward [30]. The results indicate that in the case of land-cover types with different texture values, segmentation is significantly improved by using texture data [20]. Woodcock and Harward [30] conclude that sparsely stocked forest stands which have high internal variance texture bands derived by an adaptive filter improve the segmentation results. In the worst case, a texture band will not, [20], or rarely [31], degrade the quality of the segmentation. The adaptive filtering technique applied includes a 3x3 moving window selecting the lowest value to be assigned for the pixel considered.

The second step is the filtering and merging of regions. Regions below a certain size should be merged with the spectrally closest spatial neighbor. Spatial adjacency is defined by Burger, *et al.* [3] to be the overlapping regions after delineating the region of interest. Both size and spectral measures are used to control merging. A minimum size can be applied in the first step to avoid creating single pixel regions [20]. Two different maximum sizes can be applied: one size for which merging is not allowed for that specific region, and a viable size which does not allow two regions over this size to merge. For certain features such as, small lakes and meadows which can be spectrally delimited, a user specified spectral limit can be used to avoid these features disappearing when merged with other features [20].

Hagner [10] uses the t-ratio as the spectral measure for merging regions. The t-ratio is defined as:

$$t - ratio = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{2}$$

where:

$\bar{x}_i$ = mean of region i
s_i^2 = variance of region i
n_i = number of pixels in region i

Hagner [10] decides the appropriate level of the t-ratio on a trial and error basis. The visual interpretation of aerial photographs and field checking are the basis for the comparison. In the case of multiple bands the measure used is the square root of summed, squared t-ratios. The t-ratio is not defined for single pixel regions and they are merged only on the basis of the distance to the closest spectral neighbor in feature space. Merging is an iterative process where the threshold increases until the user defined threshold is reached. The regions to be merged should be each other's reciprocal nearest neighbor in the n-dimensional feature space [20].

Additional information sources can be included in the segmentation process. Hagner [10] guides the stand delineation (segmentation) by map information and digital elevation models.

With further improvement in sensor technology and hence increased spatial resolution, individual trees will be recognizable. A forest stand is thus no longer a homogenous region and segmentation cannot provide the seperation of forest stands. To solve this problem, it is suggested to first apply a low pass filter to remove the high texture information of the image and then apply the segmentation algorithms on the blurred image [3].

2.1.2 Classification

The classification algorithms are executed on the images themselves or on partly processed images, for example, a segmented image. However, the classification principles are the same if one deals with segment- or per-pixel classification.

Classification is a two-step process:

1. The training stage:
   - unsupervised;
   - supervised.
2. The classification stage.

The Training Stage. Training is the process of defining the criteria for which the patterns are recognized (Hord, 1982, reffered to in [6]).

Unsupervised training or clustering is a more computer automated process which basically recognizes spectral clusters defined on a solely statistical basis. Different clustering methods exist; among them ISODATA (Iterative Self Organizing Data Analysis) and RGB clustering.

The first iteration of ISODATA clustering determines the means of the maximum number (N) of clusters arbitrarily. The spectral distance between all the pixels and the cluster means are calculated and the candidate pixel is assigned to the closest spectral cluster. When all the pixels are assigned to a cluster the means are recalculated and used for defining clusters in the next iteration. The process continues until a user specified portion of the pixels does not change from one cluster to another between two iterations. This number is defined by the convergence threshold. The user also sets the maximum number of iteration and clusters to be processed.

Good *a-priori* knowledge of the area is not required. However, the clustering will be useless if no appropriate interpretation of the clusters is given. Mayer and Fox [17] interpret the unsupervised signatures by comparing spectral mean plots to similar plots derived from a supervised approach. They conclude that such comparison, in conjunction with photo-interpretation, is an effective way of assigning reliable labels. Woodcock, *et al.* [31] apply unsupervised classification to determine the vegetation growth forms for the regions derived by segmentation. They do one classification for poorly illuminated areas and another for well illuminated areas.

Supervised training is a method in which the user defines a set of classes of interest. Signatures for these classes are selected by the user, i.e., the user supervises the computer in the training process. This requires some *a-priori* knowledge about the area (reference/ground truth data) and their spatial distribution. This can be a problem for remote sensing applications since the results are usually very dependent on the quality and quantity of the reference data. One of the primary advantages of remote sensing is thus questioned. Mayer and Fox [17] attribute the high accuracy for certain variables to the large amounts of training data and the low accuracy for other variables to insufficient training data.

A sampling between 10n and 100n pixels per class is recommended, n being the number of channels or bands used in the classification. It is also recommended to take several smaller homogenous samples than fewer larger ones (Swain and Davis, 1978 and Campbell, 1981 referred to in [21]).

The result of the training process is a set of signatures for the classes selected or recognized. These signatures can be either parametric or non-parametric. Parametric signatures describe pixels in the training sample by a set of statistical parameters which are the input in a statistical based classification. A non-parametric signature is delimited by boundaries in a feature space/scatter diagram rather than described by stastical parameters. However, it is possible to calculate statistics for one type of feature space objects; the ellipse of which fulfills the normal distribution requirements of a parametric classifier. The most crucial part of the training step is the evaluation of the signatures. If signatures overlap the spectral information of the data itself cannot support the classes of interest. This is often one of the main prob-

lems with a per-pixel classification which only uses the spectral information content of the image. Some different methods for evaluating signatures are presented below.

Alarm. This is a visual method and produces an image file which displays all the pixels within the minimum–maximum boundaries in the feature space of the selected signature.

Ellipse or rectangles. The separability of signatures is evaluated by displaying the ellipses or rectangles describing the signatures in a two-dimensional feature space.

Contingency matrix. The pixels in the sample/training regions are classified and the resulting distribution across different classes are displayed in a contingency matrix which contains the number and percentages of pixels that are classified as expected.

Separability/Divergence. Separability is a statistical measure of the distance between two signatures. Some of the decision rules use distance in the classification algorithm and separability can thus predict the result of the final classification. There are three options for calculating separability:

- Divergence;
- Transformed Divergence; and
- Jeffries-Matusita Distance.

The latter two have lower and upper boundaries of which the upper boundary indicates total separability.

Divergence, D_{ij} is calculated by the formula:

$$D_{ij} = \frac{1}{2}tr\left((C_i - C_j)(C_i^{-1} - C_j^{-1})\right) + \frac{1}{2}tr\left((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T\right) \quad (3)$$

where:

i and j	= the two signatures being compared
C_i	= the covariance matrix of signature i
μ_i	= mean vector of signature i
tr	= the trace function (matrix algebra)
T	= transposition function

Transformed Divergence, TD_{ij} is a further calculation using the Divergence equation and is described by the following formula:

$$TD_{ij} = 2\left(1 - exp\left(\frac{-D_{ij}}{8}\right)\right) \quad (4)$$

TD_{ij} ranges between 0 and 2000.

Jeffries-Matusita Distance is calculated by the formula:

$$JM_{ij} = \sqrt{2(1 - e^{-\alpha})} \tag{5}$$

where:

$$\alpha = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2}\right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} ln \left(\frac{|(C_i + C_j)/2|}{\sqrt{|C_i| \times |C_j|}}\right) \tag{6}$$

and

i and j	= the two signatures being compared
C_i	= the covariance matrix of signature i
μ_i	= mean vector of signature i
ln	= the natural logarithm function
$\|C_i\|$	= the determinant of C_i (matrix algebra)
T	= transposition function

Values of JM_{ij} ranges between 0 and 1414.

The Jeffries-Matusita distance is related to the pairwise probability of error which is the probability that a pixel assigned to class i is actually class j. Within a range this probability can be estimated according to:

$$\frac{1}{16}\left(2 - JM_{ij}^2\right)^2 \leq P_e \leq 1 - \frac{1}{2}\left(1 + \frac{1}{2}JM_{ij}^2\right) \tag{7}$$

Weight Factors can be included in the separability measure and influences the best average and the best minimum separability. The weight factors are used to compute a weighted divergence with the following calculation:

$$W_{ij} = \frac{\sum_{i=1}^{c-1}\left(\sum_{j=i+1}^{c} f_i f_j U_{ij}\right)}{\frac{1}{2}\left[\left(\sum_{i=1}^{c} f_i\right)^2 - \sum_{i=1}^{c} f_i^2\right]} \tag{8}$$

where:

i and j	= the two signatures being compared
U_{ij}	= the unweighted divergence between i and j
W_{ij}	= the weighted divergence between i and j
c	= the number of signatures (classes)
f_i	= the weight factor for signature i

Ayanz and Biging [21] include these separability measures in a iterative classification procedure so that for each iteration the best combination of bands are selected as input to the classification. They find no significant difference in the band combinations obtained when using either TD or JM. They recommend JM as it requires less computing time.

The Classification Stage. The signatures created by the training stage are the input to the classification stage. Pixels and/or regions are then assigned to classes based on

a decision rule. Different decision rules are developed and these are either parametric or non-parametric rules. The last part of the classification step is the accuracy assessment which is not described in this report.

Minimum distance (to means) classifier. This decision rule is considered to be the simplest one. The distance between the candidate pixel, the pixel to be classified, and the mean of the classes defined, is considered. The pixel is assigned to the class which has the closest mean. The minimum distance is based on the equation for Euclidean distance (Swain and Davis, 1978, referred to in [6]):

$$SD_{xyc} = \sqrt{\sum_{i=1}^{n} (\mu_{ci} - X_{xyi})^2} \tag{9}$$

where:

n	= number of bands (dimensions)
i	= a particular band
c	= a particular class
X_{xyi}	= data file value of pixel x,y in band i
μ_{ci}	= mean of data file values in band i for the sample for class c
SD_{xyc}	= spectral distance from pixel x,y to the mean of class c

The only parameter used to decide the class for the candidate pixel is the mean and other important parameters, such as covariance and correlation, are not considered. Misclassification is likely if there are big differences in the variance of the classes and if the candidate pixel is closer to the mean of the low variance class.

Mahalanobis distance classifier. This method assumes a normal distribution of the band histograms. In addition to the mean value of the signature class the covariance matrix (covariance and variance) is included in the calculation which tends to over-classify signatures with relatively large values in the covariance matrix.

$$D = (X - M_c^T)(cov_c^{-1})(X - M_c) \tag{10}$$

where:

D	= Mahalonobis distance
c	= a particular class
X	= the measurement vector of a candidate pixel
M_c	= the mean vector of the signature of class c
Cov_c	= the covariance matrix of the pixels in the signature of class c
Cov_c^{-1}	= the inverse of Cov_c
T	= transposition function

Maximum likelihood/Bayesian classifier. This method requires a normal distribution of the bands included in the classification and is based on the probability that a pixel belongs to a particular class. The maximum likelihood equation assumes that the probabilities for all the classes are equal. If the class probabilities are known it is possible to include these by using the Bayesian classifier. In the same way as the

Mahalanobis classifier, this method also tends to over-classify signatures with large values in the covariance matrix.

$$D = ln(a_c) - [0.5ln(|Cov_c|)] - [0.5(X - M_c)T(Cov_c^{-1}(X - M_c)] \quad (11)$$

where:

D	=	weighted distance (likelihood)
c	=	a particular class
X	=	the measurement vector of the candidate pixel
a_c	=	percent probability that any candidtate pixel is a member of class c (defaults to 1.0, or is entered from *a priori* knowledge)
Cov_c	=	the covariance matrix of the pixel in the sample of class c
$\|Cov_c\|$	=	determinant of Cov_c (matrix algebra)
ln	=	natural logarithm function
T	=	transposition function (matrix algebra)

Several authors apply maximum likelihood [8, 12, 16, 21, 22]. Kadro [12] and Schneider [22] use maximum likelihood to differentiate damage levels in boreal forest in TM imagery. Ayanz and Biging [21] explore an iterative maximum likelihood classifier. One or two cover types are classified in each iteration according to the band combination score on the different separability indices. Test classification using only signature pixels, is used to list cover types according to their accuracy. Those with the highest accuracies are classified first and are then masked. A new iteration selects the next cover class to be classified and the optimal band combination. This type of classification is compared with the conventional type, both when using and omitting prior probability. However, including prior probability does not increase the overall accuracy but there is a significant improvement by using the iterative method; accuracies increased from 58 to 66%.

Parallelepiped and feature space classifier. The parallelepiped and the feature space decision rule can deal with both parametric and non-parametric signatures. Basically both of these methods decide whether a candidate pixel is within a set of given limits. In the parallelepiped method the limits are described by constants (non-parametric) or by mean values and standard deviations. The feature space method allows the user to draw these limits in a scattergram. The only geometric feature which gives parametric signatures is the ellipse. Depending on the kind of limits given or drawn, it is possible to include correlation/covariance into the decision rule. The parallelepiped classifier leaves a group of unclassified pixels.

This method is used by several authors for different purposes [2, 3, 21, 28]. Burger and Steinwendner [2] and Burger, *et al.* [3] use a decision tree to classify defined segments and finally extract the forest. Williams, *et al.* [28] use parallelepiped limits to distinguish three different insect damage levels in a Landsat MSS image. Ayanz and Biging [21] use the parallelepiped classifier as a first step of the classification to quicken the next steps which are done by the maximum likelihood classifier.

Characteristics of the different classification methods are summed up in *Table 2.*

Table 2: Some characteristics of the different classification methods.

	Parallel-epiped	*Feature Space*	*Minimum Distance*	*Mahalanobis Distance*	*Maximum Likelihood*
Parametric Signature	possible	possible	yes	yes	yes
Non-parametric Signature	yes	yes	no	no	no
Variance	yes	depends	no	yes	yes
Covariance/ Correlation	no/ stepped borders	depends/ yes	no	yes	yes
Unclassified	yes	yes	no	no	no
Normal Distribution	no	no	no	yes	yes

The trend in remote sensing studies of landcover classification is towards combining methods, including original and derived bands (vegetation indices, principal components, tasseled cap transformation and filtered bands) as well as auxiliary data from GIS [21].

Woodcock, *et al.* [31] represents a good example. This investigation comprises both segmentation, unsupervised and supervised classification, inversion of the Li-Strahler canopy model and terrain based rules to differentiate species associations, and a fuzzy set to evaluate the map accuracy. One of their conclusions is, however, that this system is too complicated and requires too much equipment and training to be implemented by small mapping projects and is only recommended for the national or regional level.

Fuzzy membership rules are also applied to improve the classification itself. Stolz and Mauser [25] use such rules to better distinguish land cover types in steep terrain. The combined probability from a maximum likelihood classification and the membership functions assigns the final class value. They conclude that the classification results can be considerably increased by using this approach and suggest further investigations into geofactors and the fuzzy membership functions.

A three stage classifier is applied by Wilson and Franklin [29]. This algorithm uses spatial statistics, minimum distance to mean tests and spectral and geomorphomtric curve test. The geomorphometric variables are derived from a elevation model.

Another type of combination of methods is given by Donchenko, *et al.* [5] who also combine different sensors. They use both opticals, NOAA/AVHRR and Landsat MSS, and data from the radar ERS/SAR. They conclude that the appearance of ecological anomalies in boreal forests as a result aerotechnogenic pollution can be effectively revealed, identified and investigated making synergistic use of remotely sensed data.

2.2 Results of classification and segmentation

In contrast to the methods, the variables of interest have been quite uniform in most of the investigations. Some of the biometric variables studied are crown closure, volume (per area), species composition, age classes, and stand density [1, 3]. Different sensors and band combinations have been tried and the results vary considerably.

Schneider [22] applies TM data and recognizes channels 5 and 4 as being most important and valuable for separating forest categories. Brockhaus and Khorram [1] find that the basal area is significantly correlated to XS3 and TM bands 2-5 and 7 but the correlation coefficients are low. In addition, these TM bands are significantly, but weakly, correlated to age classes [1]. The same investigation finds no correlation between age class and SPOT XS bands. However, both SPOT XS and Landsat TM bands show a general negative correlation with wood volume and volume related parameters [27]. The near infrared band is, perhaps, an exception showing positive, negative or flat correlation (referred to in [27]). The trend depends upon shadowing, subcanopy and water content. The mid-infrared band is recognized by many authors as being the most sensitive to changes in forest wood volume.

Forest damage has been studied by several authors [12, 13, 22, 28]. However, monitoring damage is not necessarily related to damage symptoms [13]. The best relation is found between individual damage classes and TM bands 5 and 7. These are positive correlations. The results also show that ratios and indices are poorer estimates than the single bands themselves. A ratio is, however, applied by some authors [12, 22] using the ratio between TM5 and TM4 to separate damaged areas from non-damaged areas. TM5 has a higher reflection and TM4 a lower reflection for a damaged forest. Kadro [12] concludes that three different damage levels can be distinguished: healthy/slightly damaged, damaged, and severely damaged stands. Schneider [22] uses multitemporal data, consisting of three TM5/TM4 ratios. A correlation between crown closure and crown damage is found in the data and he concludes that it is impossible to say whether the spectral correlation is due to crown condition or crown closure. Several investigators have found a relationship between spectral data and crown cover (referred to in [31]). Williams, *et al.* [28] also apply a ratio, the IR/R vegetation index, to distinguish damage levels in a defoliated forest image. This vegetation index is used by Nelson (1981) (referred to in [28]) with 75% accuracy.

The slope between the green and red MSS band is studied by Mayer and Fox [17]. This slope, which is similar to a simple vegetation index, is found to be a significant factor in determining resource identity. The digital number is related to tree size and stocking level. For low density stands (e.g., less than 40% for lodgepole and knobcone pine) it is difficult to identify unique spectral responses. In these cases the background contributes significantly to the pixel reflectance. A similar effect is identified by Schneider [22] where the main problem is to distinguish between broadleaved trees and young coniferous stands.

Franklin and McDermid [7] use TM data to discriminate regeneration sites which are highly dominated by *Kalmia angustifolia*. They apply six of the TM bands and manage to accurately separate *K. angustifolia* from other land cover classes by 85%.

In fact, 85% accuracy is suggested to be the minimum acceptable when classifying remotely sensed data (Andersen, *et al.,* 1976, referred to in [1]). However, a broad number of studies do not operate with this level of accuracy [11]. Mayer and Fox [17] identify that the accuracy (including both omission and commission errors) for tree species, size and density ranges from 56 to 96%. Brockhaus and Khorram [1] evaluate the reliability of satellite data alone. Three different images are classified, SPOT XS, Landsat TM only using bands comparable to SPOT, and Landsat TM applying all bands. The overall classification accuracy for seven forest cover types were 73.4%, 70.8% and 88.5% respectively.

Wilson and Franklin [29] experience an increase in the overall accuracy from 85%, when applying a maximum likelihood classification, to 88% using the three stage algorithm.

Burger, *et al.* [3] apply a segmentation algorithm and thus completes a region based classification. The overall accuracy is 85% correctly classified pixels for differentiation between forests and non-forests. Differentiating between species composition is concluded as acceptable, although the percentages are not convincing: pure spruce 52.8%, mixed conifer 60.2%, and alpine pine 26.8%. The accuracies for the age classes are 84.2% for mature/old and 41.9% young. The authors judge the low accuracies of the image subset as not being representable. They also suggest that the results can be improved by lowering the threshold for segmentation.

Hagner [10] quantifies the accuracy of the stand delineation as being the proportion of the stand characteristic variance explained by the delineation. He finds very similar accuracy for the t-ratio segmentation and manual delineation.

Woodcock, *et al.* [31] apply a fuzzy set to assess the accuracy of their segmentation based approach. The fuzzy set was based on two questions:

1. How frequently is the category assigned in the map the best choice for the site?

2. How frequently is the category assigned in the map acceptable?

The accuracy connected to the questions are the maximum and the right operator, respectively. They also assess their results by a more commonly used confusion matrix [15]. The overall accuracy of the map measured by the maximum and the right operator was 79% and 88% respectively. The species associations have generally good accuracies, the numbers for hardwood are: maximum 78% and right 89%, and for conifer species: maximum 68% and right 80%.

3 Estimation of Forest Variables

Classification is assigning pixels or regions to land cover types, but on a discontinuous scale. However, most important stand variables are continuous and in order to retain the information they should be expressed on a continuous scale [10, 11, 26].

Several investigations endeavor to estimate the parameters on a continuous scale, usually applying different regression and correlation techniques. A N-dimensional K-nearest neighbor classifier (NK-classification) is also used [26, 27].

The NK-classifier is used to give a continuous estimate of the parameter considered [26, 27]. This requires ground truth data for the parameter and the companion pixel value. One assumption to this method is that similar forests exist within the large reference area covered by a satellite image. In addition, the reflectance response of the pixels in the reference area should be independent of their location, and only dependent on the state of the forest (Kilkki and Päivänen, 1987 referred to in [26]). The candidate pixel is considered in its spectral neighborhood, and its Euclidean distance to each pixel in the ground truth dataset estimated. Only the K nearest neighbors are considered for further calculations and these are weighted by a factor according to the distance. Two values are thus available for the K nearest ground truth pixels: the weight factor and the parameter itself. The parameter is then estimated for the pixel to be classified by summing the product of the weight factor and the parameter over all the nearest K neighbors. Tokola, *et al.* [26] test two different distance functions and allocation of weights. This approach has been successfully applied as part of the national scale boreal forest inventory in Finland.

Trotter, *et al.* [27] examines the relation between wood volume and Landsat TM data by linear regression analysis, non-parametric line-fitting and the NK-classification. Estimations were done both at pixel-scale and by stands which were approximately 40 hectares. Standwise estimations were also done on a $50m^3ha^{-1}$ increment scale. Regression methods are tested with volume as both a dependent and independent variable. Estimates should be within $50m^3$ of the ground based estimates and pixel scale estimates are concluded to be unacceptable. The NK-estimates (N=7 and K=15) are biased. The best performed model on the stand scale is a simple linear regression using band 4, giving the RMS error of $41m^3ha^{-1}$. On the increment scale a simple regression applying band 4 gives a RMS error of $39m^3ha^{-1}$. One conclusion of this study is that Landsat TM only provides an acceptable data source for estimating wood volume in plantation forests for areas of about 40 hectares and larger.

Franklin and McDermid [7] use two regression procedures to show the relation between spectral variables, filtered and texture data and the forest stand parameters. They use simple bivariate correlations to explore the relationship among pairs of variables, and stepwise multiple regression to illustrate the relationship between the individual forest stand parameters and several combinations of spectral data. Bivariate correlation results for HRV data and stand parameters show a poor correlation to the infrared band and green band is only significantly correlated to the mean diameter at breast height, height, age and volume. All parameters are most highly correlated to the red band.

Hagner [10] estimates stand characteristics as volume/hectare, mean diameter, mean age, and tree species mixture. This is done by applying regression functions to the spectral signatures within each delineated stand. The regression function is already derived from inventory data and spectral data. Stand edge pixels are excluded from the analysis.

Lambert, *et al.* [13] use logit regression and the non-parametric Kruskal-Wallis test to differentiate forest damage classes. Three categories of damage classes are classified by regression equations with 71–75% accuracy for stands older than 40 years.

Cohen and Spies [4] developed linear regression models to estimate stand attributes such as tree bole diameter, tree height, tree density, age and a structural complexity index. The absolute value of the correlation coefficient for the HRV panchromatic texture is between 0.45 and 0.88 and for TM wetness the range is 0.51 to 0.90; in both cases, the lowest correlation is for tree height. For texture, all variables, except tree density, are positively correlated which is exactly the opposite for TM wetness.

Woodcock, *et al.* [31] use the inverse Li-Strahler forest canopy reflectance model [14] to estimate the size and cover of trees. However, the success was limited and they conclude that estimating tree size using a single date TM image is unreliable. They suggest to improve the results by applying finer spatial resolution, multiple look angles or possibly multiple sun angles.

4 Conclusions

Reflectance from a forest stand is the total of many factors. Guyot and Riom [9] divide these into external and internal factors. One of the main challenges for remote sensing of forests is to efficiently extract the internal information content from the total information content. Holmgren and Thuresson [11] recognize this complex relationship as being overwhelming. Firstly, several different ground coverages can give the same spectral signature. This is especially problematic when the canopy coverage is low. Secondly, similar ground coverages can give different spectral responses. This can be due not only to factors such as different atmospheric conditions but also to less obvious ground cover differences such as those recognized by Burger and Steinwendner [2]. This study finds differences between ecoregions and concludes that further investigations of forest stand parameters should take ecoregions into account. In both cases, however, it should be remembered that auxiliary data can be added and therefore exclude some of the confusion.

One important external factor is the topography of the area. It often causes problems [2] and results are improved on flat terrain [22]. Cohen and Spies [4] find that SPOT HRV (panchromatic) texture and Landsat TM wetness exhibit a strong relationship with stand attributes which is primarily due to their insensitivity to topographically induced illumination angle.

An additional problem, which is inherent for all remote sensing techniques, is the fact that interesting forestry parameters are not necessarily correlated to canopy cover/total ground cover which is monitored by the sensor. Another problem which, however, should be possible to solve with improved technology is the low dynamic range of the sensors [22, 27].

Recognition of forest variables and regions by thematic classification is considered as a dead end by Holmgren and Thuresson [11]. Acceptable accuracies can be achieved

but the results are not useful for either forestry planning or management [10]. Perhaps the most important is the difficulty of matching discrete class definitions with other information sources. Hagner [10] also recognizes the selection of training samples to be laborious and often subjective.

Despite these facts several promising methods are developed. Perhaps the most promising is the combination of segmentation techniques with methods of continous estimation. The topographic correction algorithms also improve the results and segmentation results show high potential for forestry mapping [3]. However, Holmgren and Thuresson [11] have some remarks to that technique. The precision of the delineation is influenced by the pixel size and thus the mixed border pixels. They also point out that the basic assumption for segmentation — homogeneity of land cover, in this case the canopy cover — is not necessarily the same assumption used to delimit a forest stand. They also question the future applicability of the stand concept within forestry. For remote sensing in general this will not be a problem as future sensors will have a far better spatial resolution. Also the positioning accuracies are getting better as the differential techniques of GPS (Global Positioning System) improves. Extracting forest information from pixels and positioning/locating this pixel is possible with a high degree of accuracy.

When it comes to current sensors used, it appears that TM data is better correlated to certain variables than SPOT XS [1]. May, *et al.* [16] explain the greater effectiveness of TM data resulting from the greater spectral resolution provided by additional infrared bands. However, the new SPOT4 sensors provide a new mid-infrared band. The panchromatic SPOT band is recognized as a useful texture measure by several authors. Perhaps the most crucial problem for optical sensors is not the radiometric, spatial or spectral resolution, but their limited use in boreal areas due to the high probability of clouds occurring.

Within forestry groups criticizing the remote sensing applications exist, e.g., [11]. However, more optimistic views also occur. Burger, *et al.* [3] state that remote sensing is capable of delivering results with high spatial resolution and good geometric accuracy in a cost-effective way. The statement by Trotter, *et al.* [27] perhaps sums up the situation and the capability very well by explaining that the data is not sufficient for detailed forest inventory purposes but still provides useful information for areas not visited by a forest inventory. Satellite remote sensing data has to be considered as additional information which should be used in harmony with other information sources, and not as an inexhaustible information source itself.

References

[1] Brockhaus, J.A. and Khorram, S. (1996) "A Comparison of SPOT and Landsat-TM Data for Use in Conducting Inventories of Forest Resources." *International Journal of Remote Sensing,* 13(16):3035–3044.

[2] Burger, H. and Steinwendner, J. (1992) "Study of Forest Mask Generation from Satellite Images Using Image Segmentation Algorithms." Proceedings of the FIRS-Workshop, Vienna, Austria.

[3] Burger, H., Steinwendner, J., Bartl, R. and Schneider, W. (1997) "Region Based Classification of Forest Stand Structure." Proceedings of the ASPRS-ACSM-RT '97 Conference, 7–10 April 1997, Seattle, Washington.

[4] Cohen, W.B. and Spies, T.A. (1992) "Estimating Structural Attributes of Douglas-fir Western Hemlock Forest Stands From landsat and SPOT Imagery." *Remote Sensing of Environment,* 41(1):1–17.

[5] Donchenko, V.V., Goltsova, N.I., Bobylev, L.P., Johannessen, O.M., Kritsuk, S.G. and Pitulko, V.M. (1997) "Synergistic Use of ERS-1/SAR and Other Satellite Data for Polluted Areas Identification for Boreal Forests of St. Petersburg Region." Space at the Service of our Environment, Florence, Italy, ESA.

[6] **Erdas Field Guide** (1997) Fourth Edition, Revised and Expanded. ERDAS ©Inc., Atlanta, Georgia.

[7] Franklin, S.E. and McDermid, G.J. (1993) "Empirical Relation Between Digital SPOT HRV and CASI Spectral Response and Lodgepole Pine *(Pinus contorta)* Forest Stand parameters." *International Journal of Remote Sensing,* 14(12):2331–2348.

[8] Franklin, S.E., Gillespie, R.T., Titus, B.D. and Pike, D.B. (1994) "Aerial and Satellite Sensor Detection of *Kalmia angustifolia* at Forest Regeneration Sites in Central Newfoundland." *International Journal of Remote Sensing,* 15(13):2553–2557.

[9] Guyot, G. and Riom, J. (1988) "Review of Factors Affecting Remote Sensing of Forest Canopies." In: Proceedings of the Seminar on Remote Sensing of Forest Decline Attributed to Air Pollution, EA-5715, Research Project 2661-19, Laxenburg, Austria. 8-1–8-26.

[10] Hagner, O. (1990) "Use of Digital Satellite Data for Stand Delineation and Estimation of Stand Variables by Regression Analysis and Field Inventory." ISPRS Commission VII, Mid-term Symposium, Victoria, Canada.

[11] Holmgren, P. and Thuresson, T. (1998) "Satellite Remote Sensing for Forestry Planning — A Review." *Scandinavian Journal of Forest Research,* 13(1):90–110.

[12] Kadro, A. (1989) "Use of Landsat-TM Data for Forest Damage Inventory." In: Monitoring the Earth's Environment: A Project Campaign on Landsat/TM Applications. ESA SP-1102, Noordwijk:261–270.

[13] Lambert, N.J., Ardo, J., Rock, B.N. and Vogelmann, J.E. (1995) "Spectral Characterization and Regression-based Classification of Forest Damage in Norway Spruce Stands in the Czech Republic Using Landsat Thematic Mapper Data." *International Journal of Remote Sensing,* 16(7):1261–1287.

[14] Li, X. and Strahler, A.H. (1985) "Geometric Optical Modelling of a Conifer Forest Canopy." *IEEE Transactions on Geoscience and Remote Sensing,* GE-23(5):705–721.

[15] Lillesand, T.M. and Kiefer, R.W. (1994) **Remote Sensing and Image Interpretation.** Third Edition. John Wiley & Sons, Inc.

[16] May, A.M.B., Pinder III, J.E. and Kroh, G.C. (1997) "A Comparison of Landsat Thematic Mapper and SPOT Multi-spectral Imagery for the Classification of Shrub and Meadow Vegetation in Northern California, USA." *International Journal of Remote Sensing,* 18(18):3719–3728.

[17] Mayer, K.E. and Fox, L. (1981) "Identification of Conifer Species Groupings from Landsat Digital Classifications." *Photogrammetic Engineering and Remote Sensing,* 48(11):1607–1614.

[18] Nilsson, S. and Shvidenko, A. (1998) "Is Sustainable Development of Russian Forest Sector Possible?" Occasional Paper No. 11, IUFRO, Vienna, Austria.

[19] Roy, P.S., Kaul, R.N., Sharma Roy, M.R. and Garbyal, S.S. (1985) "Forest Type Stratification and Delineation of Shifting Cultivation Areas in the Eastern Part of Arunchai Pradesh Using Landsat MSS Data." *International Journal of Remote Sensing,* 6(3–4):411–418.

[20] Ryherd, S. and Woodcock, C. (1996) "Combining Spectral and Textural Data in the Segmentation of Remotely Sensed Images." *Photogrammetic Engineering and Remote Sensing,* 62(2):181–194.

[21] San Miguel Ayanz, J. and Biging, G.S. (1996) "An Iterative Classification Approach for Mapping Natural Resources from Satellite Imagery." *International Journal of Remote Sensing,* 17(5):957–981.

[22] Schneider, W. (1989) "Forest Damage Mapping in Austria with the Aid of Landsat TM Image Data." In: Monitoring the Earth's Environment: A Project Campaign on Landsat/TM Applications. ESA SP-1102, Noordwijk:223–231.

[23] Shvidenko, A. and Nilsson, S. (1996) "Expanding Forests but Declining Mature Coniferous Forests in Russia." Working Paper 96-59, International Institute for Applied System Analysis, Laxenburg, Austria.

[24] Shvidenko, A. and Nilsson, S. (1997) "Are the Russian Forests Disappearing." *Unasylva,* 188(48-1997):57–64.

[25] Stolz, R. and Mauser, W. (1996) "A Fuzzy Approach for Improving Landcover Classification by Integrating Remote Sensing and GIS Data." Progress in Environmental Remote Sensing Research and Applications. Parlow, Rotterdam, Balkema.

[26] Tokola, T., Pitkanen, J., Partinen, S. and Muinonen, E. (1996) "Point Accuracy of a Non-parametric Method in Estimation of Forest Characteristics with Different Satellite Materials." *International Journal of Remote Sensing,* 17(12):2333–2351.

[27] Trotter, C.M., Dymond, J.R. and Goulding, C.J. (1997) "Estimation of Timber Volume in a Coniferous Plantation Forest Using Landsat TM." *International Journal of Remote Sensing,* 18(10):2209–2223.

[28] Williams, D.L., Ross, F.N. and Dottavio, C.L. (1985) "A Georeferenced Landsat Digital Database for Forest Insect-damage Assessment." *International Journal of Remote Sensing,* 6(5):643–656.

[29] Wilson, B.A. and Franklin, S.E. (1992) "Characterization of Alpine Vegetation Cover Using Satellite Remote-sensing in the Front Ranges, St. Elias Mountains, Yukon-territory." *Global Ecology and Biogeography Letters,* 2(3):90–95.

[30] Woodcock, C. and Harward, V.J. (1992) "Nested Hierarchical Scene Modes and Image Segmentation." *International Journal of Remote Sensing,* 13(16):3167–3187.

[31] Woodcock, C.E., Collins, J.B., Gopal, S., Jakabhazy, V.D., Li, X., Macomber, S., Ryherd, S., Harward, V.J., Levitan, J., Wu, Y. and Warbington, R. (1994) "Mapping Forest Vegetation Using Landsat TM Imagery and a Canopy Reflectance Model." *Remote Sensing of the Environment,* 50:240–254.