Dynamic Stratification for Estimating Pointwise Forest Characteristics

Gintautas Mozgeris

Mozgeris, G. 1996. Dynamic stratification for estimating pointwise forest characteristics. Silva Fennica 30(1): 61–72.

This paper deals with the testing of dynamic stratification for estimating stand level forest characteristics (basal area, mean diameter, mean height and mean age) for a 117 ha study area in Finland. The results do not show possibilities to achieve more accurate estimates using only Landsat TM principal components as auxiliary data opposed to static stratification. It was found that in dynamic stratification non-measured observations should be assigned the mean characteristics of the measured observations that belong to the same cube (class) instead of randomly selected ones. Stratification errors tend to decrease with the lessening of stratification variable classes until a certain limit. If only one principal component is used the number of classes has however little influence. Low field values are overestimated and high values underestimated.

The only successful results were obtained using two variables of different origin – the qualitative development stage class and the quantitative 1st principal compund. The lowest root mean square error in estimating basal area was 6.40 m^2/ha, mean diameter 3.34 cm, mean height 2.65 m and mean age 14.06 years. This increase of stratification accuracy is mainly resulted by the use of development stage class as an auxiliary variable.

Keywords SMI forest management planning system, static stratification, dynamic stratification.
Author's address Department of Forest Management, Lithuanian Agricultural Academy, 4324 Kaunas - Akademija, Lithuania **Fax** +370 7 296 531 **E-mail** pedm@nora.lzua.lt
Accepted March 22, 1996

List of Symbols

PC	principal component
$\bar{e}$	error mean
S_e^2	error variance
rms	root mean square error

Symbols used in equations

x_i	vector of observation i and
$\bar{x}_j$	centroid vector of cluster j
S_j	covariance matrix of cluster j
$prior_j$	a-prior probability proportion of cluster j

1 Introduction

The SMI (shortened from Satelliittikuvat Metsien Inventoinnissa – satellite imagery in forest inventory) forest inventory and management system, developed in the Helsinki University Department of Forest Resource Management, can be best described as an application of two phase sampling for gathering data in order to solve forest management problems (Waite 1990). Forest characteristics for arbitrary points are commonly estimated using stratification. Usually a given point is first classified into a predefined auxiliary data stratum using some classification criterion. Then it is assigned the mean forest characteristics of those field sample plots belonging to the same stratum. This procedure is static by nature because the auxiliary data strata are fixed during the process. An alternative to this is dynamic stratification in which the auxiliary data strata change during the process.

The utility of dynamic stratification for estimating pointwise forest characteristics in the SMI system has not been tested so far. The principal questions to be studied and solved are:

(i) Is it possible to achieve more accurate estimates using dynamic stratification opposed to static stratification and
(ii) What are the most appropriate dynamic stratification strategies.

During the study empirical data set will be used for estimating forest characteristics by both static and dynamic stratification. Different aspects of dynamic stratification are studied: the use of 1 and 2 auxiliary data variables, methods to define the upper class limits and tactics for changing these limits during the stratification process. Stratification errors will be calculated and analyzed.

2 Methods and Material

2.1 Static Stratification

Static stratification consists of (i) clustering, which involves the formation of clusters (or strata) and (ii) classification or the assignment of observations into clusters. K-Means unsupervised clustering consists of three phases (i) the definition of initial cluster centroids, (ii) the iterative derivation of the final cluster centroids and (iii) the derivation of the covariance matrices. Classification involves the assignment of observations to the strata they most probably belong to. Four different classifiers (or classification criteria) were used in the study: Euclidean distances, Mahalanobis distances, Gaussian probabilities and Bayesian probabilities.

The Euclidean distance e_{ij} from observation i to cluster j is defined as:

$$e_{ij} = \left(x_i - \bar{x}_j\right)'\left(x_i - \bar{x}_j\right) = d_{ij}'d_{ij} \tag{1}$$

The Mahalanobis distance m_{ij} from observation i to cluster j is defined as:

$$m_{ij} = d_{ij}'S^{-1}d_{ij} \tag{2}$$

The Gaussian probability g_{ij} of observation i belonging to cluster j is defined as:

$$g_{ij} = (2\pi)^{-p/2}\left|S_j\right|^{-1/2} e^{-1/2\left(d_{ij}'S_{ij}^{-1}d_{ij}\right)} \tag{3}$$

The Bayesian probability b_{ij} of observation i belonging to cluster j is defined as:

$$b_{ij} = prior_j \frac{g_{ij}}{\sum_j g_{ij}} \tag{4}$$

2.2 Dynamic Stratification

In dynamic stratification contrary to static stratification, the auxiliary data strata change during the process. Dynamic stratification has been used in the Swedish NFI for assigning sample tree characteristics to tallied trees (Hägglund 1981). Observations are divided into two groups: A-observations and B-observations. Both input and output data are known for the A-observations (e.g. spectral principal components (PC) as x (input) variables and field characteristics as y (output) variables) but only input data for the B-observations. First, a large number of small p-dimensional boxes (classes) are defined into which all A and B observations are classified. Second, all boxes are checked. If any given box includes both non-measured (B) and measured (A) observations, the non-measured observations are assigned randomly selected or mean characteristics of the measured observations belonging to the same box and removed from the set of non-measured observations. If the set of non-measured observations is now empty the procedure is terminated. Otherwise the boxes are enlarged and the process is returned to the second phase. Classification is performed by user-defined input variable class upper limits, which are set dynamically during the process.

Dynamic stratification seems to be very similar to a nearest neighbor classification, used by Finnish Forest Research Institute for National Forest Inventory. The estimates of the field variables for every pixel are defined using the Euclidean distance, computed in the feature space from the pixel to be classified to each pixel whose ground truth is known (Tomppo 1993). The main distinctive features of dynamic stratification are the possibility to set interactively input variable class upper limits and iterative nature of the procedure.

The following issues should be considered in dynamic stratification: (i) auxiliary data to be used, (ii) the maximum number of auxiliary data strata or classes, (iii) the methods to be used to define class limits and (iv) the tactics for changing these limits during the process. Stratification was started with the 1st PC because of its highest correlation with field data. The maximum class number implemented was 20. Non-measured observations (B) were assigned (i) randomly selected or (ii) mean values of the measured observations (A). Two approaches were used in defining the class limits: (i) class widths were defined so that the number of observations falling in each class was approximately equal. For that purpose Gaussian cumulative proportions of auxiliary variable were used. The class widths were not equal – narrower classes could be found in the central part of the PC's distribution. (ii) The class widths were set to be equal ranging from the minimum until maximum values. It was decided to start with 20 narrow 1st PC classes and to lessen their number by changing the class limits until all B-observations were assigned A-observation values.

The definition of upper class limits when two x variables are used is more complicated. The following approach was used in this study: the class number of one x variable was kept constant and the class number of the other variable was gradually lessened by expanding class widths until all B-observations were assigned field data. Then the stratification procedure was terminated and repeated with a number of stable x variable classes lessened by 1. The 1st and 2nd PC's were used alternately as the first and second x variables.

2.3 Computer Programmes

The SMI system currently includes programmes for implementing both dynamic and static stratification, statistical tests, accuracy assessments and variable relationship studies. The programmes run on PC-compatible micro-computer with the MS-DOS operating system.

2.4 Material

The study area consists of an 117 ha forest lot, located near the forest station of Helsinki University in Hyytiälä. A systematic set of relascope plots (BAF 2 m²/ha, 8 plots/ha) were measured in the area in summer 1989. The plots were divided into three parts: (i) *1p* plots (totally 874 plots), (ii) *2p* plots (175) and (iii) *tst* plots (699). The *2p* and *tst* plots were acquired by dividing

Table 1. Pearson's correlation coefficients for *2p* TM PC–field variable pairs.

Field variable	Principal component 1	2	3	4	5	6
Basal area	–0.536	0.068	0.010	–0.119	0.049	0.110
Mean diameter	–0.583	0.228	–0.005	–0.042	0.022	0.040
Mean height	–0.582	0.221	0.008	–0.048	0.035	0.036
Mean age	–0.460	0.300	–0.060	–0.020	0.030	0.010

Table 2. Assessment of static stratification.

Stratification method	Basal area $\bar{\varepsilon}$	S_ε^2	*rms*	Mean diameter $\bar{\varepsilon}$	S_ε^2	*rms*	Mean height $\bar{\varepsilon}$	S_ε^2	*rms*	Mean age $\bar{\varepsilon}$	S_ε^2	*rms*
Euclidean distances	0.79	8.12	8.16	0.71	8.04	8.07	0.67	6.38	6.41	3.87	31.54	31.78
Bayesian probabilities	0.69	7.96	7.99	0.68	8.00	8.03	0.65	6.35	6.39	3.84	31.45	31.69
Gaussian probabilities	0.70	7.97	8.00	0.63	7.96	7.99	0.63	6.33	6.37	3.53	31.17	31.37
Mahalanobis distances	0.79	7.92	7.96	0.55	7.91	7.93	0.57	6.30	6.32	3.13	31.17	31.33

the *1p* plots into two unique groups. This was done in order to use *2p* plots as A-observations and *tst* plots as B-observations in dynamic stratification and for assessing stratification accuracy.

The area was segmented into compartments by the Forest and Park Service in 1986.

A LANDSAT 5 TM satellite image covering the area and taken in June 1989 was available for the study. Bands 1, 2, 3, 4, 5 and 7, i.e. the bands describing the intensities of visible and infrared light were used. The image had been resampled to a pixel size of 25 × 25 meters.

Principal components, derived from the LANDSAT TM spectral variables were used as auxiliary data in the stratification process. The *2p* and *tst* PCs are normally distributed and the differences between the distributions of PCs to be used are not significant. The following stand data were available for the *2p* and *tst* plots: basal area (m^2/ha), mean diameter (cm), mean height (m), mean age (years), as well as development stage class of the main storey. Most of quantitative variables are also normally distributed (except mean age). No *2p* and *tst* field means differ significantly from each other. The distributions of mean diameter and mean height differ from each other. It should also be mentioned, that the 1st PC correlated most with the field variables being analysed and that the other PCs correlated weakly with them (Table 1).

3 Results

3.1 Static Stratification

First, forest characteristics (basal area, mean diameter, height and age) for *tst* plots were estimated using static stratification. The clusters were formed using K-Means unsupervised clustering and *1p* PCs (the 1st, 2nd and 3rd). Both *2p* and *tst* observations were then classified into these clusters (totally 20). Four different classifiers were used: Euclidean distances, Bayesian probabilities, Gaussian probabilities and Mahalanobis distances. Each cluster was assigned the mean field values of those *2p* plots belonging to the same cluster. The forest characteristics for all *tst* plots were estimated using the calculated average basal area, mean diameter, height and age values for each cluster. The accuracy statistics – root mean square errors (*rmse*), error means ($\bar{\varepsilon}$), and variances (S_ε^2) – were calculated to assess the stratification (Table 2).

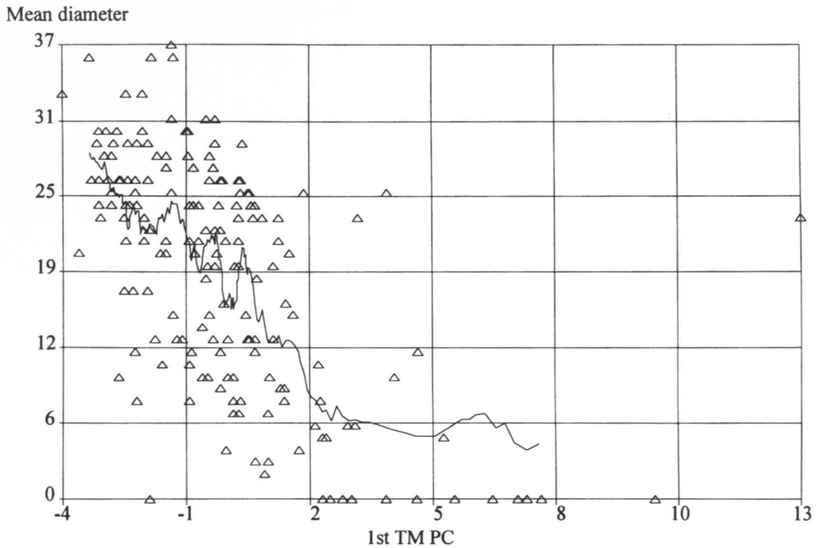

Fig. 1. Relationship between *2p* mean diameters and 1st TM PC values.

The results are similar regardless of the classification criteria. The root mean square errors and biases are however somewhat higher for Euclidean distances.

3.2 Dynamic Stratification with One Principal Component

Dynamic stratification was started with one auxiliary variable – the 1st PC. Using the equal interval lengths approach in defining the class limits the relationship between the 1st PC and the field variables was examined. The mean diameter appeared to be the most correlated field variable with the 1st PC (Pearson's correlation coefficient –0.58). For plots with mean diameters ranging from 8 to 37 cm the relationship is nearly linear (Fig. 1). For smaller diameters and clear cut plots there is practically no relationship (these comprise 15 % of all observations). So the class widths for 1st PC values ranging from the minimum value –3.5 until 2.5 were set to be equal for all classes. These classes should comprise about 85 % of the total class number. The class limits for the rest of observations (the 1st PC ranges from 2.5 until the maximum value 13.2) were calculated separately.

The results of dynamic stratification using only one *x* variable are presented in Table 3.

The advantage of assigning the mean values of measured observations instead of randomly selected ones can be very clearly seen – the average root mean square error of the former way is about 25 % less. Thus the latter method was excluded from the further investigations. The number of *x* variable classes has practically no influence over the stratification procedure, because in all cases all B-observations received values during the first run. The error statistics are very similar except for the broadest classes,

Table 3. Assessment of dynamic stratification with TM PCs.

Stratification method	Basal area $\bar{e}$	Basal area S_e^2	Basal area *rms*	Mean diameter $\bar{e}$	Mean diameter S_e^2	Mean diameter *rms*	Mean height $\bar{e}$	Mean height S_e^2	Mean height *rms*	Mean age $\bar{e}$	Mean age S_e^2	Mean age *rms*
Random assignment, 1st method of class limit definition	0.88	11.05	11.09	0.18	10.20	10.20	0.45	8.18	8.19	1.97	42.50	42.54
Random assignment, 2nd method of class limit definition	0.67	10.46	10.48	0.43	10.53	10.54	0.56	8.31	8.33	1.51	43.18	43.21
Mean assignment, 1st method of class limit definition	0.64	8.00	8.02	0.50	8.01	8.01	0.26	6.34	6.35	0.63	32.02	32.02
Mean assignment, 2nd method of class limit definition	0.73	8.03	8.07	0.30	7.99	7.99	0.44	6.34	6.35	1.70	31.78	31.83
2 PCs, number of 1st PC classes changes	0.58	8.36	8.38	0.19	8.40	8.40	0.32	6.61	6.62	0.94	33.47	33.47
2 PCs, number of 2nd PC classes changes	0.07	9.27	9.27	−0.03	9.04	9.04	0.02	7.16	7.16	0.66	35.59	35.60
Selected *x* variables class combinations	0.69	8.16	8.19	0.29	8.23	8.23	0.42	6.54	6.56	1.74	32.06	32.11

(Field variable error statistics)

Table 4. Pearson's correlation coefficients between estimation errors using the 1st PC and principal components for the *tst* plots.

Estimation errors of	Principal components 1	2	3	4	5	6
Basal area	−0.054	−0.022	0.001	0.019	−0.008	−0.134
Mean diameter	−0.034	−0.141	0.011	0.013	−0.053	−0.083
Mean height	−0.035	−0.144	0.016	0.014	−0.039	−0.096
Mean age	−0.005	−0.187	0.015	0.014	−0.034	−0.029

where they tend to increase (Fig. 2). The approach to use Gaussian cumulative proportions to define the upper class limits resulted in somewhat smaller errors. In most cases static stratification resulted however in a better accuracy.

Pearson's correlation coefficients between field variable errors and all PCs for the *tst* plots are presented in Table 4. They were used to select the next PC to be included in further stratification processes (a large absolute correlation coefficient indicates that the PC explains the residual stratification error well). The 2nd PC seemed to be most correlated with the errors and it was included in subsequent analyses.

3.3 Dynamic Stratification with Two Principal Components

In order to define the class limits of the 2nd PC the *2p* field variable distributions were analyzed (Fig. 3). There seems to be no clear relationship between basal area, mean diameter, mean height and the 2nd PC – the moving averages of these variables lay parallel to the x-axis (PC values). Mean age observations are distributed in two clusters. The correlation coefficients between the 2nd PC and field variables are 0.068, 0.228, 0.221 and 0.300 respectively. The best way to define upper limits of the classes could be the

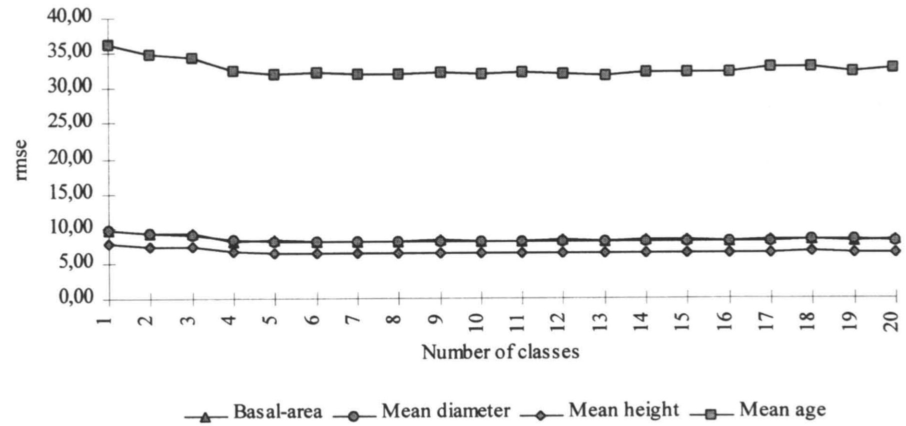

Fig. 2. Relationship between dynamic stratification errors and number of *x* variable classes (mean assignment, 2nd method of class limit definition) (values for basal area and mean diameter overlap).

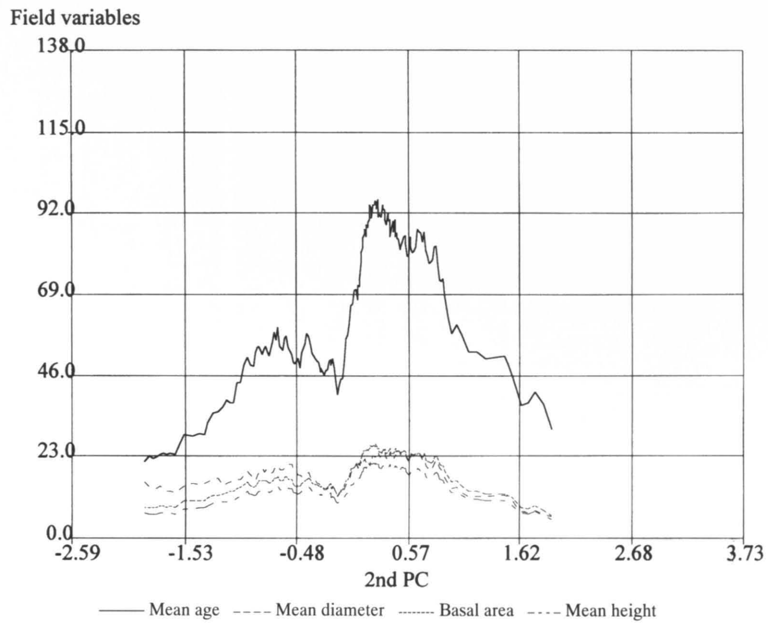

Fig. 3. Relationship between $2p$ field variable and the 2nd PC values.

division of the 2nd PC values into equal width classes in the interval between –2 and 1.5 because the remaining PC values cover less than 10 % of the field plots being examined.

Table 5 presents the dynamic stratification procedures analysed – the total number of assigned B-observations after each changing of class limits. In case (a) the 2nd PC classes were left stable and the 1st PC classes were gradually lessened and in case (b) the 1st PC classes were constant and the 2nd PC classes were altered.

It can be noticed, that the stratification was completed with a rather small number of changeable classes – usually 1 or 2. For smaller numbers of stable classes less iterations to end the procedure were required.

The average stratification errors in both (a) and (b) cases are very similar and tend to decrease with the lessening of total x variable classes – the correlation coefficient between the root mean square error and the number of the $x1$ classes ranges from 0.70 to 0.93 and from 0.63 to 0.81 respectively. In all cases the root mean square errors are higher than those resulted by static and even dynamic with one x variable stratifications (Table 3). Only by using very few (2–4) stable classes of the 2nd PC was it possible to achieve a better accuracy (*rmse* 8.38–8.89).

The final class numbers of both x variables characterizing the best stratification results were used as initial ones and new stratifications were performed. In most cases they could not be completed but in some the results were a little bit surprising. The root mean square errors were similar to the results of stratification using only one x variable (the 1st PC) and less than those achieved using the same final values of two x variable class numbers but higher than in static stratification (Table 3).

Estimation errors were correlated with the "true" field variables (Table 6). Lower values tend to be overestimated and higher values underestimated.

Table 5. Dynamic stratification using different x variable class combinations.

a) Number of the 1st PC classes changes while the number of 2nd PC classes remains stable.

Number of 1st PC classes	Number of 2nd PC classes: 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	% of assigned B-observations																		
1										100			100	100	100	100	100	100	100
2					100	100	100	100	100	99.6	100	100	99.9	99.6	99.3	99.4	99.1	99.3	99.3
3			100	100	98.4	99.0	99.3	99.3	99.3	97.3	98.7	98.7	98.4	97.1	97.6	98.1	96.1	95.0	96.1
4	100		99.9	99.7	97.9	98.9	98.6	99.0	98.7	96.9	98.0	98.3	97.7	95.7	95.9	97.7	95.6	94.6	95.1
5	99.7		99.6	99.6	97.7	98.7	98.4	98.6	98.6	96.7	97.9	97.9	97.4	95.6	95.6	97.1	95.1	94.3	94.8
6	99.7		99.6	99.6	97.7	98.6	98.4	98.4	98.6	96.4	97.1	97.6	96.1	94.7	95.3	95.7	93.4	92.7	93.8
7	99.7		99.6	99.3	97.3	98.6	98.1	97.9	97.7	95.7	96.1	96.9	95.0	93.7	94.3	95.0	91.8	91.6	92.7
8	99.7	100	99.6	99.3	97.3	98.6	97.6	97.6	97.7	95.4	95.1	96.0	94.6	93.1	92.6	94.1	91.0	90.7	92.0
...																			
20	99	95.4	92	92.1	84.8	88.0	80.5	78.3	78.8	74.1	71.8	69.8	72.1	63.8	67.8	63.8	62.4	61.9	62.1

b) Number of the 1st PC classes is stable while the number of 2nd PC classes changes.

Number of 2nd PC classes	Number of 1st PC classes: 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	% of assigned B-observations																		
2	100	100		99.7	99.9			99.7	99.7	99.9	99.9	99.7	99.7	99.7	99.9	99.9	99.7	99.7	99.7
3	99.9	99.9	100	99.7	99.9	100	100	99.7	99.7	98.7	99.7	99.6	99.6	98.6	97.6	99.7	99.0	97.9	98.6
4	99.3	99.6	99.9	99.7	99.6	99.9	99.9	98.3	99.4	97.6	99.1	98.7	98.1	97.9	96.3	97.1	96.7	96.1	96.7
...																			
20	97.1	95.3	92.7	92.8	88.4	83.5	84.3	83.1	78.8	79.4	76.0	76.5	67.8	70.0	68.7	64.9	65.8	62.7	62.1

Table 6. Pearson's correlation coefficients between the estimation errors and field variables (final values of the 1st and 2nd PC class numbers 4 and 2).

Estimation errors of:	Basal area	Mean diameter	Mean height	Mean age
Basal area	–0.541	–0.382	–0.435	–0.360
Mean diameter	–0.373	–0.576	–0.551	–0.542
Mean height	–0.434	–0.562	–0.578	–0.535
Mean age	–0.322	–0.495	–0.478	–0.603

Table 7. Relationship (R^2) between the development class, 1st TM PC and field variables.

Parameters	1st TM PC	Basal area	Mean diameter	Mean height	Mean age
	Degree of determination (R^2)				
Development class	0.40	0.60	0.90	0.89	0.85

Table 8. Dynamic stratification procedures using development stage and the 1st TM PC as classification variables (all development stage classes).

Number of the 1st PC classes																
20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4
% of assigned B-observations																
93.7	95.8	96.4	96.9	97.0	97.0	97.0	97.3	97.3	97.7	97.7	98.4	98.7	98.7	98.9	99.3	100

3.4 Dynamic Stratification with the Development Stage

It was expected that the inclusion of other than TM PC auxiliary data could improve the results of dynamic stratification. The development class of the main storey was chosen as such data. It can be quite objectively and easy estimated e.g. on aerial photographs. This qualitative variable is described by the following 9 classes: open stand, seed tree stand, small saplings under 1.3 m, saplings above 1.3 m, first thinning stand, second and third thinning stand, mature stand, shelterwood stand and uneven aged stand. The closeness of the relationship of the development class to field variables and the 1st TM PC values can be characterized by the degree of determination R^2 (Table 7). The high relationship with field variables and comparatively less with TM PC values indicates better stratification results. The development class was used as the $x1$ and the 1st TM PC as the $x2$ variable. The stratification procedure is illustrated in Table 8.

The errors (Table 9) are significantly less than those of previous dynamic stratification variants. The root mean square errors of mean diameter, height and age (the variables most correlated with development stage) – 3.34 cm, 2.65 m and 14.06 years are less than half of that in static stratification.

In order to compare the errors correctly the development stage class was included in the PC transformation. New PCs, based on quantitative TM values and qualitative development stage classes were calculated. The errors of static stratification using the new PCs are presented in Table 10. They are similar for estimating basal area but noticeably higher for mean diameter, height and age.

The increase of stratification accuracy is mainly resulted by the use of the development stage class as an auxiliary variable. The root mean square errors when the development stage class is used alone in dynamic stratification are similar to those with development stage and 1st TM PC as auxiliary variables. To interpret all nine classes with high level of accuracy may be rather complicated. Using larger development stage

Table 9. Assessment of dynamic stratification with the development stage and 1st TM PC as auxiliary variables.

Stratification method	No. of development stage classes[1]	Basal area $\bar{e}$	Basal area S_e^2	Basal area rms	Mean diameter $\bar{e}$	Mean diameter S_e^2	Mean diameter rms	Mean height $\bar{e}$	Mean height S_e^2	Mean height rms	Mean age $\bar{e}$	Mean age S_e^2	Mean age rms
Development stage and 1st PC	9	0.63	6.37	6.40	–0.09	3.34	3.34	0.10	2.64	2.65	0.29	14.06	14.06
Development stage	9	0.36	6.31	6.32	–0.16	3.29	3.29	0.03	2.71	2.71	0.24	13.48	13.48
Development stage and 1st PC	7	0.71	6.41	6.45	0.16	4.84	4.85	0.28	4.01	4.02	1.58	21.79	21.85
Development stage	7	0.46	6.50	6.51	–0.02	4.79	4.79	0.16	4.12	4.12	0.93	20.44	20.46
Development stage and 1st PC	6	0.68	6.37	6.40	0.16	4.85	4.86	0.29	4.04	4.05	1.66	21.83	21.89
Development stage	6	0.45	6.45	6.48	0.01	4.84	4.84	0.18	4.17	4.17	1.04	20.47	20.50
Development stage and 1st PC	4	0.54	6.52	6.54	0.07	4.99	4.99	0.22	4.11	4.11	1.41	22.09	22.14
Development stage	4	0.38	6.55	6.56	–0.05	4.91	4.91	0.14	4.20	4.21	0.87	20.61	20.63
Development stage and 1st PC	3	0.52	6.56	6.58	0.03	5.01	5.01	0.19	4.14	4.15	1.18	22.21	22.24
Development stage	3	0.33	6.65	6.65	–0.10	4.99	4.99	0.10	4.27	4.27	0.68	20.89	20.90

Field variable error statistics.

[1] Grouping of the development stage classes:

9 classes Every class forms a separate group
7 classes 1 – open stand, 2 – seed tree stand, 3 – small saplings under 1.3 m, 4 – saplings above 1.3 m, 5 – thinning stands, 6 – mature and shelterwood stands, 7 – uneven aged stands
6 classes 1 – open stand, 2 – seed tree stand, 3 – small saplings under 1.3 m, 4 – saplings above 1.3 m, 5 – thinning stands, 6 – mature, shelterwood and uneven aged stands
4 classes 1 – open stands, 2 – seed tree stand and saplings, 3 – thinning stands, 4 – mature, shelterwood and uneven aged stands
3 classes 1 – open stand, seed tree stand and saplings, 2 – thinning stands, 3 – mature, shelterwood and uneven aged stands

Table 10. Assessment of static stratification using PCs based on TM and development stage class values.

Stratification method	Basal area $\bar{e}$	Basal area S_e^2	Basal area rms	Mean diameter $\bar{e}$	Mean diameter S_e^2	Mean diameter rms	Mean height $\bar{e}$	Mean height S_e^2	Mean height rms	Mean age $\bar{e}$	Mean age S_e^2	Mean age rms
Euclidean distances	0.74	6.48	6.52	0.43	4.89	4.91	0.51	3.87	3.91	2.43	20.05	20.20
Bayesian probabilities	0.69	6.59	6.62	0.37	4.89	4.90	0.45	3.86	3.88	1.83	20.14	20.22
Gaussian probabilities	0.64	6.61	6.65	0.34	4.95	4.96	0.42	3.88	3.90	1.61	20.26	20.32
Mahalanobis distances	0.81	6.64	6.69	0.53	5.13	5.16	0.56	4.01	4.04	2.27	21.48	21.66

Field variable error statistics.

classes the inclusion of TM variables can improve the dynamic stratification results, except the mean age and in particularly cases mean diameter estimations (Table 9). Only three classes – (i) open, seed tree and sapling stands, (ii) thinning stands and (iii) other – which can be more easily defined on aerial photographs allow us to hope similar results in basal area and a little bit lower in other characteristics evaluation comparing to the use of nine development stage classeses. Practically there is no need to use larger number of development stage class groups (except all nine), because the estimation errors are rather similar.

4 Discussion

The results of the study do not show promise to achieve more accurate estimates using only principal components as auxiliary data in dynamic stratification opposed to static stratification. On the contrary – the errors are somewhat larger. Extrapolation of the results may be limited because of the small study area. But, if methods are unsuccessful at such a small site, they are unlikely to be more successful across broader regions. Using x variables of different origin (qualitative development stage class and quantitative PC) dynamic stratification appears to be much more advanced as static stratification. This increase of stratification accuracy is mainly resulted by the use of development stage class. It was chosen of the estimation objectivity and simplicity using remote sensing.

Using only one x variable (the 1st TM PC) the minimum number of classes should be no less than 4, above that it has practically no influence over the stratification procedure – all B-observations are assigned values. The average stratification errors using two PC's tend to decrease with the lessening of the total number of x variable classes. It is possible to achieve the same or even better results using certain classes of both x variables as the initial ones in the stratification procedure. Low values are overestimated and high values underestimated.

Significant differences in the error statistics lead to the conclusion, that in dynamic stratification the non-measured observations should be assigned the mean characteristics of the measured observations that belong to the same cube and not randomly selected ones.

Acknowledgements

The study was carried out in the Department of Forest Resource Management, University of Helsinki under direct supervision of Mr. Mark-Leo Waite and Prof. Simo Poso.

References

Hägglund, B. 1981. Forecasting growth and yield in established forests. Swedish University of Agricultural Sciences, Department of Forest Survey, Report 31: 39–43.

Lillesand, T.M. & Kiefer, R.W. 1994. Remote sensing and image interpretation. Third edition. John Wiley & Sons, Inc.

Tomppo, E. 1993. Multi-source National Forest Inventory of Finland. In: Proceedings of Ilvessalo Symposium of National Forest Inventories. Finnish Forest Research Institute, Research Papers 444: 52–60.

Waite, M-L. 1990. The SMI forest management system. Proc. SNS/IUFRO Workshop in Umeå, 26–28 February, 1990. Remote Sensing Laboratory, Swedish University of Agricultural Sciences, Report 4: 106–113.

— 1993. Documentation of the SMI/92 Forest Management Planning System.

Total of 5 references