SILVA FENNICA 1984, vol. 18 n:o 3: 261–292

A METHOD OF ESTIMATING THE STAND CHARACTERISTICS OF A FOREST COMPARTMENT USING SATELLITE IMAGERY

SIMO POSO, TUOMAS HÄME AND RAITO PAANANEN

Seloste

MENETELMÄ KUVIOITTAISTEN METSIKKÖTUNNUSTEN ESTIMOINTIIN SATELLIITTIKUVIA KÄYTTÄEN

Saapunut toimitukselle 23. 8. 1984

The paper presents a forest inventory method which can be regarded as a special application of two phase sampling. The first phase estimates are obtained from satellite imagery. Second phase estimates are measured in the field. The method is flexible and also applicable to compartmentwise forest inventories. The experiments were based on six study areas with 439 relascope plots. The correlation coefficients between first and second stage estimates varied largely according to the study area.

1. INTRODUCTION

When looking at a satellite picture one can recognize a distinct correlation between spectral values and properties of the forest growing stock. The applicability of satellite pictures to forest inventories has been studied in numerous papers. In Finland, a picture from the weather satellite ESSA 8 was first applied to forest inventory of extensive areas by Kuusela and Poso in 1970. Later, the same authors tested the practicability of Landsat material for a national forest inventory (1975). The computer aided numerical interpretation of satellite imageries started in Finland 1974 when the Technical Research Center of Finland acquired equipment for the purpose.

Numerical interpretation can be divided into supervised and unsupervised methods. In supervised method ground truth areas are located and measured in the field. This material gives rules for the classification of each pixel (a picture element of the satellite imagery referring to the smallest area unit of separate spectral value) to a class defined in forestry terms. For example, a pixel may be estimated for a class the main tree species of which is spruce with a volume varying from 100 to 160 cubic meters per hectare. Supervised interpretation has been applied to delineation of site type compartments for forest taxation purposes (Häme and Saukkola 1982, Häme 1984). Supervised classifition is often based on the maximum likelihood principle, cf. Duda and Hart (1973).

The unsupervised method corresponds to stratified sampling. First, the pixel population is divided into some 5–100 strata which are fairly homogeneous in respect of the spectral values of all of the bands or channels used in the satellite. Later, each stratum is sampled and measured for ground truth. To obtain the final results, the areal proportions of the strata must be studied from satellite imagery and the stratumwise results are obtained by respective field measurements. Saukkola (1982) has applied the method to the estimation of forest tree volumes.

In fact, the supervised method can also be applied to stratification. In that case field work should be made twice, first for stratification and then for the calculation of stratumwise results.

The great advantage in the use of satellite imagery is that the forest population can be easily defined and stratified. Work is going on extensively in the world to develop effective systems for high quality forest inventories. For example, a joint Nordic research project concerning the development of forest inventory methods was started in 1983.

Forest inventories are required for forestry planning in business management or for regional or national policy-making. The planning of separate forest properties has increased its importance in Finland. Today, some 5 % of the forest area is covered annually by a compartmentwise forest inventory. The National forest inventory is based on systematic sampling and it is repeated every seven years.

The objective of the present investigation is to develop methods for the effective combination of satellite imagery with field work and other information sources, such as maps and aerial photographs. Special attention is to be paid to compartmentwise forest inventories.

The presently applicable satellite imagery originates from Landsat 1, 2, 3, or 4 Multispectral scanner (MSS) corresponding a pixel size of 79 m × 79 m in the field and with effective area of some 79 m × 56 m. This resolution power must be regarded as rather poor for compartmentwise forest inventories because a high proportion of pixels fall on compartment boundaries (so called mixed pixels). This handicap will probably decrease in the near future when Landsat 5 material with 30 m × 30 m resolution and SPOT material with 20 m × 20 m and 10 m × 10 m resolution become available.

2. STUDY MATERIAL AND METHODS

2.1 Study material

2.1.1 Study areas, maps and remote imagery

For this study, six areas were chosen. The areas are situated in Southern Finland in the districts of Loppi, Mäntsälä and Renko, and were numbered 13, 16, 17, 18, 19 and 21, figure 1. A general description of the areas is given in 2.1.3. The field inventory was carried out in the first 5 areas in 1980–81 and in area 21 in 1982. 1:10 000 black and white infrared photographs and topographic maps were used for the field work.

The satellite data for this study is from Landsat-3. The imagery (path 204, row 17) was taken on 10th June 1980. Each picture element contains information from an area of 79 × 79 square metres. For each picture element or pixel, the digital image contains the spectral values of four multispectral scanner (MSS) bands, wavelengths 500–600 nm (green), 600–700 nm (red), 700–800 nm and 800–1100 nm (near infrared). The bands are numbered respectively 4, 5, 6 and 7.

2.1.2 Field inventory

Field information was collected using uniform systematic sampling with a plot distance of 50 meters. Each dot, representing a relascope plot with basal area factor 2, was first fixed on aerial photographs and then located on the field. The plots were supplied with x- and y-coordinates for identification. All trees were measured for diameter at breast height and stem height. Stem volume per hectare was computed for each tree using volume equations with diameter at breast height and height as independent variables (Laasasenaho 1982). The tree of each plot with the median basal area was used for age calculations. Site and development classes were assessed visually. The volume and age of each plot were also adjusted by visual estimations to represent the immediate surroundings of the plot. Species composition, timber assortment proportions and treatment requirements for each plot were also assessed. If a plot was situated on the border of two stand compartments, it was divided according to the area of each compartment on the plot. The information was given separately for each part of a divided plot.

Table 1. General statistics of the study areas.

area number	size ha	site class (mode and its proport.)	development class (mode and its proport.)	age years (mean)	pine proportion of volume in % (mean)	spruce proportion of volume in % (mean)	broadleaves proportion of volume in % (mean)	mean volume cu.m/ha	number of plots complete	number of plots complete+ divided
13	7.75	1B (45%)	4 (39%)	83	36	57	7	189	24	38
16	21.75	1B (42%)	4 (69%)	95	27	66	7	260	74	100
17	17.0	2 (39%)	4 (35%)	71	3	90	7	155	53	83
18	11.25	1A (54%)	4 (51%)	77	41	52	7	225	35	55
19	22.6	2 (70%)	2 (57%)	47	42	47	11	139	78	103
21	49.75	1A (54%)	3 (55%)	64	16	63	21	174	175	223
total	130.1	1A (33%)	3 (35%)	70	24	63	13	185	439	602

Site class codes (Kuusela 1978):

0 non-forest land; 1A very rich and rich sites; 1B damp sites; 2 sub-dry sites; 3 dry and barren sites and spruce and birch swamps on forest land; 4 pine swamps on forest land; 5 poorly productive land or waste land

Development class codes:

0 open area; 1 seedling stand; 2 young thinning stand; 3 advanced thinning stand; 4 mature stand; 5 shelterwood stand; 6 low-yielding stand (e.g. understocked)

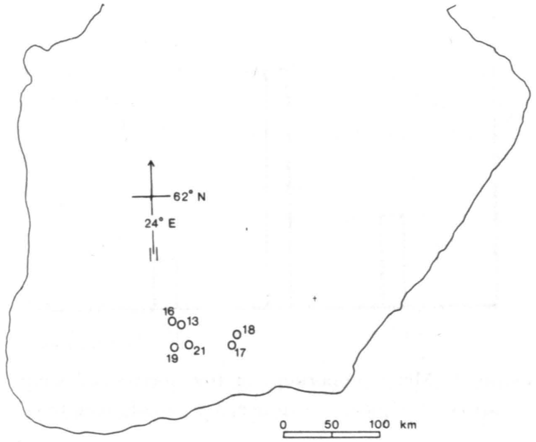


Figure 1. Location of the study areas.

2.1.3 General description of the study areas

The distributions of site class (classification of the productivity potential of forest land), development class, age and volume and mean proportions of tree species for all the study areas together are shown in figures 2–6. The means or modes of the stand characteristics are given in table 1 for each area separately and for all areas together.

The size of the areas varies from 7.75 to 49.75 hectares. Area 21 is the largest. The mean age of the growing stock on the study areas is 70 years and the mean volume per hectare is 185 cu.m. The mean volume of the growing stock in Southern Finland, according to the latest national forest inventory, is 101 cu.m./ha (Yearbook of forest statistics 1982). Over half of the growing stock consists of spruce-dominated stands. There exists no pi-

---

From the authors Poso is responsible for the overall planning of the project and for writing of chapters one and four. Häme worked as a specialist in combining of the satellite and field data in the form which made the further calculations and final analyses possible. He also wrote some parts of chapters two and four. Paananen has made full day work for the project from February 1983 on. He made most of the computer calculations and wrote chapters two and three. It should be emphasized, however, that all chapters are finally based on cooperative handling by all the authors.

The study was made at the Department of Forest Mensuration and Management of the University of Helsinki and at the Technical Research Centre of Finland. Eljas Heikkinen, MF, and forestry students Liisa Lundmark and Aaro Mikkola contributed in data handling and in field measurements. Help in data handling and computer programming was obtained by Yrjö Rauste from the Technical Research Centre of Finland and by Visa Rauste from the Computing Centre of University of Helsinki. The text was checked by Dr. Ashley Selby.

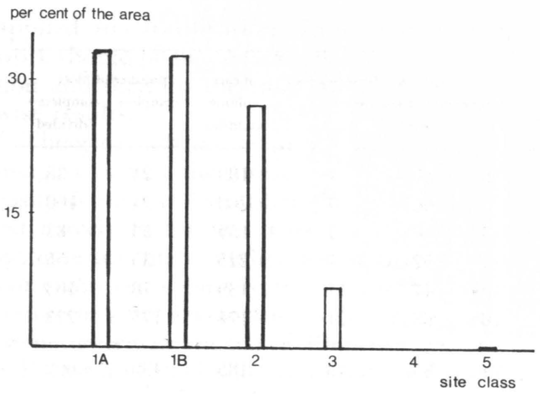


Figure 2. The distribution of site classes, all study areas. (Class codes are the same as in table 1.)

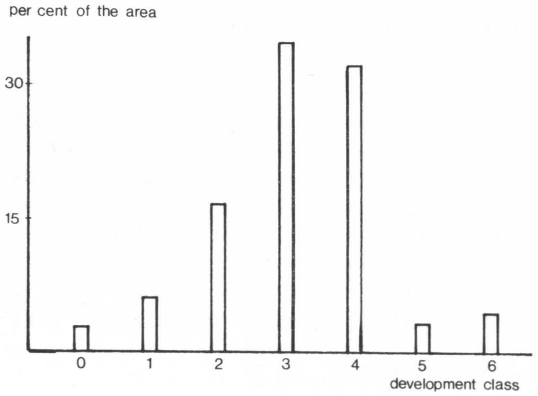


Figure 3. The distribution of development classes, all study areas. (Class codes are the same as in table 1.)

ne in over 40 % of the plots. The proportion of broadleaved species is zero in nearly 50 % of the plots. The dominant site-type is Oxalis-Myrtillus-type, comprising one third of the total area. The Oxalis-Myrtillus site type had the highest productivity potential of forest soils in the study areas.

Area 13 is the smallest and represents the forests typical for this study. Nearly 40 % of the stands are mature and mainly spruce-dominated. There are no plots classified as development class 2, so there is a remarkable lack in young thinning stands. Area 16 possesses the greatest stand volumes, the mean volume amounting to 260 cu.m/ha. The stands are mainly mature spruce stands. Area 17 consists of rather poor spruce stands, nearly 40 % of the plots belong to site class 2 (Vaccinium-type or equivalent) and 13 % of the areas are classified as low-yielding. Area 18 is also rather typical. Half of the area consists of mature mixed pine-spruce stands on Oxalis-Myrtillus-type sites. The proportion of pine is rather high and so is the mean volume. Rather young stands (mean age 47 years, with no mature stands) are characteristic of area 19. The sites are rather poor, mainly Vaccinium-type or ditched peatlands. The proportion of pine is high and the mean volume is understandably the lowest of all the study areas. The proportion of broadleaved species is greatest in area 21. Over a half of the plots are classified as site class 1A, so the sites are mainly rich. Area 21 is over two times larger than the next largest, so the mean values of the stand characteristics for all areas are strongly weighted by the values of area 21.

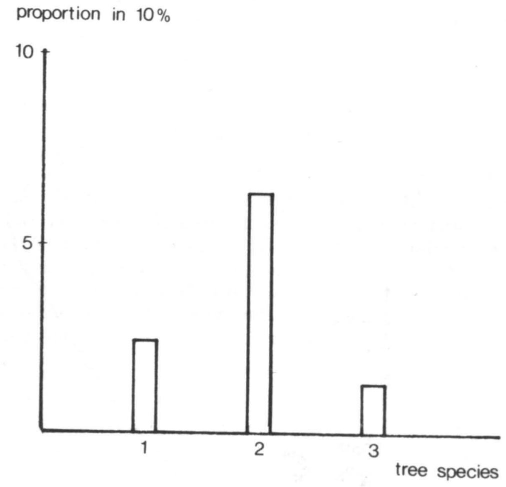


Figure 4. Mean proportions of tree species, all study areas. (1 = pine, 2 = spruce, 3 = broadleaved trees)

2.2 Creating the files

2.2.1 Basic files, combining the field and satellite information

There were two types of basic files, plot-based and pixel-based. Both files were created in the same coordinate system. Those pixels that represented the study areas and their surroundings were printed by Versatec grey scale plotter onto greymaps (see appendix 1) at the same scale (1:10 000) as the aerial photographs. The coordinates of the pixel centers were determined manually by comparing the greymaps and the aerial photographs.

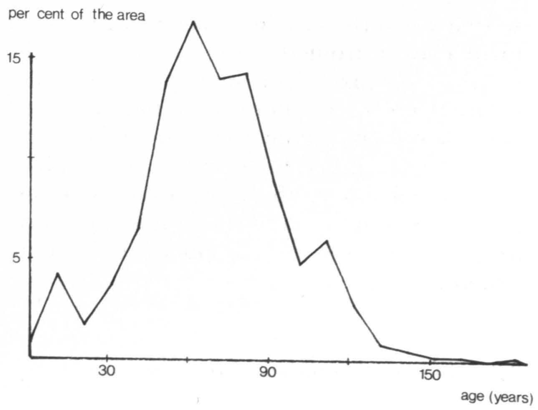


Figure 5. The distribution of age, all study areas.

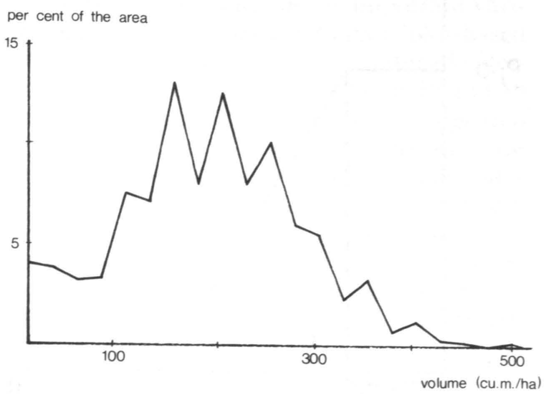


Figure 6. The distribution of volume, all study areas.

Using the coordinates of the field plots and satellite pixels it was possible to calculate the spectral values for each field plot and similarly, the field information for each pixel. In the plot-based system, spectral values, q, for each plot or part of it were computed using distance weighted means of pixels falling near the plot. The following procedure was applied:

1. For each plot, a maximum number of four pixels were sought, the centers of which were located not further than 55 metres (or, in the other experimental case, 85 metres) from the plot center. Consequently, the number of pixels for a field plot, n, varied from 1 to 4.
2. Weight w_i for each pixel was calculated by

$$w_i = \frac{d_1}{d_i}, i = 1, \ldots, n \qquad (1)$$

where w_i is weight for the i:th nearest pixel
d_1 distance of the nearest pixel center from the plot
d_i distance of the i:th nearest pixel from the plot center.

3. For each plot, the means of spectral values were calculated with equation

$$q_j = \frac{\Sigma\,(w_i \cdot q_{i,j})}{\Sigma\,w_i} \qquad (2)$$

where q_j is the weighted mean of spectral values of band j
$q_{i,j}$ the spectral value of band j on pixel i

In the pixel-based files stand variable values were calculated for each pixel as distance weighted means of the nearest plot values. The following rules were applied:

1. For each pixel, n plots or parts of them (the maximum number was 8) were sought, located not further than 55 metres (or, in the other experimental case, 65 metres) from the pixel center.
2. Weights $w_{k,i}$ for each plot or part of it were calculated, as follows

$$w_{1,i} = p_i\,f_i, i = 1, \ldots, n \qquad (3)$$

$$w_{2,i} = p_i\,f_i\,t_i, i = 1, \ldots, n \qquad (4)$$

where f_i is the relative size of a part of a plot (if the plot is undivided, $f_i = 1$)
p_i weight depending on the distance (figure 7)
t_i stand volume (cu.m./ha) of the plot or part of it

3. The weights w_1 and w_2 were scaled so that the sum of the weights was the number of the plots or parts of plots found.
4. The weighted means for the stand variable values were calculated by

$$\bar{x}_j = \frac{\Sigma\,(w_{k,i} \cdot x_i)}{\Sigma\,w_{k,i}} \qquad (5)$$

where x_i is the value of the stand variable on plot i
$\bar{x}_j$ weighted mean of the stand variable values
w_k is valued either w_1 or w_2 depending on whether the value of the variable is weighted with volume or not. In these calculations, the proportions of pine, spruce and broadleaved trees and proportion of sawlog timber were weighted with volume.

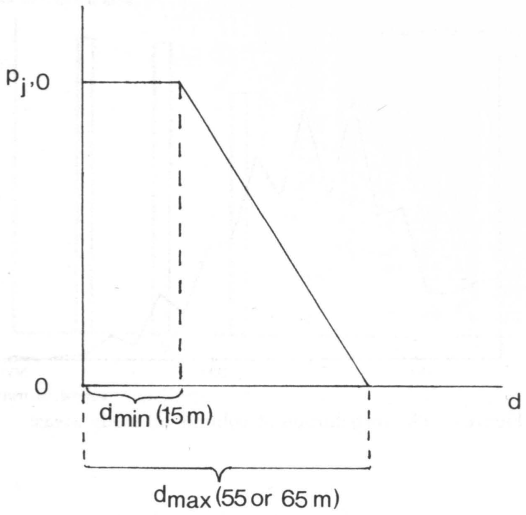


Figure 7. Weight p_i as function of distance from the pixel center.

In the calculation of development class means, plots which were classified to be 'low yielding' were ignored. In addition to means the modes of development classes were computed. If the frequency of development class 'low yielding' exceeded half of the whole frequency, the mean was replaced by mode.

After the aerial photo coordinates were given for pixels their accurracy was tested. First, the correlation coefficient was calculated between the value of a stand variable on a plot and the spectral value of the pixel nearest to the plot. The variables used were the spectral values of band 5 and volume, or the spectral values of band 7 and the proportion of broadleaved trees. Then the whole coordinate system of the satellite imagery was shifted to the direction of the x- or y- axis on the coordinate system of the aerial photographs. The shift might be 5 metres, for example. After the shift, a new correlation coefficient was calculated, and the the satellite imagery was shifted again. The result of the shifts was a matrix of correlation coefficients. The greatest shift from the original position was 250 metres. The hypothesis was that there should be a higher correlation at the correct position of the satellite imagery than in other positions.

No clear maxima in the correlation coefficients was found. Only the pixels falling to the area in the original position were in use, so that when the satellite imagery was shifted further away from the original position, the number of pixel-field plot pairs was decreased. The correlations could have changed because one edge of an area, where the correlations were either high or low, was subsequently left out of the calculation. The test gave no reason to change the original position of pixels in the aerial photo coordinate system.

2.2.2 Additional variables of the files

The study material was transformed using principal component analysis. The idea of principal component analysis is to replace a large number of highly correlated variables with a relatively small number of uncorrelated variables, with little loss of information. In this study, the purpose was to investigate the basis for using 2 or 3 principal components instead of the original variables.

First, the principal components were computed from the spectral values of bands 4–7 using the plot-based files with either 55 or 85 m maximum distance of pixels or values of the nearest pixel. All study areas were used in the analyses. The computation employed either all the plots or, in the other experimental case, only complete (undivided) plots.

Principal component analysis was also carried out for some stand variables. Principal components were computed using mean height, mean diameter, volume, age, site class and proportions of spruce and broadleaved trees as variables. The same alternatives were used as in the computation of principal components of spectral values.

After the principal component scores had been computed, the following stand variables and spectral value variables were used in the analyses of the plot-based files:

Stand variables
1. site class (codes 0–4)
2. development class (codes 0–4)
3. age (years)
4. proportion of pine (values 0–10)
5. proportion of spruce (0–10)
6. proportion of broadleaved trees (0–10)
7. volume (cubic metres per hectare)
8. principal component 1 of the stand variables (SF1)
9. principal component 2 of the stand variables (SF2)

Spectral value variables
1. weighted mean of the spectral values, band 4
2. weighted mean of the spectral values, band 5
3. weighted mean of the spectral values, band 6
4. weighted mean of the spectral values, band 7
5. value of the nearest pixel, band 4
6. value of the nearest pixel, band 5
7. value of the nearest pixel, band 6
8. value of the nearest pixel, band 7
9. principal component 1 of the spectral value variables 1–4 (PF1)
10. principal component 2 of the spectral value variables 1–4 (PF2)
11. principal component 1 of the spectral value variables 5–8 (LPF1)
12. principal component 2 of the spectral value variables 5–8 (LPF2)

In the subsequent analyses the plot based files were combined as follows:

type of spectral value	field plots used	
	complete and divided	complete
value of the nearest pixel	AL	FL
weighted mean, maximum distance 55 m	A55	F55
weighted mean, maximum distance 85 m	A85	F85

Files A55, A85, F55 and F85 contain spectral value variables 1–4 and 9–10. Files AL and FL contain spectral value variables 5–8 and 11–12.

2.3 Correlation calculations

Mainly the files with only complete plots (files FL, F55 and F85) were applied in the correlation calculations. Where both complete and divided plots are used, they are mentioned separately. The use of divided plots is problematic. One part of a plot may consist of seedlings and the other part of mature forest, but they obtain the same spectral values, which in fact represent neither of the parts of the plot, but something between them.

Simple correlation matrices and correlation diagrams were used to investigate the dependence between spectral values and stand characteristics. The correlation matrices were computed for all the important variables for all study areas together (plot-based files) and for areas 16 and 19 combined (plot-based and pixel-based files). Areas 16 and 19 were used because the correlations appeared to be highest there. Correlation diagrams for the plot-based files were printed using spectral values computed on the basis of a maximum pixel distance of 55 m. Some diagrams were computed from the pixel-based files to indicate the differences between the alternative ways of combining the field and satellite information.

2.4 Estimation of stand variables to predelineated compartments by means of spectral values

The aim of the delineation of compartments was to obtain fixed borderlines for management compartments. Depending on their size the compartments can be more or less heterogeneous. To study the effect of the compartment size, 2–3 different delineations were applied. The estimation procedure aims to determine the distributions of separate stand variables for the delineated compartments by means of spectral values. This can be done by using some model areas having a sufficient number of sample plots with both field and satellite data. In these model areas, those stand variable distributions are estimated which represent different combinations of spectral values. (In the ideal case, one spectral value combination represents one combination of stand variable values.) By means of the model distributions stand variable distributions can be estimated for any other area if only the frequencies of spectral value combinations on that area are known. Saukkola (1982) has used a method in forest inventory which is relatived to this method. He has stratified the area to be inventoried by two band parallelepiped classification. (The procedure is described e.g. by Lillesand and Kiefer 1979.) Then he has inventoried the strata in the field and reported the contents of every stratum as a distribution. Saukkola used distributions only for discrete variables such as development class.

In the present study, the first step was to take some of the study areas as model areas

(areas 16 and 19 were selected). The selection of the model areas was made on the basis of correlations. For the model areas, a stratification of plots (unsupervised classification) was carried out using principal components 1 and 2 of the spectral values (another trial was made using weighted means of the spectral values of bands 5 and 7). The number of strata was decided beforehand. The stratifying computer program BMDPKM (see BMDP Statistical Software 1981) classified each plot to one stratum according to the euclidean distances between the plot values and the stratum centers. Each plot belonged to the stratum whose center was nearest to the plot values. In this study, stratification was made either to 15 or 20 strata. The variable values used in the process were unstandardized.

The composition of the strata (principal component stratification) for the model areas are listed in appendix 2 (15 strata) and appendix 3 (20 strata). The stratification procedure is also illustrated in appendix 4, where the boundary values of principal components 1 and 2 for each stratum are printed to form rectangules.

The result of the clustering can be presented in matrix form, where each stratum h forms a separate matrix A_h.

$$A_h = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & \cdots & \cdots & a_{m,n} \end{bmatrix} \tag{6}$$

where $a_{i,j}$ is the value of stand variable j in plot i
m the number of plots in stratum h (varies by strata)
n the number of stand variables measured in each plot of stratum h (equal to each stratum)

From the A_h matrices, the proportion of each variable value or variable class $p(a_j)$ in the area to be estimated can be calculated using the equation

$$p(a_j) = \sum_{h=1}^{k} \left(\frac{e_h}{m_h} \cdot \frac{f_h}{\sum_{h=i}^{k} f_h}\right) \tag{7}$$

where $p(a_j)$ is the proportion of variable class a_j in the area to be estimated
e_h the number of field plots in stratum h with a certain value of a_j
m_h the number of field plots in stratum h
$\sum_{h=1}^{k} f_h$ the number of plots belonging to stratum h according to the spectral values of the plots (in the area to be estimated)
f_h the number of plots in the area to be estimated

In equation (7), the first part is the proportion of the plots in stratum h with the certain stand variable value. The second part is the proportion of the plots belonging to stratum h in the area to be estimated.

Using the above procedure, it is possible to estimate distributions for any stand variable measured in the field. In this study, the respective real distributions could be calculated directly on the basis of field measurements. The accurracy of estimations could be based on comparisons of the estimated and real distributions.

3. RESULTS

3.1 Correlation between stand characteristics and spectral values

3.1.1 *Loadings of the principal components of spectral values and stand variables*

The loadings of the principal components of spectral values are shown in table 2. The loadings are computed from the plot-based file (F55) using all study areas. Principal component 1 (PF1) explains 66 % and PF1+PF2 together 87 % of the total variance between the spectral values of the four bands.

The principal component loadings of the stand variables are shown in table 3. Principal component 1 (SF1) explains 44 % of the total variance. Age, volume, mean diameter and mean height are strongly loaded on SF1. Site class and proportion of spruce are strongly loaded on SF3. The three first principal components together explain 79 % of the total variance.

3.1.2 *Correlation tables*

Correlation coefficients of the most important variables are shown in table 4 for all study areas with complete plots and maximum distance of 55 m (file F55).

The highest correlation coefficient, −0.50, is found between principal component 1 (PF1) and age. In general, stand age seems to be more strongly correlated with spectral values than other stand variables. The correlation of stand variables with the weighted means of the spectral values is slightly better than with the values of the nearest pixel.

The correlation of principal component 1 (PF1) was good with some variables, e.g. age, development class and volume. Principal component 2 (PF2) showed a rather high correlation with site class and with proportions of pine and broadleaves. Development class correlates best with weighted mean of

Table 2. Loadings of the principal components of spectral values, file F55, all study areas.

	PF1	PF2	PF3	PF4
weighted mean, band 4	0.49	0.46	−0.02	−0.02
band 5	0.73	0.50	−0.46	0.02
band 6	0.90	−0.37	0.00	0.23
band 7	0.86	−0.46	0.00	−0.22
variance explained	2.63	0.84	0.43	0.10

Table 3. Loadings of the principal components of stand variables, file F55, all study areas.

	SF1	SF2	SF3
site class	−0.42	0.73	−0.31
age	0.70	0.55	−0.13
volume	0.76	−0.09	−0.04
proportion of spruce	0.34	−0.65	−0.49
proportion of broadleaves	−0.14	−0.06	0.89
mean height	0.95	0.06	0.15
mean diameter	0.89	0.15	0.15
variance explained	3.07	1.30	1.18

Table 4. Correlation table, all study areas, file F55.

	site class	development class	age	proportions of pine	proportions of spruce	proportions of broadleaves	volume	principal components SF1	principal components SF2
weighted mean, band 4	−0.01	−0.31	−0.37	0.20	−0.22	0.02	−0.25	−0.33	−0.08
band 5	0.22	−0.32	−0.37	0.39	−0.36	−0.07	−0.34	−0.45	−0.11
band 6	−0.19	−0.21	−0.45	−0.04	−0.14	0.31	−0.29	−0.32	−0.26
band 7	−0.24	−0.21	−0.43	−0.05	−0.15	0.34	−0.24	−0.27	−0.27
nearest pixel value, band 4	−0.02	−0.29	−0.34	0.19	−0.20	0.00	−0.23	−0.30	−0.08
band 5	0.22	−0.30	−0.32	0.37	−0.33	−0.08	−0.33	−0.41	0.13
band 6	−0.16	−0.22	−0.43	−0.02	0.15	0.29	−0.29	−0.31	−0.23
band 7	−0.22	−0.20	−0.41	−0.05	−0.14	0.33	−0.24	−0.26	−0.26
principal component PF1	−0.09	0.32	−0.50	0.13	−0.26	0.21	−0.34	−0.41	−0.17
principal component PF2	0.34	−0.17	−0.01	0.39	−0.20	−0.35	−0.09	−0.17	0.28

Table 5. Correlation table, areas 16+19, file F55.

	site class	development class	age	proportions of pine	proportions of spruce	proportions of broadleaves	volume	principal component SF1
weighted mean, band 4	0.36	−0.75	−0.73	0.37	−0.41	0.04	−0.67	−0.74
band 5	0.40	−0.73	−0.73	0.56	−0.59	−0.01	−0.71	−0.77
band 6	0.36	−0.71	−0.70	0.36	−0.46	0.17	−0.70	−0.74
band 7	0.33	−0.68	−0.70	0.31	−0.41	0.17	−0.64	−0.70
principal component PF1	0.41	−0.82	−0.82	0.47	−0.54	0.10	−0.78	−0.84
principal component PF2	0.28	−0.52	−0.50	0.44	−0.41	−0.14	−0.46	−0.53

the pixel values of band 5 and with principal component PF1. The coefficient is −0.32. Correlation with the weighted mean of the spectral values of band 4 is only slightly lower. Only classes 1−4 were used in calculating the correlations for development class.

Volume is often regarded as the most important stand variable to be estimated. In this material, the best correlations with volume (−0.34) occurred with principal component PF1 and with the weighted mean of the spectral values of band 5.

The highest correlations for the proportion of conifers (pine and spruce) were found with the weighted mean of the spectral values of band 5, principal component PF2 and the nearest pixel value of band 5. Coefficients are, however, smaller than 0.40. Broadleaved species can best be separated on infrared wavelengths. However, the highest correlation of the proportion of broadleaves is with principal component PF2. The first principal component calculated from the stand variables, SF1, correlates better with spectral value variables than the second, SF2. Band 5 correlates best with SF1 (−0.45). Principal components SF1 and PF1 also show a rather strong correlation (−0.41).

When the study areas were examined separately, it was noticed that the best correlations between stand variables and spectral values appeared in areas 16 and 19. The most interesting correlations for areas 16+19 are printed in table 5. The coefficients are about 0.3−0.4 units higher than in the material as a whole. When, for example, the highest correlation coefficient for the whole material is −0.50, it is −0.84 for areas 16+19 (between principal components PF1 and SF1). In general, correlations which are highest in the whole material are highest also in areas 16+19. The best correlation of volume appears with principal component PF1, with a coefficient of −0.78.

Table 6 shows the correlation coefficients when the spectral value variables are computed using weighted means of the spectral values with either 55 (F55) or 85 (F85) metres maximum distance, or values of the nearest pixel (FL). It can be seen that 85 meters maximum distance gives somewhat better correlations than 55 meters maximum distance. The lowest correlations appear in file FL. The differences between the files are, however, rather small.

Table 6. Correlation coefficients for files F55, F85 and FL, all study areas.

		site class	age	proportion of pine	volume	principal components SF1	principal components SF2
band 4	F55	−0.01	−0.37	0.20	−0.25	−0.32	−0.08
	F85	−0.03	−0.42	0.21	−0.27	−0.35	−0.10
	FL	0.02	−0.34	0.19	−0.23	−0.30	−0.08
band 5	F55	0.22	−0.37	0.39	−0.34	−0.45	0.11
	F85	0.22	−0.40	0.42	−0.35	−0.48	0.12
	FL	0.22	−0.32	0.37	−0.33	−0.41	0.13
band 6	F55	−0.19	−0.45	−0.04	−0.29	−0.32	−0.28
	F85	−0.21	−0.45	−0.04	−0.29	−0.31	−0.27
	FL	−0.17	−0.43	0.02	−0.29	−0.31	−0.23
band 7	F55	−0.24	−0.43	−0.05	−0.25	−0.27	−0.27
	F85	−0.26	−0.44	−0.07	−0.24	−0.26	−0.29
	FL	−0.22	−0.41	−0.05	−0.24	−0.26	−0.26
PF1	F55	−0.09	−0.50	0.14	−0.35	−0.42	−0.17
	F85	−0.10	−0.51	0.14	−0.34	−0.41	−0.18
LPF1	FL	−0.08	−0.48	0.13	−0.34	−0.40	−0.16
PF2	F55	0.34	−0.01	0.40	−0.01	−0.17	−0.28
	F85	0.36	−0.03	0.43	−0.10	−0.21	−0.30
LPF2	FL	0.30	0.00	0.37	−0.08	−0.16	−0.26

Table 7. Some correlations of the pixel-based file, areas 16+19.

	band 4	band 5	band 6	band 7
weighted mean of age	−0.62	−0.58	−0.52	−0.44
weighted mean of volume	−0.53	−0.50	−0.46	−0.39

To illustrate the significance of the location of pixels in relation to the plot, some correlations between volume and separate spectral value variables (file F55) for all study areas are shown in figure 8. The correlations are computed using plots which are classified according to the distance of the nearest pixel center from the plot. It can be seen that the correlations in bands 4 and 5 are highest when the distance to the nearest pixel is 10−30 meters. In that case, usually more than two pixels fall inside the maximum distance and their spectral values are weighted rather evenly. When the distance to the nearest pixel is under 10 meters the weights of the farthest pixels are small and their spectral values become almost meaningless in the calculation of the weighted means. As opposed to bands 4 and 5, the correlations with bands 6 and 7 are highest when the nearest pixel center falls nearer than 10 meters to the plot.

The correlations of the pixel-based file with maximum distance of 65 meters are shown in table 7. Only areas 16+19 are used. Correlation coefficients are not as high as those obtained in the plot-based files. Band 4 correlates best with age and volume. As with the plot-based files, age correlates better with spectral values than volume.

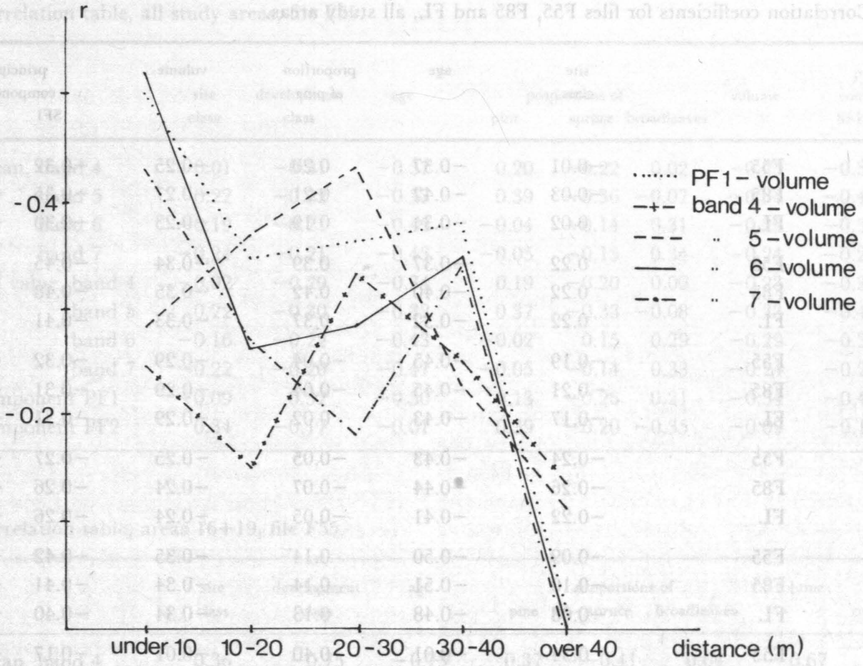


Figure 8. Correlation coefficient r between stand volume and some spectral value variables as a function of distance of the nearest pixel from the plot. (Spectral values of bands 4–7 are calculated as weighted means of 2–4 pixels with maximum distance of 55 m.)

3.1.3 Correlation diagrams

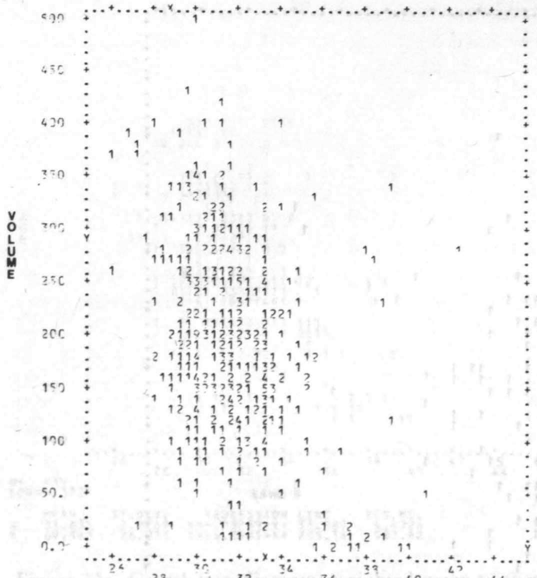
Correlation diagrams between weighted mean of the spectral values of band 5 and volume for all areas and areas 16+19 are plotted on figures 9a and 9b. Figure 9a shows, for example, that for spectral value 31.0 the volume varies from 0 to 430 cu.m./ha. and correspondingly, for one volume value, spectral values can vary by over 10 units. Areas 16+19 show a better correlation, but there also some plots deviate drastically from the general pattern of the diagram. The use of logarithmic transformations of volume did not improve the correlations. Figures 10a and 10b show the correlation between principal component PF1 and volume for all areas and areas 16+19. Values of the principal component scores range from −2.5 to 4.0. Figure 10b shows a rather good correlation.

Correlation diagrams between weighted mean of the spectral values of band 5 and age are plotted in figures 11a and 11b. Again, outputs for areas 16+19 show distinctly higher correlation than those for all areas. The diagrams between principal component PF1 and age look similar to diagrams 11a and 11b, and they are not presented here. Diagrams 11a and 11b show that the correlation is better when age is less than 70 years. The coefficient for all areas and age under 70 years is −0.55.

Figure 12 shows the correlation between principal components PF1 and SF1 for areas 16+19. The diagram shows that the use of principal components, especially principal components of the spectral values, is useful in this kind of study.

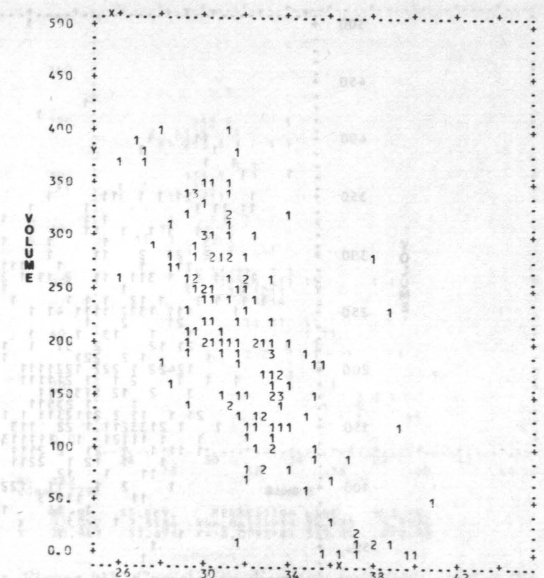
The correlation diagram for all areas between principal component PF2 and site class is plotted in figure 13. The diagram shows that the principal component scores representing one class may vary from −2.5 to 2.5.

The correlation diagram calculated from the pixel-based file with 65 metres maximum distance between pixel value of band 5 and weighted mean of volume is plotted in figure 14. Only areas 16+19 were used in the computation. The pattern of the diagram looks rather good, but some pixels receive unexpected values due to pixel location on the border of forest and agricultural land. In general, the diagram is more scattered than those obtained from the plot-based files. When the four outliers on the right side of the diagram were removed, the correlation increased to −0.68.



```
N=   439
COR=-.3382
     MEAN      ST.DEV.     REGRESSION LINE       RES.MS.
X   31.635     2.2769   X=-.00848*Y+ 33.270     4.5994
Y   192.88     91.001   Y=-13.543*X+ 621.31     7347.1
```

Figure 9a. Correlation diagram: weighted mean of the spectral values, band 5 – volume, all study areas.



```
N=   152
COR=-.7075
     MEAN      ST.DEV.     REGRESSION LINE       RES.MS.
X   32.057     2.8586   X=-.02020*Y+ 36.061     4.1085
Y   198.22     100.11   Y=-24.777*X+ 992.48     5038.7
```

Figure 9b. Correlation diagram: weighted mean of the spectral values, band 5 – volume, study areas 16+19.

3.2 Results based on experimentation

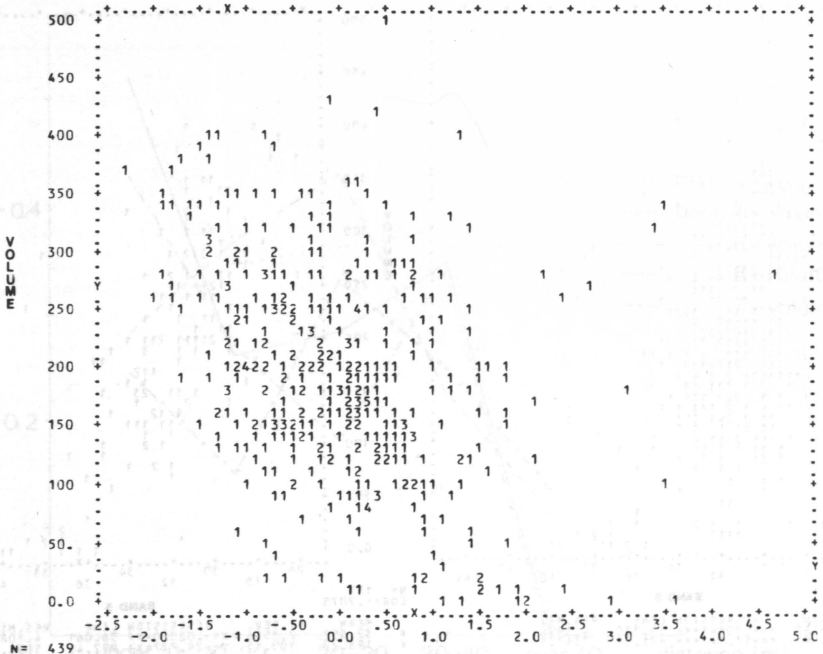
3.2.1 Variation within strata

Some statistics concerning the 15 strata of the model areas 16+19 (principal component stratification) are given in table 8. The number of plots in a stratum varies from 2 to 26. The coefficient of variation inside one stratum can be as high as 198 %. The difference between the smallest and largest variable values inside one stratum varies for volume from 10 (stratum 15) to 210 (strata 3,10 and 11) cu.m./ha. The proportion of the within strata variance of the total variance is 22 % for age and 27 % for volume.

Stratum 1 contains plots belonging to site class 3 and all pure pine stands. The other strata seem to be more heterogeneous. Stratum 6 contains only development class 2, stratum 8 only development class 3 and no broadleaved trees. Stratum 13 contains only site class 3 and no broadleaved trees. The best strata for volume are 7 and 15 (stratum 15 having only 2 plots).

3.2.2 Results calculated for compartments

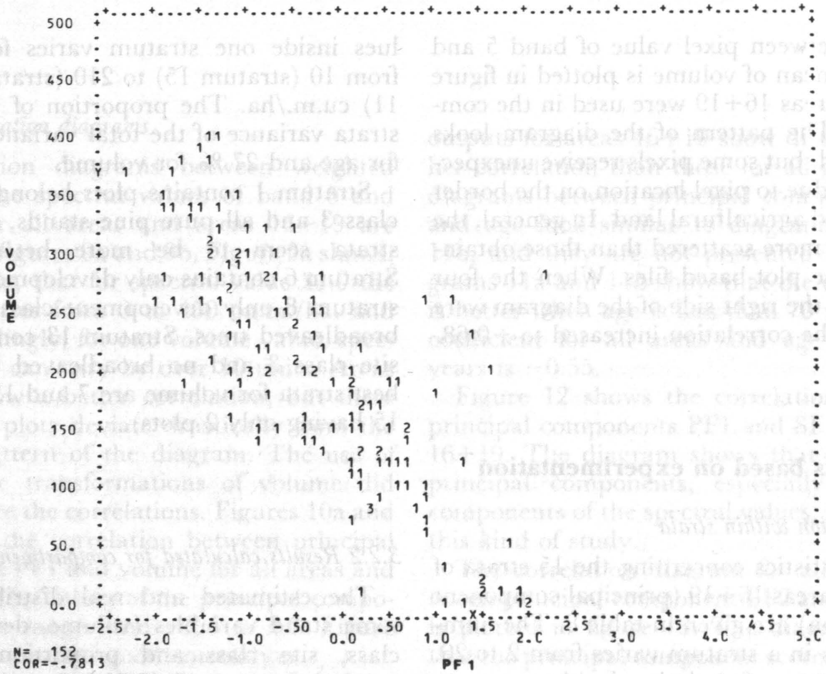
The estimated and real distributions of some stand variables (volume, development class, site class and proportions of tree species) for area 13 (7.75 ha) and one compartment of area 21 (4.25 ha) are given in figures 15–22. Table 9 shows the estimates



```
N=  439
COR=-.3452

   MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X  33E-12     1.0000   X=-.00379*Y+ .73175    .88282
Y  192.88     91.001   Y=-31.417*X+ 192.88    7310.8
```

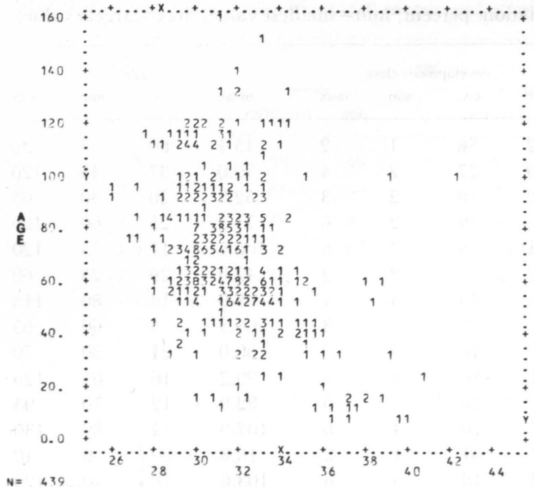
Figure 10a. Correlation diagram: principal component PF1 -volume, all study areas.



```
N=  152
COR=-.7813

   MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X  -.27294    1.0328   X=-.00806*Y+ 1.3249    .41830
Y  198.22     100.11   Y=-75.731*X+ 177.55    3930.0
```

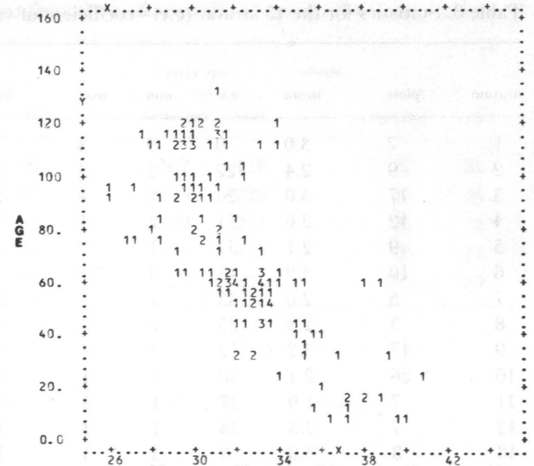
Figure 10b. Correlation diagram: principal component PF1 -volume, study areas 16+19.



```
N=  439
COR=-.3731

   MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X  31.635    2.2769   X=-.03239*Y+ 33.889    4.4726
Y  69.581    26.231   Y=-4.2985*X+ 205.56    593.62
```

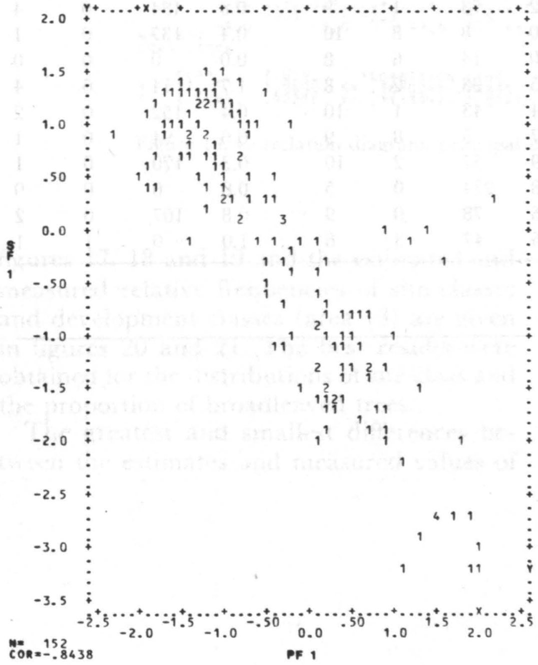
Figure 11a. Correlation diagram: weighted mean of the spectral values, band 5 – age, all study areas.



```
N=  152
COR=-.7292

   MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X  32.057    2.8586   X=-.06581*Y+ 36.725    3.8524
Y  70.947    31.674   Y=-8.0795*X+ 329.95    472.98
```

Figure 11b. Correlation diagram: weighted mean of the spectral values, band 5 – age, study areas 16+19.



```
N=  152
COR=-.8438

   MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X  -.27294    1.0328   X=-.68974*Y-.46016    .30933
Y  -.27144    1.2634   Y=-1.0322*X-.55316    .46289
```

Figure 12. Correlation diagram: principal component PF1 -principal component SF1, study areas 16+19.

and (field) measured values of means and standard deviations of age and volume in some other compartments.

The relative frequency distributions of volume in area 13 (figure 15) show that the distribution of the estimated values is wider than for the field measured values. The estimated mean volume amounts to 216 cu.m./ha (with a standard deviation of 37 %) while the mean volume measured in the field is 192 cu.m./ha (with a standard deviation of 30 %). The over estimation is therefore +12 %.

The relative frequency distributions of volume in one compartment of area 21 are shown in figure 16. The estimated mean volume is 157 cu.m./ha (with a standard deviation of 39 %) and the measured mean volume is 229 cu.m./ha (with a standard deviation of 41 %). The underestimation is therefore −31 %. However, this is partly due to the difference in the measurement time of the model areas (1980–81) and area 21 (1982). If the growth of two years is added to the estimated total volume, the difference is decreased to −25 %.

The estimated and real relative frequencies of the proportions of tree species are given in

Table 8. Statistics for the 15 strata. (c.v.=coefficient of variation, percent, min=smallest value, max=largest value)

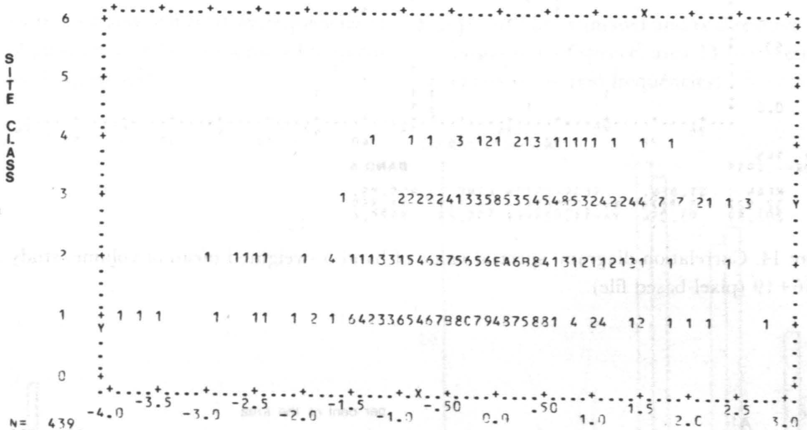
stratum	plots	site class mean	site class c.v.	site class min	site class max	development class mean	development class c.v.	development class min	development class max	age mean	age c.v.	age min	age max
1	7	3.0	0	3	3	1.2	38	1	2	15.1	60	7	30
2	9	2.4	22	2	3	2.9	27	2	4	70.0	37	45	120
3	17	3.0	20	2	4	2.2	18	2	3	52.3	20	30	65
4	12	2.6	31	2	4	3.5	39	2	6	71.3	25	60	120
5	9	2.1	37	1	3	4.1	19	3	6	100.6	17	75	120
6	10	3.2	13	3	4	2.0	0	2	2	42.0	29	25	60
7	5	2.0	35	1	3	3.6	15	3	4	95.0	13	80	115
8	3	2.3	25	2	3	3.0	0	3	3	63.3	5	60	65
9	17	3.2	12	3	4	2.1	16	2	3	50.0	21	30	70
10	26	2.1	38	1	4	4.1	16	3	6	99.2	16	65	120
11	7	1.9	37	1	3	3.7	20	3	5	82.9	12	70	95
12	7	2.3	33	1	3	4.7	20	4	6	107.9	14	80	130
13	8	3.0	0	3	3	1.1	31	1	2	16.8	59	9	40
14	11	2.5	44	1	4	4.2	18	3	6	103.6	18	70	120
15	2	2.5	28	2	3	4.0	35	3	5	77.5	23	65	90

stratum	plots	proportion of pine mean	proportion of pine c.v.	proportion of pine min	proportion of pine max	proportion of spruce mean	proportion of spruce c.v.	proportion of spruce min	proportion of spruce max	proportion of broadleaves mean	proportion of broadleaves c.v.	proportion of broadleaves min	proportion of broadleaves max
1	7	10.0	0	10	10	0.0	0	0	0	0.0	0	0	0
2	9	2.9	94	0	9	6.6	43	0	10	0.6	182	0	3
3	17	3.2	77	0	8	4.8	40	2	8	2.0	85	0	7
4	12	2.4	74	0	6	5.7	37	1	8	1.9	114	0	8
5	9	2.2	112	0	7	7.6	31	3	10	0.2	198	0	1
6	10	5.0	74	0	9	4.2	73	1	9	0.8	184	0	4
7	5	0.6	149	0	2	9.0	8	8	10	0.4	137	0	1
8	3	3.0	33	2	4	7.0	14	6	8	0.0	0	0	0
9	17	2.8	85	0	8	5.5	33	2	8	1.7	74	0	4
10	26	3.2	89	0	9	6.4	43	1	10	0.4	152	0	2
11	7	0.7	68	0	1	8.7	5	8	9	0.6	94	0	1
12	7	2.9	79	0	7	6.9	37	2	10	0.3	170	0	1
13	8	9.2	19	5	10	0.8	234	0	5	0.0	0	0	0
14	11	4.5	85	0	9	4.6	78	0	9	0.8	107	0	2
15	2	4.5	47	3	6	4.6	47	3	6	1.0	0	1	1

stratum	plots	volume mean	volume c.v.	volume min	volume max
1	7	26.7	160	0	115
2	9	221.7	24	150	320
3	17	131.5	41	10	220
4	12	205.8	26	130	280
5	9	308.9	17	250	390
6	10	124.0	36	70	190
7	5	321.0	6	300	345
8	3	240.0	14	200	260
9	17	129.4	31	60	200
10	26	250.4	22	140	350
11	7	332.9	22	190	400
12	7	317.1	21	200	400
13	8	32.2	187	0	180
14	11	223.6	21	150	320
15	2	205.0	3	200	210

Table 9. Estimates and (field) measured means and standard deviations of age and volume for five compartments.

area	size ha	age mean estimated	age mean real	age st. dev. % estimated	age st. dev. % real	volume mean estimated	volume mean real	volume st. dev. % estimated	volume st. dev. % real
17	1.5	59	48	35	46	162	118	22	81
13	2.6	84	99	34	11	239	226	20	26
17	3.1	79	94	36	30	220	211	26	32
16	3.6	99	87	35	21	311	337	20	17
18	11.2	59	75	28	44	128	236	105	30



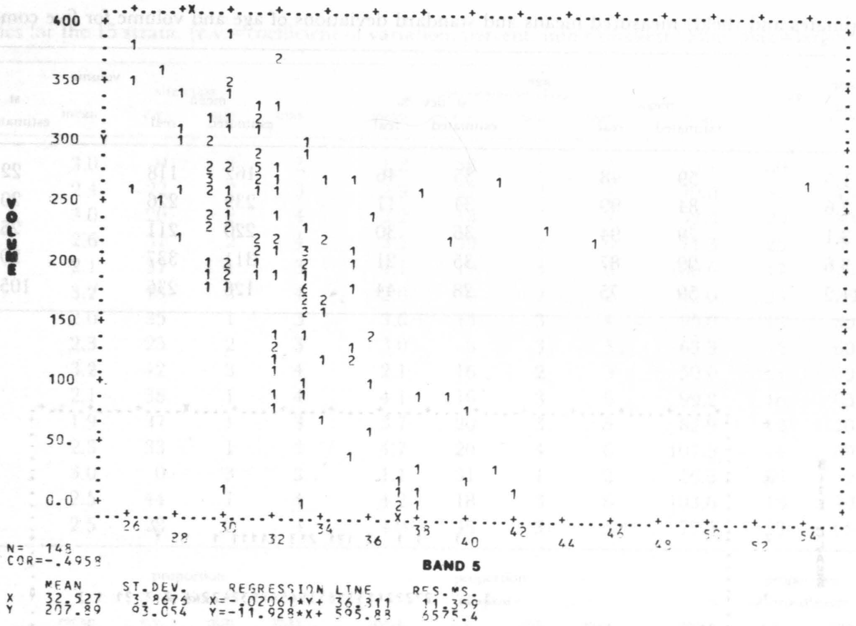
```
N=  439
COR= .3396

     MEAN        ST.DEV.     REGRESSION LINE        RES.MS.
X  -51E-12      1.0000    X= .36802*Y-.73855      .88669
Y   2.0068      .92280    Y= .31339*X+ 2.0068     .75506
```

Figure 13. Correlation diagram: principal component PF2 -site class, all study areas.

figures 17, 18 and 19 and the estimated and measured relative frequencies of site classes and development classes (area 13) are given in figures 20 and 21. The best results were obtained for the distributions of site class and the proportion of broadleaved trees.

The greatest and smallest differences between the estimates and measured values of mean ages are +23 % and −15 % and the greatest and smallest differences between the estimates and measured values of mean volumes are −46 % and +4 %. The estimated and real volume distributions are also illustrated in figure 22, where the means and standard deviations of the distributions are presented as patterns for each compartment.



```
N=  148
COR=-.4958

    MEAN      ST.DEV.     REGRESSION LINE      RES.MS.
X   32.527    3.8676    X=-.02061*Y+ 36.311    11.359
Y   207.89    93.054    Y=-11.928*X+ 595.88    6575.4
```

Figure 14. Correlation diagram: spectral value of band 5 -weighted mean of volume, study areas 16+19 (pixel-based file).

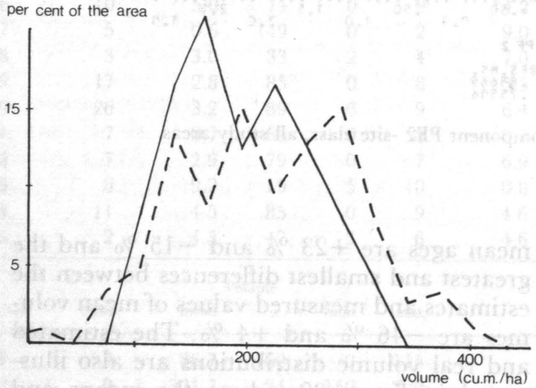


Figure 15. The estimated and real distributions of volume, area 13 (– – – estimated distribution, —— real distribution).

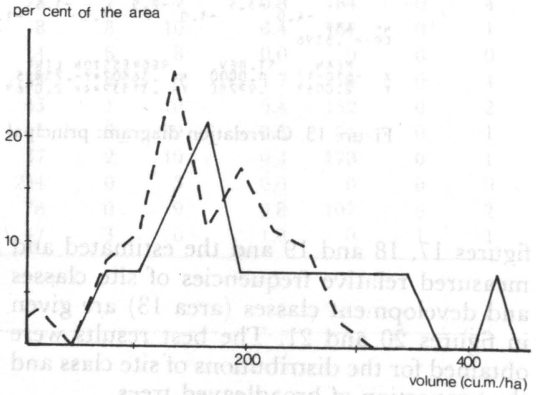


Figure 16. The estimated and real distributions of volume, the compartment of area 21 (– – – estimated distribution, —— real distribution).

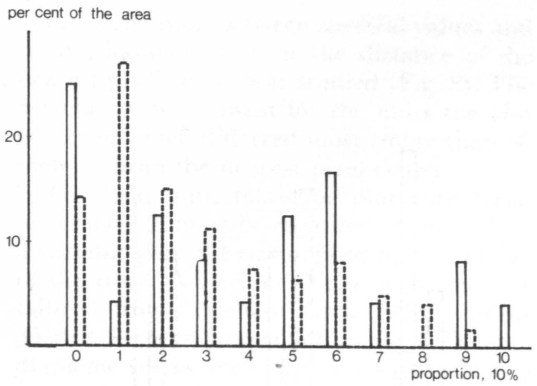


Figure 17. The estimated and real relative frequencies of proportion of pine, area 13 (– – – estimated frequencies, —— real frequencies).

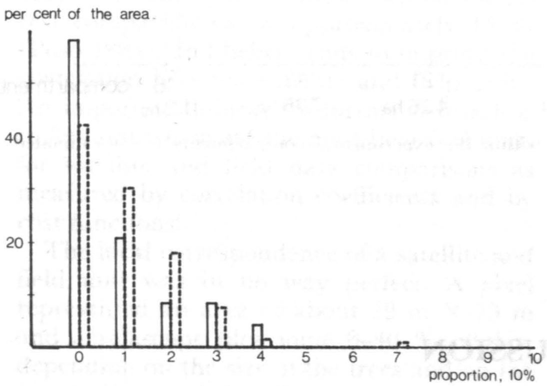


Figure 19. The estimated and real relative frequencies of proportion of broadleaved trees, area 13 (– – – estimated frequencies, —— real frequencies).

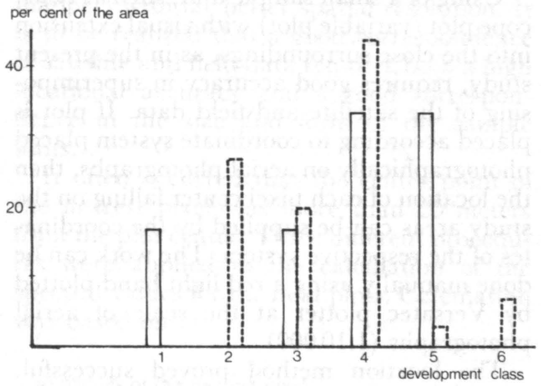


Figure 21. The estimated and real relative frequencies of development classes, area 13 (– – – estimated frequencies, —— real frequencies, class codes are the same as in table 1).

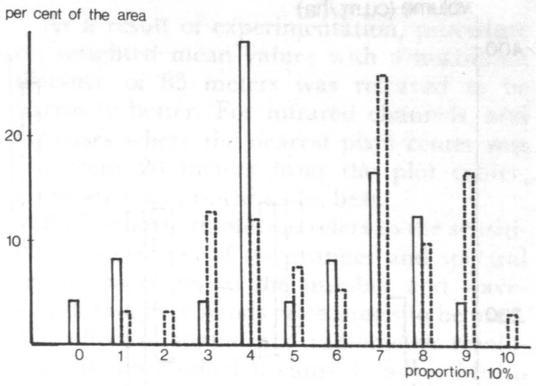


Figure 18. The estimated and real relative frequencies of proportion of spruce, area 13 (– – – estimated frequencies, —— real frequencies).

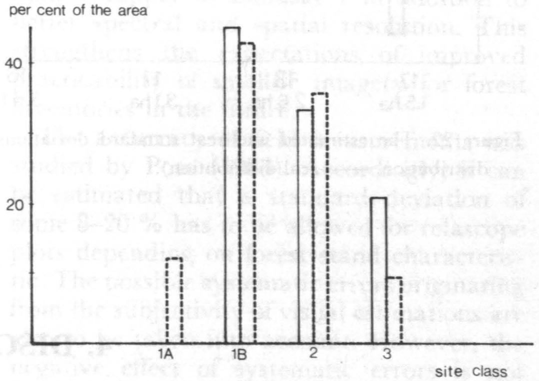


Figure 20. The estimated and real relative frequencies of site classes, area 13 (– – – estimated frequencies, —— real frequencies, class codes are the same as in table 1).

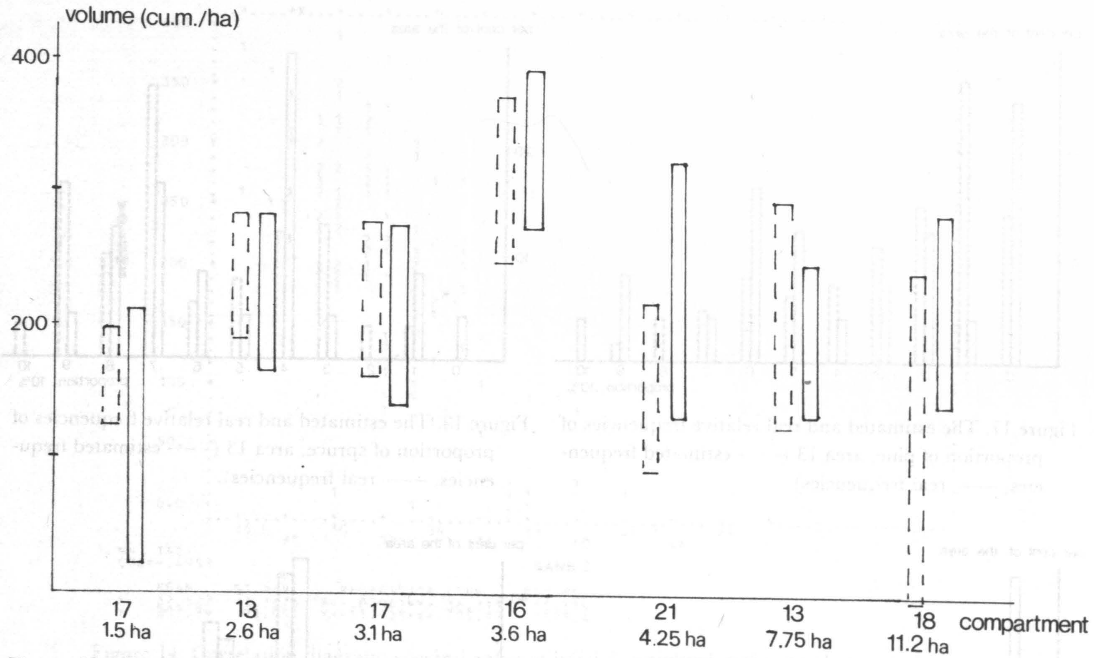


Figure 22. The estimated and real standard deviations within the experimental compartments (– – – estimated distribution, —— real distribution).

4. DISCUSSION

4.1 Definition and measuring of sample units

One possible unit for forest inventory purposes when numerical interpretation of Landsat material is applied is a pixel. However, this alternative proved to be less appealing than the sample plot located on the basis of aerial photographs. In principle, the plots can be selected with uniform systematic sampling. For a forest inventory, a spacing of 50 m × 50 m, for example, offers sufficient accuracy.

One reason for preferring the plots over pixels is that the estimation of the field variables (stand characteristics) is easier for a plot than for a pixel. The calculation of spectral values for a plot is easy. It can be based either on the values of the nearest pixel or on the weighted means of some 2–4 nearest pixels. Other reasons for favouring the plots over pixels are the ease in comparing multidate data when permanent sample plots are used. A pixel, on the other hand, often includes two or even more stands which create difficulties in data handling.

Using of a small sample unit, such as relascope plot (variable plot) with visual extension into the close surroundings, as in the present study, requires good accuracy in superimposing of the satellite and field data. If plot is placed according to coordinate system placed photographically on aerial photographs, then the location of each pixel center falling on the study areas can be supplied by the coordinates of the respective system. The work can be done manually using a red light band plotted by Versatec plotter at the scale of aerial photographs (1:10 000).

The location method proved successful. According to an experiment in an extra study area, a standard deviation of roughly 20 meters was obtained. An indication of accurate locating was also found when the dependence of the correlation between spectral values and stand characteristics on the distance of the nearest pixel center was studied (Fig. 8). The correlation was lowest for the units the plot center of which differred most (more than 40 meters) from the nearest pixel center.

The field unit, relascope plot with basal area factor 2, must be regarded as small. It is a variable plot, the size depending on the size of the trees. A tree of 30 cm of diameter is tallied from circular area with radius 35.36×30 cm = 10.6 m and a tree of 15 cm of diameter respectively with radius 35.36×15 cm = 5.3 m. The accurrate measurements made by relascope plots were enlarged subjectively by visual estimation of the plot surroundings. This procedure decreased random variation (standard deviation) within the forest compartments by approximately 15 % (Poso 1983), and hence tends to improve the correlation between satellite and field units. An important subject for further research is to find out which are the most benefical units for satellite and field data comparisons as measured by correlation coefficients and by cost functions.

The local correspondence of a satellite and field unit was in no way perfect. A pixel represented an area of about 79 m × 79 m and a relascope plot some 0–40 % of this, depending on the size of the trees and on the forest characteristics. Decreasing of the discrepancy in size by enlarging the size of field units would be costly. Better possibilities for the improvement are offered by new satellites, such as SPOT and Landsat 5 (Thematic mapper) with far better spatial resolution. It is to be realized that a good correspondence of satellite and field data requires both a high locational accuracy and a good correspondence in the size and form of the sample units.

It often occurred that the center point of the nearest pixel was more than 20 meters from the plot center. Three different procedures were applied to the calculation of the spectral values for the field plots. Calculation was based on

a) values of the nearest pixel
b) weighted mean values of pixels with a maximum distance of 55 meters
c) weighted mean values of pixels with a maximum distance of 85 meters.

As a result of experimentation, procedure c), weighted mean values with a maximum distance of 85 meters was releaved to be narrowly better. For infrared channels, and for cases where the nearest pixel center was less than 20 meters from the plot center, alternative a) proved to be best.

Radiometric resolution refers to the sensitivity of detectors of the scanner and spectral resolution refers to the number and wavelength ranges of bands or channels to be used. The effect of spatial and radiometric resolutions is not studied because Landsat Multispectral scanner data was the only spectral data used. Williams and Stauffer (1983) and Brass et al. (1983) emphasize the significance of better radiometric resolution of the Thematic Mapper of Landsat 4 in addition to better spectral and spatial resolution. This strengthens the expectations of improved practicability of satellite imagery for forest inventories in the future.

The accuracy of field measurements was studied by Poso (1983). Accordingly, it can be estimated that a standard deviation of some 8–20 % has to be allowed for relascope plots depending on forest stand characteristic. The possible systematic errors originating from the subjectivity of visual estimations are also to be taken into account. However, the negative effect of systematic errors is not found to be particularly harmful in the comparison of satellite and field data.

4.2 The correlation between satellite and field data

The highest coefficients of correlation between stand variables of a plot and the spectral values of the four MSS bands were obtained when spectral values were calculated as a weighted means of the one to four closest pixels with a maximum distance of 85 or 55 meters. The differences in correlation coefficients between distances of 85 and 55 meters can be regarded as unimportant. On the other hand, correlations with spectral values of the closest pixel were only marginally lower. It was found, however, that if the pixel center fell close to that of the field plot, the use of the closest pixel alone led to highest correlations with the infrared bands. (It seems possible that the spectral values of a plot should be calculated by different procedures depending on the MSS band.)

The application of a pixel as a sample unit produced distinctly lower correlations than obtained by using a plot as a basic unit. The correlation coefficients for volume were 0.47 and 0.68 and for age 0.54 and 0.71 when studied on the basis of the two study areas 16+19, which showed the highest correlations.

The correlations varied largely according to stand characteristic and study area. The highest correlations (0.50 to 0.82) were registered for stand age and volume and for the two study areas (16+19), the characteristics of which differred from each other. The correlations for all six study areas as a whole were much lower (0.34 to 0.50 for stand age and volume). High correlations were also found for development class, while site and proportions of tree species showed rather low values (about 0.3 for all study areas and 0.45 for the combined area of 16 and 19).

The use of principal component values instead of original MSS spectral values generally increased the correlations between satellite-based and field-based data. The correlations using the first principal component of spectral values and alternatively, using the best separate MSS band values were for stand age 0.82 and 0.73 and for stand volume 0.78 and 0.71 respectively when material for study areas 16+19 were used. The highest correlation (0.84) was obtained from the first principal component of spectral values and the first principal component of all of the stand characteristics.

The practicality of principal components should be studied in greater detail using larger materials. It is probable that the principal component technique is practical when there exists random variation in the spectral values and when sample units are to be stratified on the basis of spectral values.

The coefficients for the combined study areas showed only modest correlation between spectral values and stand characteristics (0.50 for age, 0.34 for volume). The high correlations of study areas 16+19 are not to be generalized. The areas were fairly homogeneous but differred largely from each other. The results simply indicate that in some conditions, at least, the correlations are high.

4.3 Applicability of the design

The most important objective of this study was to find a suitable methodology for combining the data from satellite imagery with field measurements in order to obtain inventory results for forest compartments. The compartments may be small, a few hectares in size, or may, even be large geographic regions. It was regarded as important that the method would be flexible enough to allow an opportunity for using additional data, such as data from maps, from earlier inventories and from aerial photographs.

The delineation of forest compartments is usually based on the stand characteristics. One of the most important objective is to obtain homogeneous units. The delineation of boundaries, however, have shown to be subjective, and the boundaries also change with time. The delineation is not made easier by the poor ground resolution of satellite imagery. These were the reasons why the "permanent" delineation on the basis of topographic maps and ownership boundaries was preferred in this study. The borderlines can then be noted by uniform coordination system.

The use of permanent compartments means that it is easy to define the sample units which belong to a compartment and it is also easy to make comparisons for multitemporal inventories. Some difficulties may rise from the fact that the compartments are often much more heterogeneous than in ordinary delineation. One part of a compartment may be an open area and another a mature forest, for example.

In the methodological experiments the material of areas 16+19, which showed the best correlations, was used to obtain field data for each stratum. The field data measured in other areas was only used for controlling the compartmentwise estimates.

As the correlations between spectral values and stand characteristics were relatively high for the areas (16+19), the within strata variances were less than 30 % of the total variance of the stand characteristics age and volume. This corresponds to the accuracy obtainable from a high quality photo interpretation and it must be regarded as exceptional. The tests with other material gave a within strata variance of some 85 % of the total variance, which is hardly useful for forest inventories. The use of materials with lower correlations (i.e. areas other than 16+19) was not tried.

The design developed and tried in this paper seems fully operational and well suited to compartmentwise forest inventory and updating purposes. With only a part of the computer programs having been made, the tests required plenty of manual work. This is why the experiments have mainly concerned operationality, whereas experiments concerning the optimization of the application are to be made later. Distributions and mean values of stand characteristics were only calculated for five compartments. A comparison of estimated and controlled characteristics show a positive dependence, but it is not possible to draw any significant conclusions on the basis of a sample of only five units.

The design offers the possibility of calculating the results for a compartment in the form of n-dimensional distributions, where n refers to the number of stand characteristics measured in the field for a sample unit. It is not advisable, however, to apply this possibility to a very detailed output if the accuracy of the estimates is not high. Only one dimensional distributions, such as distributions of stand volume over volume classes, were studied in this paper. The results seem to justify the conclusion that the estimated compartmentwise distributions are usually wider than those obtainable directly from field measurements. This means that the method tends to decrease the differences between separate compartments.

4.4 Suggestions for further studying

Optimization of the application of the design for specific forest inventory purposes has not yet been studied. Special attention should be given to the stratification of the satellite units. One line of investigation in this connection is the advantage of using principal component analysis. Another line is the use of map data, as referred to previously. Saukkola (1982) for example, has noted that there is a very large difference in tree volumes on mineral soil and on bog soil although the spectral values were very similar. With the use of map data the within strata variance could have been decreased.

The lines of further investigation suggested by the present are listed as follows.

1. The establishment of optimum satellite and field unit
2. Determining the optimum stratification from satellite information
3. Stratification aided by map- and aerial photo information
4. Determining the optimum allocation of field plots
5. Numerical interpretation aided by information of spatial variation
6. Comparing stratified sampling (with stratumwise field-variable matrices) and traditional classification methods.
7. The problem of divided plots and stand boundaries.
8. The suitability of the design for the inventory of changes, i.e. multitemporal application
9. The optimization of field measurements
10. The quality of separate satellite imageries
11. The economics of the application of satellite imagery

Finally, the authors are of the opinion that it is very important that the design should be experimented on a practical scale for ready adoption in practical forestry. This is the quickest way to introduce new technology for the common good.

SUMMARY

The paper describes a method suitable for multiobjective forest inventories. The method combines field measured data with measurements obtainable from satellite imagery or from other additional sources such as aerial photographs and maps. The key objectives are flexibility and efficiency, also applicability to compartmentwise forest inventories.

The method can be regarded as two phase sampling. In the first phase a large systematic dot sampling is placed according to uniform coordinate system. Each dot defines a sampling unit, a relascope or circular plot for example. The first phase data of the sampling units consists of radiation values calculated from satellite imagery and possibly other data, measured e.g. from maps.

The first phase units are to be stratified into homogeneous strata on the basis of the first phase data. In the second phase, some sample units are drawn from the first phase sample for each stratum and measured in the field for all the forest stand characteristics which are to be included in the inventory. The stratumwise information is then conformed by the matrix of m_h field sample plots with n forest stand characteristics in each.

The results of a forest area inventory, say of a compartment, are calculated by taking all of the first phase sample units belonging to the forest area and by studying their distribution by strata. Each first phase sample unit is supplied by the data of the respective stratum measured in the field. The weight of a field unit in the calculation is inversely proportional to the number of field units in the stratum.

The experiments consisted of six study areas totalling 130 hectares (325 acres). The first phase sample was formed by a systematic grid of 50 m × 50 m, producing about 500 relascope plots. The radiation values were calculated from Landsat 3 imagery acquired on June 10, 1980. The ground resolution of the imagery was 79 m × 79 m. If the plots did not fall very close to a pixel center the radiation values were calculated on the basis of the 2–4 nearest pixels. For testing purposes, all of the plots were measured in the field in 1980 and 1981. Some were located close to borderlines of distictly different compartments and were eliminated. The final number of field plots in the tests was 439. The field estimates were combinations of accurrate measurements according to the relascope principle with distance/diameter ratio of 35.35 and ocular estimations.

The correlations between the radiation values and stand characteristics of the plots varied according to study area and stand characteristic. The correlation coefficient between the first principal component of the four band radiation values and stand age was 0.82 and between the first principal component and stand volume 0.78 for two best areas. The respective values for the total material were 0.50 and 0.34. The two best study areas consisted of large and relatively homogeneous stands whereas the others were smaller and less homogeneous.

A factor affecting the correlation is the accuracy in supplying the pixels with geographic coordinates. Tests in an extra study areas showed a standard error of about 25 m. When the ground resolution of satellite imagery is improved the accuracy of pixel location will be increased. This will improve correlations especially in forests consisting of small stands with high variation.

The stratification of plots experiment was attempted only with the material of the two study areas of the highest correlations. The within strata variance was demonstrated to be 22 % from the total variance for stand age and 27 % for stand volume. This indicates that the stratifications was very succesful. However, the percentages for the total material were much higher.

The design was shown to be operational and it can be regarded to be suitable for compartmentwise forest inventories and updating purposes. It offers a possibility of calculating the results in the form of n-dimensional distributions, where n refers to the number of stand characteristics measured in the field for a sample unit. However, only one dimensional distributions, such as stand volume classes, were studied in this paper. The design is flexible and can employ additional

information, such as from maps, for stratification purposes. More experiments and studies are, however, regarded as necessary, especially in connection with the methods application to real practice.

REFERENCES

BMDP Statistical Software 1981. University of California Press.

BRASS, J. A., SPANNER, M. A., ULLIMAN, J. J., PETERSON, D. L., AMBROSIA, V. G. & BROCKHAUS, J. 1983. Thematic mapper simulator research for forest resource mapping in the Clearwater National Forest, Idaho. 17th International Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, May 9–13: 1–10.

DUDA, R. O. & HART, P. E. 1973. Pattern classification and Scene Analysis. Wiley-Interscience.

HÄME, T. 1984. Landsat-aided forest site type mapping. Photogrammetric Engineering and Remote sensing 50(8): 1175–1183.

HÄME, T. & SAUKKOLA, P. 1982. Satelliittikuvat Pohjois-Suomen metsäveroluokituksessa. Abstract: Satellite imagery in forest taxation in Northern Finland. Helsinki. Technical Research Centre of Finland, Research Reports 112.

KUUSELA, K. 1978. Suomen metsävarat ja metsien omistus 1971–76. Summary: Forest Resources and Ownership in Finland 1971–76. Helsinki. Comm. Inst. For. Fenn. 93(6).

— & POSO; S. 1970. Satellite pictures in the estimation of the growing stock over extensive areas. Helsinki. The Photogrammetric Journal of Finland 4(1): 3–9.

— & POSO, S. 1975. Demonstration of the applicability of satellite data to forestry. Seloste: Havaintoja satelliittikuvien käyttökelpoisuudesta metsätaloudessa. Helsinki. Comm. Inst. For. Fenn. 83(4).

LAASASENAHO, J. 1982. Taper curve and volume functions for pine, spruce and birch. Seloste: Männyn, kuusen ja koivun runkokäyrä- ja tilavuusyhtälöt. Helsinki. Comm. Inst. For. Fenn. 108.

LILLESAND, T. M. & KIEFER, R. W. 1979. Remote sensing and image interpretation. John Wiley & sons.

POSO; S. 1983. Kuvioittaisen arvioimismenetelmän perusteita. Summary: Basic Features of Forest Inventory by Compartments. Helsinki. Silva Fenn. 1983(4): 313–349.

SAUKKOLA, P. 1982. Satelliittikuviin perustuva puuston inventointi. Abstract: Timber inventory based on satellite imagery. Helsinki. Technical Research Centre of Finland, Research Reports 85.

WILLIAMS; D. L. & STAUFFER, M. L. 1983. What can the forestry community expect from Landsat-D thematic mapper data? Earth Resources Branch, NASA/Goddard Space Flight Center, Greenbelt, Maryland.: 475–492.

Yearbook of Forest Statistics 1982. Metsätilastollinen vuosikirja 1982. The Finnish Forest Research Institute. Helsinki. Folia For. 550.

Total of 14 references

SELOSTE

MENETELMÄ KUVIOITTAISTEN METSIKKÖTUNNUSTEN ESTIMOINTIIN SATELLIITTIKUVIA KÄYTTÄEN

Satelliittikuvien käyttömahdollisuuksien tutkiminen metsätaloudessa sai Suomessa alkunsa jo 1960-luvun lopulla, jolloin Kuusela ja Poso selvittivät ESSA 8-sääsatelliitin kuvien käyttökelpoisuutta laaja-alaisissa metsänarvioinneissa. Vuodesta 1972 lähtien maata on kiertänyt Landsat-ohjelman tekokuita, uusin, 1. 3. 1984 lähetetty, on järjestyksessä viides. Digitaalisten Landsat-kuvien numeerinen tulkinta alkoi vuonna 1974 Valtion teknillisen tutkimuskeskuksen saatua tulkintaan sopivan laitteiston.

Käsillä olevan tutkimuksen tavoitteena on kehittää lähinnä kuvioittaiseen arviointiin soveltuvaa menetelmää, jossa pyritään mahdollisimman tehokkaasti ja joustavasti yhdistämään satelliittikuvasta saatava tieto mitattuun maastotietoon ja muuhun lisäinformaatioon.

Tutkimus kohdistui 6 Etelä-Suomesta mitattuun alueeseen, joiden kokonaispinta-ala on 130 ha. Maastotiedot mitattiin systemaattisella relaskooppikoeala-arvioinnilla 50 metrin koealavälein. Koealoilta mitattiin tavallisimmat metsikkötunnukset. Havainnot laajennettiin silmävaraisesti koskemaan myös koealan näkymäympäristöä. Kaikkiaan mitattiin 520 koealaa.

Satelliittitiedot saatiin Landsat 3:n MSS- keilaimen kuvasta, joka on otettu 10. 6. 1980. Digitaalisen kuvan kuvanelementit eli pikselit vastaavat 79 m × 79 m maa-alaa, säteilyarvot saadaan kullekin elementille neljältä aallonpituuskanavalta: 500–600 nm (kanava 4), 600–700 nm (kanava 5), 700–800 nm (kanava 6) ja 800–1100 nm (kanava 7).

Satelliittikuvan pikselit paikannettiin maastokoealojen koordinaatistossa satelliittikuvasta tulostetun Versatec-harmaasävykartan ja ilmakuvan avulla. Tämän jälkeen koeala- ja satelliittikuvatiedostot yhdistettiin kahdella eri menetelmällä. Koealaperusteisessa yhdistämisessä laskettiin kullekin maastokoealalle säteilyarvot joko koealaa lähinnä sijainneen kuvanelementin arvona tai lähimpien kuvanelementtien säteilyarvoista etäisyydellä painotettuna keskiarvona. Maksimietäisyytenä, jota lähempänä sijainneet kuvanelementit otettiin mukaan koealan säteilyarvojen laskentaan, käytettiin vaihtoehtoisesti 55 tai 85 metriä.

Pikseliperusteisessa yhdistämisessä kullekin kuvanelementille laskettiin maastotiedot joko elementin keskipistettä lähinnä sijainneen koealan arvoina tai lähellä sijainneiden koealojen arvoista etäisyydellä painotettuna keskiarvona. Maksimietäisyytenä käytettiin joko 55 tai 65 metriä.

Satelliittikuvan paikantamisen tarkkuutta tutkittiin siirtämällä satelliittikuvaa systemaattisesti maastokoealakoordinaatiston akselien suunnassa ja laskemalla kunkin siirron jälkeen eräiden metsikkö- ja säteilyarvotunnusten väliset korrelaatiot. Selvää maksimikorrelaation antavaa kohtaa ei löydetty, joten satelliittikuva jätettiin alkuperäiselle paikalleen maastokoealakoordinaatistossa.

Satelliittikuvainformaatiota tiivistettiin laskemalla neljän aallonpituuskanavan säteilyarvoista pääkomponentit. Kaksi ensimmäistä pääkomponenttia selittivät 87 % kanavien kokonaisvarianssista. Myös maastotunnuksista laskettiin pääkomponentit.

Koealaperusteisessa yhdistämisessä saatiin korkeammat korrelaatiot kuin pikseliperusteisessa yhdistämisessä. Korrelaatiot vaihtelivat muuttujittain ja tutkimusalueittain (taulukot 4 ja 5). Kaikki tutkimusalueet yhdistettyinä korrelaatiokertoimet olivat parhaimmillaan –0.50, kahden parhaan alueen yhdistelmässä ne nousivat –0.84:ään. Parhaiten metsikkötunnusten kanssa korreloivat säteilyarvoista laskettu l. pääkomponentti ja kanavan 5 painotettu keskiarvo. Maastotunnuksista parhaiten säteilyarvojen kanssa korreloi ikä. Myös tilavuuden korrelaatiot säteilyarvojen kanssa olivat parhailla alueilla hyviä. Eri laskentatavoista todettiin, että laskettaessa säteilyarvot lähimpien pikselien arvoista painotettuina keskiarvoina saatiin korkeammat korrelaatiot kuin pelkästään lähimmän kuvanelementin arvoja käytettäessä. Pikseliperusteisessa tarkastelussa saadut korrelaatiot olivat kahden parhaan tutkimusalueen yhdistelmässä –0.60 luokkaa.

Tutkimuksessa laaditulla laskentamallilla estimoidaan kiinteärajaisille kuvioille metsikkötunnusten jakaumat käyttäen ns. mallialueilta mitattujen säteilyarvojen ja metsikkötunnusten välisiä riippuvuussuhteita. Mallialueiksi valittiin kaksi korrelaatioiltaan parasta tutkimusaluetta. Näiden alueiden maastokoealat ositettiin niille laskettujen säteilyarvojen perusteella. Tässä ns. ohjaamattomassa luokituksessa käytettiin säteilyarvoista laskettujen kahden ensimmäisen pääkomponentin pistearvoja. Osituksen tuloksena saatiin mallialueille 15 tai 20 sävyarvoiltaan homogeenista ositetta sekä niitä vastaavat maastotunnusten jakaumat. (Ihannetapauksessa yhtä sävyarvo-ositetta vastaa yksi maastotunnusten yhdistelmä.)

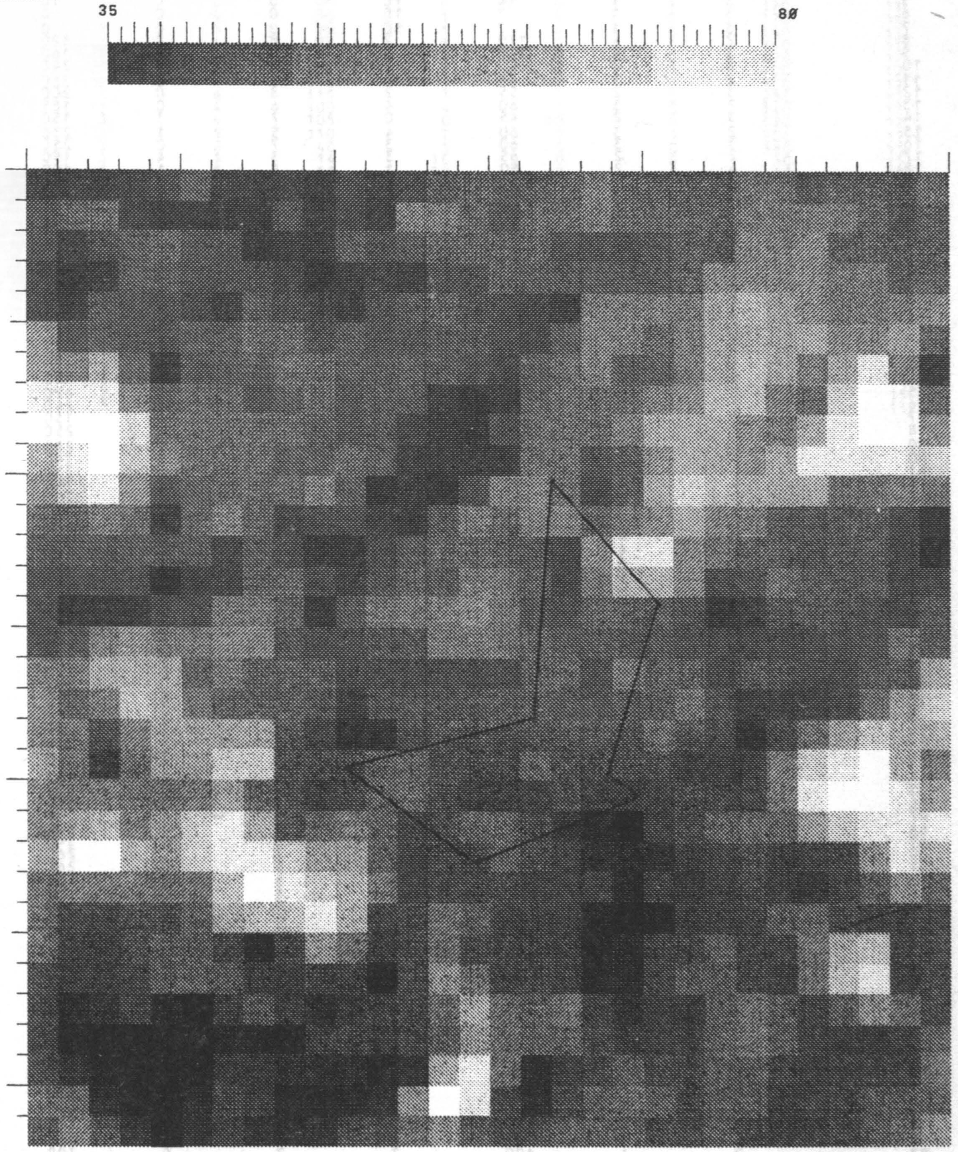
Mallin avulla voitiin halutulle alueelle estimoida maastotunnusten jakaumat selvittämällä ensin alueen koealojen jakauma mallista saatuihin spektrisiin ositteisiin. Kukin tiettyyn ositteeseen kuuluva koeala sai maastotunnukset mallin antamana tämän ositteen maastotunnusten jakaumana. Koealoittaiset jakaumat voitiin lopuksi yhdistää koko alueen estimoiduksi maastotunnusten jakaumaksi. Estimoitua jakaumaa ja sen tunnuslukuja verrattiin maastossa mitattuihin arvoihin estimoinnin tarkkuuden selvittämiseksi.

Laskelmat antoivat vaihtelevia tuloksia. Osituksesta saatiin ositteiden sisäisen varianssin osuudeksi kokonaisvarianssista iän osalta 22 % ja tilavuuden osalta 27 %. Kyseessä oli kuitenkin parhaiden tutkimusalueiden yhdistelmä eikä näin hyvä ositustulos ole yleistettävissä. Seitsemälle kuviolle lasketut jakaumat antoivat sekä ali- että yliarvioita. Estimoidun ja todellisen keskiarvon ero vaihteli iällä –15 +23 prosenttiin ja tilavuudella –46 +37 prosenttiin.

Kuvatun menetelmän odotetaan mm. joustavuutensa vuoksi soveltuvan hyvin esim. kuviotietojen muutosten seurantaan. Menetelmän käyttökelpoisuus tullee paranemaan, kun vuonna 1984 saadaan käyttöön erotuskyvyltään parempia (30 m × 30 m) Landsat 5:n Thematic mapperin kuvia. Menetelmän tutkiminen ja kehittäminen on tarpeen mm. maastomittausten optimoinnin ja sävyarvojen perusteella tapahtuvan osituksen sekä peruskartoilta ja ilmakuvilta saatavan lisäinformaation käytön osalta. Menetelmää on myös testattava laaja-alaisemmin käytännön metsätaloudessa.

Appendix 1. Versatec- greymap of area 19, scale 1:10 000.

```
PIIRRETTAVA TIEDOSTO:        LOPPI4
KUVAN KOKO:          32   LINJAA   X        30   SARAKETTA
KUVA-ALUE:          LINJAT      1   -      32          SARAKKEET:      1    -     30
HARMAA-ASTEIKKO:
```



Appendix 2. The composition of the 15 strata, principal component stratification.

	site class	development class	age	proportion of			volume	stratum
				pine	spruce	broadleaves		
1	3	1	14	10	0	0	10	1
2	3	2	30	10	0	0	115	1
3	3	2	25	10	0	0	50	1
4	3	1	9	10	0	0	2	1
5	3	1	14	10	0	0	10	1
6	3	1	7	10	0	0	0	1
7	3	1	7	10	0	0	0	1
8	2	4	110	2	8	0	320	2
9	3	4	120	9	0	1	160	2
10	2	2	55	2	8	0	240	2
11	2	3	65	0	10	0	260	2
12	3	2	55	2	8	0	200	2
13	3	3	60	4	6	0	185	2
14	3	2	45	0	7	3	150	2
15	2	3	60	4	5	1	230	2
16	3	3	60	3	7	0	250	2
17	2	2	65	3	3	4	190	3
18	4	2	50	1	8	1	100	3
19	3	2	30	8	2	0	75	3
20	3	2	45	5	4	1	115	3
21	3	2	59	4	4	2	145	3
22	4	2	45	1	7	2	95	3
23	3	2	50	4	4	2	130	3
24	3	2	50	1	7	2	120	3
25	3	2	55	0	8	2	10	3
26	4	2	60	2	5	3	110	3
27	3	2	60	5	4	1	160	3
28	2	3	65	0	3	7	220	3
29	3	3	55	5	5	0	200	3
30	3	2	60	1	6	3	170	3
31	2	3	60	3	5	2	190	3
32	3	2	30	8	2	0	75	3
33	3	2	50	4	4	2	130	3
34	4	3	80	5	5	0	260	4
35	2	6	65	1	1	8	220	4
36	2	4	75	4	4	2	240	4
37	2	6	65	1	7	2	150	4
38	3	4	120	2	7	1	250	4
39	2	4	85	2	8	0	280	4
40	3	2	60	0	7	3	140	4
41	2	3	60	6	3	1	250	4
42	4	2	60	1	7	2	140	4
43	2	3	65	2	7	1	200	4
44	3	3	60	3	7	0	210	4
45	2	2	60	2	5	3	130	4
46	2	4	100	3	7	0	330	5
47	2	4	110	1	8	1	340	5
48	2	3	75	0	9	1	390	5
49	3	4	110	7	3	0	260	5
50	2	4	75	1	9	0	350	5
51	3	4	100	0	10	0	340	5
52	3	6	120	5	5	0	260	5
53	1	4	120	3	7	0	250	5
54	1	4	95	0	10	0	260	5
55	3	2	50	3	7	0	170	6
56	3	2	30	9	1	0	150	6
57	3	2	35	8	2	0	180	6
58	4	2	60	3	7	0	90	6
59	4	2	45	1	9	0	125	6
60	3	2	30	9	1	0	85	6
61	3	2	40	8	2	0	70	6
62	3	2	45	0	6	4	100	6
63	3	2	25	8	1	1	80	6
64	3	2	60	1	6	3	190	6
65	1	4	90	2	8	0	310	7
66	2	4	115	0	9	1	310	7
67	2	4	95	1	9	0	300	7
68	2	3	80	0	9	1	345	7
69	3	3	95	0	10	0	340	7
70	3	3	65	3	7	0	260	8
71	2	3	60	4	6	0	260	8
72	2	3	65	2	8	0	200	8
73	4	2	40	2	8	0	60	9
74	3	2	50	6	4	0	140	9
75	4	3	70	3	6	1	190	9
76	3	2	60	0	7	3	200	9
77	3	2	50	1	7	2	145	9
78	3	2	50	1	7	2	145	9
79	3	2	55	3	5	2	170	9
80	3	2	55	3	4	3	125	9
81	3	2	45	1	5	4	100	9
82	3	2	45	1	7	2	120	9
83	3	2	30	8	2	0	65	9
84	3	2	60	0	7	3	120	9
85	3	2	55	4	6	0	145	9
86	4	2	50	4	3	3	110	9
87	3	2	45	0	8	2	120	9
88	3	2	30	4	5	1	75	9
89	3	3	60	6	3	1	170	9
90	2	3	65	0	8	2	210	10
91	3	6	120	7	3	0	350	10
92	2	3	65	3	6	1	200	10
93	2	4	100	3	7	0	300	10
94	2	4	110	6	4	0	290	10
95	1	4	90	2	7	1	280	10
96	1	4	95	2	7	1	300	10
97	2	4	115	7	3	0	260	10
98	3	4	115	7	3	0	240	10
99	2	4	110	2	8	0	300	10
100	2	4	110	1	9	0	140	10
101	1	4	90	0	10	0	320	10
102	1	4	90	2	7	1	300	10
103	3	4	115	0	10	0	280	10
104	2	5	85	5	4	1	160	10
105	1	4	80	0	9	1	250	10
106	2	4	75	1	9	0	275	10
107	2	4	90	1	9	0	280	10
108	2	4	115	1	9	0	270	10
109	2	4	110	1	9	0	270	10
110	2	4	115	1	9	0	290	10
111	1	6	100	4	4	2	150	10
112	4	4	110	6	3	1	175	10
113	3	4	110	9	1	0	200	10
114	3	4	90	3	7	0	230	10
115	3	4	110	8	2	0	190	10
116	1	3	75	1	8	1	280	11
117	2	3	80	0	9	1	400	11
118	2	3	75	1	8	1	380	11
119	3	4	95	0	9	1	370	11
120	2	4	90	1	9	0	370	11
121	1	5	70	1	9	0	190	11
122	2	4	95	1	9	0	340	11
123	1	4	80	1	9	0	320	12
124	2	6	130	2	7	1	400	12
125	2	4	115	3	7	0	300	12
126	3	6	115	3	7	0	280	12
127	3	4	100	4	6	0	340	12
128	3	4	105	0	10	0	380	12
129	2	5	110	7	2	1	200	12
130	3	2	40	5	5	0	180	13
131	3	1	14	10	0	0	10	13
132	3	1	13	10	0	0	10	13
133	3	1	10	9	1	0	0	13
134	3	1	20	10	0	0	25	13
135	3	1	14	10	0	0	15	13
136	3	1	9	10	0	0	3	13
137	3	1	14	10	0	0	15	13
138	1	3	70	0	8	2	280	14
139	4	4	120	0	8	2	320	14
140	2	4	110	1	9	0	220	14
141	3	4	105	6	4	0	200	14
142	1	4	85	0	9	1	240	14
143	3	6	110	5	3	2	200	14
144	3	4	110	8	1	1	150	14
145	3	4	120	9	1	0	250	14
146	1	5	75	3	7	0	220	14
147	3	4	115	9	1	0	180	14
148	4	4	120	9	0	1	200	14
149	2	5	90	6	3	1	200	15
150	3	3	65	3	6	1	210	15

Appendix 3. The composition of the 20 strata, principal component stratification.

	site class	development class	age	proportion of			volume	stratum
				pine	spruce	broadleaves		
1	3	2	30	10	0	0	115	1
2	3	2	25	10	0	0	50	1
3	3	1	9	10	0	0	2	1
4	3	1	14	10	0	0	10	1
5	2	3	60	6	4	0	230	1
6	3	1	7	10	0	0	0	1
7	3	1	7	10	0	0	0	1
8	2	4	110	2	8	0	320	2
9	2	2	55	2	8	0	240	2
10	2	3	65	0	10	0	260	2
11	3	2	55	2	8	0	200	2
12	2	3	60	4	5	1	230	2
13	3	3	60	3	7	0	250	2
14	4	2	40	2	8	0	60	3
15	2	2	65	3	3	4	190	3
16	4	2	50	1	8	1	100	3
17	3	2	45	5	4	1	115	3
18	3	2	59	4	4	2	145	3
19	2	3	65	0	3	7	220	3
20	3	3	55	5	5	0	200	3
21	3	2	60	1	6	3	170	3
22	2	3	60	3	5	2	190	3
23	4	3	80	5	5	0	260	4
24	2	6	65	1	1	8	220	4
25	2	4	75	4	4	2	240	4
26	2	6	65	1	7	2	150	4
27	3	4	120	2	7	1	250	4
28	2	4	85	2	8	0	280	4
29	3	2	60	0	7	3	140	4
30	2	3	60	6	3	1	250	4
31	4	2	60	1	7	2	140	4
32	2	3	65	2	7	1	200	4
33	3	3	60	3	7	0	210	4
34	2	2	60	2	5	3	130	4
35	3	2	40	5	5	0	180	5
36	3	1	10	9	1	0	0	5
37	3	2	30	9	1	0	85	5
38	3	1	20	10	0	0	25	5
39	3	2	40	8	2	0	70	5
40	3	2	45	0	6	4	100	5
41	3	2	25	8	1	1	80	5
42	2	3	65	0	8	2	210	6
43	2	4	110	1	8	1	340	6
44	2	3	75	0	9	1	390	6
45	3	4	110	7	3	0	260	6
46	3	4	100	0	10	0	340	6
47	3	4	105	0	10	0	380	6
48	3	6	120	5	5	0	260	6
49	1	4	120	3	7	0	250	6
50	1	4	95	0	10	0	260	6
51	1	4	90	2	8	0	310	7
52	2	4	115	0	9	1	310	7
53	2	4	95	1	9	0	300	7
54	2	3	80	0	9	1	345	7
55	3	3	95	0	10	0	340	7
56	3	2	50	6	4	0	140	8
57	4	3	70	3	6	1	190	8
58	3	2	60	0	7	3	200	8
59	3	2	50	1	7	2	145	8
60	3	2	50	1	7	2	145	8
61	3	2	55	3	5	2	170	8
62	3	2	55	3	4	3	125	8
63	3	2	45	1	5	4	100	8
64	3	2	45	1	7	2	120	8
65	3	2	30	8	2	0	65	8
66	3	2	60	0	7	3	120	8
67	3	2	55	4	6	0	145	8
68	4	2	50	4	3	3	110	8
69	3	2	45	0	8	2	120	8
70	3	2	30	4	5	1	75	8
71	3	3	60	6	3	1	170	8
72	3	3	65	3	7	0	260	9
73	2	3	65	2	8	0	200	9
74	3	6	120	7	3	0	350	10
75	2	4	110	6	4	0	290	10
76	1	4	95	2	7	1	300	10
77	3	4	115	7	3	0	240	10
78	4	4	120	0	8	2	320	10
79	2	4	110	2	8	0	300	10
80	2	4	110	1	9	0	140	10
81	1	4	90	0	10	0	320	10
82	1	4	90	2	7	1	300	10
83	2	5	85	5	4	1	160	10
84	1	4	80	0	9	1	250	10
85	2	4	75	1	9	0	275	10
86	2	4	90	1	9	0	280	10
87	1	5	75	3	7	0	220	10
88	4	4	110	6	3	1	175	10
89	3	4	110	8	2	0	190	10
90	1	4	80	1	9	0	320	11
91	2	6	130	2	7	1	400	11
92	2	4	115	3	7	0	300	11
93	3	4	105	6	4	0	200	11
94	1	4	85	0	9	1	240	11
95	3	6	110	5	3	2	200	11
96	2	5	110	7	2	1	200	11
97	3	1	14	10	0	0	10	12
98	3	1	14	10	0	0	10	12
99	3	1	13	10	0	0	10	12
100	3	1	14	10	0	0	15	12
101	3	1	9	10	0	0	3	12
102	2	3	60	7	3	0	280	12
103	3	1	14	10	0	0	15	12
104	1	3	75	1	8	1	280	13
105	2	3	80	0	9	1	400	13
106	2	3	75	1	8	1	380	13
107	3	4	95	0	9	1	370	13
108	2	4	90	1	9	0	370	13
109	1	5	70	1	9	0	190	13
110	2	4	95	1	9	0	340	13
111	1	3	70	0	8	2	280	14
112	2	4	110	1	9	0	220	14
113	3	4	110	8	1	1	150	14
114	3	4	120	9	1	0	250	14
115	3	4	115	9	1	0	180	14
116	3	4	120	9	0	1	160	14
117	4	4	120	9	0	1	200	14
118	2	3	60	4	6	0	185	14
119	3	2	45	0	7	3	150	14
120	2	5	90	6	3	1	200	15
121	3	3	65	3	6	1	210	15
122	3	2	50	3	7	0	170	16
123	3	2	30	9	1	0	150	16
124	3	2	35	8	2	0	180	16
125	4	2	60	3	7	0	90	16
126	4	2	45	1	9	0	125	16
127	3	2	60	1	6	3	190	16
128	2	3	65	3	6	1	200	17
129	2	4	100	3	7	0	300	17
130	1	4	90	2	7	1	280	17
131	2	4	115	7	3	0	260	17
132	3	4	115	0	10	0	280	17
133	2	4	115	1	9	0	270	17
134	2	4	110	1	9	0	270	17
135	2	4	115	1	9	0	290	17
136	1	6	100	4	4	2	150	17
137	3	4	110	9	1	0	200	17
138	3	4	90	3	7	0	230	17
139	2	3	60	4	6	0	260	18
140	3	2	30	8	2	0	75	19
141	4	2	45	1	7	2	95	19
142	3	2	50	4	4	2	130	19
143	3	2	50	1	7	2	120	19
144	3	2	55	0	8	2	10	19
145	4	2	60	2	5	3	110	19
146	3	2	60	5	4	1	160	19
147	3	2	30	8	2	0	75	19
148	3	2	50	4	4	2	130	19
149	2	4	100	3	7	0	330	20
150	3	6	115	3	7	0	280	20
151	3	4	100	4	6	0	340	20
152	2	4	75	1	9	0	350	20

Appendix 4. The principal component scores (PF1, PF2) of the 15 strata (Mean volumes of the strata are printed inside the boxes).