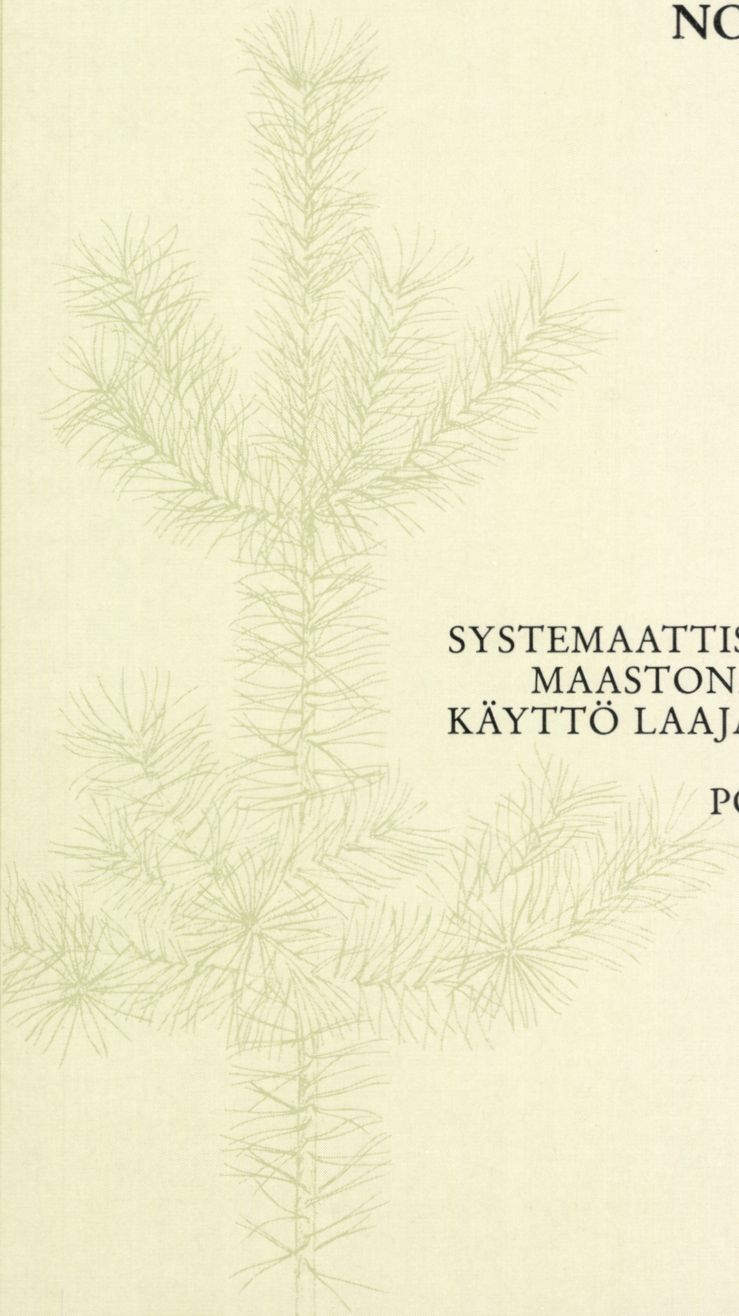
COMMUNICATIONES INSTITUTI FORESTALIS FENNIAE 131

THE COMBINED USE OF SYSTEMATIC FIELD AND PHOTO SAMPLES IN A LARGE-SCALE FOREST INVENTORY IN NORTH FINLAND

EERO MATTILA



SELOSTE

SYSTEMAATTISEN ILMAKUVA- JA MAASTONÄYTTEEN YHTEISKÄYTTÖ LAAJAN METSÄALUEEN INVENTOINNISSA POHJOIS-SUOMESSA

HELSINKI 1985

COMMUNICATIONES INSTITUTI FORESTALIS FENNIAE

THE FINNISH FOREST RESEARCH INSTITUTE
(METSÄNTUTKIMUSLAITOS)

Unioninkatu 40 A
SF-00170 Helsinki 17
FINLAND

telex: 125181 hyfor sf
attn: metla/

phone: 90-661 401

Director:
Professor Aarne Nyyssönen

Head of Information Office:
Olli Kiiskinen

Distribution and exchange of publications:

The Finnish Forest Research Institute
Library
Unioninkatu 40 A
SF-00170 Helsinki 17
FINLAND

Publications of the Finnish Forest Research Institute:

— Communicationes Instituti Forestalis Fenniae (Commun. Inst. For. Fenn.)
— Folia Forestalia (Folia For.)
— Metsäntutkimuslaitoksen tiedonantoja

Cover (front & back): Scots pine (*Pinus sylvestris* L.) is the most important tree species in Finland. Pine dominated forest covers about 60 per cent of forest land and its total volume is nearly 700 mil. cu.m. The front cover shows a young Scots pine and the back cover a 30-metre-high, 140-year-old tree.

COMMUNICATIONES INSTITUTI FORESTALIS FENNIAE

131

EERO MATTILA

THE COMBINED USE OF SYSTEMATIC FIELD AND PHOTO SAMPLES IN A LARGE-SCALE FOREST INVENTORY IN NORTH FINLAND

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public criticism in Metsätalo, auditorium II, on October 18th, 1985, at twelve o'clock noon.

SELOSTE

SYSTEMAATTISEN ILMAKUVA- JA MAASTONÄYTTEEN YHTEISKÄYTTÖ LAAJAN METSÄALUEEN INVENTOINNISSA POHJOIS-SUOMESSA

HELSINKI 1985

MATTILA, E. 1985. The combined use of systematic field and photo samples in a large-scale forest inventory in North Finland. Seloste: Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa. Commun. Inst. For. Fenn. 131: 1—97.

The study concerns the two-phase systematic sampling design used in the 7th National Forest Inventory in the Forestry Board District of Koillis-Suomi, North Finland. The inventory was carried out in the study area in 1982—83.

The systematic field sample permits preliminary estimates independent of the photo sample. The total field sample consists of remeasured semi-permanent plots of the previous inventory as well as new temporary and semi-permanent plots. The remeasured plots are used to study the changes of the forest characteristics during the inventory period. They can also be used for the preliminary estimation of the present status together with the new field sample.

The new field sample plots are a part of the new photo sample. The photo plots can be interpreted both before and after the measurement of the field sample which alters the method flexible. The samples are combined by communes six in number. The field data are extended to the whole photo sample by homogenous photo strata. Results can be calculated with ease from the combined sample because all photo plots obtain formally complete field data.

Expressed as relative standard error, the new field sample gives a total volume estimate in the study area with a sampling error of 4—5 %. The photo sample reduces the mean square error to about one third. Taking into account the systematic allocation of the samples, the sampling error of the estimated total volume in the combined sample is 2,1—2,5 %. The results concern a land area of 25 700 sq.km.

The potential uses of the samples are discussed in the study. The two main alternatives employ the original samples and the additional sampling. Both still include several possibilities. The final purpose of the study was to find those details in the inventory method to be improved in the future. These are: photo material, type and size of the plot and field tract and the extension of field data to the photo sample.

Tutkimuksessa tarkastellaan kaksivaiheista systemaattista otantamenettelyä, jota käytettiin valtakunnan metsien inventoinnissa Koillis-Suomen piirimetsälautakunnan alueella vuosina 1982—83.

Systemaattisesta otannasta johtuen pelkästä maastonäytteestä voidaan laskea alustavia harhattomia estimaatteja. Koko maastonäyte sisältää edellisen inventoinnin uudelleen mitatut puolipysyvät koealat ja uuden maastonäytteen, jossa on sekä tilapäisiä että puolipysyviä koealoja. Uudelleen mitattuja koealoja käytetään inventointijakson aikana tapahtuneiden muutosten analysointiin. Niitä voidaan myös käyttää yhdessä uuden maastonäytteen kanssa alustavien estimaattien laskentaan.

Uuden maastonäytteen koealat sisältyvät uuteen ilmakuvanäytteeseen. Ilmakuvakoealat voidaan tulkita osittain tai kokonaan maastonäytteen mittaamisen jälkeen, mikä tekee menetelmästä ajallisesti joustavan. Näytteet yhdistetään kunnittain, joita tutkimusalueella on kuusi. Maastotiedot laajennetaan koko ilmakuvanäytteeseen homogeenisten ilmakuvaositteiden sisällä. Kaikki tulkintakoealat saavat muodollisesti täydelliset maastotiedot, mikä tekee tulosten laskennan yhdistetystä näytteestä helpoksi.

Puuston runkotilavuuden estimaatin otantavirhe ilmaistuna suhteellisena keskivirheenä on tutkimusalueella uuden maastonäytteen perusteella 4—5 %. Ilmakuvanäyte pudottaa keskineliövirheen noin kolmannekseen. Tutkimusalueen puuston runkotilavuuden estimaatin suhteellinen otantavirhe on yhdistetyn näytteen perusteella arvioituna ja ottaen huomioon näytteiden systemaattisuus 2,1—2,5 %. Maapinta-ala tutkimusalueella on 25 700 km^2.

Tutkimuksessa käsitellään myös näytteiden käyttömahdollisuuksia. Kaksi päävaihtoehtoa ovat alkuperäisten näytteiden hyödyntäminen ja lisäotanta. Nämä molemmat sisältävät vielä useita mahdollisuuksia. Eräs tutkimuksen tavoite oli saada selville ne inventointimenetelmän yksityiskohdat, jotka ovat kehittämisen tarpeessa tulevaisuudessa. Näitä voidaan luetella ilmakuvamateriaali, koealan ja maastolohkon koko ja rakenne ja maastotiedon laajentamismenettely.

ODC 585 + 524.41 + 524.6 + 905.2 + (480.99)
ISBN 951-40-0702-6
ISSN 0358-9609

Helsinki 1985. Valtion painatuskeskus

CONTENTS

1. INTRODUCTION 5
   11. General 5
   12. Grouping method 5
   13. Developments of field based methods in Scandinavia 7
   14. Some observations on NFS outside Scandinavia 9
   15. The setting in Finnish Lapland 10
   16. Study objects 12
2. INVENTORY DESIGN 14
   21. Goal setting 14
   22. Sampling procedure 15
      221. Photo sample 15
      222. Field sample 17
   23. Study material 20
3. PRELIMINARY ESTIMATION ON THE BASIS OF THE FIELD SAMPLE .. 21
   31. Characteristics and formulae 21
   32. Results and conclusions 24
4. USE OF THE PHOTO SAMPLE 36
   41. Statistical basis 36
   42. The accuracy of assessments from aerial photographs 38
      421. Qualitative variables 38
      422. Volume 43
   43. Stratification 46
   44. Extension of the field data 51
      441. The basis 51
      442. Practical realization 52
   45. Statistical precision 55
      451. Sampling of the photo points 55
      452. Error due to the interpretation 60
      453. Total sampling error 64
5. POTENTIAL USES OF THE SAMPLES 71
   51. Original samples 71
      511. Sub-area or a set of sub-areas as target area 71
      512. Target area other than sub-area or a set of sub-areas 72
         5121. New combination of the samples 72
         5122. Original combination of the samples 73
   52. Additional sampling 79
      521. Increase of the photo sample 79
         5211. Effect on the error of the photo sample 79
         5212. Effect on the error due to the interpretation 81
         5213. Order of the possible applications 82
      522. Increase of the field sample 82
6. GENERAL CONSIDERATIONS 84
7. SUMMARY 88
REFERENCES 91
SELOSTE 95

PREFACE

Prof. Kullervo Kuusela, head of the Finnish National Forest Inventory, at the beginning of the 1980's defined the headlines of the method to be used in the next forest inventory in Finnish Lapland. My duty has been to plan the practical realization and to carry out a statistical analysis of the forest inventory in question. The project has advanced according to the timetable, mainly due to competent co-workers. I appreciate having worked with following persons at different phases of the work.

Maps: Lauri Vuorisalo. Aerial photographs: Pertti Virtanen, Aarne Alanen and Lauri Silander. Photointerpretation: Jouni Kulju, Juhani Kumpuniemi and Pertti Virtanen. Field work: Aulis Heino, Ossi Kivistö, Ilkka Kohmo, Matti Kujala, Jouni Kulju, Juhani Kumpuniemi, Sakari Salminen, Pertti Virtanen, Lauri Vuorisalo and Hannu Yli-Kojola. Data recording: Irmeli Häjänen, Salme Säreströniemi and Raija Vainio. Calculations: Matti Kujala, Juhani Kumpuniemi and Sakari Salminen. Edp-consultation: Pirjo Kaila, Erkki Kaila and Mauri Timonen. Statistical consultation: Jaakko Heinonen, Risto Häkkinen, Simo Poso, Sakari Salminen and Sven Svensson. Typing: Sirpa Hast, Eila Kuirinlahti, Anja Leskinen and Aulikki Kylänpää. Figures and tables: Annikki Ryynänen. General functions: Anja Leskinen, Pirkko Mattila, Pekka Oksman, Annikki Ryynänen, Yrjö Sihvo and Lea Suhonen.

The manuscript has been read and criticized with a constructive purpose by professors Kullervo Kuusela, Jouko Laasasenaho and Yrjö Vuokila. Their comments have been taken into account in the final formulation of the text. The language has been checked by Ashley Selby. Gratefully for the patience of all You and others.

Rovaniemi, May 1985

Eero Mattila

1. INTRODUCTION

11. General

The practice of national forest inventories began in the Scandinavian countries in the 1920's. Sampling was preferred to complete surveys due to the relatively large forest areas. Cruising in the field was at that time the only way to acquire information of the forests. The sampling designs were initially linear, in the form of systematic strip and line-plot surveys. The estimation of the statistical precision was discussed by Langsaeter (1926), Lappi-Seppälä (1924) and Lindeberg (1923). The surveys have been repeated several times with varying sampling designs. The methodology will be briefly discussed later in this chapter.

The national forest inventory methods in Finland, Sweden and Norway have been developed, until quite recently, on the basis of one-phase field sampling. There are many reasons for this state of affairs (Nyyssönen et al. 1968, p. 5). An inventory system, once established, tends to resist abrupt changes — which is a good thing when considering cost and continuity. Perhaps the main reason for the dominance of field estimation, however, is the quality of the information required. Timber production is fairly intensive and forest planning needs detailed data of the inner characteristics of the growing stock and the sites. The forests are relatively accessible, which also favours field cruising. The areal distribution of forest land is fairly even which lessens the efficiency of using maps and aerial photographs as the first-phase source of information in a large-area forest inventory.

There is a considerable change in the forest conditions between the south and the north, for example, of Finland. Forestry is more extensive and the forests are less accessible in Lapland. In addition, the structure of the growing stock is more simple which makes photo-interpretation more accurate. These factors together increase the efficiency of aerial photographs as one source of information in the sampling.

Research concerning the combined use of photo-interpretation and field estimation in the National Forest Inventory in Finland was started at the end of the 1960's. The method developed was used in the northernmost part of Lapland in 1970 and 1978 and in 1974—1976 in the other parts of Lapland. This so called grouping method is described in the following section. developments of the field-based methods in the Scandinavian countries will be discussed thereafter.

12. Grouping method

The grouping method (Poso 1972) is a special application of two-phase sampling for stratification (Neyman 1938). A systematic photo sample is first interpreted from small-scale black and white aerial photographs. The main variables interpreted are land use class, sub-class (division into mineral soils and peatland), dominant tree species and development class as well as volume of the growing stock. The purpose of interpreting the photo sample is to estimate the stratum weights.

The second step gives the name to the method. The photo plots are divided into internally homogeneous groups on the basis of the interpreted data. The goal in forming of the groups is to minimize the variation of the true field data within the groups. One plot is randomly selected from each group to be measured in the field. In this way we get the most representative field sample as to the variation present in the photo sample.

The fraction between the field and photo samples determines the average size of the groups. Allocation of the field sample can be made more efficient by using different group sizes in different photo strata. This, however, leads to a varying field sampling intensity which causes problems when wishing to calculate results from the bare field sample. The group size by interpreted land use class and inventory year during 1974—1976 was as follows:

Inventory year	Interpreted land use class[1]) 1	2	3	4—7
1974	7,7	8,0	10,2	9,4
1975	7,0	7,5	7,3	7,1
1976 S[2])	7,5	8,1	10,3	8,2
1976 N[2])	7,0	7,6	11,4	10,4

[1]) 1 = Forest land 2 = Scrub land 3 = Waste land 4—7 = Other land area
[2]) S = South Lapland N = North Lapland

The field sampling intensity mainly decreases with decreasing potential timber production capacity. There appears, however, some irrational deviations from this rule.

The field measured data is extended to the whole photo sample by groups (Figure 1). In this way all photo plots receive formally complete field data which makes it possible to calculate results for sub-areas independently of the actual field sample. The statistical precision of the results in a subarea depends on the number of the photo plots and on the accuracy of the photo-interpretation.

The size of the groups also varies a little within the photo strata because the photo sample can rarely been divided into groups of an exactly equal size by stratum. This causes an additional variation of the estimates when the results are calculated from the combined field and photo sample. Let us use notations

n = number of photo plots
n_h = number of photo plots in stratum h
m_h = number of field plots in stratum h
x_{hj} = volume of field plot j in stratum h
$\bar{x}_h = (\sum_{j=1}^{m_h} x_{hj})/m_h$; mean volume in stratum h
$w_h = n_h/n$; weight of stratum h
l = number of strata

In double sampling for stratification mean volumes are calculated as

$$\bar{x} = \sum_{h=1}^{l} w_h \bar{x}_h$$

The field plots are of an equal weight by stratum because the $\bar{x}_h$'s are calculated as arithmetical means. After the extension of the field data to the photo sample, mean volumes are calculated as

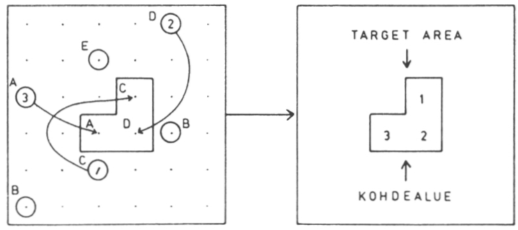
$$\bar{x}' = (\sum_{i=1}^{n} x_i)/n = \sum_{h=1}^{l} w_h \bar{x}'_h$$

where x_i = true or extended volume of photo plot i

and $$\bar{x}'_h = (\sum_{i=1}^{n_h} x_{hi})/n_h.$$

Deviations occur between $\bar{x}$ and $\bar{x}'$ because n_h/m_h (= theoretical group size) is rarely an even number and all x_{hi}'s are not equal to $\bar{x}'_h$. It is important from the statistical point of view that the deviations are random. This additional source of variation is in fact, the only "price" of the extension of the field data. It can be concluded, however, that the extra sampling error is insignificant in relation to the total variance with the sampling design used in the National Forest Inventory in Lapland.

The efficiency of the grouping method depends on the quality and price of the data interpreted from the aerial photographs. With the photo material used in Lapland, the estimation of the forest land area and the total growing stock is more efficient using the grouping method than when using



A - E = group sign - *ryhmän 'merkki'*
o = field plot - *maastokoeala*
1 - 3 = value of a measured quantity - *maastossa mitatun suureen arvo*

Figure 1. Extension of the field data to the photo sample by grouping method. A schematic illustration. The estimated mean in the target area is 2.
Kuva 1. Maastotiedon laajennus tulkintanäytteeseen ryhmitysmenetelmällä — kaavamainen esimerkki. Keskiarvon estimaatti kohdealueella on 2.

one-phase systematic field sampling (Poso 1972, p. 86). The grouping method is not efficient for estimating the inner characteristics of the growing stock and sites, not to mention the necessary silvicultural measures done or to be done.

The grouping method is flexible with regard to the calculation of results for subareas. Applied as described above, however, it is rigid in the sense that the photo sample must be finished before the field sample can be selected. Another inconvenience is that it is difficult to calculate unbiased results from the field sample alone if the field sampling intensity is variable.

The grouping method was applied in its basic form in the northernmost Lapland in 1970. The photo sample covered the whole study area but the field sample was concentrated to a part of the photo tracts (Poso and Kujala 1971, p. 8). The field sampling intensity varied considerably depending on the interpreted land use class. In the following inventory in 1978 the field sample was increased in the best timber production areas while the photo sample was kept unchanged (Mattila and Kujala 1980, p. 6). In the southern part of Lapland in 1974—1976, the photo sample was concentrated to tracts which covered 50 % of the study area (Poso and Kujala 1977, p. 42—43). The photo plots were further concentrated to some degree within the tracts. Field plots were measured only in every third photo tract. The main deviation from the basic application was that some 35 % of the field sample was systematic and selected independently of the photo-interpreted data.

13. Developments of field based methods in Scandinavia

The three first National Forest Inventories in Finland were carried out with 15 years intervals (1921—1924, 1936—1938 and 1951—1953). The data was often considered, and also proved to be out-of-date in the latter part of the inventory period. The inventories gave results which could not always be expected on the basis of the earlier information (Ilvessalo 1962 p. 8—9). For this reason it was decided to carry out an intermediary survey of the forests in the beginning of the 1960's. As in 1951—1953, the inventory was made along continuous lines which ran from the southwest to the norheast (Ilvessalo, ibid., p. 13, Tiihonen 1966, p. 8). Due to the character of the survey, the sampling intensity was relatively low. The Fourth National Forest Inventory initially aimed to cover only the southern part of Finland. During the inventory, however, adequate financial resources made it possible to extend the inventory to the whole country. The survey was made during 1960—63. A very important practical arrangement was realized in 1962 as the inventory was made continuous. Every year from 1964 onwards the inventory would proceed stepwise from region to region. The yearly grants for the purpose have determined an inventory period of 7—8 years.

The Fifth National Forest Inventory in Finland was carried out during 1964—1970. Two-phase photo and field sampling was then used in the northernmost part of Lapland as described earlier. The method used in the other parts of the country was tractwise line-plot survey with a systematic layout of the tracts (Kuusela and Salminen 1969, p. 10). The field survey area was divided into two parts, with a differing layout and structure for the tracts. The statistical precision of the results for the field survey area was presented by Salminen (1973).

A remarkable reform in the 5th nationwide forest inventory was the adoption of the detached tracts instead of the continuous survey lines. The change was favoured by social developments and improved communications. Another basic decision concerned the kind of field plots. The idea of the angle gauge technique had been presented in the late 1940's (Bitterlich 1948). The use of relascope plots in the fourth inventory in Finland was considered but the final choice was still for fixed-area plots (Ilvessalo 1962, p. 14). In the Fifth National Forest Inventory, however, the situation was reversed. The calculational principles of the relascope technique employed in the forest inventories have been presented by Kuusela (1960, 1966).

The sixth National Forest Inventory in Finland was finished in 1976. Two-phase photo and field sampling was used in Lapland. The layout and structure of the tracts was the same in the whole field survey area. The distance between the tracts was 8 km.

Each tract consisted of a 4 100 m long survey line including 41 plots at 100 m intervals (Salminen 1981, p. 8—9).

The Seventh National Forest Inventory was completed in South-Finland in 1982 (Kuusela and Salminen 1983). The layout of the tracts was the same as in the preceding inventory with the exception that the new tracts were located between the old ones. In the earlier inventories, the survey line had been measured for the estimation of the stratum proportions, but in the 7th inventory, line measurements were made only for the location of the plots. The stratum proportions are estimated from the distribution of the plots. In this respect the method is the same as that used in Lapland. The line is not measured in Lapland because the plots there are located with the aid of aerial photographs.

The results of the seven nationwide forest inventories of Finland discussed above form a 60 year long time series. For a couple of communes in South Finland the period is even longer. Verner Cajanus carried out an experimental survey of Sahalahti and Kuhmalahti in 1912 to find an appropriate method for large-area forest inventories (Ilvessalo 1923, p. 2). An experiment for the same purpose was also made in Sweden in 1911—1912 in the county of Värmland. The experiences of Sweden were also utilized in Norway (Taksering . . . 1970, p. 91). However, the line-survey had been studied in practice in Norway as early as 1909 (Lappi-Seppälä 1924, p. 12).

The two first national forest surveys of Sweden were carried out duing 1923—1929 and 1938—1952. These inventories progressed county by county as is the case nowadays in South-Finland. From 1953, a sparse sample covering the whole country has been measured every year in the inventory in Sweden. The yearly sampling intensity was adjusted for an inventory period of 10 years. So the 3rd inventory was finished in 1962, the 4th in 1972 and the 5th in 1982 (e.g. Ohlsson 1975, p. 6, Hägglund 1983, p. 1).

Continuous survey strips and lines were used in Sweden until 1952. The inventory introduced the tract as the primary sampling unit in 1953. The growing stock has been measured within circular plots since 1938. The plots have also been the basis for the main part of the area information since 1953.

It soon became evident in Sweden that new information about forest conditions was needed more frequently than 10 years. Following this requirement Arman (1969) presented results which were calculated from samples from the years 1958—1967 i.e. the samples were from the latter half of the 3rd inventory and the first half of the 4th inventory. The results based on the latter half of the 4th inventory were published in 1975 (Eriksson and Janz 1975). The 5-year period, with a poorer statistical precision of the results, was accepted because of the rapidly changing state of the forests (Svensson 1980, p. 2).

The demand for a faster data production was taken into account in the planning of the 5th inventory. The sampling intensity was increased to make the results for 5-year periods adequately accurate, cf. Ohlsson (1975), Bengtsson (1978) and Svensson (1980). The results for the first half of the 5th inventory (1973—1977) have been published by Svensson (ibid.). The latest information from the years 1978—1982 has for the present been presented as annual statistics (Skogsdata 83). The statistical aspects concerning the 5th National Forest Survey of Sweden have been discussed by Ranneby (1981) and Svensson (1983).

A project called NYTAX was started in 1978 for planning the 6th Swedish inventory, starting 1983. The several subjects of the project extend from goal setting and statistical aspects to data gathering techniques. The project is presented in nine NYTAX-reports which are listed in the references of the paper of Hägglund (1983). The last report (Jonasson 1982) deals with experimentation of the new inventory in practice.

In the 6th National Forest Survey of Sweden, temporary and permanent tracts are measured every year troughout the whole country. The permanent tracts will be established during 1983—1987 and remeasured during 1988—1992. The country is divided into five regions, each with a different structure and layout of the tracts (Hägglund 1983, p. 10, 14, 69—74). The division into the regions was changed to some degree from the earlier inventory. Fixed-area plots are still used, but the radius is different for different purposes (Hägglund, ibid., p. 13). As to the data management aspects, all data is now gathered in digital form in the field.

The First National Forest Survey of Norway was carried out during 1919—1930 and the second one during 1937—1956. The inventories were carried out county by county using the strip-survey method (Taksering . . . 1933, 1970). Circular plots and separate tracts were introduced to the inventory in the last year of the 2nd survey. Tractwise line and plot survey was used in the 3rd (1957—1963) and 4th (1964—1976) inventories of the norwegian forests (Taksering . . . 1970, p. 12, Nersten 1981, p. 23). The last inventory consists of a series of sparse yearly samples each of which covers the whole country. Since 1970, the trees for the volume and growth estimation have been selected by the angle gauge method.

The 5th National Forest Survey was started in Norway in 1982. It was considered to be necessary to return to the county-by-county principle in the progression of the inventory (Nersten 1981, p. 33). The layout and structure of the tracts is variable, as in Sweden (Instrukt . . . 1982, p. 2). Relascope plots are now used instead of fixed-area plots, as in Finland (Instrukt . . . ibid., p. 10). Data recording in the field is still done by the conventional method using pen and blank forms.

Finland, Sweden and Norway comprise an unique geographical region with respect to national forest surveys (NFS). The results obtained by objective sampling extend over a relatively long time period (Figure 2). The inventories can be considered of high frequency, expecially at present. The third prominent feature is the dominance of the methods based on field sampling only. The reasons for this have been discussed earlier in this paper.

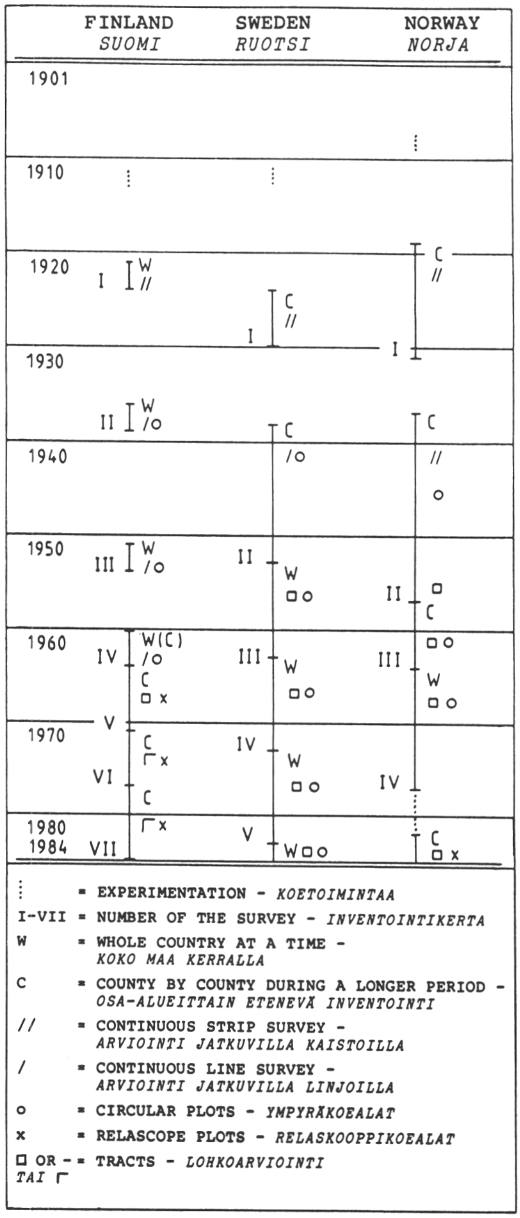


Figure 2. Timetable and method of the National Forest Surveys based on field sampling in Finland, Sweden and Norway.

Kuva 2. Maasto-otantaan perustuvien valtakunnan metsien inventointien aikataulu ja menetelmä Suomessa, Ruotsissa ja Norjassa.

14. Some observations on NFS outside Scandinavia

Only in the Scandinavian countries in Europe have nationwide statistical forest surveys been conducted before the 1960's. For the present, objective data concerning the national forest resources is available also from England, Spain, The Democratic Republic of Germany, Austria and France (Kuusela 1984, p. 73). A first inventory has also been started in Switzerland. In cerning the growth, drain and total volume Europe, a system of recurrent inventories has been introduced outside Scandinavia only in Austria. Only a time series of repeated inventories over a longer period reveals the actual reliability of the data con-of the forests.

Those countries which started large-area forest inventories after the second world war, very often utilize aerial photographs in some way or other. The First National Forest Survey of Switzerland (1983—1986), for example, includes interpretation of stereo aerial photos employing modern equipment (Schmid-Haas 1983, p. 53). The modern techniques have not, however, made ground checking unnecessary in objective forest surveies. Field plots can be reduced but, at the same time, the plots must be located and the measurements must be made more carefully.

Another prominent feature in the post-war large-area forest inventories is the establishment of permanent plots in the field. The idea is also being introduced in the "old" inventory countries, as for example in Finland (Mattila 1983, Päivinen and Yli-Kojola 1983) and Sweden (Hägglund 1983). Reports concerning the use of permanent field plots as a part of a national forest inventory are available from Switzerland (Schmid-Haas ibid.), Taiwan (Yang 1981) and Brazil (Netto 1981).

Permanent plots usually have a fixed area and the trees measured on the initial occasion are mapped carefully within the plots to avoid confusion on the successive occasions. Much attention is also paid to the marking of the plots for furure relocation. The plots and the trees are intended to be recognizable for tens of years. This course of action is in many respects different from the experiments made in North-Finland. Kuusela (1979) presented the principles of using permanent relascope plots. The plots are remeasured only once, after 5—7 years, and are therefore called semi-permanent. The marking of the plots has been reduced to a minimum level.

Use of aerial photographs and permanent plots is very usual in North-America. A part of the permanent plots have been remeasured four to five times (Ek 1983, p. 132). CFI and SPR (p.) are utilized, especially in management plan inventories. Statewide forest inventories are conducted by the U.S. Forest Service (for further information see U.S.D.A. . . . 1982). According to Ek (ibid), permanent plots are also utilized in the statewide inventories.

Use of satellite imagery is the most modern feature in natural resource surveys. By this means it has become possible to survey vast remote areas with feasible costs. The tropics and the northern sub-arctic and arctic regions are the most typical study objects in this context. A tremendous amount of literature concerning the subject is available. The modern trends include an interest in all vegetation and in the physical environment as a whole. In the United States and in Canada there are large survey projects using satellite imagery which are called integrated multi-resource inventories.

Some tests have also been made with satellite imagery in the national forest surveys of the Scandinavian countries (Kuusela and Poso 1975, Green et al. 1981). These have not, however, led to practical applications. The main reason is the relatively poor resolution power of the available imagery in relation to the plot and stand size in the northern countries. The idea of multi-resource inventories has, however, been adopted in this geographical area. Country-wide separate studies have recently been carried out integrated in NFS (Eriksson et al. 1979, Rühling and Skärby 1979, Mattila 1981).

15. The setting in Finnish Lapland

Industrial timber harvesting was occasional and extensive in Finnish Lapland as late as the 1950's. Cuttings then increased rapidly both in the private and in the state-owned forests and a situation arose where over large areas the drain continuously exceeded the growth. As a consequence the volume of the growing stock in Lapland decreased some 10 % in the course of 15—20 years. Regeneration of the forests has also altered the age structure to a noticeable degree. The following figures are based on the National Forest Inventories of 1951—53 (Ilvessalo 1957), 1974—76 and 1978 (Kuusela and Salminen 1978, Mattila and Kujala 1980) in the two forestry board districts in question:

Characteristic	*Koillis-Suomi* F.b.d.		*Lappi* F.b.d.		Lapland	
	I[1])	II[2])	I[1])	II[2])	I[1])	II[2])
Age of stand:	Per cent of forest land					
0—40 years	10,2	26,2	10,8	17,2	10,6	20,2
161+ „	35,4	24,4	31,0	26,0	32,4	25,5
Total volume:						
mill. cu.m.	109	87	201	194	301	281
change, %	—20		—4		—10	

[1]) I = 1951—53 II = 1974—76, 1978

The figures are not entirely consistent, but this does not confuse the comparisons. The perceptible trends in Lapland are typical of areas where the initial forest consists primarily of aged stands. A rapid regereneration of the forests first decreases both the total volume and the growth. However, regeneration of the old forests increases the relative growth rate. In a regulated forestry, the total volume sooner or later turns to an increase, even though the cuttings are not reduced. It is important to realize that this kind of development can take place only on a sustained-yield basis. As to the conception of sustained yield, a policy of undiminishing growing stock can be regarded only as a special case of it (cf. Parry et al. 1983).

The Forestry Board District of *Koillis-Suomi* was established in 1954 by dividing the district of *Lappi* into two parts. The greater decrease of the volume in *Koillis-Suomi* can be explained by the greater initial stocking of that sub-area (see Ilvessalo 1957, p. 44). It can be concluded that the utilization of the forests in the beginning of the 1950's was not so far advanced in *Koillis-Suomi* than in the rest of the study area. Despite this, the latest inventory results indicate that the lowest volume may have already been passed in *Koillis-Suomi*. An increase in volume was already noticed in *Lappi* in the previous inventory as can be seen from the following estimated total volumes:

Inventory	Sub-area *Koillis-Suomi*	*Lappi*	Lapland
	mill. cu. m.		
1951—53	109	201	310
1962—63	—	—	260
1969	88	181	269
1974—78	87	194	281
1982—84	88	—	—

The estimates for the years 1962—63 and 1969 have been taken from the papers of Tiihonen (1966) and Kuusela and Salovaara (1971). As before, the figures are not quite consistent. Owing to the differences in the methodology, the volumes of the three first inventories should be increased by some 2—3 %. However, it is evident that the minimum volume of the forests of Finnish Lapland occurred in the late 1950's or in the early 1960's.

Forest industry expanded its activities in North Finland after the Second World War to be able to utilize the whole potential timber production capacity. Forestry soon made an important basic livelihood especially in the remote rural districts. The expanding timber production was both a reason for and a consequence of the increasing population of the peripheries. The cuttings were naturally first directed to the most accessible and valuable forests and to sites easy to regenerate. An other typical feature was a heavy concentration of the cuttings in the state-owned forests to reduce the logging costs.

The estimates of the allowable cut for Lapland made in the 1950's and 1960's concerned the entire forest area irrespective of the actual cutting practice. The allowable cut was determined on the assumption that the cuttings would sooner or later be extended to the economically less favourable forests. The virgin uncut forests formed a reserve supporting the heavy cuttings elsewhere and were thus an essential part of sustained-yield management.

The establishment of nature conservation areas in the 1960's and especially in the 1970's has a retroactive effect on the character of the earlier cutting policy. A considerable area of reserve forests has been protected against cuttings, at least for some time. As a consequence, the allowable cut on the remaining area must be decreased to ensure an even flow of timber in the future. On the other hand, the forest industries, once established, need a certain minimum guantity of raw material to maintain operation, and employment.

The area of the forest land under protection in Lapland in 1981 was 0,88 mill. ha (Kuusela 1982, p. 6). This is 16 % of the total forest land area of Lapland. The protection can be regarded as permanent for an area which includes 0,32 mill. ha of forest land. The whole nature conservation reduction in the allowable drain is 1,15 mill. m^3 or 18 % of a total 6,5 mill. m^3. The corresponding reduction of timber assortments is 0,80 mill. m^3.

Timber must be imported to Lapland to avoid over cuttings. To lessen the timber production deficit in the future, silviculture and amelioration have been intensified in the remaining timber production areas. Import cause unemployment in forestry. On the other hand, heavy silvicultural measures and amelioration cause conflicts with some other forms of forest use. The original facts have become obscure to the general public and public opinion has turned against forestry and the forest industries. The phenomenon has attained absurd proportions, one of which is to argue against the available results of forest research.

The effects of air pollution are the latest threat to the forests of Lapland. It is suspected that the growth is first increased due to the higher nitrogen content of the rain water. There are fears, however, that the trees will later be damaged by the sulphur compounds. According to the latest (unpublished) inventory results from North Finland, growth has already increased more than can be predicted on the basis of the achieved silviculture and amelioration. For example in *Koillis-Suomi,* the relative growth rate increased during 1976—1982 by some 30 %. One consequence of this phenomenon is that the volume of the growing stock has not changed in the way presupposed by the actual drain and the earlier growth estimate.

The setting described above is a challenge to the present forest research in Lapland. The main question emerging from the turmoil may be considered to be: what is the sustained-yield allowable cut in Lapland and what is the extent of the reduction due to the nature conservation areas? Another important question concerns the forest balance at different area levels and by ownership categories. The structure of the growth and drain deserves special concern in this context. Further information is required on the actual effects of climate, air pollutants, silviculture and amelioration on growth. The interrelations of the different forest use forms are appearing as a new and important group of study objects. For example, the effects of forestry on reindeer ranges and berry crops form a concret problem.

Forest balance is calculated as the difference between the allowable cut and the actual total drain. A task of the National Forest Survey is to yield reliable estimates of the volume and growth of the growing stock by forestry board district for the derivation of the allowable cut. The sample size is increased in the new inventory in Lapland to permit estimation by sub-areas and ownership categories. For example, it is now possible to calculate independent results for the forests in the nature conservation areas. The inventory method, a combination of photo-interpretation and field measurements, permits a subsequent increase of the photo sample in sub-areas, if necessary. The structure of the growth and the drain is studied on permanent field plots. The field sample is systematic thereby making a frame for separate studies. Special investigations concerning, for example, the reindeer ranges and occurrence of cloudberry are integrated into the new National Forest Inventory in Lapland.

The investigation in hand is a statistical analysis of the large-area forest inventory method being used in Lapland. Little or no attention will be paid to the inventory results in itself. The analysis concerns the characteristics which are important from the standpoint of timber production. The sampling design used is described in the following chapter. For the sake of comparison, the details of the design used in North-Sweden are also given. Then the study proceeds through an analysis of the field sample. The fourth chapter deals with the combination of the field sample and the photo sample. The usability of the samples for different purposes are discussed in the fifth chapter. And finally, some general conclusions concerning the inventory method as a whole complete the investigation.

16. Study objects

The new inventory method permits us to calculate unbiased results in several ways. The two main categories are the estimation on the field sample basis and the estimation from the combined photo and field sample. The bare field sample further affords a few extra alternatives. So it is possible to calculate a series of results which, as such, indicate the degree of accuracy of the inventory.

The purpose of this investigation is to examine the usability of a systematic field sample in connection with two-phase sampling from aerial photos and from the field. To begin with, estimations are made on the basis of the bare field samples. This establishes the foundation for further scrutinizing. The following step is to combine the photo and field samples and to observe the effect on the statistical precision. The third main study object is to determine the optimum use of the samples at different area levels.

The accuracy of the photo-interpretation must be studied before the field data can be extended to the photo sample and the statistical precision estimated. The precision of the main results is estimated in the first place by communes and by main reference areas. An interesting question is whether the precision goal set before the inventory has been attained.

The other part of the precision analysis concerns results for areas smaller than a commune. An important task is to assess the precision of the results when data extended from outside the target area is used in the calculations.

The fundamental reason for the investigation is to examine the rationality of the sampling procedure at different levels. If a better solution is found it must be taken into account in the future inventories.

2. INVENTORY DESIGN

21. Goal setting

The Seventh National Forest Inventory was started in Lapland in 1982. *Kuusamo* commune (Figure 3) was chosen to be the test area for the remodelled inventory method. The resulting method is to be used unchanged in the whole county of *Lappi* except for the northernmost communes of *Utsjoki, Inari* and *Enontekiö*. The primary goals set and the decisions made before the test inventory concerned the reference areas, precision requirements and type of field sample.

The importance of the commune as a reference area has been increased. It was decided that the main results must be usable at the commune level. For this purpose, the study area was divided into 17 sub-areas which usually comprise one or in some case a couple of communes. The new inventory is executed and the results are calculated by sub-areas.

The optimum inventory design has earlier been considered at the forestry board district level (Poso and Kujala 1977). Little or no attention has been paid to the results by commune. That is why the results are not adequate as such in the new situation.

The main reference areas remain forestry board districts or a part of them. The Forestry Board Districts in Lapland are *Lappi* and *Koillis-Suomi*. *Lappi* is further divided into two main reference areas in the National Forest Inventory. This partition is needed because the forest characteristics of the southern part are very different from those of the northern part.

The statistical precision requirement concerns the estimated total volume of the growing stock. It was decided that the relative standard error must not exceed 4 % in the three main reference areas defined above.

The two goals — usable results at the commune level and the precision requirement in the main reference areas — are naturally closely related. Usability of the results implies accuracy which is a function of the statistical precision and occurrence of bias. Bias being eliminated, the estimated total volume and its variance in a certain area is the sum of the corresponding volumes and variances in the sub-areas. So with increasing precision at the commune level the results also become more precise at the forestry board district level.

There are two main approaches in determining the sample size (Husch 1971, p. 44). In the ideal situation, with sufficient prior information, the optimum sampling design can be mathematically outlined a priori. This was the course of action in planning the new national forest inventory in Sweden. The other way is to obtain a representative sample within the available time and budget without predetermined precision goal. In fact, the approach used in this investigation is a mixture of both these principles.

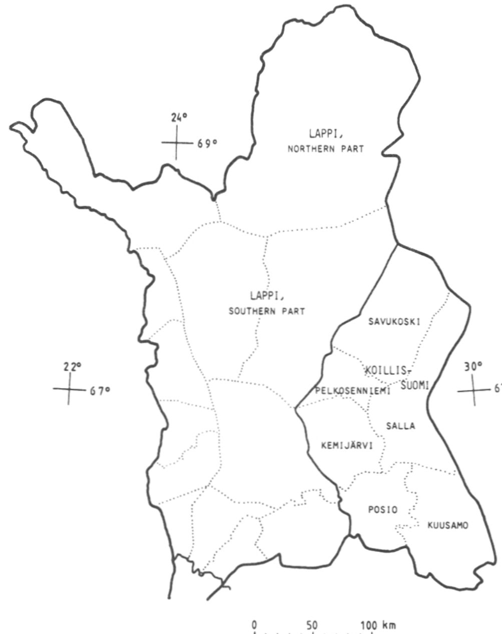


Figure 3. Forestry board districts of Lappi and Koillis-Suomi and the division into sub-areas.
Kuva 3. Lapin ja Koillis-Suomen piirimetsälautakunnat sekä niiden jako osa-alueisiin.

The course of approach in planning the new sampling design is different compared with the earlier inventory. The only criteria in planning the sample then was precision at the forestry board district level. So the samples could be concentrated to a high degree, because the results at the commune level were only of minor importance. In addition, the boundaries of the grouping areas did not follow the boundaries of the communes which leads to slightly biased results by communes.

The usability of results in sub-areas can be readily increased in two ways. First, by making the samples areally more representative, and secondly, by combining the information from photos and from the field by sub-areas. In the former case, there two further possibilities exist. The most implicit solution is to lessen the concentration of the samples. In this connection, but hardly as a single measure, it is necessary to consider an increase in the number of the sample plots per unit area.

Statistically the most important decision concerning the new inventory is associated with the selection of the field sample. The whole field sample is now systematic. This means that the information from airphotos does not affect the selection of the field plots. So the two sampling phases are independent of each other, except that the field sample is a part of the photo sample.

The systematic field sample has certain favourable practical consequences. Firstly, unbiased pilot results can be calculated on the basis of the bare field sample. Secondly, photo interpretation can be made before and/or after the field work. Further, the results can be made more precise in small areas afterwards by interpreting more photo plots.

A field sample chosen systematically results in a poorer statistical precision than when chosen according to the grouping method. This is due to the fact that the former sample represents the variation within the photo sample less well than the latter. The difference is the greater the smaller is the field sample. It is therefore desirable to increase the field sample to ensure representativeness.

Despite the systematic field sample, an extension of the field information to the whole photo sample is to be made. This will enable the flexibility of calculations associated with the grouping method. Following this, the systematic field sample leads to groups of varying size.

The systematic part of the field sample of the earlier inventory is remeasured in connection with the new inventory. The aim is to secure a better understanding of the growth and drain in the time between the two inventories. At the same time, a part of the new field sample is permanently established for remeasurement in the future.

The remeasured old field plots are not used as a part of the new two-phase sample. This is due to problems in photointerpretation. The old field plots should be located manually on the new aerial photos. This would be a time consuming task and would possibly lead to a systematically different, inconsistent interpretation. It is very important that all photo plots are interpreted consistently independent of whether they are field plots or not.

The remeasurement produces a systematic field sample with up-to-date field data. These plots can be utilized together with the new field plots to calculate results on the basis of the bare field sample. In this way we obtain usable, unbiased inventory results in the main reference areas without the photo sample.

Multiple-use of forests has become an object for scientific investigation in Finland as elsewhere (e.g. Saastamoinen 1982). In the conditions of Lapland, it is desirable to integrate multiple use aspects into the National Forest Inventory. A survey of the winter ranges of reindeer was carried out in this way in the 1970's (Mattila 1981). It was decided to make the survey an integrated part of the forest inventory. Remeasurement of old field plots makes possible an efficient analysis of changes. In general, a systematic field sample promotes quantitative studies into natural resources of this kind.

22. Sampling procedure

221. *Photo sample*

The whole inventory area is divided into squares of 8 km x 8 km. The location of the squares is defined in a plane coordinate system. A part of the squares is systematically selected to the first-phase sample as photo tracts. The sampling unit is in all respects the same as in the earlier inventory. The placing and the number of the photo tracts, however, are different in the new inventory.

Each photo tract includes 40 photo plots of two points (Figure 4). The distance between the points within a plot is 40 meters in the field on a south-north axis. The shortest distances between the plots within a tract in the eight main quarters are 300, 1 410 and 2 000 meters.

The net of the photo tracts include both continuous strips in the south-north direction as in the earlier inventory and separate tracts between the strips (Figure 5). The new strips are located between the old ones. The separate photo tracts are on the old field tracts (every third tract of the old strips). In this way 2/3 of the original squares are

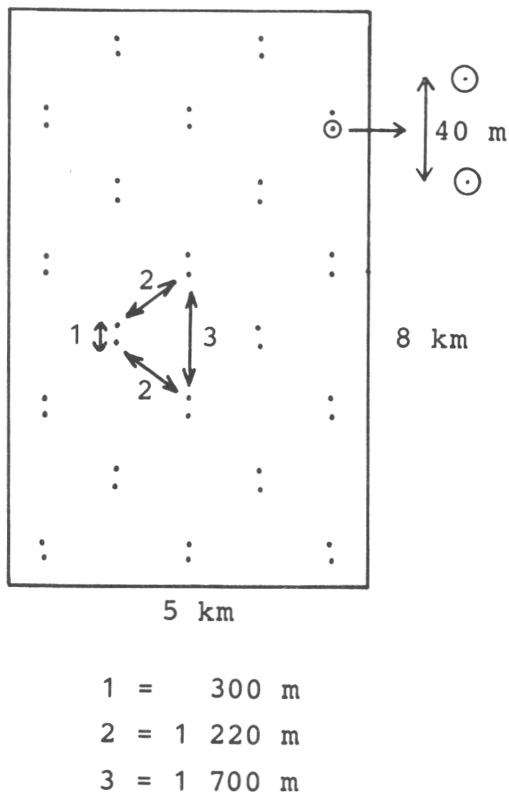


Figure 4. Structure of the photo tract. One tract includes 40 plots to be interpreted.

Kuva 4. Ilmakuvalohkon rakenne. Lohko sisältää 40 tulkintakoealaa.

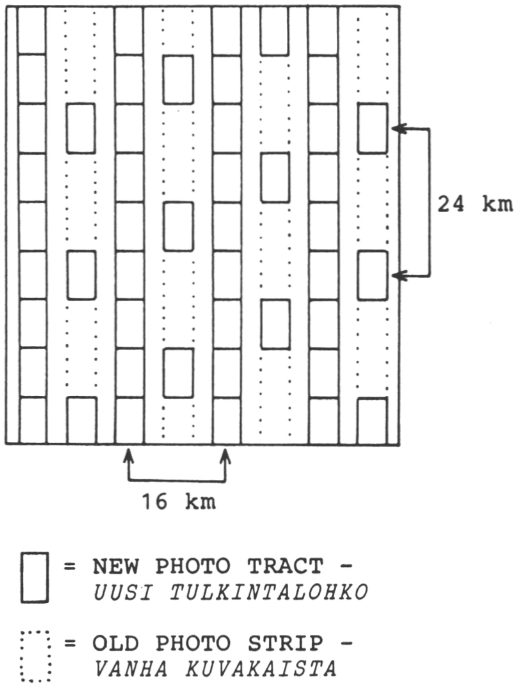


Figure 5. Structure of the photo sample.
Kuva 5. Ilmakuvanäytteen rakenne.

included in the new photo sample.

The reason for the transfer of the strips is that the whole inventory area is to be sampled in the course of successive inventories. The time series of the results will be more realistic because the whole variation of the population is present in them. On the other hand, the causal relationships will be more obscure.

The single tracts form a link to the old photo sample. The correlation between the successive inventory results will be increased to some extent. The airphoto material from the two points of time makes effective comparative studies possible. The main function of the single tracts, however, is to increase the areal representativeness of the photo sample.

The layout of the photo sample results in four tracts and 160 plots in an area of 24 km x 16 km. So the average area per photo plot is 2,4 sq.km. The corresponding area was 3,2 sq.km. in the earlier inventory. The increase in the number of the photo plots per unit area is 33 %.

Conventional small-scale black-and-white aerial photographs are used in the interpretation of the first-phase sample plots. Satellite imagery has not been seriously considered in this context because of the relatively poor spatial resolution. A usual size of pixels is 79 x 79 m. In the inventory, however, the field plots must be located with an accuracy of a few meters. Kuusela and Poso (1975) studied ERTS-1 material as a complement for conventional airphotos in the first phase of sampling in Lapland. They found that this material is useful primarily in a coarse stratification. Only the land use class distribution was estimated moderately well in a study in Sweden (Green et al. 1981). A large-area forest inventory of an other nature than the national forest inventory, however, can successfully employ satellite imagery (Kuusela and Poso 1970, Saukkola 1982).

The scale and type of aerial photographs affects the costs and the quality of the interpreted data. Efficiency being the criteria, small-scale black and white material is justified in the first-phase sampling of a large-area forest inventory. Poso et al. (1968) support this conclusion.

Availability is also a weighty criteria when choosing the photo material. Special flights to acquire photographs for the National Forest Inventory have not been seriously considered. The aerial photographs must cover the whole of Lapland and they must not be older than 5—10 years. The material should also be as homogeneous as possible.

Practically, only one photo material fulfils the qualifications. It comes from the Defence Forces taken from an altitude of about 9 300 meters at a scale of 1 : 60 000. This high altitude imagery is enlarged to the scale of 1 : 50 000 for use in the National Forest Inventory in Lapland.

Stereoscopy is utilized in the photo-interpretation. The dot grid has been photographed into the other photo of each stereo pair. The only auxiliary means used is the simple lens stereoscope which enlarges the view 2,8 times. So the scale of the final view after the two enlargements is about 1 : 18 000. A greater magnification is not reasonable considering the resolution power of the photo material.

The variables interpreted from the photos and the corresponding classifications are as follows (Ilmakuvatulkinnan . . . 1982):

Exposition:
— plane terrain
— inclined terrain (eight quarters)
— hilltop sites

Land use class:
— forest land
— scrub land
— waste land
— other land of forestry
— field
— built-up area
— roads etc.
— sea, lakes and rivers

Sub-class:
— mineral soil
— peatland

Drainage:
— undrained
— drained
— old drainage area (peatland) and ploughed or harrowed area (mineral soil)

Treatment class:
— open area and seedling stand
— stand consisting of seed trees or other standards
— young closed or nearly closed stand
— old closed stand
— shelterwood stand

Dominant tree species:
— Scots pine
— Norway spruce
— deciduous trees

Volume:
— in classes of 10 m³/ha

Crown coverage:
— in classes of 5—10 %

Structure of the growing stock:
— homogeneous
— slightly irregular
— irregular

Distance to the nearest stand boundary:
— less than 40 m
— 40—100 m
— more than 100 m

Direction from the nearest stand boundary

The four last mentioned variables are interpreted only if the two points within a plot are in the same stand. These variables are to facilitate studies concerning the practicability of satellite imagery. Drainage and dominant tree species are interpreted only in the main stand when the points are located in separate stands. All other variables, except volume, concern the whole stand. Volume is interpreted in a relascope plot, the centre of which coincides with the sample point.

Clarity and ease of interpretation were sought when choosing the variables and forming the classifications. The task was not very difficult due to the earlier experience with the same photo material. The principle of the classifications is that the classes must be readily distinguishable on aerial photographs.

222. Field sample

The field sample is concentrated to every other photo tract in the continuous photo strips. The number of field tracts is 50 % greater than in the earlier inventory (Figure 6). The ratio between the field and the photo tracts is 9:24, i.e. nine tracts of 24 (27,5 %) are field tracts. The corresponding figures earlier were 8:24 or 33 %.

The number of the field plots within a field tract is 14. Six of them are established as semi-permanent. The location of the semi-permanent field

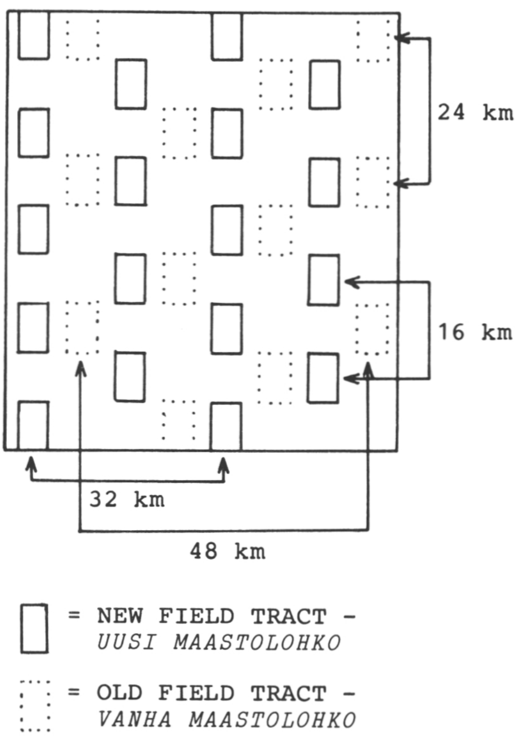


Figure 6. Location of the field tracts in the two successive inventories.
Kuva 6. Maastolohkojen sijainti kahdessa peräkkäisessä inventoinnissa.

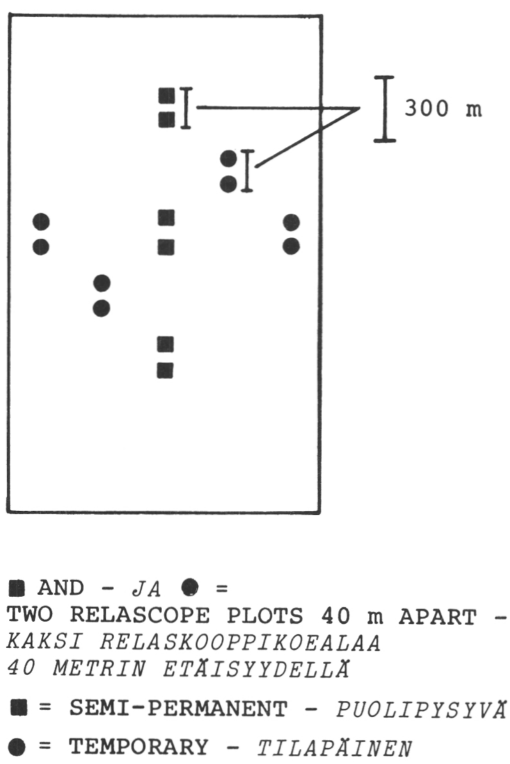


Figure 7. Location of the temporary and the semi-permanent field plots within the field tracts.
Kuva 7. Pysyväisluonteisten ja vain kerran mitattavien maastokoealojen sijainti maastolohkolla.

2 408500946V

plots within a tract is the same as in the earlier inventory (Figure 7). The ratio between the field and the photo plots is 42:320 or 13 %. This is about the same as in the earlier inventory.

The field plots within a field tract are located in a quadratic area of 3,04 km x 3,04 km (Figure 8). The straight distances between the plots range from 300 to 4 300 m. The corresponding figures are 200 m and 2 828 m within the field tract used in South Finland (Salminen 1981, p. 8). The plot used there, however, consisted of only one relascope plot which increased the variation between the plots. The number of the plots for volume estimation per tract was 21 in South Finland.

A wider spread of field plots in the north is desirable because the spatial correlation of forest characteristics increases from the south to the north. Considerable attention was paid in Sweden to the correlation functions of the main forest characteristics in planning the new national forest inventory there (see Ranneby 1981). It was concluded that 200 m — 600 m is the appropriate minimum distance between the plots for volume estimation (Hägglund 1983, p. 12). The following details have been taken from the paper of Hägglund (ibid.).

The size of the tracts and the number of the plots per tract are variable in Sweden. The number of the plots for volume estimation is 12 in the temporary tracts in the main part of the country. The size of the tract is 1,5 km x 1,5 km in that part of Sweden which is probably most similar to the finnish study area. The distance between the plots is 480 m — 520 m.

A question arises whether the new field tract in Lapland is more informative than the tract of Sweden described above. An ocular comparison of the tracts (Figure 9) suggests an affirmative answer. However, one cannot be sure without a statistical analysis (see Ranneby 1979). The tract in Lapland is wider and it includes more plots but the plots are clustered to a considerable degree. The nature of the plots is also different which further confuses the comparison. In Sweden, fixed-area circular plots with a radius of 7,07 m are used for the volume estimation in the temporary field tracts.

The field tract in Lapland is so constructed that its measurement is possible in two days. In Sweden, the tract comprises one day's work, except in the extreme south of the country. So in principle, to do well in the efficiency comparison the tract of Lapland ought to be twice so informative as the tract of North Sweden. This is probably not the case. One must note, however, that a part of the difference may arise from the more advanced data gathering technique in Sweden.

Permanent field plots are measured both in Lapland and North-Sweden. All field tracts include plots of a permanent nature in Lapland, while in Sweden special permanent tracts are established. It cannot be stated with any certainty which of the two tracts is spatially more representative in this case (Figure 10). The time spent in the measurement of one tract is one day both in Lapland and in North Sweden.

The time interval between the successive inventories is about seven years in Lapland. Sampling is continuous in Sweden in that a sparse sample is measured every year. The results are calculated there on the basis of five successive yearly samples. Quantitative data concerning the sampling are given

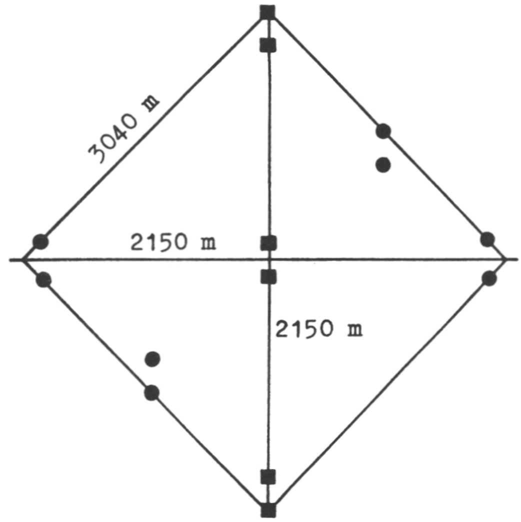


Figure 8. Distances between the field plots within the field tracts.
Kuva 8. Maastokoealojen väliset etäisyydet maastolohkolla.

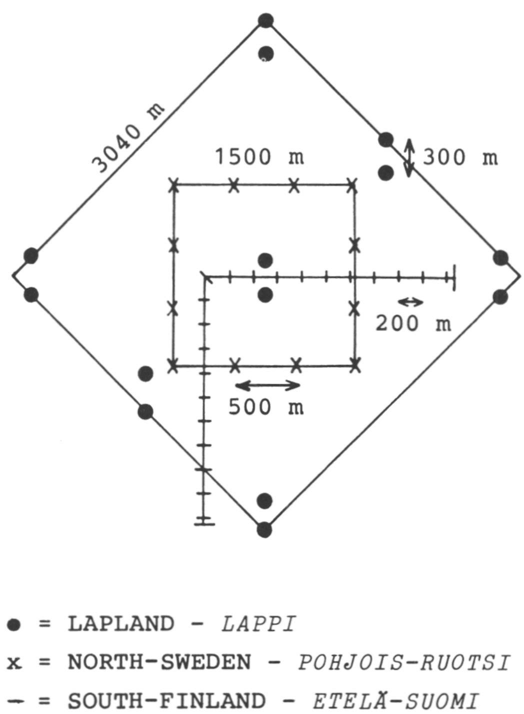


Figure 9. Structure and size of the field tracts used in Lapland, North-Sweden and South-Finland.
Kuva 9. Maastolohkon koko ja rakenne Lapissa, Pohjois-Ruotsissa ja Etelä-Suomessa.

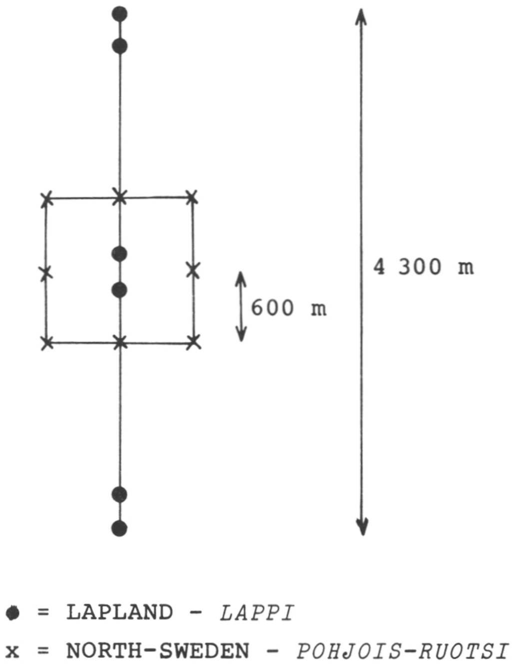


Figure 10. Location of the permanent field plots within the field tracts in Lapland and in North-Sweden.
Kuva 10. Pysyvien maastokoealojen sijainti maastolohkolla Lapissa ja Pohjois-Ruotsissa.

in Table 1. The reference acreage used is 4 608 sq. km which corresponds to an average county in Lapland. The figures concerning North Sweden have been derived from the report of Hägglund (1983, p. 14, 15, 17 and 70).

The numbers of the tracts has been counted in a framework of a 64 km x 72 km rectangular area. In this way, the minimum possible number of tracts per fixed acreage is obtained. In practice, areas are irregular in form and the actual number of tracts is significantly greater. So for example, 74 photo tracts wholly or partly fall in *Kuusamo* instead of about 50 (whole) tracts estimated on the basis of the acreage. It is readily seen from Table 1 that the total field sample is now much more representative than in the earlier inventory. This is particularly due to the greater number of field tracts. Field work and the corresponding cost on the other hand has doubled. Fifty percent of the additional work comes from the remeasurement of the semi-permanent field plots.

The bare field sample of Lapland does not of course, stand comparison with the samples of South Finland and North Sweden. The comparison becomes more even after the photo sample has been included. The combined field and photo sample of Lapland includes less tracts but more plots per unit area than the field samples of South Finland and North Sweden. The photo tract covers a greater area and includes more plots than the different field tracts. However, the photo plots which have not been measured in the field are inferior to the real field plots. The magnitude of the additional error depends on the correlation between the interpreted data and the reality of the field.

Table 1. Number of different tracts and plots and field working days in an area of 4 608 sq.km. (64 km × 72 km).
Taulukko 1. Erilaisten lohkojen ja koealojen määrät sekä työmenekki 4 608 km²:n (64 km × 72 km) alueella.

	Lapland *Lappi* 1982+	1976	South Finland *Etelä-Suomi*	North Sweden[1]) *Pohjois-Ruotsi*
Photo sample *Ilmakuvanäyte*				
Tracts — *Lohkoja*	48	36		
Plots — *Koealoja*	1 920	1 440		
Field sample *Maastonäyte*				
Temporary — *Tilapäinen*				
Tracts — *Lohkoja*	18	12	72	49
Plots — *Koealoja*	252	204[2])	1 512	584
Workdays — *Työpäiviä*	36	24	72	49
Permanent — *Pysyvä*				
Tracts — *Lohkoja*	12			48[3])
Plots — *Koealoja*	72			377
Workdays — *Työpäiviä*	12			48
Total — *Yhteensä*				
Tracts — *Lohkoja*	30	12	72	97
Plots — *Koealoja*	324	204	1 512	961
Workdays — *Työpäiviä*	48	24	72	97

1) In a period of seven years — *Seitsemän vuoden aikana*
2) Including 72 plots established semi-permanent — *Sisältää 72 puolipysyvää koealaa*
3) 24 tracts which are measured twice — *24 lohkoa, jotka mitataan kahdesti jakson aikana*

23. Study material

The new inventory design was tested in *Kuusamo* in 1982 (Mattila 1983). The sampling procedure was found practical and so it was applied as such in the rest of *Koillis-Suomi* in 1983. The tract net used there is illustrated in Figure 11. A quantitative description of the samples is given in Table 2.

Each plot represents a fixed area because the samples are systematic. Thus the sample sizes can be estimated a priori if the acreage is known. It appears that the number of the photo plots can be pre-estimated fairly well by sub-area. The estimation becomes less accurate with decreasing acreage and/ or with increasing concentration of the plots. The form of the area in relation to the sampling scheme also affects the accuracy.

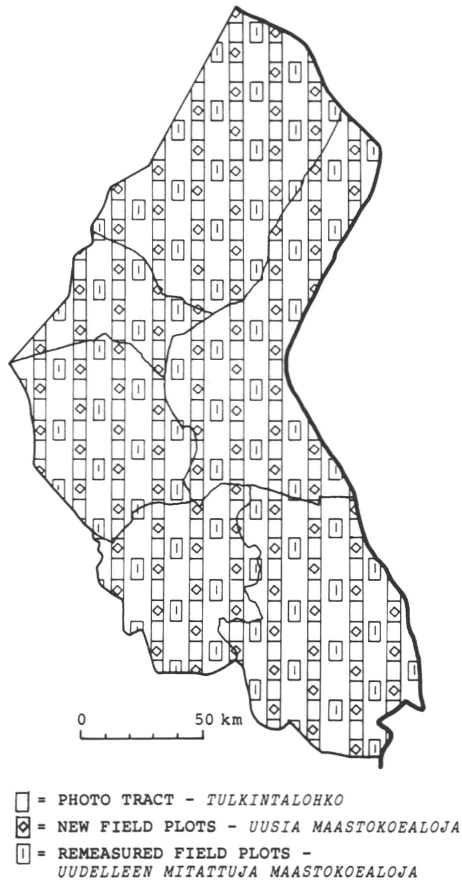


Figure 11. Location of the different samples measured in forestry board district of *Koillis-Suomi* during 1981—83.
Kuva 11. Erilaisten vuosina 1981—83 mitattujen näytteiden sijainti Koillis-Suomen piirimetsälautakunnan alueella.

Table 2. The number of the tracts and the actual (A) as well as the estimated (E) numbers of the plots in *Koillis-Suomi* by sub-area. The numbers refer to the total area including waters.
Taulukko 2. Lohkomäärät ja todelliset (A) sekä pinta-alan perusteella arvioidut (E) koealamäärät Koillis-Suomessa osa-alueittain. Pinta-aloihin ja lukumääriin sisältyvät myös vedet.

Variable *Muuttuja*		Kuusamo	Posio	Sub-area — *Osa-alue* Salla	Kemijärvi	Pelkosen-niemi	Savukoski	Total *Yhteensä*
Acreage *Pinta-ala*	km²	5 806	3 531	5 879	3 943	1 862	6 372	27 493
Photo sample *Ilmakuvanäyte*								
Tracts — *Lohkoja*		75	48	74	51	28	79	315
Plots-	A	2 412	1 464	2 427	1 697	830	2 750	11 580
Koealoja	E	2 419	1 471	2 449	1 643	776	2 697	11 455
Field sample *Maastonäyte*								
Remeasured *Uud. mitattu*								
Tracts — *Lohkoja*		16	11	15	12	6	19	76
Plots-	A	94	57	90	65	28	112	446
Koealoja	E	91	55	92	62	29	101	430
New — *Uusi*								
Tracts — *Lohkoja*		24	17	25	18	10	30	115
Plots-	A	314	211	317	212	119	352	1 525
Koealoja	E	317	193	322	216	102	354	1 504

3. PRELIMINARY ESTIMATION ON THE BASIS OF THE FIELD SAMPLE

31. Characteristics and formulae

Inventory results can be divided into area information and area related tree information. The former includes proportions of forest strata, the latter characteristics of the growing stock by stratum. The area of forest land and total volume of the growing stock are generally the most important characteristics to be estimated in large scale forest inventories. The statistical precision is estimated for these results at all events. Other characteristics of main interest concern the structure of the growing stock and certain site conditions. For the purpose of this investigation, the following choice of characteristics has been made.

Area of:
(1) Mineral soils
(2) Forested land[1])
(3) Forest land
(4) Mature stands[2])

Mean volume (m^3/ha) of:
(5) Growing stock on the total land area
(6) Growing stock on forest land
(7) Scots pine on the total land area
(8) Saw timber on the total land area

1) Forest land and scrub land together
2) Mature stands can occur only on forest land

Economic, ecological and demonstrative aspects have been considered when compiling the above list.

Total volume of the growing stock V is a product of the total land area A and the corresponding mean volume $\bar{x}$:

$$(1) \quad V = A * \bar{x} \; (m^3)$$

The land area is assumed to be known. The mean volume must be estimated from the sample as

$$(2) \quad \bar{x} = \frac{\sum_{i=1}^{m} x_i}{m}$$

where m = number of plots on land
x_i = volume measured in plot i

One can proceed in the calculations in two steps. The tract being the sampling unit, it is logical to estimate the tract values first. So for example, the mean volume on the total land area can be written

$$(3) \quad \bar{x} = \frac{\sum_{j=1}^{k} m_j \bar{x}_j}{m}$$

where k = number of tracts on the total land area
m_j = number of plots on land in tract j
$\bar{x}_j$ = mean volume on land in tract j

Further progress towards the tractwise approximation gives

$$(4) \quad \bar{x} = \frac{\sum_{j=1}^{k} \left(\frac{m_j}{\bar{m}} * \bar{x}_j\right)}{k}$$

where $\bar{m}$ = average number of plots on land per tract

In addition to the mean volume on the total land area and its constituents, whichever stratum proportion can be estimated according to the principle of formula 4. For that purpose, tract values must first be estimated from the distribution of the sample points for the strata in question. It is worth noting that the resulting values are on the land area basis.

Estimation of the proportion of a sub-stratum (p'_h) within a stratum requires one further operation:

$$(5) \quad p'_h = \frac{1}{p_{h1}} * p_{h2}$$

where p_{h1} = proportion of stratum h_1 on land
p_{h2} = proportion of sub-stratum h_2 on land.

For example, proportion of mature stands on forest land can be calculated in this way when the proportion of forest land on land (p_{h1}) and the proportion of mature stands on land (p_{h2}) are known.

The formula for the estimation of the mean volume within a stratum is analogous to formula 4:

$$(6) \qquad \bar{x}_h = \frac{\sum_{j=1}^{k_h} \left(\frac{m_{hj}}{\bar{m}_h} * \bar{x}_{hj}\right)}{k_h}$$

where k_h = number of tracts where at least one relascope plot falls in stratum h
$\bar{m}_h$ = average number of plots per tract in stratum h
m_{hj} = number of plots in stratum h in tract j
$\bar{x}_{hj}$ = mean volume in stratum h in tract j

All estimated proportions and mean volumes calculated by the principle presented in formulae 4 and 6 are ratio estimates. They are biased because the size of the tracts is variable (cf. Cochran 1977, p. 31). The magnitude of the bias depends on the size of the sample (number of tracts) and the variation of the tract size. According to Cochran (ibid., p. 162), the magnitude of the bias in ratio estimate $\hat{R}$ follows a rule

$$\frac{|\text{bias in } \hat{R}|}{\sigma_{\hat{R}}} \leq \text{CV of } \bar{s}$$

where $\bar{s}$ refers to the mean size of the tracts. CV (coefficient of variation) of $\bar{s}$ can be easily estimated from the sample to get the upper limit for the bias. It is a general working rule that bias is negligible if CV of $\bar{s}$ $\leqq$ 0,1.

The actual bias of a ratio estimate depends on the correlation between the tract size and the quantity (p or $\bar{x}$) to be estimated. the maximum possible bias implies a correlation coefficient of 1 or —1. It is not reasonable to assume that any mean volume or proportion is directly correlated with the tract size. So in this case, the actual bias is probably negligible.

The estimation of characteristics in the National Forest Inventory in Lapland is based on multi-stage sampling. The concept of multi-stage sampling must not be confused with that of multi-phase sampling. In the present case, the tract serves as the primary sampling unit. Each tract consists of 64 secondary sampling units of 1 km^2. Finally, the relascope plot can be regarded as the tertiary sampling unit. Concerning the new field sample, the sampling fraction of the first stage (f_1) is 1/4 and that of the second stage (f_2) 7/64 (p. 17). As to the third sampling stage, it seems realistic to suppose that the four relascope plots within a sampled secondary unit represent a negligible part of it.

The unbiased estimation of the error variance is rather complicated in multi-stage sampling, even in the case of random sampling in each stage when f_1 is not negligible (see Cochran 1977, p. 287). Independent of the number of the stages, a conservative estimate can be calculated on the basis of the variation of the tract values. The variance estimator for mean volumes then reads (Cochran ibid. p. 279, 288):

$$v(\bar{x}) = \frac{\sum_{j=1}^{k} (\bar{x}_j - \bar{x})^2}{k(k-1)}$$

This formula is totally valid only in the case that the location of the tracts is random, that f_1 is negligible and that the tracts are of equal weight. The number of the plots on land per tract being variable, the variance of an estimated proportion p_h can be approximated as (cf. Loetsch and Haller 1973, p. 184)

$$v(p_h) = \frac{1}{\bar{m}^2 k} * \frac{\sum_{j=1}^{k} (m_{hj} - p_h * m_j)^2}{k-1}$$

From this we can derive

$$v(p_h) = \frac{1}{k(k-1)} * \sum_{j=1}^{k} \left(\frac{m_{hj}}{\bar{m}} - \frac{p_h * m_j}{\bar{m}}\right)^2$$
$$= \frac{1}{k(k-1)} * \sum_{j=1}^{k} \left(\left(\frac{m_j}{\bar{m}}\right)^2 * \left(\frac{m_{hj}}{m_j} - p_h\right)^2\right)$$

But because $m_{hj}/m_j = p_{hj}$ we can write

$$(7) \; v(p_h) = \frac{\sum_{j=1}^{k} \left(\left(\frac{m_j}{\bar{m}}\right)^2 * (p_{hj} - p_h)^2\right)}{k(k-1)}$$

In this form the estimator has a logical connection with formulae 4 and 6. With an appropriate application of it, one can approximate the variances of $\bar{x}$ (and the constituents of $\bar{x}$), p_h and $\bar{x}_h$.

Let as suppose a two-stage sampling design with relascope plot as the secondary sampling unit. This would actually be the case if the 28 relascope plots within a sampled tract were randomly located. The variance estimator for an estimated mean volume reads (Cochran 1977, p. 278)

$$(8) \quad v(\bar{x}) = \frac{1-f_1}{k} s_1^2 + \frac{f_1(1-f_2)}{km'} s_2^2$$

where m' = number of plots per tract

$$s_1^2 = \frac{\sum_{j=1}^{k} (\bar{x}_j - \bar{x})^2}{k-1}$$

and
$$s_2^2 = \frac{\sum_{j=1}^{k} \sum_{i=1}^{m'} (x_{ij} - \bar{x}_j)^2}{k(m'-1)}$$

Concerning proportions, the estimator for s^2_1 is analogous to that presented above (p_{hj} and p_h instead of $\bar{x}_j$ and $\bar{x}$) and the estimator for s^2_2 reads

$$s_2^2 = \frac{m'}{k(m'-1)} * \sum_{j=1}^{k} p_{hj}(1-p_{hj})$$

Let us denote the two terms in the right in formula 8 as $v(\bar{x})_B$ and $v(\bar{x})_W$. B refers to the variation between tracts and W refers to the variation within tracts. Consequently, the variance of an estimated mean consists of two components, here called *variance between tracts* and *variance within tracts:*

$$v(\bar{x}) = v(\bar{x})_B + v(\bar{x})_W$$

The above expressions of s_1^2 and s_2^2 presuppose a constant number of plots m' per tract. The number of the plots on land being variable, s_1^2/k is estimated according to the principle present in formula 7. s_2^2 can be regarded as an weighted mean of the population variances within the tracts. Using the number of the plots on land as the weight gives the following expressions: (f_2 is negligible):

$$(9) \quad v(\bar{x})_W = \frac{f_1}{(k\bar{m})^2} * \sum_{j=1}^{k} \{\frac{m_j}{m_j-1} * \sum_{i=1}^{m_j} (x_{ij} - \bar{x}_j)^2\}$$

$$(10) \quad v(p_h)_W = \frac{f_1}{(k\bar{m})^2} * \sum_{j=1}^{k} \{\frac{m_j^2}{m_j-1} p_{hj} (1-p_{hj})\}$$

If field plots were measured in all tracts ($f_1 = 1$) the first term on the right in formula 8 would be equal to zero. On the other hand, the value of the latter term would be zero if the sampled tracts were inventoried totally ($f_2 = 1$). The former case is equivalent to a kind of stratified sampling with the tracts as strata. One can conclude that a heavy concentration of the plots within the tracts probably leads to an underestimation of the population variances within the tracts. If this happens, the variance of an estimated mean or proportion will also be underestimated. The bias of the estimated sampling error increases with increasing value of f_1 although the concentration of the plots remains unchanged. In the sampling design used in Lapland, the four relascope plots within 300 m are, no doubt, correlated. So for reasons of caution, it is better to estimate the sampling errors by formula 7. The subject will be dealt with in other context later in this investigation.

An area estimate includes a sampling error which arises from the estimation of the corresponding proportion. So these two quantities are associated with the same relative standard error. The same is the case with respect to the characteristics' mean and total volume on land. Here, mean and total can refer to the growing stock as a whole, or only a part of it, for example Scots pine or saw timber etc.

Area and mean volume must both be estimated when deriving total volumes within strata. According to the principles of error propagation

$$(11) \quad (S_{Vh}\ \%)^2 = (S_{Ah}\ \%)^2 + (S_{\bar{x}h}\ \%)^2$$

where S = relative standard error in per cent.

If the precision of p'_h (formula 5) is to be estimated, one can proceed by the tract values p'_{hj} and apply the principle present in formula 7.

The systematic tract net is likely to lead to an overestimation of the sampling variance (Loetsch and Haller 1973, p. 165). Valid use of formula 7 presupposes that the tracts are located randomly or that the spatial distribution of the forest is random (Husch et al. 1982, p. 182) The latter is very unlikely in a large forest area. Systematic changes in the forest characteristics occur within, for example, the forestry board districts. A better estimate of the variance is obtained with techniques based on the differences within groups of tracts (e.g. Salminen 1973, Ranneby 1981 b).

The net of field tracts is quite sparse in

Lapland, so the number of the tracts in one commune is small. This is the main obstacle to using groups of tracts in the estimation of the precision. But, on the other hand, we can assume that the systematic change within communes is small. This, however, may not be true in the sub-areas of *Salla* and *Savukoski* which are very long in the southwest — northeast (Fig. 3. p. 14).

The results for a set of sub-areas can be obtained in two ways. Firstly, the samples of the sub-areas can be put together before the calculations. In the alternative solution, one sums the estimated quantities of the sub-areas. The two approximations lead to slightly different results because the actual area represented by one sample plot is not quite equal in the sub-areas.

The variance of a summed area or total volume is the sum of the corresponding variances in the sub-areas. When using the combined data in the calculations, a consistent variance estimate is obtained with the aid of the total tract net. The number of the tracts then becomes great enough to allow the use of tract groups in the estimation.

The principles of the estimation described above apply to the remeasured and to the new field sample separately, as well as to the samples combined. The two components of the total field sample give us two independent estimates. These can be weighted in different ways in the combined estimate. Sample size (number of plots) is the weight when the total field sample is used as a single data. An other and more general solution is to take the statistical precision of the sub-samples into account. The revised estimator of, for example, mean volume on land $\overline{x}^*$ can be written (Burk and Ek 1982):

(12)
$$\overline{x}^* = \frac{\sum_{i=1}^{l} (w_i \overline{x}_i)}{\sum_{i=1}^{l} w_i}$$

where l = number of the sub-samples
$\overline{x}_i$ = mean volume estimated from sub-sample i
$w_i = 1/v(\overline{x}_i)$
and $v(\overline{x}_i)$ = variance of the mean volume estimated from sub-sample i.

The variance of $\overline{x}^*$ may be estimated by formula (Meier 1953):

(13)
$$v(\overline{x}^*) = \frac{1}{w} + \frac{4}{w} * \sum_{i=1}^{l} \left\{\frac{w_i}{n_i w}\left(1 - \frac{w_i}{w}\right)\right\}$$

where $w = \sum_{i=1}^{l} w_i$
$n_i = k_i - 1$
and k_i = number of field tracts in sub-sample i.

According to Burk and Ek (ibid.) formula 13 gives good approximations if all n_i are > 10. Only in one case is this condition not fulfilled. In the sub-area of *Pelkosenniemi* remeasurement of permanent plots has only been made in six tracts (Table 2). The number of the new field tracts (10) is also rather small there.

32. Results and conclusions

The field sample measured in *Koillis-Suomi* comprised of 3 651 relascope plots on land in 196 tracts (Table 3). The remeasured field sample included 39 % of the tracts and 23 % of the plots. The semi-permanent relascope plots were remeasured during 1981—83. More than half of this sample was acquired in 1981. The rest were remeasured in 1982—83 simultaneously with the measurement of the new field sample.

The numbers of the tracts and the plots given in Table 3 indicate directly the occurrence of forest land and mature stands by sub-area and sub-sample. For example, the new field sample indicates that mature stands occupy 19,36 % of the total land area in *Koillis-Suomi* (546 points from 2 820). The corresponding percentage on forest land is 28,65 % (546 points from 1 906).

There can occur more than one sub-area within a tract. In that case one and the same tract is included in the number of the tracts in more than one sub-area. This is the reason why the number of the tracts in *Koillis-Suomi* is often smaller than the sum of the corresponding numbers in the sub-areas.

The data of the remeasured field sample cannot be put into exactly the same format as the data of the new field sample. A considerable number of the clusters of the old field sample were originally measured as single units without dividing them into two separate relascope plots, as is the case with the new field sample. Neither was the division made when the semi-permanent field

Table 3. Field sample in *Koillis-Suomi*. Number of tracts (T) and relascope plots (P) on land, on forest land and in mature stands by sub-area and sub-sample.
Taulukko 3. Koillis-Suomen maastonäyte. Lohkojen (T) ja relaskooppikoealojen (P) lukumäärä maalla, metsämaalla ja uudistuskypsissä metsissä näytteen eri osissa osa-alueittain.

Sub-area *Osa-alue*	Sample[1]) — *Näyte*[1]) 1 T	1 P	2 T	2 P	1 + 2 T	1 + 2 P
			On land — *Maalla*			
Kemijärvi	12	125	18	379	30	504
Kuusamo	16	169	24	523	40	692
Pelkosenniemi	6	49	10	229	16	278
Posio	9	90	17	368	26	458
Salla	15	176	26	622	41	798
Savukoski	19	222	30	699	49	921
Koillis-Suomi	77	831	119	2 820	196	3 651
			On forest land — *Metsämaalla*			
Kemijärvi	12	83	18	262	30	345
Kuusamo	16	119	24	350	40	469
Pelkosenniemi	5	30	10	147	15	177
Posio	9	59	17	257	26	316
Salla	15	119	25	415	40	534
Savukoski	19	145	28	475	47	620
Koillis-Suomi	76	555	116	1 906	192	2 461
			In mature stands[2]) — *Uudistuskyps. metsissä*[2])			
Kemijärvi	7	22	8	32	15	54
Kuusamo	9	31	21	87	30	118
Pelkosenniemi	2	2	9	31	11	33
Posio	3	6	14	73	17	79
Salla	8	34	22	148	30	182
Savukoski	14	63	25	175	39	238
Koillis-Suomi	43	158	95	546	139	704

[2]) Including low-yielding mature stands — *Sis. vajaatuottoiset uud. kypsät metsät.*
[1]) 1 = Remeasured field sample — *Uudelleen mitattu maastonäyte*
2 = New field sample — *Uusi maastonäyte*

sample was remeasured. The different format of the data does not of course affect the results of the inventory. The problem is in the estimation of the statistical precision of the results.

An estimator of the sampling error (formula 8) consists of two terms which express the variance between the tracts and the variance within the tracts separately. The variation of the tract values (s_1^2) can be estimated consistently from the remeasured and the new field samples despite the different format of the data. This is not the case with the variation within the tracts (s_2^2) . A consistent estimation requires either cluster or relascope plot as the measurement unit in both samples. The single relascope plot is a logical choice because its center point can fall in one and only one forest stratum.

The undivided clusters of the remeasured field sample can be artificially divided into parts by giving the same volume to the two relascope plots. This, however, would lead to a considerable underestimation of the variation of the volumes within the tracts. A relascope plot with factor two (one counted tree equals to two sq.m. in basal area) is so small a unit that considerable deviations of the measured volumes within clusters occur even when the two points are in the same stand.

From the new field sample it will be worth examining how much the variation within the tracts affects the estimated sampling error. The values of s_1^2/k are in general greater than those of s_2^2/km (k = number of tracts and m = number of relascope plots per tract). This being the case, the bipartitate estimator gives a smaller estimate of the sampling error than does s_1^2/k. In other words, incorporating the variation within the tracts into the calculations decreases the estimated sampling error. The following set of figures concerns the proportion of forest land (p) and the mean volume on land ($\overline{x}$) estimated from the new field sample. Expressed in per cent, the ratio between the smaller and the larger CV-values of the error estimates by sub-area and sub-sample is:

Sub-area	Area	Sample[1)] 1		2		1+2	
	km²	p	$\overline{x}$	p	$\overline{x}$	p	$\overline{x}$
		%					
Pelkosenniemi	1 826	89	100	89	96	88	94
Posio	3 096	96	89	92	91	93	89
Kemijärvi	3 578	92	90	96	93	92	91
Kuusamo	5 006	91	90	94	89	91	89
Salla	5 745	91	90	92	90	90	89
Savukoski	6 423	91	90	93	90	91	90

[1)] 1 = Semi-permanent
2 = Temporary

The smaller variances have been calculated as $v = (1\text{-}f_1)s^2_1 + f_1 s_2^2/k\overline{m}$ and the larger ones as $v = s_1^2/k$. The value of f_1 is 0,25 in the new field sample. The number of the tracts k is the same in the semi-permanent and temporary samples by sub-area because semi-permanent plots are established in every field tract. The semi-permanent component takes 43 % of the total field sample.

Incorporating the variation within the tracts into the estimation of the sampling error decreases the estimated relative standard error by some 10 % in the new field sample. The result seems to be independent of the size of the sub-area and the number of the relascope plots per tract. This indicates that the ratio between the values of s_1^2/k and $s_2^2/k\overline{m}$ is fairly stable and the final outcome depends mainly on the value of f_1. A small value of f_1 results in a small decrease of the estimated sampling error and vice versa. So the decrease in the remeasured field sample ($f_1 = 1/6$) should be smaller than in the new field sample. This can be experimentally studied only with the proportions. Concerning the remeasured field sample, the percentages corresponding to the first column in the set of figures above are 97, 99, 93, 101, 92 and 94. The result is in accordance with the conclusions.

The value of f_1 is 5/12 in the combined field sample (the new and the remeasured field samples together). Concerning the proportions, the ratio between the two estimates of the standard error in the combined field sample by sub-area and characteristic is:

Sub-area	Characteristic[1)] 1	2	3	4
	%			
Pelkosenniemi	80	82	80	91
Posio	94	95	90	80
Kemijärvi	80	86	85	81
Kuusamo	87	87	85	84
Salla	80	83	82	80
Savukoski	83	85	84	82

[1)] 1 = Proportion of mineral soils
2 = ,, forested area
3 = ,, forest land
4 = ,, mature stands

Using the bipartitate variance estimator decreases the estimated standard error by 10—20 % in the combined field sample. The percentages calculated earlier from the new field sample suggest that the reduction is about the same with the mean volumes.

The comparison of the two variance estimators indicates that the approaches lead to significantly different estimates with large values of f_1 (say larger than 0,20). We know that formula s_1^2/k is conservative. But, on the other hand, using the bipartitate variance estimator is questionable because the relascope plots are in clusters, which probably leads to an underestimation of the variation within the tracts (p. 23). In addition, the variation within the tracts cannot be completely estimated from the remeasured field sample. For these reasons, the sampling errors dealt with in the following discussion have been calculated on the basis of the variation of the tract values only.

The estimated relative errors are presented by characteristic and sub-area in Table 4. As can be expected, the remeasured field sample generally gives the poorest statistical precision. Both the size and the location of the remeasured sample contribute to this outcome. Some deviations from the general trend occur in the small sub-areas. One can conclude that the inconsistencies are accidents in the sense that the true population variation is not present in the samples.

Combining the remeasured and the new field samples increases the statistical precision, with a few exceptions. The inconsistencies in question are associated with a seemingly too low variation present in the new field sample. Adding the remeasured plots increases the variation so much that despite the increased size of the sample, the estimated statistical precision decreases. This, of course, is an accident caused by the small size of the field samples.

Table 4. Estimated value e and the corresponding relative standard error s_e % in the different samples by variable and sub-area in Koillis-Suomi.
Taulukko 4. Muuttujien arvot e ja vastaavat suhteelliset keskivirheet s_e % arvioituina eri näytteistä Koillis-Suomen osa-alueilla.

Variable[1]) *Muuttuja*[1])	Sub-area — *Osa-alue* Kemijärvi		Kuusamo		Pelkosenniemi		Posio		Salla		Savukoski	
	e	s_e	e	s_e	e	s_e	e	s_e	e	s_e	e	s_e
	Remeasured field sample — *Uudelleen mitattu maastonäyte*											
1	53,60	19,53	56,81	5,54	38,78	29,45	45,56	6,99	63,64	11,93	67,57	7,01
2	83,20	7,24	89,94	3,01	71,43	15,71	77,78	7,72	82,39	7,35	87,39	3,21
3	66,40	11,70	70,41	4,86	61,22	13,11	65,56	8,07	67,61	11,95	65,32	8,75
4	17,60	35,22	18,34	31,29	4,08	62,06	6,67	63,22	19,32	33,23	28,38	23,76
5	33,92	22,09	34,83	18,44	13,92	48,91	22,81	17,93	32,46	20,16	30,62	10,74
6	48,04	16,56	47,65	19,31	21,40	47,40	33,14	18,52	45,35	18,41	42,43	9,95
7	22,57	29,64	14,90	23,59	4,36	68,19	17,69	19,59	15,52	34,37	12,71	12,92
8	10,73	33,94	10,40	26,00	1,99	54,46	5,93	37,85	8,32	28,27	9,80	15,38
	New field sample — *Uusi maastonäyte*											
1	56,46	10,12	52,58	6,80	43,23	19,83	55,98	5,64	62,70	8,21	69,53	5,12
2	87,34	2,81	85,09	2,90	86,03	5,39	85,87	2,09	84,73	3,12	85,27	2,79
3	69,13	5,34	66,92	5,57	64,19	13,14	69,84	4,89	66,72	5,67	67,95	4,58
4	8,44	40,17	16,64	16,61	13,54	21,50	19,84	27,37	23,79	18,64	25,04	13,10
5	32,97	10,72	40,03	12,04	33,20	11,01	44,13	14,33	34,89	11,37	37,63	8,06
6	44,89	9,41	55,71	10,19	47,38	7,48	59,19	14,26	48,04	10,72	51,49	8,29
7	24,93	14,09	20,47	16,50	12,60	24,87	22,73	16,20	13,90	16,64	19,51	11,49
8	8,06	18,16	13,73	16,21	6,46	26,49	13,36	16,38	9,15	18,83	10,71	13,11
	Total field sample I[2]) — *Koko maastonäyte I*[2])											
1	55,75	8,90	53,61	5,24	42,45	16,91	53,93	5,00	62,91	6,84	69,06	4,21
2	86,31	2,74	86,27	2,31	83,45	5,16	84,28	2,26	84,21	2,89	85,78	2,23
3	68,45	4,88	67,78	4,32	63,67	10,89	69,00	4,21	66,92	5,09	67,32	4,04
4	10,71	28,00	17,05	14,60	11,87	21,10	17,25	26,18	22,81	16,26	25,84	11,43
5	33,21	9,60	38,76	10,13	29,80	11,97	39,94	13,40	34,36	9,85	35,94	6,82
6	45,65	8,07	53,67	9,00	42,97	9,12	54,32	13,27	47,44	9,22	49,37	6,92
7	24,34	12,66	19,11	13,99	11,14	23,87	21,74	13,90	14,26	14,93	17,87	10,01
8	8,72	16,36	12,92	13,89	5,67	25,29	11,90	15,71	8,96	15,92	10,49	10,67
	Total field sample II[3]) — *Koko maastonäyte II*[3])											
1	55,80	9,42	54,96	4,51	41,63	18,61	50,81	4,77	63,00	7,07	68,82	4,30
2	86,75	2,71	87,29	2,19	83,90	5,48	85,21	2,07	84,36	2,95	86,16	2,19
3	68,63	5,06	68,81	3,85	62,63	10,58	68,58	4,47	66,88	5,28	67,35	4,18
4	10,55	29,55	16,96	15,14	8,15	26,65	11,62	30,85	22,35	17,07	25,68	11,80
5	33,14	10,04	38,16	10,58	28,89	12,26	29,10	12,61	34,24	10,32	34,41	6,76
6	45,58	8,57	53,49	9,40	44,55	7,89	41,25	12,65	47,30	9,65	46,91	6,67
7	24,42	13,32	17,80	14,38	8,26	29,75	20,06	13,61	14,16	15,41	15,08	9,13
8	8,43	16,63	12,39	14,55	3,27	31,25	9,74	17,40	8,86	16,43	10,29	10,42

[1]) 1—4 Proportions (%): *Osuusluvut:*
1 = Mineral soils — *Kangasmaat*
2 = Forested area — *Metsää kasvava maa*
3 = Forest land — *Metsämaa*
4 = Mature stands — *Uudistuskypsät metsät*

5—8 Mean volumes *Keskitilavuudet* (m³/ha)
5 = Tree stock on land — *Puusto maalla*
6 = Tree stock on forest land — *Puusto metsämaalla*
7 = Pine on land — *Mänty maalla*
8 = Saw-timber on land — *Tukki maalla*

[2]) Weighted by the numbers of the relascope plots in the sub-samples
Painotettu osanäytteiden koealojen lukumäärillä
[3]) Weighted by the statistical precision of the sub-samples
Painotettu osanäytteiden tilastollisella tarkkuudella

The inconsistencies mentioned above are perceivable in Figure 12. It is quite evident that the remeasured field sample overestimates the statistical precision of the estimated proportion of forest land in *Kuusamo*. As to the estimated mean volume on land in *Pelkosenniemi,* both sub-samples probably give a poor estimate of the statistical precision. This is particularly surprising as the estimations go in opposite directions.

The true population variation is not, of course, similar in the different sub-areas.

This does not suffice to explain the most striking inconsistensies of the results. One can conclude that the estimates of the statistical precision derived from the field samples are fairly inprecise. A single sample, especially small one, can lead to a considerable under- or overestimation of the statistical precision. The most apparent example of this concerns the mean volumes in *Pelkosenniemi*.

Total land area can be used as an indicator of the magnitude of the sampling variance when using a certain sampling design. Assuming a similar pattern of the population variation, the sampling variance follows a rule

$$v_b = \frac{A_a}{A_b} * v_a$$

where A_a = land area in sub-area a
A_b = land area in sub-area b
v_a = sampling variance in sub-area a
v_b = sampling variance in sub-area b

So, for example, the sampling design being the same, sampling variance is halved when the land area is doubled. This is simply because the size of the sample is doubled. As to the relative standard errors s_a % and s_b %, the corresponding formula reads

$$s_b \% = \sqrt{\frac{A_a}{A_b}} * s_a \%$$

When concerning mean volume on land, this formula is valid as such. As to the proportions and mean volumes of strata, a more complex approximation is needed if the proportions in sub-area a are distinctly different from those in sub-area b. The relative standard error of the proportion of forest land in sub-area b can be calculated as (cf. Svensson 1983, p. 28)

$$s_b \% = \frac{AF_a}{AF_b} * \sqrt{\frac{A_b}{A_a}} * s_a \% = \frac{p_a}{p_b} * \sqrt{\frac{A_a}{A_b}} s_a \%$$

where the AF's and p's refer to the forest land area and the corresponding proportion by sub-area. Concluding from Svensson (ibid.), this formula is also applicable to the mean volumes. Coefficient p_a/p_b has a little effect on the value of s_b % if the p's are of equal class of magnitude, as is generally the case with the proportion of forest land in *Koillis-Suomi*.

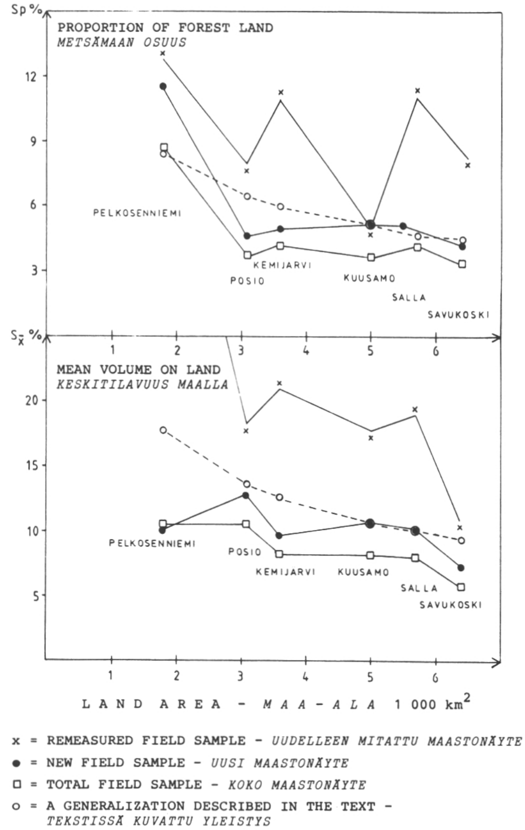


Figure 12. Relative standard error of the mean volume on land ($s_{\bar{x}}$ %) and the proportion of forest land (s_p %) estimated from the different field samples by sub-area.
Kuva 12. Puuston keskitilavuuden ja metsämaan osuuden suhteellinen keskivirhe ($s_{\bar{x}}$ % ja s_p %) arvioituna eri maastonäytteistä osa-alueittain.

Let us assume that the statistical precision estimated from the new field sample in *Kuusamo* is a good approximation. Proceeding from this sub-area and using the formulae presented above, the new field sample gives following relative standard error of the proportion of forest land and the mean volume on land:

Sub-area	Land area, sq. km.	s_p %	$s_{\bar{x}}$ %
Pelkosenniemi	1 826	9,22	19,94
Posio	3 096	7,08	15,31
Kemijärvi	3 578	6,59	14,24
Kuusamo	5 006	5,57	12,04
Salla	5 745	5,20	11,24
Savukoski	6 423	4,92	10,63

The values of s_p % have been calculated assuming that $p_a/p_b = 1$. The new estimates are illustrated by the dashed line in Figure 12. The impression is further confirmed that one must not trust too deeply error characteristics estimated from a single small sample. So, for example, it is evident that the mean volume on land in *Pelkosenniemi* estimated from the new field sample is associated with a relative standard error of about 15—20 % which is considerably more than estimated from the actual sample (11 %).

The procedure discussed above can be applied downwards from the error characteristics of *Koillis-Suomi,* too. For that purpose, the new field samples are put together and the results are calculated for the total area. The error characteristics of *Koillis-Suomi* and those derived for the subareas are:

Sub-area	s_p %	$s_{\bar{x}}$%
Pelkosenniemi	8,89	17,47
Posio	6,82	13,42
Kemijärvi	6,35	12,48
Kuusamo	5,37	10,55
Salla	5,01	9,85
Savukoski	4,74	9,32
Koillis-Suomi	2,37	4,66

The values of s_p % and $s_{\bar{x}}$ % are now smaller than in the previous approximation. One can conclude that the population variation present in the new field sample is, in *Kuusamo,* greater than the average in *Koillis-Suomi.* As to the volumes, the forests probably are more variable in *Kuusamo.* The southern location of this sub-area and the ownership distribution of the forests there support the statement. If the population variation in the total area were the same as in *Kuusamo,* the value of $s_{\bar{x}}$ % in *Koillis-Suomi* would be 5,32 instead of 4,66. The corresponding values of s_p % are 2,46 and 2,37.

The two sets of figures above justify the general conclusions that the new field sample yields an estimate for the total volume with a relative standard error of 4—5 % in *Koillis-Suomi* and from 9—10 % *(Savukoski)* to 17—19 % *(Pelkosenniemi).* The corresponding values of the area of forest land are 2—3 % and from 4—5 % to 8—9 %. As to the validity of these error characteristics, one must remember that the estimates are biased, probably overestimates, because the sample is systematic. Ignoring the variation within the tracts also causes some overestimation of the sampling error.

It is assumed here that the estimated errors include the effect of the bias of ratio estimates on the accuracy. The effect is regarded as zero if the ratio between the bias and the standard error is less than one tenth (Cochran 1977, p. 14 and 162). Accuracy is expressed as mean square error which is calculated as

$$\mathrm{MSE} = v(e) + B^2$$

where B refers to the absolute bias and v(e) to the sampling variance. A bias increases the probability that the true mean volume falls outside the range $e \pm t * s_e$. This is avoided by using $\sqrt{\mathrm{MSE}}$ instead of s_e. The upper limit of the ratio between $|B|$ and s_e in the new field sample by sub-area is (the estimator has been presented on p. 22):

Sub-area	Characteristic	
	$\bar{x}_p$	others
Pelkosenniemi	0,116	0,067
Posio	0,087	0,073
Kemijärvi	0,106	0,084
Kuusamo	0,073	0,060
Salla	0,091	0,063
Savukoski	0,087	0,067
Koillis-Suomi	0,032	0,021

Two sets of values are needed because the tract size is different in the estimation of the mean volume on forest land than in the estimation of the other characteristics. Only in a couple of cases does the ratio exceeds the critical value 0,1. Considering that the values refer to the maximum possible ratio, it is evident that the bias due to the ratio estimation does not have a significant effect on the accuracy of the results calculated on the basis of the new field sample. For interest's sake, assume that the estimated errors refer to the s_e %-values. The relation between s_e % and $\sqrt{\mathrm{MSE}}$ % is:

Sub-area	Characteristic[1])		
	1	2	2/1
Pelkosenniemi	7,48	7,53	1,007
Posio	14,26	14,31	1,004
Kemijärvi	9,41	9,46	1,005
Kuusamo	10,19	10,22	1,003
Salla	10,72	10,76	1,004
Savukoski	8,29	8,32	1,004
Koillis-Suomi	4,43	4,43	1,000

[1]) 1 = s_e % and 2 = maximum possible $\sqrt{\mathrm{MSE}}$ %

The relative difference is less than 1 %, which is not of practical importance.

The remeasured field sample and the new field sample are combined in two ways in this investigation (p. 24). The two approaches, weighting either by the sample size or by the estimated statistical precision, generally leads to fairly close estimates, as can be seen from Table 4. The differences increase with decreasing statistical precision. With a few exceptions, however, the difference is less than one standard deviation. Conserning the estimated proportion of mature stands and the mean volumes on land and on forest land (variables 4—6), the greatest deviations occur in *Posio*. The two latter variables are, of course, strongly dependent on the first.

The seemingly inconsistent results in *Posio* are explained by the fact that it is not the relative standard error, but the the absolute sampling variance that affects the final outcome when the samples are combined and weighted by the estimated statistical precision (formula 13, p. 24). This affects the result especially as to the proportion of mature stands. The remeasured field sample evidently gives too low a sampling variance for the characteristics in question, and consequently this sample obtains a too great weight in combining the samples. Once again, it proves to be that a small single sample gives unprecise estimates for the error characteristics. This severely limits the use of the formula presented by Burk and Ek (1982).

The estimated error characteristics in the combined samples are not significantly different in the two approaches. Weighting by the sample size gives a smaller relative error more frequently than weighting by the statistical precision. Considering the accuracy of the estimated error characteristics, however, this cannot be regarded as a decisive factor. A more essential observation is that weighting by the statistical precision does not yield any marked gain by sub-area. Consequently, being a more simple approach, weighting by the sample size is the preferable alternative. This is all the more appropriate as the error characteristics of the sub-samples cannot be estimated with a sufficient accuracy.

The results of *Koillis-Suomi* can be calculated by summing the estimated quantities of the sub-areas. For each quantity (area or total volume) we have four estimates, as discussed above (Table 5). As stated before, the two approaches in combining the sub-samples give about the same relative standard error by sub-area. Summing the results of the sub-areas further levels the randomlike deviations. Practically speaking, samples 3 and 4 have about the same relative standard errors in *Koillis-Suomi.*

The results in Table 5 reveal a significant detail concerning the mean volumes. The mean volumes on land and on forest land as well as the mean volume of pine on land (characteristics 5—7) estimated from the remeasured field sample are much lower than those estimated from the new field sample. It can be seen from Table 4 that the phenomenon is distinctly apparent in four sub-areas. The subject will be discussed in more detail later in this chapter.

Sample 4 gives distinctly lower volume estimates for *Koillis-Suomi* than does sample 3. Consequently, weighting by the precision gives a greater weight to the remeasured field sample than does weighting by the sample size. This inconsistency is caused by the low absolute sampling variances in the remeasured field sample.

The other way to calculate the results for *Koillis-Suomi* is to combine the samples of the sub-areas before the calculations. Let us use notation II for this procedure and notation I for the summing of the results of the sub-areas. The outcome with the different samples is presented in Table 6. The two ways (I and II) of calculating the results for *Koillis-Suomi* give very similar estimates using samples 1—3. This is true also for the proportions estimated from sample 4. As to the volumes, the inconsistency, discussed above, between samples 3 and 4 is much smaller in case II than in case I. The result is in accordance with the trends present in Table 4.

The dominant impression from the experiences in the calculations discussed above is that the easiest way is the best way. Combining the sub-samples weighting inversely by variance is not recommendable in this context. The main reason for this is the fact that the error characteristics estimated from the sub-samples are unprecise, which leads to inconsistent outcomes. The procedure is a beneficial way to combine estimates obtained by quite different sampling designs.

The results for the main reference areas can be calculated by combining the samples of the sub-areas before the calculations. This

Table 5. Four different estimates for the eight main characteristics in Koillis-Suomi.
Taulukko 5. Kahdeksan pätunnusluvun neljä eri estimaattia Koillis-Suomen piirimetsälautakunnan alueella.

Sample[1]) *Otos*[1])	Estimate[2]) *Estimaatti*[2])	1	2	3	4	5	6	7	8
		Characteristic[3]) — *Tunnusluku*[3])							
1	e	57,94	83,89	66,71	18,55	30,18	42,47	15,15	8,69
	s_e %	4,76	2,52	4,15	14,47	8,09	7,60	11,72	11,70
2	»	59,37	85,53	67,60	19,36	37,30	51,28	19,09	10,60
		3,26	1,29	2,41	8,40	4,75	4,43	6,42	7,23
3	»	59,10	85,25	67,42	19,31	35,80	49,44	18,19	10,20
		2,73	1,13	2,07	7,27	4,16	3,89	5,64	6,23
4	»	58,90	85,78	67,52	18,18	33,89	47,38	16,82	9,55
		2,79	1,13	2,08	7,69	4,23	3,90	5,80	6,51

[1]) 1 = Remeasured field sample — *Uudelleen mitattu maastonäyte*
2 = New field sample — *Uusi maastonäyte*
3 = Combined field sample — *Yhdistetty maastonäyte*
4 = The estimates from samples 1 and 2 combined weighting by precision
Näytteiden 1 ja 2 antamat estimaatit yhdistetty painottaen tilastollisella tarkkuudella
[2]) e = Estimated proportion, % (1—4) or mean volume, m³/ha (5—8)
Arvioitu osuusluku, % (1—4) tai keskitilavuus, m³%ha (5—8)
s_e % = Relative standard error — *Suhteellinen keskivirhe*
[3]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4, alaviitta 1*

Table 6. Proportion of forest land (3) and mean volume on land (5), as well as on forest land (6) in *Koillis-Suomi* estimated by summing the results of the sub-areas (I) and by combining the samples prior to the calculations (II).
Taulukko 6. Metsämaan osuus (3) ja puuston keskitilavuus maalla (5) sekä metsämaalla (6) Koillis-Suomen piirimetsälautakunnassa arvioituna yhdistämämällä osa-alueiden tulokset (I) ja yhdistämällä osa-alueiden näytteet ennen laskentaa (II).

Sample[1]) *Otos*[1])	Estimate[1]) *Estimaatti*[1])	Characteristic — *Tunnusluku*					
		3I	3II	5I	5II	6I	6II
1	e	66,71	66,79	30,18	30,53	42,47	42,89
	s_e %	4,14	4,04	8,09	8,06	7,60	7,53
2	»	67,60	67,59	37,30	37,33	51,28	51,33
		2,41	2,37	4,75	4,66	4,43	4,43
3	»	67,42	67,41	35,80	35,78	49,44	49,42
		2,07	2,00	4,16	4,16	3,89	3,89
4	»	67,52	67,38	33,89	35,06	47,38	48,53
		2,08	2,06	4,23	4,09	3,90	3,87

[1]) As in Table 5 — *Kuten taulukossa 5*

procedure gives estimates for the main reference area which do not equal the sums of the corresponding estimated quantities in the sub-areas. In our case, however, the differences are small and definitely within the sampling error. The reason for the deviations is that the sampling intensity (the average area represented by one field plot) varies slightly from one sub-area to another. The random-like variation decreases with increasing sampling intensity.

The alternatives discussed above give about the same estimated sampling error. So one can, on good grounds, select the most practical way of calculating the results. The only inconvenience seems to be that the results are not, in all respects, compatible.

The remeasured field sample gives a good opportunity for studying the changes during the inventory period. The subject will be discussed in a general way in the following text. A more detailed change analysis falls outside the scope of this investigation.

The remeasured field sample indicates that the classification of forest land has not been applied entirely consistently in the two inventories in question (Table 7). Taking into account that the estimations have been made on the same plots, the changes of the estimated proportion of forest land in three sub-areas are too great to be real. The changes go to both directions which indicates that different persons apply the classification in different ways. One can suspect that the same problem also concerns the other nominal scale variables.

There are considerable changes in the mean volumes, too. Unlike proportions, mean volumes are measured quantities and the changes are real in the sample. The statistical significance of the change depends on the representativeness of the sample and the correlation between the two measurements. The variance of the difference is calculated as (Zöhrer 1980, p. 104)

$$s_{\bar{x}d}^2 = s_{\bar{x}1}^2 + s_{\bar{x}2}^2 - 2rs_{\bar{x}1}s_{\bar{x}2} \tag{14}$$

where $s_{\bar{x}1}$ and $s_{\bar{x}2}$ refer to the standard error of the estimated mean volume on occasion 1 and 2 and r is the correlation coefficient calculated on the basis of the two successive measurements. Consequently, the statistical significance of an observed difference increases with increasing sample size and/or with increasing correlation. The standard error of the mean volume estimated from the old systematic field sample on occasion 1 and 2 by sub-area is (m³/ha):

Sub-area	$s_{\bar{x}1}$	$s_{\bar{x}2}$
Pelkosenniemi	4,88	6,81
Posio	3,11	4,09
Kemijärvi	7,42	7,49
Kuusamo	4,13	6,42
Salla	5,34	6,54
Savukoski	3,46	3,29
Koillis-Suomi	2,14	2,46

The errors have been calculated on the basis of the variation of the tract values. The estimated mean volume on land decreased

Table 7. Proportion of forest land (%) and mean volume on land (m³/ha) in *Koillis-Suomi* estimated from the field sample by sub-area and sub-sample.
Taulukko 7. Metsämaan osuus ja puuston keskitilavuus maalla Koillis-Suomessa arvioituna maastonäytteen eri osista osa-alueittain.

Sample — *Näyte*	Sub-area¹) — *Osa-alue*¹) 1	2	3	4	5	6	Koillis-Suomi
				%			
Occasion 1 — *1. invent.*							
Permanent — *Pysyvä*	62,40	74,56	57,72	65,17	67,05	65,92	66,59
Occasion 2 — *2. invent.*							
Old perm. — *Vanha pysyvä*	66,40	70,41	61,22	65,56	67,61	65,32	66,79
New sample — *Uusi näyte*	70,24	65,63	65,69	67,76	69,52	68,24	68,04
Permanent — *Pysyvä*	68,25	67,89	62,99	71,30	64,59	67,74	67,25
Temp. — *Tilap.*	69,13	66,92	64,19	69,84	66,72	67,95	67,59
Total new — *Koko uusi*	68,45	67,78	63,67	69,00	66,92	67,32	67,41
				m³/ha²)			
Occasion 1 — *1. invent.*							
Permanent — *Pysyvä*	30,98	41,61	16,93	20,15	31,88	30,35	31,16
Occasion 2 — *2. invent.*							
Old perm. — *Vanha pysyvä*	33,92	34,83	13,92	22,81	32,46	30,62	30,53
New sample — *Uusi näyte*	37,51	42,15	31,37	40,82	35,77	38,21	38,05
Permanent — *Pysyvä*	29,36	38,45	34,67	46,45	34,22	37,21	36,79
Temp. — *Tilap.*	32,97	40,03	33,20	44,13	34,78	37,63	37,33
Koko uusi — *Koko uusi*	33,21	38,76	29,80	39,94	34,36	35,94	35,78

¹) 1 = Kemijärvi 2 = Kuusamo 3 = Pelkosenniemi
4 = Posio 5 = Salla 6 = Savukoski
²) Including usable dead trees — *Sisältää käyttökelpoiset kuolleet puut*

considerably from occasion 1 to occasion 2 in *Kuusamo* and *Pelkosenniemi*. However, the corresponding standard errors are much greater on occasion 2. The same contradictory trend is also apparent in *Koillis-Suomi*. The decrease of the statistical precision is greater in relative than in absolute terms, from 6,91 % to 8,06 % in *Koillis-Suomi*. These estimates have been obtained by combining the samples of the sub-areas before the calculations.

Studying the statistical significance of the observed changes in the volumes requires an estimation of the r-values. In the following, the calculations are based on the tract-values using the formula

$$r^2 = \frac{\{(\sum_{i=1}^{k} w_i^2 (x_i - \bar{x})(y_i - \bar{y})\}^2}{\sum_{i=1}^{k} w_i^2 (x_i - \bar{x})^2 * \sum_{i=1}^{k} w_i^2 (y_i - \bar{y})^2}$$

where k = number of tracts
x_i = mean volume of tract i on occasion 1
y_i = mean volume of tract i on occasion 2
$\bar{x}$ = mean volume on occasion 1
$\bar{y}$ = mean volume on occasion 2
w_i = weight of tract i = $m_i/\bar{m}$

The w_i's are necessary because the number of the plots on land per tract (m_i) is variable. Here it is supposed that the w_i-values are the same on both occasions. The outcome of the calculations is presented in Table 8. The correlation varies from fairly low in *Kuusamo* to very high in *Kemijärvi*. The time interval between the two measurements is 5—7 years. The permanent plots were established in 1976 and all of them in *Kuusamo* and *Posio* were remeasured in 1981. In the other sub-areas, half of the plots were remeasured in 1981 and the rest in 1983. The results are inconsistent in the sense that the correlation is lowest in those areas where the time interval between the measurements is shortest. This indicates a faster rate of change in the forests of *Kuusamo* and *Posio*. Such a result is feasible considering the southern location of these sub-areas and the ownership distribution of the forests there.

All the permanent plots could not be located on the second inventory occasion. The plots missed were substituted with new ones which were located and measured independently of the first occasion. This procedure tends to decrease the correlation calculated on the basis of the tract values. The number of the plots measured on land in 1981 were 523, including 14 new stocked plots, which hardly have a marked effect on the correlation. The more rapid change of the forests in the south becomes apparent, in that the new plots were more frequent in the south than in the north.

The observed differences d and the corresponding standard errors s_d by sub-area are (m^3/ha):

Sub-area	d	s_d
Pelkosenniemi	—3,01	2,65
Posio	+2,66	2,46
Kemijärvi	+2,94	1,25
Kuusamo	—6,78	4,88
Salla	+0,58	2,89
Savukoski	+0,27	0,94
Koillis-Suomi	—0,63	1,28

The signs before the values indicate the

Table 8. Data concerning the estimation of the precision of the change in the mean volume on land observed in the remeasured field-sample.
Taulukko 8. Laskennan perustiedot uudelleen mitatuilla maastokoealoilla havaitun puuston keskitilavuuden muutoksen tilastollisen tarkkuuden arviointia varten.

Sub-area *Osa-alue*	k	$\bar{m}$	r	$s^2_{\bar{x}1} + s^2_{\bar{x}2}$	$2rs_{\bar{x}1}s_{\bar{x}2}$	$s^2_{\bar{x}d}$
Pelkosenniemi	6	8,17	0,950	70,21	63,18	7,03
Posio	9	10,00	0,800	26,42	20,37	6,05
Kemijärvi	12	10,42	0,986	111,21	109,65	1,56
Kuusamo	16	10,56	0,650	58,34	34,52	23,82
Salla	15	11,73	0,901	71,32	62,96	8,36
Savukoski	19	11,68	0,962	22,76	21,87	0,89
Koillis-Suomi	77	10,79	0,854	10,65	9,01	1,64

3 408500946V

direction of the change from occasion 1 to occasion 2. The statistical interpretation of the standard error of the difference is as follows. In *Koillis-Suomi,* the true change of the mean volume is between —0,63 ± 1,28 m^3/ha with the probability of 68 %. Accordingly, even this narrow confidence interval includes the possibility that the true change is positive. It can be seen from the tabulated normal distribution that the total probability of a decreased mean volume in *Koillis-Suomi* is 69 %. The corresponding percentages in *Pelkosenniemi* and *Kuusamo* are 87 % and 92 %. An increased mean volume is more probable than a decreased one in the rest of the study area. The percentages are by sub-area: *Salla* 57 %, *Savukoski* 61 %, *Posio* 86 % and *Kemijärvi* 99 %.

The bias of the ratio estimates can affect the validity of the conclusions. The size of the remeasured field sample is small and therefore the maximum possible ratio between |B| and s_e is large. The ratio in question is:

Sub-area	Characteristic	
	$\bar{x}_p$	others
Pelkosenniemi	0,270	0,249
Posio	0,122	0,097
Kemijärvi	0,140	0,068
Kuusamo	0,059	0,046
Salla	0,127	0,044
Savukoski	0,094	0,037
Koillis-Suomi	0,047	0,023

The ratio exceeds the critical value of 0,1 in several cases. The sensitivity of the estimated s_d-values can be studied by using mean square errors ($= s^2_{\bar{x}} + B^2$) instead of $s^2_{\bar{x}1}$ and $s^2_{\bar{x}2}$. The following comparison reveals that the results are not very sensitive to the bias due to the ratio estimates even with small samples:

Sub-area	s_d	s'_d 1)
Pelkosenniemi	2,65	2,73
Posio	2,46	2,47
Kemijärvi	1,25	1,25
Kuusamo	4,88	4,89
Salla	2,89	2,89
Savukoski	0,94	0,95
Koillis-Suomi	1,28	1,28

1) 'Bias-corrected' estimate.

Consequently, the probabilities concerning the changes of the total volumes presented above are fairly stable.

The changes can be studied from two or more successive samples. The statistical significance of an observed difference always depends on the precision of the results and the correlation between the successive measurements. When permanent plots or tracts are not measured the samples are supposed to be independent and the variance of the difference is estimated as

$$s^2_{\bar{x}d} = s^2_{\bar{x}1} + s^2_{\bar{x}2}$$

Consequently, the changes are not estimated as reliably by independent samples as by dependent ones. So for example, if the remeasured field sample were uncorrelated with the initial measurement the standard error of the difference of the estimated mean volume in *Koillis-Suomi* would be 3,26 m^3/ha instead of 1,28 m^3/ha. The corresponding probabilities of a decreased mean volume would be 58 % instead of 69 %.

The case discussed above is an example of total remeasurement of a permanent sample. When repeated several times on the same plots, this kind of forest inventory is called Continuous Forest Inventory (CFI). According to Shain and Rudolph (1965), CFI has been applied in a very wide scale in the United States. Sampling with Partial Replacement of the samples (SPR) combines the advantages of permanent and temporary plots in forest inventory. Some of the plots are established as semipermanent on the first inventory occasion. These plots are remeasured and a new temporary sample is acquired on the following inventory occasion etc. The theory behind the method has been discussed in detail by Ware and Cunia (1962).

The practical applications of the SPR-technique include several modifications. In addition, even with the same sampling desing, one can use several estimators (Omule and Kozak 1982). One usual application is to remeasure only the permanent plots on the second inventory occasion. The mean volume can then be estimated as (cf. Nyyssönen 1967, p. 207)

$$\bar{x}_r = \bar{x}_{2p} + b(\bar{x}_1 - \bar{x}_{1p})$$

where $\bar{x}_{1p}$ = mean volume on permanent plots on occasion 1
$\bar{x}_{2p}$ = mean volume on permanent plots on occasion 2
$\bar{x}_1$ = mean volume on occasion 1
b = coefficient of regression of $\bar{x}_{2p}$ on $\bar{x}_{1p}$

The combined mean volume estimate on the first occasion ($\bar{x}_1$) must be unbiased in this approach. This is not the case with the field sample of the previous inventory in North Finland. The temporary field plots were selected from the photo sample with varying probability. So for example, there are more field plots per unit area on forest land than on scrub and waste land. Sampling intensity also varies areally. Consequently, avoiding bias requires a complex weighting of the temporary field plots. This is the reason why the approach in question is not used here.

The situation will be different in the next inventory. Sampling intensity is constant in all respects in the present field sample. In the future, the whole information of the two successive sets of both the permanent and the temporary samples can be utilized (see Zöhrer 1980, p. 108).

The permanent field plots are remeasured only once to avoid cumulative biases. The establishment cost is lower because the plots and the trees need not to be marked very carefully. A new set of semi-permanent field plots is established on each inventory occasion. This course of action excludes the use of the CFI-technique in the inventory.

The mean volumes on the old permanent field plots are on both occasions considerably lower than the mean volumes on the new permanent and temporary field plots. However, the mean volumes estimated from the combined field and photo samples in the two inventories are fairly close to each other:

Sub-area	Inventory 1976	Inventory 1982—83
	m^3/ha	m^3/ha
Pelkosenniemi	30,5	31,4
Posio	36,2	41,1
Kemijärvi	34,0	33,6
Kuusamo	44,3	40,6
Salla	31,0	34,0
Savukoski	36,9	37,2
Koillis-Suomi	36,2	36,7

The two successive estimates are to some degree correlated because the field tracts of the previous inventory are included in the photo sample of the new inventory. The statistical aspects are not dealt with in more detail in this context. The conclusion must be drawn that the old systematic field sample accidentally represents very poorly the population under study. The old temporary field sample and the old photo sample level down the deviation. Owing to the way in which the temporary field plots have been selected, it is fairly laborious to unravel which sample component makes the greatest contribution to the final outcome.

The usability of the field samples depends on the kind of infromation required and the size of the reference area. The permanent field sample is too small for the purposes of a single inventory both at the subarea level and at the main reference area level. The permanent and the temporary field samples together give informative pilot results by sub-area and fairly precise estimates for the main reference area. The statistical precision cannot be estimated reliably from the field sample by sub-area. An equal sampling intensity is recommendable in the whole field sample because it makes a full utilisation of the SPR-technique possible.

The permanent field sample is advantageous in the estimation of the changes between successive inventories. Let us assume that the mean volume on land in *Koillis-Suomi* has been estimated from the combined field and photo sample with a relative standard error of 2,1 % in 1976 (cf. Poso and Kujala 1978, p. 46) and 2,3 % in 1982—83. If the estimates were not correlated, the standard error of the observed difference would be 1,29 m^3/ha. This is about the same as estimated on the basis of the remeasurement of the old permanent field sample (1,28 m^3/ha).

4. USE OF THE PHOTO SAMPLE

41. Statistical basis

The purpose of sampling is to provide information about variables of interest. Under certain conditions, information about auxiliary variables can be utilized to make the estimation more efficient. Stratified sampling as well as regression and ratio estimation can be used when the additional information required is available without error. This not being the case, double or two-phase sampling still offers means to utilize the auxiliary variables.

Double sampling implies a relatively large but cheap first-phase sample to provide the extra information required for stratification, regression or ratio. Let x refer to an auxiliary variable and y to a variable of primary interest. A good estimate of the population mean $\bar{X}$ is needed for ratio and regression estimates $\bar{y}$ (Cochran 1977, p. 327). Stratification being the goal, the information needed comprises the distribution of the x_i-values into the strata. The efficiency of two-phase sampling increases with decreasing relative cost of the first-phase sample and with increasing correlation between x and y.

The principle of double sampling is well suited to natural resource surveys based on photo-interpretation and field measurements. The auxiliary variables are interpreted from aerial photographs in the first-phase sample and a part of the interpreted plots is measured in the field as the second-phase sample. The best correlations are generally achieved by using the same variables both as x- and y-variables. So, for example, the assessment of volume from photos explains the field measured volume better than the other x-variables.

The goals set for the National Forest Survey in Lapland include extension of the field data to the photo sample. The technical realization of the extension requires a division of the photo sample into homogenous strata on the basis of the interpreted data. Consequently, the inventory method used actually is based on double sampling for stratification. Regression and ratio estimators can be considered in this context only for some main variables.

Double sampling for stratification (Neyman 1938) implies estimation of the stratum weights by the first-phase sample. Stratification lessens the more expensive sampling in the second phase. The final goal is to cut down the total sampling cost needed to reach a predetermined statistical precision. Within a fixed budget, the goal is to maximize the precision of the information acquired by sampling.

The pure statistical application of the method implies random samples in both phases. For practical reasons, however, systematic photo samples are very common in this context. The field sample, on the contrary, is usually selected in a restricted random way by photo stratum. The field sample is then allocated into the strata in an optimal way.

The photo sample is used in principle only for the estimation of the stratum weights. The forest characteristics within each stratum are estimated from the field sample. So for example, mean volume on land is estimated as

$$(15)\qquad \bar{x} = \sum_{h=1}^{L} w_h \bar{x}_h$$

where w_h and x_h refer to the weights and the mean volumes by photo stratum and L is the number of the strata. Mean volume in a field stratum, for example on forest land, is estimated as

$$(16)\qquad \bar{x}_p = \frac{\sum_{h=1}^{L} w_h p_h \bar{x}_{ph}}{p}$$

where $\bar{x}_{ph}$ = mean volume on forest land in photo stratum h

p_h = proportion of forest land in photo stratum h

$p = \sum_{h=1}^{L} w_h p_h$ = proportion of forest land

The following variance estimators imply that the photo sample covers the whole study area.

The population from which the photo sample is chosen in Lapland is infinite. Number of photo plots is great, generally many thousands and always more than 50. So an almost unbiased sample estimator for the variance of $\bar{x}$ can be written (Cochran 1977, p. 333):

$$(17) \quad v(\bar{x}) \cong \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h} + \frac{1}{n} \sum_{h=1}^{L} w_h (\bar{x}_h - \bar{x})^2$$

where s_h^2 = estimated population variance in stratum h

m_h = number of field plots in stratum h

The divisor in the first term on the right is m_h-1 when the variance of an estimated proportion is in question (see Cochran ibid., p. 52). Comparison with the corresponding formula for stratified random sampling (e.g. Cochran ibid., p. 95) reveals that the second term in the right expresses the error due to the estimation of the stratum proportions. Assuming that the values of w_h, s^2_h and $\bar{x}_h$ have been estimated accurately enough, the effect of the sample size on the value of $v(\bar{x})$ can be simulated with ease by changing the values of m_h and n in formula 17.

It can be concluded from formulae 16 and 17 that the variance of an estimated mean volume in a field stratum can be approximated as:

$$(18) \quad v(\bar{x}_p) \cong \frac{1}{p^2} * \sum_{h=1}^{L} \frac{w_{ph}^2 s_{ph}^2}{m_{ph}} + \frac{1}{n} * \sum_{h=1}^{L} w_h (\bar{x}_{ph} - \bar{x}_p)^2$$

where $\bar{x}_{ph}$ = mean volume in the field stratum in photo stratum h

m_{ph} = number of field plots correspondingly

w_{ph} = $w_h * p_h$

and $$s_{ph}^2 = \frac{\sum_{i=1}^{m_{ph}} (x_{phi} - \bar{x}_{ph})^2}{m_{ph} - 1}$$

x_{phi} = volume of plot i in the field stratum in photo stratum h

Formulae 17 and 18 can also be used in connection with the grouping method (Poso 1972) with certain specifications. The groups must be combined into group sets which are used as strata in the calculations. The weight of a group set equals the sum of the weights of the groups included. As to the proportions, the artificial strata should be internally homogenous in relation to the expected values of p. This requirement is generally not very difficult to fulfil with one proportion. The problem is that estimation of the variances of several proportions requires several divisions of the photo sample into group sets.

The group sets formed for the proportions are generally fairly heterogenous as to the volumes. In the case that the groups are homogenous in relation to the interpreted volume and there exists a distinct correlation between the photointerpreted volumes and the field measured volumes, the variances of mean volumes estimated by formulae 17 and 18 are overestimates.

This overestimation is avoided by estimating the variances within the group sets through regression analysis. Assuming linear regression, the dependence of the field measured volumes y_i on the photointerpreted volumes x_i can be expressed as

$$(19) \quad \hat{y}_i = a + bx_i$$

where $\hat{y}_i$ = field volume of plot i estimated through regression

a and b = regression coefficients

The variance of the $\hat{y}$-values is calculated as

$$(20) \quad s_{\hat{y}x}^2 = \frac{\sum_{i=1}^{m} (y_i - \hat{y}_i)^2}{m - 2}$$

where m = number of the field plots

Regression coefficients a and b and the variance of y are estimated for each group set separately. It is readily understood that with a close correlation between the x- and y-values the deviations of the y-values from the regression line are smaller than the deviations from the mean value of y.

Forming the group sets afterwards for the estimation of the sampling error is somewhat clumsy. It helps, however, to understand that the grouping method is simply a double sampling for stratification. It reveals also that the gain from selecting the field sample by grouping becomes apparent in the difference of the s^2_h-values from the $s^2_{\hat{y}xh}$-values. A groupwise approach in the estimation of the sampling error requires more complex algebraic expressions for the estimators (see Poso and Kujala 1977, p. 31).

The field sample is systematic in the

new National Forest Inventory in Lapland. Despite this, an objective is to extend the field data to the whole photo sample. For that purpose, the photo sample is first divided into homogenous photo strata. Then the strata are mechanically cut into parts which correspond to the groups in the grouping method. The partition is now needed only for the extension of the field data which is done by groups. The technical realization will be described in more detail later in this paper.

42. The accuracy of assessments from aerial photographs

421. Qualitative variables

The main emphasis of the photo-interpretation is on the assessment of qualitative variables. Classifications *sub-class, treatment class, dominant tree species* and *drainage status* are done purely on the nominal scale basis. *Land use class* inclines to the ordinal scale in that the classification is partly based on the potential timber production capacity of the sites.

Similarity or error matrices are a visual way to present the photo-interpretation accuracy of the nominal scale variables. An example concerning *land use class* in *Kuusamo* is in Table 9. An error matrix is a square array expressing the "in the field distributions" in classes derived from the photo-interpretations. It is a matter of taste whether these conditioned distributions are presented in rows or columns. The total agreement of the photo-interpretation is in any case the sum of the principal diagonal elements.

The sum of the off-diagonal elements expresses the number of the wrong interpretations. These can be called either omission or commission errors depending on the point of view (Townshend 1981, p. 87). Omission means that an element belonging in class A in the field has not been allocated to class A ("poisjääntivirhe" in Finnish). A commission error occurs when an element allocated to A does not belong in A in the field ("luokitusvirhe"). A comparison of the distributions of these two types of errors gives information about the causes of the wrong allocations.

Overall accuracy is calculated as the ratio of the correct interpretations to all interpretations. The expected value of this statistic depends on the marginal distributions. So a comparison of overall accuracy values is adequate only with similar or nearly similar marginal distributions.

The contingency coefficient C takes into account the expected values of both the major diagonal and off-diagonal elements. The value of C also depends on the total number of elements. In addition, the maximum value of C is dependent on the number of classes. So, as before, this statistic is also of limited value in the comparison of different cases.

A valid solution is to test the statistical significance of the χ^2-values which are needed in the estimation of C. This, however, presupposes a certain minimum number of elements expected in each cell of the square array (e.g. Siegel 1956, p. 201). The original classes must often be combined in order to fulfil this qualification.

χ^2 tests the independence hypothesis with regard to association, not agreement (Cohen 1960, p. 39). It can be demonstrated that the assessments from photos and the field observations can be highly dependent although the agreement is poor. Cohen (ibid.) proposes a statistical measure K to test just the agreement. He concluded that the only relevant quantities in this case are the actual and the expected overall accuracies p_a and p_c:

$$(21) \quad K = \frac{p_a - p_c}{1 - p_c}$$

The expected overall accuracy is estimated from the marginal totals of a square array. The value of K is +1 when all elements are on the major diagonal. The variance of K is estimated as

$$(22) \quad s_K^2 = \frac{p_a(1 - p_a)}{m(1 - p_c)^2}$$

where m = total number of elements

This estimate should be precise enough when m is greater than 100. For testing the null hypothesis ($p_c = p_a$) the variance is estimated as

$$(23) \quad s_{K0}^2 = \frac{p_c}{m(1 - p_c)}$$

Table 9. Error matrix concerning the assessment from photos of land use classes in *Kuusamo*.
Taulukko 9. Maaluokan ilmakuvatulkinnan virhematriisi Kuusamossa.

Interpreted land use class[1]) *Tulkittu maaluokka*[1])	Land use class observed in the field[1]) *Maastossa todettu maaluokka*[1])						Errors of commission *Luokitus-virheet*	
	1	2	3	4	5	Total	n	%
1	*319*	19	1	6	1	346	27	32,1
2	27	*65*	12	1	1	106	41	48,8
3	1	11	*35*	0	1	48	13	15,5
4	3	0	0	*23*	0	26	3	3,6
5	0	0	0	0	*102*	102	0	0,0
Total	350	95	48	30	105	628	84	100,0
Errors of omission *Poisjäänti-virheet* n	31	30	13	7	3	84	Overall accuracy 544/628=86,6 % *Yleistarkkuus*	
%	36,9	35,7	15,5	8,3	3,6	100,0		

[1]) 1 = Forest land — *Metsämaa*
2 = Scrub land — *Kitumaa*
3 = Waste land — *Joutomaa*
4 = Other land — *Muu maa*
5 = Waters — *Vedet*

Dividing the value of K by the value of s_{K0} results in a test quantity the statistical significance of which is found from the normal distribution. More important than to test the null hypothesis, however, is to study the difference between two independent K's. The test quantity is then calculated as

$$(24) \quad z = \frac{K_1 - K_2}{\sqrt{s_{K1}^2 + s_{K2}^2}}$$

where s^2_{K1} and s^2_{K2} refer to the variances of K_1 and K_2 calculated by formula 22. The approach can be used to show consistency of two sets of interpretations (Congalton and Mead 1983). This kind of testing implies interpretation of the same units by one person on two occasions or by two persons on one occasion. For testing purposes, the units can be interpreted from different photo materials, as well.

The K-statistic is used in this investigation only to show the degree of agreement in the assessment of the main nominal scale variables from photographs. Consistency cannot be studied because all photo plots have been interpreted only once and by one person.

The field sample of *Kuusamo* consists of 628 points (relascope plots) of which 523 fall on land. The overall accuracy of the interpretation of land use classes is 86,6 % with the classification presented in Table 9. The agreement is only slightly poorer (84,5 %) after the class *waters* has been excluded.

It is worth noting that even the division into land or water is not quite certain. An erroneous location of points serves as an implicit explanation. Errors also arise because the borderline between land and water every now and then is vague even in the field.

The limitations concerning the expected numbers of the observations hamper the use of the coefficient of contingency C to indicate the accuracy of the interpretation. One solution to fulfill the qualifications is to combine the two classes *other land* and *waters*. This is not, however, very reasonable. Neither is the problem solved by excluding the class *waters*. Difficulties are also present with the other nominal scale variables. In addition, the solutions are different in different sub-areas which makes the C-values inconsistent. Consequently, it has beer decided here to use only the K-statistic as the measure of the accuracy of the assessment of nominal scale variables from photos.

Using the K-statistic permits a more detailed classification than presented in Table 9. A division of the class *other land* into four parts results in the following outcome:

Statistic	Case I[1])	Case II[2])
Number of points	628	523
Agreements	542	440
Overall accuracy, %	86,3	84,1
K-value	0,7838	0,6902
s_K (formula 22)	0,0217	0,0312
s_{K0} (formula 23)	0,0304	0,0427
z_0 (= K/s_{K0})	25,81	16,18

[1]) Seven classes on land + *waters*
[2]) As above but *waters* excluded

The critical z_0-value is 2,33 at the probability level of 0,01. Consequently, the observed K-values above are highly significant in both cases. Excluding *waters* decreases the agreement considerably because the lakes and the rivers can be delineated almost without error on aerial photographs.

The decrease is smaller in the other subareas of *Koillis-Suomi* because this stratum is not so frequent there.

The land use class is assessed for all points. The other nominal scale variables are assessed only within certain interpreted land use classes which are by variable (x = assessment has taken place):

Variable to be assessed	Interpreted land use class[1]) 1	2	3	4—8
Sub-class	x	x	x	
Treatment class	x			
Drainage status	x	x	x	
Dominant tree species	x	x		

[1]) 1 = Forest land 2 = Scrub land 3 = Waste land and 4—8 = other land use classes + waters

All variables are assessed on forest land.

Table 10. Accuracy of the assessment of the main nominal scale variables from photos by sub-areas in *Koillis-Suomi.*
Taulukko 10. Tärkeimpien nominaaliasteikkotasoisten muuttujien tulkintatarkkuus Koillis-Suomen osa-alueilla.

Characteristic[1]) *Tunnusluku*[1])	Sub-area[2]) — *Osa-alue*[2]) 1	2	3	4	5	6	Koillis-Suomi
			Land use class — *Maaluokka*				
n	378	523	229	366	622	699	2817
O.a.	82	84	79	80	83	87	83
K	0,64	0,69	0,61	0,60	0,67	0,74	0,67
t_o	12,6	16,2	10,0	11,4	17,5	19,2	36,4
			Sub-class — *Alaryhmä*				
	365	490	227	352	604	698	2736
”	87	90	89	88	92	94	91
	0,74	0,80	0,78	0,76	0,84	0,85	0,81
	13,9	17,4	11,7	13,9	19,3	19,2	40,3
			Treatment class — *Kehitysluokka*				
	226	319	125	217	366	443	1696
”	65	53	38	62	59	46	54
	0,43	0,35	0,17	0,44	0,44	0,28	0,36
	8,0	10,2	3,3	9,5	14,0	10,0	24,1
			Drainage status — *Ojitustilanne*				
	327	423	202	316	548	645	2461
”	77	72	78	73	83	90	80
	0,61	0,55	0,64	0,56	0,70	0,78	0,66
	13,2	14,3	12,6	12,0	18,9	18,7	37,5
			Dom. tree species — *Vallitseva puulaji*				
	300	381	165	281	465	543	2135
”	91	83	64	80	77	79	80
	0,55	0,57	0,26	0,38	0,52	0,50	0,50
	4,9	9,2	3,2	4,4	11,0	9,8	19,0

[1]) n = Number of observations — *Havaintojen lukumäärä*
O.a. = Overall accuracy, % — *Oikeita tulkintoja, %*
K = K-statistic value — *K-tunnusluvun arvo*
t_0 = K/S_{K_0} (Test quantity — *Testisuure)*
[2]) 1 = Kemijärvi, 2 = Kuusamo, 3 = Pelkosenniemi, 4 = Posio, 5 = Salla, 6 = Savukoski

The interpreted land use class being 4—8, any other assessments are not done.

The relevant data concerning the accuracy of the assessment of the main nominal scale variables from photos are presented in Table 10. As to the variable land use class, the material included comprises those relascope plots which fall on land (land use classes 4—7) both on aerial photographs and in the field. The concept of *conditioned photo stratum* has been applied to the other variables, too. The concept implies the exclusion of those plots on which the classification in question has not been made either on aerial photographs or in the field. This kind of "error" arises from an erroneous interpretation of another variable.

The statistical significance of the observed agreement is generally very high, especially in the whole material of *Koillis-Suomi*. This is natural because a good agreement was the goal when planning the classifications. In addition, the variables which were known to have a poor agreement were excluded from the photo-interpretation. These measures were used to make the work as efficient as possible with the photographic material in question.

The order of the nominal scale variables in respect to the K-value is sub-class, land use class, drainage status, dominant tree species and treatment class (Table 10). It must be kept in mind that the statistical significance of the K-values depends on the number of the observations (formula 23). So a direct comparison of the z_0-values does not give valid information. However, it is possible to calibrate the obtained z_0-values in respect to a certain number of observations assuming that the K-values are not changed. z_0 is calculated as

$$z_0 = \sqrt{m} * \sqrt{\frac{K^2(1-p_c)}{p_c}}$$

where the latter factor in the right is assumed to be a known (estimated) constant and m is the number of the observations. Under these presumptions the calibrated z_0-values z'_0 can be calculated simply as

$$(25) \qquad z'_0 = \sqrt{\frac{m'}{m}} * z_0$$

where m' refers to the chosen number of observations. Choosing m' = 2 500 gives following results in *Koillis-Suomi:*

Variable	Actual K	z_0 Actual	z_0 m' = 2 500
Land use class	0,67	36,4	34,3
Sub-class	0,81	40,3	38,5
Treatment class	0,36	24,1	29,3
Drainage status	0,66	37,5	37,8
Dom. tree species	0,50	19,0	20,6

The transformation does not change the order of the variables in respect to the z_0-values. It is noticeable that even with an equal number of observations, the accuracy of the assessment of different variables from photos cannot be compared on the basis of the K-values only. The K-value of treatment class is distinctly smaller than that of dominant tree species. However, the former K-value is statistically more significant than the latter one. This can happen when the classifications used and the distributions into the classes are distinctly different.

Comparison of the agreement meets with problems also within one and the same variable because the number of the observations is not equal by sub-area. Therefore as above, a calibration of the z_0-values is needed to compare the sub-areas with each other. Following example concerns the accuracy of the assessment of land use class with m' = 484:

Variable	Actual K	z_0 Actual	z_0 m' = 484
Kemijärvi	0,64	12,6	14,3
Kuusamo	0,69	16,2	15,6
Pelkosenniemi	0,61	10,0	14,5
Posio	0,60	11,4	13,1
Salla	0,67	17,5	15,4
Savukoski	0,74	19,2	16,0
Koillis-Suomi	0,67	36,4	15,1

m' was chosen to represent a medium-size sub-area with the exception that $\sqrt{m'}$ is an even number. The calibrated z_0-values are fairly close to each other which is an important result in itself. Assuming a nearly similar photo material and a similar distribution of the photo plots into classes, the result is an evidence of a consistent assessment of the land use class. The minimum and maximum values of the calibrated z_0-values by variable are (the number of the sub-area in parentheses):

Variable	m'	Minimum	Maximum
Land use class	484	13,1 (4)	16,0 (6)
Sub-class	441	15,3 (6)	16,5 (5)
Treatment class	289	5,0 (3)	12,4 (5)
Drainage status	400	13,5 (4)	17,7 (3)
Dom. tree species	361	4,7 (3)	9,7 (5)

The numbering of the sub-areas is the same as in Table 10. The results are consistent in the sense that the difference between the minimum and maximum values increases with decreasing photo-interpretation accuracy. It appears that the minima center in sub-areas 3 and 4 and the maxima in sub-area 5. Evidently, accuracy of the assessment of nominal scale variables is from photos lowest in *Pelkosenniemi* and *Posio* and highest in *Salla*. In this context it can be discussed only superficially to what degree the differences in the agreement are due to photo materials, interpreters and area-specific features.

Photo material is the same for the whole study area. On the other hand, the age of the photo material varies considerably. The time interval between the photo exposure and the field measurements ranges from 4 years to 14 years (Table 11). The average time intervals have been calculated on the basis of the whole photo sample weighted by the number of the relascope plots. The time interval between the photo exposure and the field measurement fails to explain the observed differences in the agreement of the photo-interpretation. In fact, the results are in this respect somewhat contradictory. One can conclude that the age of the photo material is not a very important factor in the assessment of the nominal scale variables. The photo material used is at a scale of 1 : 50 000 and 78 % of it is 5—10 years old. The outcome can be different with larger scales and/or older photographs.

Three interpreters were engaged in the interpretation of the photo sample of *Koillis-Suomi* in 1982 and 1983. The agreement expressed by the calibrated z_0-values by interpreter and variable is:

Variable	m'	Interpreter 1	2	3
Land use class	900	18,3	22,4	20,1
Sub-class	900	19,8	24,4	24,1
Treatment class	576	15,3	19,1	19,0
Drainage status	784	18,1	24,4	26,1
Dom. tree species	676	15,4	10,0	8,7

There are noticeable differences in the

Table 11. The time interval between the photo exposure and the field work by sub-area.
Taulukko 11. Ilmakuvauksen ja maastotöiden välinen aikaero osa-alueittain.

Interval *Väli*	Sub-area[1]) — *Osa-alue*[1]) 1	2	3	4	5	6	Koillis-Suomi
	Proportion of the plots[2]) — *Osuus koealoista*[2])						
Years — *Vuotta*	%						
0[3])		4			3	1	2
4	8		97		14	30	18
5		25					5
6	92	71	3	52	50	1	46
7				48			6
9					26	62	20
10[4])					3	2	1
12					3	4	2
14					1		
Total	100	100	100	100	100	100	100
Mean — *Keskiarvo*	Years — *Vuotta* 5,8	5,5	4,1	6,5	6,7	7,6	6,3

[1]) As in Table 10 — *Kuten taulukossa 10*
[2]) Including waters — *Mukaanlukien vedet*
[3]) Aerial photographs not available. The necessary data has been gathered in the field — *Ei ilmakuvia. Tarvittavat tiedot kerätty maastossa.*
[4]) As above. The necessary data has been gathered from forest maps the average age of which is 10 years — *Kuten yllä. Tarvittavat tiedot kerätty metsätalouskartoilta, joiden ikä on keskimäärin 10 vuotta.*

agreement between the three materials. However, it must be stressed that the outcome is not solely dependent on the skill of the interpreter. Differencies of the photo material (age of photos) and area-specific features also affect the result.

The results discussed above concern the correctness of the photo-interpretation by variable, sub-area and interpreter. Testing the consistency of the photo-interpretation by interpreter or by photo material requires an equal number of observations and the same distribution of the plots into the classes. These qualifications are perfectly fulfilled only by using the same set of units for interpretation. Congalton and Mead (1893) present some possible applications of consistency testing by the K-statistic.

The photo sample of *Koillis-Suomi* does not include material for valid consistency testing. The interpretation was made for practical inventory work without taking into account study aspects of this kind. However, for the sake of demonstration, the differences of the observed K-values in the materials of the three interpreters are briefly analyzed. The variances of the K-values are in this case estimated by formula 22 and the values of the test quantity by formula 24. The relevant data concerning interpretation of land use class are:

Characteristic	Interpreter 1	Interpreter 2	Interpreter 3
m	947	1 292	530
K	0,63496	0,69534	0,66097
s^2_K	0,000662	0,000368	0,0010465
s^2_K (900)	0,000697	0,000528	0,000616
z	1,725 (1–2)	1,016 (2–3)	
		0,718 (1–3)	

The variances of the K-values have been calibrated to correspond 900 observations by formula

$$s_K^2(m') = \frac{m}{m'} * \frac{p_a(1-p_a)}{(1-p_c)^2}$$

It appears that the K-values are significantly different in materials 1 and 2 and in materials 2 and 3. Let it be repeated that the differences are not only due to the interpreter. The z-values in the different cases are:

Variable	m'	Comparison 1—2	1—3	2—3
Land use class	900	1,725	0,718	1,016
Sub-class	900	1,548	1,282	0,281
Treatment class	576	1,421	1,197	0,233
Drainage status	784	0,839	1,756	0,996
Dom. tree species	676	3,119	3,778	0,696

The comparison reveals that the agreement of the assessment of the nominal scale variables from photos is fairly similar in materials 2 and 3.

It can be stated that the assessment from photos of the nominal scale variables in question is beneficial from the standpoint of the inventory goals. There are, however, differences between the variables in this respect. The total usability of the variables depends both on the accuracy of the assessments and on their importance. The latter factor cannot be quantified in objective terms. In fact, the importance of the variables is totally dependent on the goals of the inventory. The following approximation proceeds from the standpoint of practical forestry:

Variable	Photointerpretation accuracy	Importance of the variable
Land use class	High	High
Sub-class	High	Medium
Treatment class	Medium	High
Drainage status	High	Low
Dom. tree species	Low	Medium

Accordingly, land use class is of the highest usability and dominant tree species is the lowest. This kind of ranking is necessary when combining the photo and the field samples.

422. Volume

Estimation of the volume in the field is partly based on a determination of the basal area per hectare by the angle-count method. The basal area factor used in the study area is 2. Consequently, each tallied tree corresponds to 2 sq.m. of basal area per hectare. Basal area is also used as an auxiliary variable in the photo-interpretation of volume. The other stand variable needed in the volume assessment is mean height of the growing stock according to the principle of the stand volume tables presented by Nyyssönen (1954).

Crown coverage is a measure of stand density well suited to photo-interpretation. For the Finnish conditions, Nyyssönen (1955) developed an aerial stand volume table based on crown coverage and mean height. It has not, however, been considered for use in Lapland. The main reason being the very different scales of photographs used (1 : 10 000 instead of 1 : 50 000).

The assessment of volume is made on land use classes forest land and scrub land. On forest land, volume affects determination of treatment class in the sense that the treatment classes have certain minimum and/or maximum volume limits as follows (the treatment classes have been described in p. 16):

Treatment class	Volume, 10 m³/ha	Approximative age limits, years
0	0—1	0—30
1	0—3	30—
2	1—8	30—90
3	7—	70—
4	4—8	100—

Age has been added to characterize the treatment classes in more detail. It is of course not assessed from aerial photographs. Besides volume and age, treatment class depends on the structure of the growing stock. The appearance of the crown canopy is a valuable key in treatment class interpretation. For the sake of interest, the correlation between the photo volumes and the field volumes deserves to be studied by land use class and treatment class. A few concepts must be first defined to clarify the discussion.

The *photo stratum* is determined solely on the basis of the interpretation. The *pure stratum* includes only those plots which belong to the stratum also in the field. Concerning treatment classes which are determined only on forest land, a condition can be set that the actual land use class must be forest land but that the actual treatment class need not be the same as the photo stratum. Let this concept be called the *conditioned stratum* of treatment class x, where x refers to the treatment class assessed from photos. The last defined case can occur only within the pure stratum of forest land.

The correlations between the volumes assessed from photos and the field volumes in the different cases in *Kuusamo* is as follows (m = number of field points):

Stratum	Photo		Conditioned		Pure	
	m	r	m	r	m	r
0,1	49	0,576	42	0,547	20	—0,110
2	118	0,623	112	0,633	73	0,585
3	133	0,334	123	0,322	88	0,579
4	46	—0,119	42	—0,152	6	—0,913
0—4[1])	346	0,598	319	0,615	187	0,818
5[2])	106	0,494	65	0,585	65	0,585
Total	452	0,633	384	0,657	252	0,848
			(430	0,633)		

[1]) Forest land
[2]) Scrub land

3 454 points fall into the total photo stratum of forest land plus scrub land in *Kuusamo.* 452 points of these have been measured in the field. 22 points from the field sample were proved not to belong to the two land use classes in question. The remaining 430 points (= pure stratum of forest or scrub land) further include 46 st. wrong land use class interpretations. Thus the pure strata of forest land and scrub land together include 384 field points with a correlation of 0,657 between the volumes assessed from photos and the field volumes. The correlation increases to 0,848 after the points with a mis-interpretation of treatment class (132 st.) have been excluded.

The results reveal that determinations land use class and volume have only a slight connection. The exclusion of the points with a misinterpreted land use class does not significantly affect the correlation. Treatment class and volume, on the other hand, are fairly closely connected. This is to be expected because the treatment classes are partly defined on the basis of volume.

The pure stratum gives the highest correlation at the land use class level but not necessarily at treatment class level. There exists a significant positive correlation only in the closed stands (treatment classes 2 and 3). The high negative correlation in the pure stratum of shelterwood stands (4) may be an accident. The six points in question are probably not a representative sample from the stratum.

The negative correlations can also be explained by cuttings. Seed tree stands (1) and shelterwood stands (4) are established for natural regeneration. In general, all standards are removed in one cutting after the stand has been restocked by seedlings. In the case of sound silviculture, this regeneration stage is not very long. For economical reasons, removing the standards is more likely to take place in stands with a proper outturn. So a negative correlation can arise if the aerial photographs are many years old at the moment of the field measurement. The photos of *Kuusamo* were taken in 1976—1977 and the field work was done in 1982. The time interval is long enough for cuttings to be an explanation of the negative correlations.

It can be concluded that the relascope principle used in the interpretation of volume is not very satisfactory. This is because the estimation of the basal area of one point is not accurate from small-scale aerial photographs. In addition, an exact location of a photo point in the field is difficult. In sparse and heterogenous stands a deviation of some meters can lead to a significantly different estimate of the basal area. The variation increases with increasing basal area factor. Several observations are generally needed to estimate the stand basal area with an appropriate accuracy even in the field (Vuokila 1959, p. 22).

Young closed stands (treatment class 2) are relatively homogenous compared with the other treatment classes. The results are consistent, in that the correlation between the volumes assessed from photos and the field volumes is best in that treatment class. In old closed stands, the correlation is significantly greater in the pure stratum than in the conditioned stratum. The age of the photos can again serve as an explanation. About 1/6 of the points in the conditioned stratum of old closed stands has been proved to belong to treatment classes 0 and 1 as can be seen from the following similarity matrix:

Photo stratum	Field stratum 0,1	2	3	4	Total
0,1	20	18	1	3	42
2	16	73	20	3	112
3	20	10	88	5	123
4	6	11	19	6	42
Total	62	112	128	17	319

Concerning old closed stands, it is quite evident that the major part of the disagreements of this kind are due to the cuttings. Similar errors are common in the photo stratum of young closed stands, too. The effect on the correlation, however, is the opposite. The conclusion can therefore be drawn that in young closed stands errors in treatment class interpretation are mainly "real" and not caused by cuttings. It is often difficult to distinguish between treatment classes 0,1 and 2 even in the field, still more frequently from smallscale aerial photographs.

The dependence of the actual volume (y) on the assessed volume (x) by photo stratum is visualized in Figure 13. The linear regression lines in question and the corresponding standard deviations of y are as follows:

Photo stratum[1])	Equation	s_y m³/ha	s_{yx} m³/ha
1	y = 0,64x + 10,7	52,5	40,7
2	y = 0,64x + 11,4	55,4	44,5
21	y = 1,23x + 0,9	20,2	16,1
22	y = 1,08x + 0,4	39,4	31,0
23	y = 0,71x + 1,2	60,3	57,1
24	y = 0,41x + 78,1	47,0	47,0
3	y = 1,57x + 1,0	26,9	23,5

[1]) As in Fig. 13.

The latter standard deviations of y have been calculated on the basis of the measured volumes from the regression line. A very

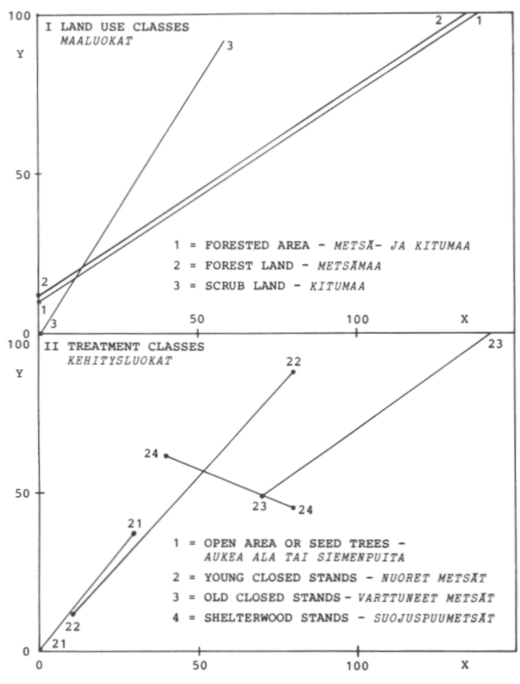


Figure 13. Dependence of the field observed volume (y) on the photointerpreted volume (x) in *Kuusamo* by photo stratum.

Kuva 13. Maastossa mitatun kuutiomäärän (y) riippuvuus ilmakuvalta tulkitusta kuutiomäärästä (x) Kuusamossa eri kuvaositteissa.

logical trend can be perceived in Fig. 13. The gradient of the regression lines decreases with increasing probability of heavy cuttings. The drain tends to be greater in heavily stocked stands which weights down the right ends of the regression lines.

The data concerning the regression of y on x by photo stratum and sub-area is presented in Table 12. The strata have been defined in a slightly different way for the table than was done above. However, this does not change the essential features in the material of *Kuusamo*. The conclusions drawn above are also valid in the other sub-areas. The correlation between the volumes assessed from photos and the field volumes is poor in the photo strata old closed stands and shelterwood stands. Besides cuttings, erroneous interpretations and locations explain the result. The forests in question are generally irregular with openings and clusters in the tree stock. This causes a great variation of the volumes measured in the relascope plots.

The difference between the two last characteristics given in Table 12 reveals the goodness of fit of the estimated linear regression lines in the material. The ratio between $s_{\bar{y}x}$ % and $s_{\bar{y}}$ % calculated as 100 * ($s_{\bar{y}x}$ %/$s_{\bar{y}}$ %) by photo stratum and sub-area can be seen in Table 13. The assessment of volume is of no use in shelterwood stands, nor in old closed stands. This is an important result considering the decisions to be made in combining the field and the photo samples.

43. Stratification

The division of the photo sample into strata affects the statistical precision of the estimates, as Frayer (1978), for example, has demonstrated. The best solution depends on population parameters, sample size and accuracy in the photo-interpretation. A detailed stratification presupposes a great sample and a good correlation between the data measured in the field and that assessed from the interpretation of aerial photos. The optimum number of the strata increases with increasing heterogeneity of the population.

Different characteristics usually have different optimum stratifications. In principle, it is possible to use a separate stratification for each of the main characteristics. This is, however, an uphill task and the resulting estimates are not compatible with each others. The latter may be a considerable annoyance when the results are applied to practical forestry. An implicit solution is to use a single stratification which fulfils the minimum precision requirement for each of the main characteristics.

Mathematical programming can be applied to the planning of the sample in multivariate forest inventories; an example is finding the optimum allocation of the sample into pre-fixed strata (Marshall and Nautiyal 1980, Mitchell and Bare 1981). The approach also suits the SPR technique with or without stratification (Kilpatrick 1981, Omule and Williams 1982). In this investigation, however, the problem is a slightly different. The problem, in fact, concerns post-stratification with a pre-fixed allocation of the samples.

The effect of the stratification on sampling error in two-phase sampling can be studied by simulation. For study purposes a computer program was prepared which permits the user define 1—30 different photo strata. The variances are estimated by formulae 17 and 18 (p. 37). The simulation concerns the eight main variables defined in chapter 3 (p. 21). The results are reduced to a more perceivable form by transformation

$$s_e\,\% = 100 * \frac{\sqrt{v(e)}}{e}$$

where e = estimated value of p or $\bar{x}$
v(e) = estimated variance of p or $\bar{x}$

The characteristic is calculated as relative standard error. However, it must not be confused with the sampling error. The reason will be discussed later in this paper.

The variables used to define the photo strata are called here *stratum variables*. All variables assessed in the photo-interpretation can be used as stratum variables but the most important ones are land use class, sub-class, treatment class, drainage situation and volume. The *best stratum variable* for a characteristic is the one which lowers the estimated variance most of all. The best stratum variables can be found by dividing the photo sample into two strata in different ways. The procedure is useful in the estimation of proportions. The following demonstration concerns area categories *mineral soils* (1), *forested land* (2), *forest land* (3) and *mature stands* (4). The two strata are de-

Table 12. Dependence of the field volume on the assessed volume by photo stratum and sub-area.

Taulukko 12. Maastossa mitatun kuutiomäärän riippuvuus ilmakuvalta tulkitusta kuutiomäärästä eri ilmakuvaositteissa ja osa-alueissa.

Characteristic[1]) *Tunnusluku*[1])	Photo stratum[2]) — *Ilmakuvaosite*[2]) 0,1	2	3	4	0–4	5	0–5
				Kemijärvi			
m	24	146	49	22	241	97	338
r	0,347	0,581	0,116	0,357	0,520	0,098	0,567
a	1,060	0,638	5,140	—1,769	2,241	1,091	1,632
b	1,325	1,400	0,279	1,422	0,588	0,316	0,652
$S_{\bar{y}}$ %	32,9	7,9	10,0	16,0	6,3	16,8	6,3
$S_{\bar{y}x}$ %	33,0	6,5	10,3	16,0	5,4	17,0	5,2
				Kuusamo[3])			
	49	118	133	46	346	106	452
	0,569	0,620	0,324	—0,167	0,599	0,464	0,630
»	—8,922	—1,642	0,261	8,683	1,166	9,446	1,114
	1,201	1,097	0,684	—0,557	0,639	1,626	0,647
	31,7	9,9	5,9	12,2	5,4	15,0	5,3
	26,9	7,8	5,7	12,4	4,3	13,5	4,1
				Pelkosenniemi			
	34	58	26	18	136	56	192
	0,463	0,285	—0,001	0,136	0,477	0,447	0,516
»	0,866	2,086	8,276	4,872	2,063	0,455	1,720
	1,566	0,547	—41,266	0,346	0,576	1,322	0,619
	21,1	15,6	13,6	10,9	8,6	16,2	8,1
	19,6	15,3	14,5	11,8	7,9	14,9	7,0
				Posio			
	15[3])	102	85	35	237	79	316
	—	0,688	0,191	0,386	0,641	0,607	0,679
»	—	—0,161	5,808	0,127	1,187	0,372	1,103
	—	0,120	0,377	1,057	0,760	1,343	0,771
	—	11,5	7,0	12,0	6,6	13,4	6,4
	—	8,5	7,0	11,6	5,1	10,8	4,7
				Salla			
	76	141	106	72	395	126	521
	0,721	0,472	0,257	0,243	0,643	0,370	0,660
»	0,131	0,536	4,090	2,702	1,530	0,737	1,360
	1,586	1,119	0,428	0,750	0,682	0,970	0,700
	21,8	9,8	6,1	6,0	5,2	10,2	4,9
	15,4	8,7	5,3	6,0	4,0	9,6	3,7
				Savukoski			
	74	150	131	112	467	129	596
	0,465	0,529	0,045	0,151	0,484	0,519	0,533
»	0,334	0,686	6,977	3,539	2,319	—8,458	1,887
	1,180	1,020	—6,847	0,544	0,510	1,364	0,558
	16,0	8,6	5,6	6,0	4,2	10,6	4,0
	14,5	7,3	5,6	6,0	3,6	9,1	3,4
				Koillis-Suomi			
	272	715	530	305	1822	593	2415
	0,546	0,532	0,198	0,138	0,575	0,437	0,610
»	0,373	0,716	4,464	3,817	1,760	0,404	1,476
	1,338	1,022	0,367	0,447	0,622	1,214	0,651
	10,0	4,1	2,9	3,6	2,3	5,4	2,2
	8,4	3,5	2,8	3,6	1,9	4,9	1,8

[1]) m = Number of observations — *Havaintojen lukumäärä*
r = Coefficient of correlation — *Korrelaatiokerroin*
a and b = Coefficient of the regression equation — *Regressioyhtälön kertoimet*

$$S_{\bar{y}}\ \% = \frac{100}{\bar{y}} * \sqrt{\frac{(y_i - \hat{y}_i)^2}{m\,(m-1)}} \text{ and/ja } s_{\bar{y}x}\ \% = \frac{100}{\bar{y}} * \sqrt{\frac{(y_i - \bar{y})^2}{m\,(m-2)}}$$

[2]) 0,1 = Open area, stand of small seedlings (height ≤ 3 m) and seed tree stand — *Aukea ala, taimikko (puiden keskipituus ≤ 3 m) ja siemenpuuala*
2 = Young closed stand — *Nuori sulkeutunut metsä*
3 = Old closed stand — *Varttunut sulkeutunut metsä*
4 = Shelterwood stand — *Suojuspuuasento*
5 = Scrub land — *Kitumaa*

[3]) The data is not quite compatible with that presented in the text and in Fig. 13. — *Luvut eivät ole aivan yhtäpitäviä tekstin ja kuvan 13 kanssa.*

[4]) All volumes equal to null — *Kaikki tilavuudet ovat nollia.*

Table 13. The ratio between the relative standard error of the mean volume estimated with regression and without regression by photo stratum and sub-area.
Taulukko 13. Keskitilavuuden regressioestimaatin ja ilman regressiota lasketun estimaatin suhteellisen keskivirheen välinen suhde eri ilmakuvaositteissa osa-alueittain.

Sub-area *Osa-alue*	Photo-stratum[1]) — *Kuvaosite*[1]) 0,1	2	3	4	0—4	5	0—5
				%			
Kemijärvi	100	82	103	100	86	101	83
Kuusamo	85	79	97	102	80	90	77
Pelkosenniemi	93	98	107	108	92	92	86
Posio	—	74	100	97	77	81	73
Salla	71	89	87	100	77	94	76
Savukoski	91	85	100	100	86	86	85
Koillis-Suomi	84	85	97	100	83	91	82

[1]) As in Table 12 — *Kuten taulukossa 12*

fined as

1. Photo points falling in the area category in question
2. All other photo points on land

The four divisions into two strata are applied to each area category. In addition, the alternative with one stratum (without stratification) is used. The simulation gives following s_e %-values in *Kuusamo:*

Stratification	Area category 1	2	3	4
1	*2,70*	1,60	2,39	9,26
2	3,84	*1,41*	2,70	9,66
3	3,14	1,61	*2,18*	9,42
4	2,96	1,65	2,55	*8,46*
Without stratification	4,16	1,83	3,08	9,80

In the case that stratification is not done the variances are calculated as

$$v(p) = \frac{p * q}{m - 1}$$

where m = total number of field points
p = estimated proportion
and q = 1 — p

Comparison of the s_e %-values with the lowest row reveals that the four stratifications always result in a smaller variance. Each area category has the lowest variance when the corresponding photo stratum has been used in the stratification. The result is consequent considering the fairly good agreement in the photointerpretation of the nominal scale variables in question.

The *minimum stratification* for a characteristic comprises a few strata which cannot be combined without a significant increase of the variance. In other words, the strata in question at least are needed to realize the major part of the potential gain from the photo sample. One can conclude that the best stratum variable and the minimum stratification always belong together. The minimum stratification need not be the same as the optimum one. On the other hand, the minimum stratification is generally more detailed than the best division into two strata.

The results above indicate that the combined minimum stratification for the four characteristics in question is a mixture of the best stratum pairs. So, at least land use class, subclass and treatment class must be used as stratum variables. The following simulation of the variance proceeds stepwise by increasing the number of the strata. The stratifications used are

0 = Without stratification
1 = The best division into two strata
2 = Combination of the best stratum pairs
3 = A full stratification on the basis of the interpreted land use class, sub-class and treatment class
4 = As above including drainage situation
5 = As above including division into coarse volume classes
6 = An "overdone" stratification

The variances of the 0-alternative are denoted as 100 and all other variances are related to that level. The resulting relative values are as follows:

Case	Number of strata	Area category 1	2	3	4
0	1	100	100	100	100
1	2	42	59	50	75
2	6	39	70	50	72
3	8	39	57	48	72
4	14	39	60	46	72
5	19	37	58	43	70
6	27	38	56	43	68

An important experience from the simulation is that after a certain basic stratification with 2—8 strata, the variances do not change significantly with increasing number of strata. There is no obvious optimum stratification on the range of 2—30 strata. We can, however, suppose that the estimated variances begin to increase when the average number of field plots per photo stratum becomes too small, say less than 15. There are 523 field points in the sample of *Kuusamo,* so the "critical" number of strata would be some 35—40.

Local minima and maxima are possible which becomes apparent with the second area category (proportion of forested area). In this case the best division into two strata serves as the minimum stratification because the result does not improve to a degree worth mentioning with the more detailed stratifications. In multivariate inventories, however, it seems favourable to use a relatively detailed stratification to avoid significant local maxima.

The estimated proportions and corresponding s_e %-values with the different stratifications are presented in Table 14. The s_e %-values of the estimated proportion of mature stands are fairly large compared with the other proportions. This is partly due to the lowest actual occurrence of this area category. The classification used in the photo-interpretation also contributes to the outcome. The pure photo stratum (misinterpretations excluded) of *old closed stands* includes both thinning stands in advanced stage and mature stands. In addition, erroneous interpretations are common in the photo stratum in question (p. 45).

The simulation also concerns the mean volumes of the growing stock on land (5) and on forest land (6) as well as the mean volumes of Scots pine (7) and saw-timber (8) on land. The different stratifications used are the same as above with the proportions. The alternative with two strata (1) has been divided into four cases:

11 = Best stratification for mineral soils
12 = „ „ forested land
13 = „ „ forest land
14 = „ „ mature stands

The relative variances for the mean volumes in the different cases are:

Case	Mean volume 5	6	7	8
0	100	100	100	100
11	86	112	94	88
12	93	101	96	97
13	86	115	93	92
14	77	91	91	77
2	74	96	91	75
3	73	94	89	75
4	72	93	89	75
5	62	81	82	63
6	62	82	82	63

Table 14. The estimated proportions (p) of some area categories and the corresponding s_e %-values calculated by the variance estimator of simple random sampling (case 0, see p. 48) or two phase sampling for stratification (case 1—6) in *Kuusamo.*

Taulukko 14. Eräiden pinta-alaositteiden arvioidut osuudet ja vastaavat s_e %-arvot Kuusamossa. s_e %-arvot on laskettu kuten suhteelliset keskivirheet yksinkertaisen satunnaisotannan varianssikaavaa (tapaus 0, ks. s. 48) tai kaksivaiheisen luokitetun otannan varianssikaavaa (tapaukset 1—6) käyttäen.

Case *Tapaus*	Area category[1]) — *Osite*[1]) 1	2	3	4
p %	52,58	85,09	66,92	16,63
0 s_e %	4,16	1,83	3,08	9,80
1 „	54,91	83,93	67,01	17,69
	2,70	1,41	2,18	8,46
2 „	54,65	84,30	66,96	17,60
	2,61	1,53	2,18	8,30
3 „	54,61	83,85	66,76	17,58
	2,61	1,38	2,14	8,30
4 „	54,51	83,74	66,64	17,59
	2,61	1,42	2,10	8,29
5 „	54,81	83,88	67,03	17,27
	2,51	1,39	2,02	8,17
6 „	54,85	84,02	67,14	17,17
	2,57	1,37	2,02	8,11

[1]) 1 = Mineral soils — *Kangasmaat*
2 = Forested area — *Metsäala*
3 = Forest land — *Metsämaa*
4 = Mature stands — *Uudistuskypsät metsät*

4 408500946V

Concerning the 0-alternative, the variances for the characteristics 5,7 and 8 (the mean volumes on land) are calculated as

$$v(\bar{x}) = \frac{\sum_{i=1}^{m} (x_i - \bar{x})^2}{m(m-1)}$$

where m = total number of the plots
x_i = volume of plot i
$\bar{x}$ = mean volume

As to the mean volume on forest land, the corresponding variance estimator reads

$$v(\bar{x}_p) = \frac{\sum_{i=1}^{m_p} (x_{pi} - \bar{x}_p)^2}{m_p(m_p-1)}$$

where m_p = number of plots on forest land
x_{pi} = volume of plot i forest land
$\bar{x}_p$ = mean volume forest land

With a detailed stratification it can happen that the number of the observations on forest land in a photo stratum is 1. Then the variance within the photo stratum in question is set equal to the observed volume in the simulation. This was done in three photo strata in the case of the "overdone" stratification. The effect of this operation on the result is small because the joint weight of the three strata is 0,084.

The variances of the mean volumes decrease with increasing number of strata. In this sense the result is the same as with the proportions. The pattern of the decrease, however, is different (Fig. 14). The divisions into two strata are not so effective as with the proportions. This is to be expected because volume is not directly used as stratum variable in these alternatives. From the alternatives with only two strata, the division into mature stands and other land area gives the smallest variances. The reason is that volume is used in the definition of the treatment classes. Volume is directly used as a stratum variable in the alternatives 5 and 6. As can be expected, direct inclusion of volume as stratum variable further lowers the variances.

An incorrect stratification can make the variance greater than in the case of no stratification. This becomes apparent with the results concerning the mean volume on forest land (characteristic 6). Here again is

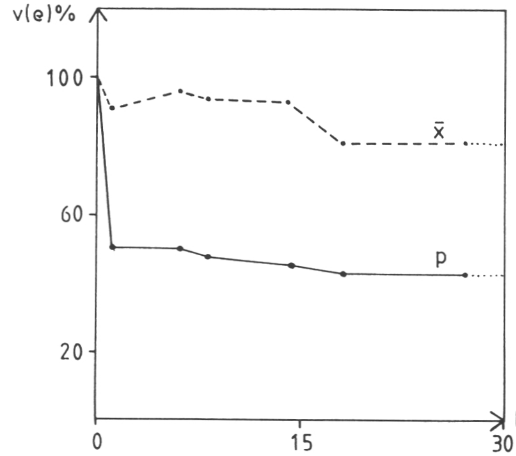


Figure 14. Relative variance of the mean volume on land ($\bar{x}$) and the proportion of forest land (p) as a function of the number of the super cells. The variances have been calculated as in double sampling for stratification.

Kuva 14. Puuston keskitilavuuden ($\bar{x}$) ja metsämaan osuuden (p) suhteelliset varianssit kuvaositteiden lukumäärän funktiona. Varianssit on laskettu kaksivaiheisen ilmakuva- ja maasto-otannan kaavoilla.

a justification for using detailed stratifications in multivariable inventories.

The potential gain from the stratification can be studied through regression analysis. For that purpose, $\bar{x}$ and $\bar{x}_p$ are replaced by regression estimates $\hat{x}_i$ and $\hat{x}_{pi}$ in the variance estimators. The other difference is that the numbers of the degrees of freedom are m-2 and m_p-2. Let the variances calculated without stratification and regression be denoted as 100. The outcome then is as follows:

Way of estimation	Characteristic $\bar{x}$ s_e %	$\bar{x}$ v(e)	$\bar{x}_p$ s_e %	$\bar{x}_p$ v(e)
Field sample[1])	5,54	100	4,65	100
Stratification[2])	4,36	62	4,22	81
Regression	4,16	56	4,22	82

[1]) Without stratification and regression
[2]) Alternative 5

It seems that the potential gain from the stratification has been attained with respect to the characteristic mean volume on forest land. However, this is not necessarily the case because the regression estimates have been calculated assuming linear regression.

We can conclude that the "potential" level defined above can be exceeded if the linear solution is a poor fit.

It was stressed in connection with the definition of characteristic s_e % that it is not the same as sampling error. The reason is that both the photo sample and the field sample have been concentrated to a considerable degree. The difference becomes apparent when the s_e %-values in the first row above (field sample) are compared with the relative standard errors estimated from the field sample of *Kuusamo* (p. 28). The sampling error of $\bar{x}$ is 11—12 % instead of 5,54 % above. The corresponding figures concerning p (proportion of forest land) are 5—6 % and 2,02 %. It seems that concentration of the samples affects more the precision of p than $\bar{x}$. The explanation is that the field points are more correlated with respect to p than $\bar{x}$.

Despite the concentration of the samples, the values of s_e % and v(e), calculated above, are informative in comparison with the different stratifications. One must keep in mind, however, that the simulation concerns only one factor affecting the sampling error. It will be proved later in this investigation that, after the extension of the field data to the photo sample, the variation of the tract values is also an important element of the sampling error. Stratification also affects the variation of the tract values as can be seen from the figures calculated from the material of private forests of *Kuusamo* (Table 15).

The CV %-values have been calculated weighting by the tract size (number of points on land or on forest land, see p.   ). The variation of the tract values increases only slightly with an increasing number of strata. The increase is almost insignificant with p. The reason may be that in groups 3 and 4 the strata have been increased mainly by volume and treatment class. The experimental result above further supports the idea that there exists no distinct optimum stratification, especially in a multivariate forest inventory.

Table 15. Coefficient of variation (CV%) of the tract values of mean volume on land ($\bar{x}$), as well as on forest land ($\bar{x}_p$) and proportion of forest land (p) with different stratifications in Kuusamo.

Taulukko 15. Puuston keskitilavuuden ($\bar{x}$) ja metsämaan puuston keskitilavuuden ($\bar{x}_p$) sekä metsämaan osuuden (p) lohkoarvojen vaihtelukertoimet eri luokituksilla Kuusamossa.

Number of strata *Ositteiden lukumäärä*	CV % $\bar{x}$	 $\bar{x}_p$	 p
	Group 1 — *Ryhmä 1*		
16	32,6	29,9	14,9
17	31,9	29,9	14,7
17	34,0	30,8	14,8
17	34,1	30,9	15,8
Mean of group 1 — *Ryhmän 1 keskiarvo*	33,2	30,4	15,1
	Group 2 — *Ryhmä 2*		
20	33,7	29,9	15,4
21	34,3	30,3	15,5
21	31,9	30,1	14,7
22	36,0	33,5	15,4
Mean of group 2 — *Ryhmän 2 keskiarvo*	34,0	31,0	15,3
	Group 3 — *Ryhmä 3*		
24	36,1	31,9	15,6
24	37,5	34,1	15,5
25	35,3	32,6	15,7
27	34,8	31,8	15,8
Mean of group 3 — *Ryhmän 3 keskiarvo*	35,9	32,6	15,7
	Group 4 — *Ryhmä 4*		
27	35,6	32,8	15,9
28	36,7	34,1	15,1
28	36,9	34,5	15,3
29	36,7	34,1	15,5
Mean of group 4 — *Ryhmän 4 keskiarvo*	36,5	33,9	15,5

44. Extension of the field data

441. The basis

The photo sample is first divided into homogeneous or nearly homogeneous photo strata here called *super cells*. The number of the super cells is 20—30 in a middle-sized sub-area. In the second stage, the super cells are divided into *cells* the number of which equals the number of field points by super cell. The cell size is nearly equal by super cell but varies between super cells. The division into cells is accomplished in such a way that each cell includes one and only one field point. After these arrangements, the field data of the field points are given to all photo points by cell.

Assuming a constant cell size by super cell, the field points within a super cell have an equal weight in the calculations. This is always the case when an extension of the field data is not made and the results are calculated according to the principles of double sampling for stratification. The weight of field point i in super cell h is

$$w_{hi} = \frac{w_h}{m_h}$$

where w_h = weight of super cell h
m_h = number of field points in super cell h

The estimate of the mean volume on land can be calculated as

$$\text{(26)} \qquad \bar{x} = \sum_{h=1}^{L} \sum_{i=1}^{m_h} \left(w_{hi} * x_{hi}\right) = \sum_{h=1}^{L} w_h \bar{x}_h$$

where x_{hi} = volume of field plot i in super cell h

In the case that the field data are extended to the whole photo sample by cell, an estimator for the mean volume reads

$$\text{(27)} \qquad \bar{x} = \frac{1}{n} * \sum_{h=1}^{L} \sum_{i=1}^{l_h} \sum_{j=1}^{n_{hi}} x_{hij}$$

where n = number of photo points
l_h = number of cells in super cell h
n_{hi} = number of photo points in super cell h in cell i
x_{hij} = extended or in the case of true field plot measured volume of point j in cell i in super cell h

Indeed, all photo points can be used in the calculations as if they had been measured in the field. If the cells are of equal size by super cell, then

$$\text{(28)} \qquad \bar{x} = \sum_{h=1}^{L} \cdot \left(\frac{n_h}{n} * \frac{1}{l_h} * \sum_{i=1}^{l_h} x_{hi}\right)$$

where n_h = number of photo points in super cell h

$$x_{hi} = \frac{1}{n_{hi}} * \sum_{j=1}^{n_{hi}} x_{hij} =$$ measured volume of field plot i in super cell h

and n_{hi} = number of photo points in super cell h in cell i

The number of the cells equals to the number of the field points. So $l_h = m_h$ and formula 28 can be rewritten as

$$\bar{x} = \sum_{h=1}^{L} \sum_{i=1}^{l_h} \left(\frac{n_h}{n} * \frac{1}{m_h} * x_{hi}\right)$$

But because $n_h/n = w_h$ and $w_h/m_h = w_{hi}$, formula 27 can finally be written in the same form as formula 26. Let it be repeated that this happens only when the cell size is constant by super cell.

The division of a super cell into cells of an exactly equal size is rarely possible. It can be demonstrated that the least variance of the cell sizes within a super cell is obtained by minimizing the maximum deviation from the constant average cell size. So, for example, with 50 photo points including six field points four cells of eight points (4 * 8) and two cells of nine points (2 * 9) should be selected instead of an alternative 5 * 8 and 1 * 10. The fact that the cells are not quite equalsized by super cell causes minor differences between the two estimates obtained with and without extension of the field data to the photo sample.

442. Practical realization

The goals of the new forest inventory in Lapland include the applicability of the main results at the commune (= sub-area) level. A necessary condition for avoiding bias is that the field samples and the photo samples are combined by sub-area. Another basic decision made concerns the photo samples interpreted by different persons. The interpreters were trained to be as consistent as possible. Special attention was paid to land use class and volume variables. It is therefore assumed that the interpretations can be combined for an extension of the field data to the photo sample. As with the measurement of the field sample, two or more persons have generally been engaged in the interpretation of the photo sample within each sub-area.

The division of the photo sample into super cells requires determining the stratum variables and their order in the strafication. Accuracy in the interpretation of photos and the area-specific features form the basis for the choice.

The fineness of the stratification depends mainly on the size of the sub-area. The number of the super cells used in *Koillis-Suomi* varies from 17 *(Pelkosenniemi)* to 38 *(Savukoski)*. The guiding principle in forming the super cells was to make them homogeneous in respect to the interpreted land use class, sub-class, drainage status, treatment class, dominant tree species and volume. The variables to be first neglected where necessary were dominant tree species and volume. In the largest sub-areas, however, even an overdone stratification was possible. Data concerning the number and the size of the super cells in the sub-areas are given in Table 16. The order of the sub-areas with respect to size is 3, 4, 1, 2, 5 and 6. The number of the super cells increases fairly regularly with increasing size of sub-area. At the same time, however, the size of the super cells also increases. So it cannot be stated with certainty whether the super cells in the small sub-areas are less homogenous than on an average.

There are a few super cells with only one field point. These are well discernible photo strata which are not readily combined with any other photo stratum. However, the combination must be made in the estimation of the statistical precision.

The results are fairly stable once a certain degree of fineness of the stratification is attained. This condition is usually reached with stratifications including all good photo variables as stratum variables (see p. 49). A simulation with different stratifications was tried in *Kuusamo* to see the effect on the main characteristics. Six groups of simulations are separated on the basis of the number of the super cells used. The estimated values of some main characteristics are presented in Table 17.

It appears that the extreme values are very often associated with a low or a high number of super cells. This result supports the use of a moderate stratification for the extension of the field data to the photo sample. The greatest deviations occur between the single estimates of the mean vol-

Table 16. The number of the photo strata here called super cells (s) and the minimum, maximum and average numbers of the photo points (n), as well as field points (m) per stratum by sub-area.

Taulukko 16. Ilmakuvaositteiden lukumäärä (s) ja kuvapisteiden (n) sekä maastopisteiden (m) pienin, suurin ja keskimääräinen lukumäärä kuvaositteissa osa-alueittain.

Characteristic *Tunnusluku*	Sub-area[1]) — *Osa-alue*[1]) 1	2	3	4	5	6
s	23	27	17	21	27	38
n minimum — *minimi*	22	12	29	21	28	14
,, maximum — *maksimi*	434	732	303	456	718	304
,, average — *keskiarvo*	148	179	98	139	180	145
m	2	2	1	3	3	1
,, ,,	67	102	35	54	80	36
,,	18	23	14	20	24	19

[1]) 1 = *Kemijärvi* 2 = *Kuusamo* 3 = *Pelkosenniemi*
4 = *Posio* 5 = *Salla* 6 = *Savukoski*

Table 17. The estimates of mean volume on land ($\bar{x}$) and on forest land ($\bar{x}_p$), as well as proportion of forest land (p) obtained in a simulation of the stratification in the sample of *Kuusamo*.

Taulukko 17. Kuusamon näytteellä tehdyn luokituskokeilun antamat estimaatit puuston keskitilavuudelle maalla ($\bar{x}$) ja metsämaalla ($\bar{x}_p$) sekä metsämaan osuudelle (p).

Simulation *Simulointi*	Super cells *Ositteita*	Characteristic — *Tunnusluku* $\bar{x}$	$\bar{x}_p$	p
		Group 1 — *Ryhmä 1*		
1	16	39,95	55,88	68,32
2	17	39,09	54,62	68,35
3	17	40,24	56,28	68,41
4	17	39,96	55,88	68,36
Mean of 1—4 — *Ryhmän 1 keskiarvo*		39,81	55,67	68,36
		Group 2 — *Ryhmä 2*		
5	18	40,38	56,42	68,52
6	18	38,94	54,35	68,46
7	19	38,85	54,24	68,38
8	19	39,89	55,78	68,34
Mean of 5—8 — *Ryhmän 2 keskiarvo*		39,52	55,20	68,43
		Group 3 — *Ryhmä 3*		
9	20	40,24	56,14	68,67
10	21	40,67	56,83	68,58
11	21	38,93	54,34	68,44
12	22	39,05	54,57	68,49
Mean of 9—12 — *Ryhmän 3 keskiarvo*		39,72	55,47	68,54
		Group 4 — *Ryhmä 4*		
13	24	40,54	56,65	68,52
14	24	40,02	55,80	68,67
15	24	39,55	55,16	68,36
16	25	39,25	54,75	68,57
Mean of 13—16 — *Ryhmän 4 keskiarvo*		39,84	55,59	68,53
		Group 5 — *Ryhmä 5*		
17	27	39,31	54,87	68,54
18	27	39,51	55,26	68,42
19	28	39,09	54,56	68,55
20	28	39,26	54,71	68,68
Mean of 17—20 — *Ryhmän 5 keskiarvo*		39,29	54,85	68,55
		Group 6 — *Ryhmä 6*		
21	29	38,91	54,30	68,58
22	29	39,52	55,11	68,64
Mean of 21—22 — *Ryhmän 6 keskiarvo*		39,22	54,71	68,61
		Groups 1—6 — *Ryhmät 1—6*		
Minimum single estimate *Pienin yksittäinen estimaatti*		38,85	54,24	68,32
Maximum single estimate *Suurin yksittäinen estimaatti*		40,67	56,83	68,68
Overall mean *Kokonaiskeskiarvo*		39,60	55,30	68,49
Minimum group mean *Pienin ryhmäkeskiarvo*		39,22	54,71	68,36
Maximum group mean *Suurin ryhmäkeskiarvo*		39,84	55,67	68,61

umes. However, assuming a relative standard error of about 6 % (p. 65), maximum differences do not exceed the sampling error. This is especially true with the proportions. This is particularly important because some poor stratifications have deliberately been included in the simulation.

The stratum variables and their emphasis used in the practical stratifications are (as to the classifications see p. 16):

Land use classes 1—3:	Land use classes 4—8:
1. Sub-class	1. Land use class

Mineral soils:	**Peatland:**
2. Land use class	2. Drainage status
3. Treatment class	3. Land use class
(41. Drainage status)	41. Volume
42. Volume	(42. Treatment class)

5. Dominant tree species

Land use class is the only stratum variable in the assessed land use classes 4—8. Drainage status is the strongest stratum variable on peatland. Volume outweighs treatment class on peatland because the assessment of both land use class and treatment class is uncertain on peatland. On forest land with mineral soils, drainage status is used instead of volume in treatment classes 0 and 1. In the other treatment classes on forest land mineral soils, however, drainage status is not used at all as a stratum variable.

The principle of the extension of the field data to the photo sample is simple but the practical realization of it is more complex. A more thorough testing, e.g. using simulation, is not possible without automatic data processing. A computer system was constructed to combine the photo and the field samples with as little human assistance as possible. The only auxiliary data the system needs comprises the number of the super cells and their sizes expressed as the number of the photo points within each super cell. Preliminary studies are required to obtain this data.

The computer system begins with a very detailed sorting of the photo sample. The sorting is made with respect to the selected stratum variables, the weights of which are different in the different parts of the photo sample as was described above. In this way, the accuracy of the photo-interpretation and the importance of the forest characteristics can best be taken into account.

It is precisely in the sorted photo file where the sizes of the super cells must be defined a priori. The computer system divides the file into cells of equal or nearly equal size by super cell. The number of the cells equals the number of the field points by super cell. The location of the field points is random in the original sorted photo file. So a mechanical division into cells produces cells with a varying number of field points within them. To avoid this, the field points are re-distributed in the sorted photo file by super cell before the division into cells. This is done in such a way that each cell obtains one and only one field point. This operation is justified assuming that the super cells are homogenous with respect to the interpreted data.

The groups of the grouping method (Poso 1972) are formed, in principle, in the same way as cells. The main difference is that the points to be measured in the field are selected from each group afterwards. So the only additional measure required due to the systematic field sample is the re-distribution of the field points within the super cells.

Grouping is accomplished using a grouping area of 600 clusters (cluster = two photo points 40 meters apart). To form homogenous super cells of sufficient size, the divisions into super cells and cells are by communes. In addition, the two points within a cluster are now separated from each other and they can fall into separate super cell and cell depending on the photo information. This is considered to increase the homogeneity of the super cells and cells.

The field points within the cells have their true field data in a separate field file. The field data are picked out from the field file and distributed to all photo points by cell. Finding the field data quickly requires that the field file is in a random access form. For that purpose, the "addresses" of the field records in the field file are in the computer's main memory during the whole operation.

The main output of the computer system is a file including all photo points furnished with formally complete field data. The points actually measured in the field have their own true field data. The file, as such, is applicable to the general calculation routines of the National Forest Inventory. The output also gives information which is required for the estimation of the precision of the results.

The whole operation is repeated until the potential advantages from the photo sample seem to be exhausted.

45. Statistical precision

451. Sampling the photo points

The photo sample comprises a systematic tract net which covers 2/3 of the study area (p. 15). Each tract of 8 km * 8 km includes 40 clusters of two relascope plots to be interpreted from aerial photographs. On the basis of the structure of the photo sample one can distinquish four sampling stages:

Stage	Sampling unit	Size	Sampling fraction
1	Tract	8 km * 8 km	2/3
2	Square	1 km * 1 km	20/64
3	Cluster	40 m[1])	~ 0
4	Plot	Variable	Variable

[1]) Two relascope plots 40 meters apart

Sampling fraction f_3 can be regarded to be about zero because the two clusters cover a very small portion of the square. On the other hand, sampling fraction f_4 can be quite high, depending on the dimensions of the growing stock within the clusters.

The sample design can also be considered as a two-stage sampling, with the tract as the primary unit and the square or cluster or relascope plot as the secondary one. In fact, even the primary unit need not necessarily be a tract. The main problem in all approaches is that the photo sample is systematic in every respect and, in addition to this, the samples form clusters in several stages. The formulae for estimating the sampling variances in a multi-stage sampling imply random sampling at all stages. However, the problem could be avoided if sampling were random in the first stage and f_1 approached zero (see Cochran 1977, p. 279 and 288). This not being the case, calculating the variances on the basis of the tract values only gives conservative estimates for the sampling error. Concerning the photo sample in Lapland, both the high value of f_1 and the systematic location of the photo tracts affects the result in the same direction.

A two-stage sampling design with the tract as the primary unit and the relascope plot as the secondary one was discussed in connection with the field sample (section 32, p. 26). It was then stated that clustering of the relascope plots within the tracts is likely to lead to an underestimation of the sampling error. For interest's sake, it is worth applying the one-stage and the two-stage approximations to the photo sample, as well. This can be done using the same calculation routines as with the field sample after the field data have been extended to the photo sample. However, the variance of the estimates calculated on the basis of the combined field and photo sample is not the final sampling error. This subject will be dealt with later in this paper. The following comparison concerns the mean volume on land ($\bar{x}$) and on forest land ($\bar{x}_p$) as well as proportion of forest land (p) by sub-area. Denoting the variances calculated on the basis of the tract values only (one-stage sampling) by 100 gives following relative variances calculated by two-stage sampling:

Sub-area	$\bar{x}$	$\bar{x}_p$	p
Pelkosenniemi	62	77	41
Posio	45	46	56
Kemijärvi	45	46	53
Kuusamo	46	47	59
Salla	46	43	44
Savukoski	43	42	51
Koillis-Suomi	45	44	48

The bipartite variance estimators (Formulae 8—10, p. 23) gives about half of the variance calculated on the basis of the variation of the tract values only (Formula 7). The difference in the relative standard error is 25—35 %. The result agrees well with the conclusions drawn on the basis of the field sample.

It appears that with the photo sample the bipartite variance estimator decreases more the variances of the mean volumes than the variance of the proportion of forest land. The explanation is obtained by comparing the different variation components by variable. For the sake of clarity, the quantities to be compared here are calculated as (cf. Formula 8):

$$B = 100 * \frac{\sqrt{s_1^2}}{e}$$

$$W = 100 * \frac{\sqrt{\frac{s_2^2}{\bar{m}}}}{e}$$

where e refers to the estimated value of $\bar{x}$, $\bar{x}_p$ or p and m is the mean number of relascope plots per tract on land ($\bar{x}$ and p) or on forest land ($\bar{x}_p$). B expresses the coefficient of variation of the tract values in percentage. The two approaches to the estimation of the variance lead to the same outcome only when the values of B and W are equal. The bipartite variance estimator generally gives smaller variance because the value of W tends to be smaller than that of B (Table 18). With the exception of *Pelkosenniemi,* the values of W/B are greater with p than with $\bar{x}$ and $\bar{x}_p$. Consequently, taking the variation within the tracts into account in the estimation of the sampling variance causes a smaller decrease in the estimate with p. It is not the absolute value of W in itself but the ratio of W to B that affects the final outcome.

The absolute values of W and B slightly decrease with increasing size of sub-area. The phenomenon is due to the tracts which partly fall outside the sub-area. Variability of the tract size increases with decreasing size of sub-area. At the same time the values of s_1^2 and s_2^2 tend to increase, although these are estimated weighting by tract size. In very small and homogeneous sub-areas the number of the degrees of freedom can also affect the result.

Coefficient of variation of the tract values (B) is a good measure for comparing the variability of the eight main variables. The value of CV in percentage by sub-area and variable is:

Sub-area	Variable[1]) 1	2	3	4	5	6	7	8
Pelkosenniemi	42	12	28	50	26	21	38	48
Posio	22	8	15	63	40	37	44	56
Kemijärvi	24	10	16	93	36	33	42	66
Kuusamo	20	10	15	52	37	33	40	54
Salla	28	11	21	56	35	36	42	50
Savukoski	18	10	16	50	34	32	38	47
Koillis-Suomi	26	10	17	64	35	33	47	56

[1]) See Table 4, footnote 1, p.

The variation becomes relatively greater with decreasing proportion or mean volume to be estimated. The coefficient of variation of the tract values is 15—25 % with p (variable 3), 25—35 % with $\bar{x}_p$ (6) and 30—40 % with $\bar{x}$ (5) in the photo sample. The corresponding variations in the field sample are 20—30 %, 40—55 % and 45—60 %. Consequently, the increase of the relascope plots from 28 to 80 per tract decreases the variation of the tract values by some 30—40 %. This is not very much considering that the relascope plots are spread more within photo tracts than within field tracts (see Fig. 4 and 7, p. 15 and 17).

The tract values of the mean volumes show a distinctly lower variation in *Pelkosenniemi* than in the other sub-areas. The case is just the opposite with the proportions, excluding the proportion of mature stands. It was earlier suspected that the true population vari-

Table 18. The values of B (CV between tracts) and W (CV within tracts divided by the mean number of the plots per tract), as well as the ratio between W and B by sub-area.
Taulukko 18. Lohkojen välinen suhteellinen vaihtelu B ja lohkojen sisäinen suhteellinen vaihteu jaettuna koealojen lohkoittaisella määrällä W sekä W:n ja B:n välinen suhde osa-alueittain.

Sub-area *Osa-alue*	Characteristic[1]) — *Tunnusluku*[1]) $\bar{x}$ W	$\bar{x}$ B	$\bar{x}$ W/B	$\bar{x}_p$ W	$\bar{x}_p$ B	$\bar{x}_p$ W/B	p W	p B	p W/B
Pelkosenniemi	17,0	25,9	0,66	16,6	20,6	0,81	9,5	27,7	0,34
Posio	16,8	39,9	0,42	16,0	37,4	0,43	9,0	15,2	0,59
Kemijärvi	15,5	36,4	0,43	14,2	32,9	0,43	8,6	16,1	0,53
Kuusamo	16,5	37,1	0,44	15,0	33,2	0,45	9,3	15,0	0,62
Salla	15,0	35,0	0,43	13,9	35,9	0,39	8,8	21,4	0,41
Savukoski	13,1	33,6	0,39	11,7	31,8	0,37	8,0	15,5	0,52
Koillis-Suomi	14,5	34,9	0,42	13,5	32,9	0,41	8,4	17,5	0,48

[1]) $\bar{x}$ = Mean volume on land — *Puuston keskitilavuus maalla*
$\bar{x}_p$ = Mean volume on forest land — *Puuston keskitilavuus metsämaalla*
p = Proportion of forest land — *Metsämaan osuus*

ation is not always present in the field sample. However, the differences of the variations are also clear in the combined field and photo sample. This can be seen from the following CV-values:

Sub-area	Variable $\bar{x}$	$\bar{x}_p$	p
	Field sample		
Pelkosenniemi	35	24	42
Others[1])	53	49	25
Ratio	0,66	0,49	1,68
	Photo sample		
Pelkosenniemi	26	21	28
Others[1])	36	34	17
Ratio	0,72	0,62	1,65

[1]) The mean of the CV-values in the other sub-areas

The photo sample significantly decreases the difference between the sub-areas in question only with respect to $\bar{x}_p$. On the other hand, the difference remains unchanged with p. The outcome supports the statement that *Pelkosenniemi* is different from the other sub-areas with respect to the population variation of its forests. This is, in fact, quite possible considering the geographical characteristics.

The sampling of the photo points introduces an element to the total sampling error. Let this partial factor be called here *error of the photo sample.* The upper and the lower limits of which, expressed as relative standard errors, are presented in Table 19. The upper limit has been calculated by formulae of one-stage sampling and the lower limit by formulae of two-stage sampilng. If the location of the tracts were random, the true error of the photo sample would be between these extreme values, possibly near the upper limit. The location of the tracts being systematic, however, the true error may actually lie nearer the lower limit (cf. Nyyssönen et al. 1971, p. 6).

It must be remembered that the relative standard errors in Table 19 are not the final sampling errors, because all photo points have not been measured in the field. The

Table 19. Error of the photo sample expressed as relative standard error (%) by characteristic and sub-area. 1 = estimated on the basis of the variation of the tract values, 2 = both variation between tracts and variation within tracts included.

Taulukko 19. Ilmakuvanäytteen virhe ilmaistuna suhteellisena keskivirheenä (%) eri tunnusluvuilla osa-alueittain. 1 = arvioituna lohkoarvojen vaihtelun perusteella, 2 = sisältäen sekä lohkojen sisäisen että välisen vaihtelun.

Characteristic[1]) Tunnusluku[1])		Pelkos.	Posio	Kemij.	Kuusamo	Salla	Savukoski	Koillis-Suomi
		Sub-area — *Osa-alue* s_e %						
1	1	7,93	3,19	3,33	2,31	3,28	2,04	1,48
	2	5,08	2,32	2,31	1,76	2,08	1,38	0,96
2	”	2,23	1,24	1,39	1,19	1,27	1,10	0,55
		1,54	0,99	0,99	0,88	0,90	0,77	0,39
3	”	5,23	2,21	2,25	1,75	2,49	1,75	0,98
		3,36	1,67	1,64	1,34	1,66	1,25	0,68
4	”	9,45	9,24	13,03	6,07	6,48	5,58	3,58
		7,43	6,33	9,00	4,46	4,25	3,71	2,34
5	”	4,89	5,82	5,10	4,32	4,07	3,78	1,97
		3,86	3,91	3,44	2,94	2,75	2,49	1,32
6	”	3,88	5,45	4,60	3,91	4,17	3,58	1,86
		3,41	3,68	3,11	2,71	2,75	2,33	1,24
7	”	7,16	6,38	5,89	4,60	4,87	4,32	2,64
		5,85	4,58	3,97	3,42	3,64	3,08	1,79
8	”	9,02	8,24	9,23	6,28	5,86	5,25	3,17
		7,35	5,97	6,38	4,51	4,27	3,78	2,21

[1]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4 (s. 27), alaviitta 1*

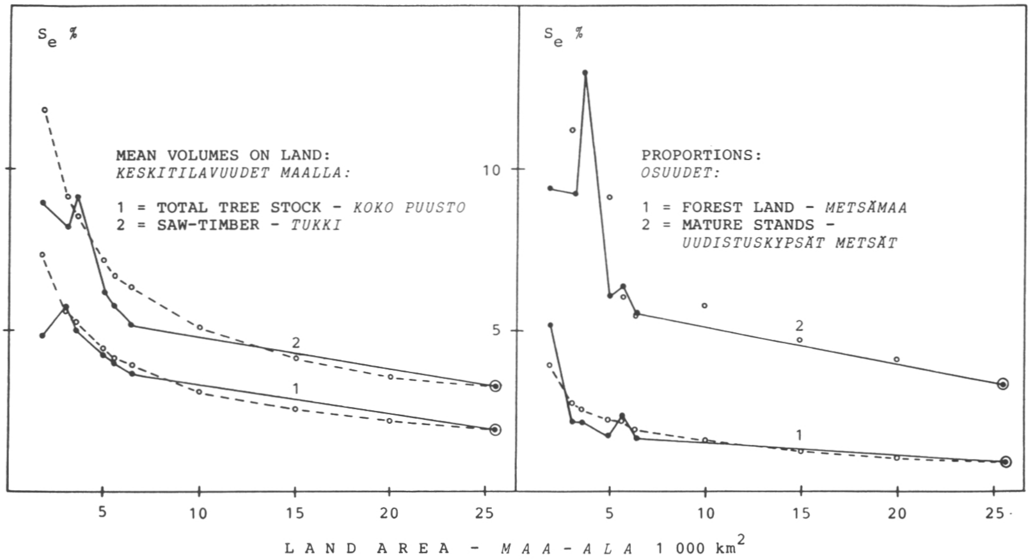


Figure 15. Observed (continuous line) and predicted (dashed line) error of photo sample as a function of land area.
Kuva 15. Todettu (yhtenäinen viiva) ja ennustettu (katkoviiva) ilmakuvanäytteen virhe maapinta-alan funktiona.

manner of presentation is justified only for the sake of clarity. It will later be proved that the absolute variances are needed when deriving the total sampling error.

The error of the photo sample is illustrated as a function of area in Figure 15. The illustration concerns the observed errors (the upper limit) of four characteristics and a generalization for three of them (cf. Fig. 12, p. 28). The procedure has been discussed earlier in connection with the field sample. The formulae used now are the same and the estimation proceeds from *Koillis-Suomi* downwards. The relevant data used in constructing the curves marked by dashed line in Figure 15 are presented in Table 20. Concerning variables $\bar{x}$, $\bar{x}_t$ and p, the observed errors agree fairly well with those estimated on the basis of the land area. This is not the case with the proportion of mature stands p_m. The prediction fails thoroughly in *Pelkosenniemi* and *Kemijärvi* giving values about twice as high as the observed ones (19,41 and 23,25 % instead of 9,45 and 13,03 %). Evidently the reason is connected with the very low proportion of mature stands in these sub-areas.

The formula used to predict s_p and s_{pm} read (see p. 28 for explanation):

$$s_b \% = \frac{p_a}{p_b} * \sqrt{\frac{A_a}{A_b}} * s_a \%$$

By certain transformations the following equivalence becomes apparent:

$$\frac{s_b'^2}{s_a'^2} = \frac{k_b}{k_a} * \frac{A_a}{A_b}$$

where $s_a'^2$ = absolute variance of the tract values in sub-area a

$s_b'^2$ = as above in sub-area b

k_a and k_b = number of tracts in sub-areas a and b

Assuming that ratio k_a/k_b approaches ratio A_a/A_b (i.e. the number of tracts decreases with decreasing area), the value of $s'_b{}^2/s'_a{}^2$ approaches 1. In other words, the right prediction presupposes that the absolute variance of the tract values is the same in sub-areas a and b. In reality, however, there are differences between the sub-areas as can be seen from the following absolute variances (10^3) of the tract values of p and p_m:

Sub-area	Characteristic			
	p	p_m	A_a/A_b	k_a/k_b
Pelkosenniemi	30,53	4,47	14,03	11,25
Posio	11,04	12,89	8,28	6,70
Kemijärvi	12,23	5,53	7,17	6,18
Kuusamo	10,20	8,14	5,12	4,26
Salla	19,55	18,57	4,46	4,26
Savukoski	11,19	15,92	4,00	3,98
Koillis-Suomi	13,76	15,25	1,00	1,00

Concerning p_m, the variance in *Pelkosenniemi* and *Kemijärvi* is no more than one third of the variance in *Koillis-Suomi,* which explains the major part of the disagreement of the predictions. The proportion of the tracts which partly fall outside the target area increases with decreasing area. As stated earlier, the variation of the tract values then also increases. At the same time, however, the number of the tracts decreases more slowly than the area which maintains the increase of the estimated error at the right proportion.

The case of *Pelkosenniemi* leads to the conclusion that land area alone does not always suffice in the prediction of the error of the photo sample. Some kind of information about the true population variation in the target area is required to adjust the predictions. However, caution is required when the variation is estimated on the basis of a sparse sample, as in the case of the field sample. The combined field and photo sample justifies the statement that the variation of the mean volumes is lower in *Pelkosenniemi* than on an average in *Koillis-Suomi.* The difference is even greater when the proportion of mature stands is considered. It is evident that a low variation of the proportion of mature stands also results in a low variation of the volumes. This is because these variables are connected in such a way that the greatest volumes usually occur in mature stands. However, *Kemijärvi* indicates that the rule of thumb concerning the variation pattern of these variables is not very definite.

The magnitude of the error of the photo sample becomes more comprehensible by comparing it with the estimated sampling error in the new field sample. The predicted errors of p (proportion of forest land) and $\bar{x}$ (mean volume on land) expressed as relative standard error in *Koillis-Suomi* and in the biggest and smallest sub-area are (one-stage sampling):

Table 20. The error of the photo sample expressed as relative standard error (s_e %) for mean volume on land ($\bar{x}$), mean volume of saw-timber on land ($\bar{x}_t$), proportion of forest land (p) and proportion of mature stands (p_m). The estimates have been derived on the basis of the land area from *Koillis-Suomi* downwards (see Fig. 15).

Taulukko 20. Ilmakuvanäytteen virhe ilmaistuna suhteellisena keskivirheenä (s_e %) osa-alueittain estimaateille puuston keskitilavuus maalla ($\bar{x}$), tukin keskitilavuus maalla ($\bar{x}$), metsämaan osuus (p) ja uudistuskypsien metsien osuus (p_m). Osa-alueiden virhe-estimaatit on johdettu maa-alan perusteella Koillis-Suomen virhe-estimaateista (ks. kuva 15).

Sub-area *Osa-alue*	Land area Maa-ala	Characteristic — *Tunnusluku*					
		$\bar{x}$	$\bar{x}_t$	p		p_m	
		s_e	s_e	e[1])	s_e	e[1])	s_e
	1000 km²	%					
Koillis-Suomi	25,67	1,97	3,17	67,2	0,98	19,4	3,58
	20	2,23	3,59	,,	1,12	,,	4,06
	15	2,58	4,15	,,	1,29	,,	4,68
	10	3,16	5,08	,,	1,58	,,	5,74
Savukoski	6,42	3,94	6,34	68,2	1,95	25,4	5,47
Salla	5,75	4,16	6,70	65,4	2,15	24,4	6,01
Kuusamo	5,01	4,46	7,18	67,1	2,24	17,3	9,10
Kemijärvi	3,58	5,28	8,49	68,7	2,59	8,0	23,25
Posio	3,10	5,67	9,12	69,4	2,76	17,9	11,17
Pelkosenniemi	1,83	7,38	11,87	63,1	3,95	13,4	19,41

[1]) Estimated proportion — *Arvioitu osuusluku*

Sub-area	Characteristic $s_{\bar{x}}$ %	s_p %
	Field sample[1])	
Koillis-Suomi	4—5	2—3
Savukoski	9—10	4—5
Pelkosenniemi	17—19	8—9
	Photo sample[2])	
Koillis-Suomi	2	1
Savukoski	4	2
Pelkosenniemi	7—8	4

[1]) The results have been presented in page 29
[2]) Error of the photo sample, not the final sampling error

The error of the photo sample is less than half of the sampling error in the new field sample. The ratio in question between the observed errors is presented in Table 21. The row for *Koillis-Suomi* reveals no significant differences between the characteristics. There are however, noticeable differences between the characteristics in the sub-areas. In addition to this, the sub-areas seems to be different as can be seen from the mean ratios of the eight characteristics. *Kuusamo* and Salla show a clearly lower mean ratio than the rest of the study area. The causes of the observed differences are, however, outside of the scope of this investigation. There are probably both accidental and real (area-specific and/or characteristic-specific) deviations from the general trend.

The number of the tracts increases from 3 to 8 per unit area of 16 * 48 km^2 in the extension of the field data to the photo sample. The mean square error of the tract values decreases more than implied by the increase of the tracts. This is because there are more plots per tract and the plots are more spread within tract in the photo sample than in the field sample. The present decrease of the mean square error would have not been reached using the structure of the field tract in both sampling phases although all possible tracts had been interpreted as photo tracts.

452. Error due to the interpretation

Results are calculated on the basis of the combined field and photo sample as if all plots had been measured in the field. However, there are deviations between the extended field data and the reality of the field, which increases the sampling error. Let this error component be called *error due to interpretation.* The magnitude of this error component depends on the accuracy of the photo-interpretation and the fraction between the field sample and the photo sample. The error due to the interpretation equals zero if all photo points have been measured in the field. The same is the case if the photo-interpretation has been achieved without error.

The field points make a sample from the population of the photo points. The means and the proportions in the super cells are estimated on the basis of the field points. Each super cell includes both clusters and single photo points because the two points within a cluster need not fall into the same super cell. The proportion of the single points in the photo sample by sub-area and interpreted land use class is:

Sub-area	Land use class[1]) 1	2	3 %	4	Total
Pelkosenniemi	18	23	11	47	18
Posio	23	35	34	53	28
Kemijärvi	19	23	21	42	21
Kuusamo	24	41	33	44	29
Salla	16	32	33	37	21
Savukoski	11	25	18	65	15
Koillis-Suomi	17	30	23	44	21

[1]) 1 = Forest land, 2 = Scrub land, 3 = Waste land and 4 = Other land area

Clustering of the photo points increases the error of the photo sample by increasing the variation of the tract values. The field sample is comprised of clusters and single points in the same way as the photo sample. The two points within a field cluster are correlated, which affects the precision of the estimated means and proportions by super cell. The mean volume in the field sample in super cell h can be expressed as

$$\bar{x}_h = \bar{x}'_h + (\bar{x}_h - \bar{x}'_h)$$

where $\bar{x}'_h$ refers to the mean volume in the photo sample. The conditional expected value of $\bar{x}_h$ equals to $\bar{x}'_h$. Then (following Cochran 1977, p. 329)

$$\text{cov}\,\{\bar{x}'_h, (\bar{x}_h - \bar{x}'_h)\} = 0$$

Assuming simple random sampling in the second phase

$$V(\bar{x}_h - \bar{x}'_h) = \left(\frac{1}{m_h} - \frac{1}{n_h}\right) E(s'^2_h)$$

where $E(s'^2_h)$ refers to the expected variance of the volumes in the photo sample in super cell h. The above equivalence

Table 21. The ratio between the error of the photo sample (expressed as relative standard error) and the relative standard error in the new field sample by characteristic and sub-area.

Taulukko 21. Ilmakuvanäytteen virheen (ilmaistuna suhteellisena keskivirheenä) ja uudesta maastonäytteestä laskettujen estimaattien suhteellisen keskivirheen välinen suhdeluku.

Sub-area *Osa-alue*	Characteristic[1]) — *Tunnusluku*[1]) 1	2	3	4	5	6	7	8	1—8
Pelkosenniemi	40	41	40	44	44	52	29	34	41
Posio	57	59	45	34	41	38	39	50	45
Kemijärvi	33	49	42	32	48	49	42	51	43
Kuusamo	34	41	31	37	36	38	28	39	36
Salla	40	41	44	35	36	39	29	31	37
Savukoski	40	39	38	43	47	43	38	40	41
Koillis-Suomi	42	43	41	41	41	41	40	43	42

[1]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4 (s. 27), alaviite 1*

implies constant n_h and f_h $(= m_h/n_h)$ in repeated samples. The variance is approximated on the basis of the field sample as

$$s_h^2 = \frac{1}{m_h - 1} \sum_{j=1}^{m_h} (x_{hj} - \bar{x}_h)^2$$

where m_h = number of field points in super cell h
and x_{hj} = volume measured on point j in super cell h

The conditional error due to the interpretation can then be approximated as (Cochran ibid., p. 329)

$$V_I(\bar{x}) = \sum_{h=1}^{L} \frac{W_h E(s_h^{'2})}{n} \left(\frac{1}{f_h} - 1\right)$$

where n refers to the total number of the photo points. An estimator for the error component in question reads

$$(29) \qquad V_I(\bar{x}) = \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h} (1 - f_h)$$

Formula 29 is valid only with a fixed f_h. However, f_h is variable in repeated sampling. This must be taken into account in the statistical evaluation of the inventory method as a whole. Taking the expectation from the conditional variance and by using an approximation (Cochran ibid., p. 134)

$$E\left(\frac{1}{m_h}\right) = \frac{1}{m} * w_h + \frac{1 - w_h}{m^2 w_h^2}$$

that is valid for simple random sampling, the error due to interpretation is

$$V_I(\bar{x}) = E\left\{\sum_{h=1}^{L} \frac{w_h^2 E(s_h^{'2})}{m_h} \left(1 - \frac{m_h}{n_h}\right)\right\}$$

$$= \sum_{h=1}^{L} \frac{W_h E(s_h^{'2})}{n} * \left(\frac{n}{m} - 1 + g(m, n, W_h)\right)$$

where g is a function the value of which depends on the values of $1/m$, $1/n$ and W_h. If the g-function is neglected the estimator reads

$$V_I(\bar{x}) \simeq (1 - f) * \sum_{h=1}^{L} \frac{W_h E(s_h^{'2})}{m}$$

The error component represented by the g-function arises from the fact that the stratification is conducted after the selection of the photo points. The technique in question is called poststratification (Cochran 1977, p. 134). The additional error is estimated by formula

(30)

$$G = \sum_{h=1}^{L} \frac{s_h^{'2}(n - n_h)}{nm^2} = \sum_{h=1}^{L} \frac{s_h^{'2}}{m_h} * \frac{m_h}{m} * \frac{1}{m} * \frac{n - n_h}{n}$$

Assuming that $m_h/m \sim w_h$ and $(n - n_h)/n \sim 1$.

$$G \simeq \sum_{h=1}^{L} \frac{w_h s_h^{'2}}{m_h} * \frac{1}{m}$$

The mean value of w_h in the sub-areas of *Koillis-Suomi* varies from 0,0625 in *Pelkosenniemi* to 0,0270 in *Savukoski.* The corresponding values of 1/m are 0,0074 and 0,0025. Consequently, the increase of the variance is about 10 % and must be taken into account in the calculations.

The final formulae for the estimation of the error due to the interpretation read

$$(31) \quad v_I(\bar{x}, p) = \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h} (1-f_h) + G$$

$$(32) \quad v_I(\bar{x}_p) = \frac{1}{p^2} * \sum_{h=1}^{L} \frac{w_{ph}^2 s_{ph}^2}{m_{ph}} (1-f_h) + G$$

The notation has been explained in connection with the formulae 17 and 18 (p. 37). Formula 31 concerns all estimates of proportions as well as mean volumes on land. Formula 32 is applied to mean volumes in forest strata, as for example mean volume on forest land (characteristic 6).

As presented earlier, the photo sample as well as the field sample consist both of clusters and of single points in each super cell. Using clusters as the sampling unit leads to the following estimator for the mean volume on land in super cell h:

$$\bar{x}_h = \frac{\sum_{i=1}^{m_h'} \frac{M_{hi}}{\bar{M}_h} * \bar{x}_{hi}'}{m_h'}$$

where m_h' = total number of whole clusters and single points in super cell h

M_{hi} = number of relascope plots in cluster i (1 or 2)

$\bar{M}_h$ = mean number of relascope plots per cluster in super cell h

$\bar{x}_{hi}'$ = mean volume in cluster i

Concerning single points, $\bar{x}'_{hi}$ equals the measured volume. The above estimator is analogous to the estimator of mean volume $\bar{x}$ calculated on the basis of the tract means (Formula 4, p. 21). Accordingly, the principle presented in formula 7 (p. 22) can be applied to the estimation of the variance of an estimated proportion or mean volume in super cell h. This approach has been applied in this investigation. So, concerning mean volume on land, the total estimator for the error due to the interpretation reads (cf. formulae 30 and 31)

$$v_I(\bar{x}) = \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h'} (1-f_h) + \sum_{h=1}^{L} \frac{s_h^2 (n'-n_h')}{n'} * \frac{1}{m'^2}$$

where $m' = \sum_h m_h'$

n' = total number of whole photo clusters and single photo points = $\sum_h n_h'$

s_h^2 = mean square of the mean volumes in the clusters and in the single points in super cell h

It can be seen from the above expressions that clustering also affects the value of the g-function. As with the tract values, the estimated means and proportions in the super cells are ratio estimates which are biased. As to the possible magnitude of the bias see pages 29, 34 and 68.

The error due to the interpretation, expressed as relative standard error, is presented in Table 22. As with the error of the photo sample, the relative standard error in question is not the final sampling error. Comparing the error due to the interpretation with the error of the photo sample reveals that these are generally in the same class of magnitude when the latter error component is estimated on the basis of the tract values only (estimate 1 in Table 9, p. 39). There are however, some interesting deviations from this rule.

The estimated error of the photo sample concerning the mean volumes and the proportion of mature stands is relatively low in *Pelkosenniemi.* It was concluded earlier that the phenomenon may well be, due to a low population variation in this sub-area (p. 57). As to the error due to the interpretation, however, *Pelkosenniemi* does not differ from the other sub-areas when its surface area taken into account. With constant accuracy of photo-interpretation, the interpretional error increases with decreasing area mainly because the number of the field points per super cell decreases. The final outcome depends on the variation of the values measured in the field points within each super cell. Consequently, the concept of population variation is different with error from the interpretation than with the error of the photo sample. The variation of the cluster values is decisive in the former case, the variation of the tract values in the latter one. In other words, the pattern of the population variation depends on the relevant sampling unit.

The case of *Pelkosenniemi* indicates that

Table 22. Error due the interpretation expressed as relative standard error (s_e %) as well as value of the estimated proportion or mean volume (e) by characteristic and sub-area.

Taulukko 22. Tulkintavirhe ilmaistuna suhteellisena keskivirheenä (s_e %) sekä osuusluvun tai keskitilavuuden estimaatin arvo (e) eri muuttujilla osa-alueittain.

Characteristic[1]) *Tunnusluku*[1])		Sub-area — *Osa-alue* Pelkos.	Posio	Kemij.	Kuusamo	Salla	Savukokski
1	s_e %	5,51	3,34	3,35	2,47	1,92	1,49
	e[2])	42,92	55,73	56,40	54,72	63,00	69,12
2	,,	2,62	1,37	1,38	1,38	1,42	0,95
		84,37	84,88	86,16	84,04	83,34	86,15
3	,,	3,79	2,65	2,73	1,94	1,94	1,62
		63,09	69,36	68,71	67,14	65,35	68,21
4	,,	17,04	10,54	18,64	9,69	6,57	6,62
		13,37	17,93	7,99	17,27	24,44	25,44
5	s_e %	7,61	5,45	5,77	4,83	4,10	3,83
	e[3])	31,38	41,09	33,64	40,62	34,04	37,21
6	,,	7,72	5,61	5,76	4,75	4,31	4,00
		45,25	55,30	46,17	56,96	47,50	50,76
7	,,	12,99	9,70	7,18	8,33	8,33	6,95
		11,46	21,54	25,47	21,22	12,70	19,20
8	,,	15,56	8,77	9,98	7,41	8,45	6,62
		6,39	12,36	8,04	14,05	8,80	10,52

[1]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4 (s. 27), alaviitta 1*
[2]) Proportion, % — *Osuusluku, %*
[3]) Mean volume, m³/ha — *Keskitilavuus, m³/ha*

the defined error components are informative as to the structure of the total sampling error in different sub-areas. This being the case, they can be used in finding the most efficient sampling procedure, assuming that the population variation of the different sampling units is known. However, one must keep in mind that the main concern of the inventory is in the total study area of *Koillis-Suomi,* not a (small) sub-area within it.

The estimated error of the photo sample with the proportion of mature stands is also relatively low in *Kemijärvi.* But as in *Pelkosenniemi,* the estimated error due to the interpretation is high, even higher than implied by the surface area. This is evidently associated with the very low value of p_m in *Kemijärvi.*

There is no significant difference between the mean volumes on land and on forest land (characteristics 5 and 6) as to the estimated relative values of the both error components. The main reason is that mean volume on forest land is closely connected with mean volume on land. The volumes occurring outside forest land make only a little part of the mean volume on land. Thus a clear difference of the relative error values between these characteristics presupposes that, in relative terms, the population variation of the tract values (error of the photo sample) or of the single field clusters (error from the interpretation) is distinctly different on forest land than on the other land areas.

The error components show a different behaviour with the constituents of the mean volume on land (pine and saw-timber, characteristics 7 and 8). As to the error of the photo sample, the relative error increases with decreasing magnitude of the constituent. Accordingly, pine has a smaller relative error than does saw-timber. But this is not always the case with the error from the interpretation. Despite the lower mean volume of saw-timber, it is associated with a lower relative error than for the mean volume of pine in three sub-areas. Consequently, the relative variation of the saw-timber volumes measured at the field points is in some cases smaller than that of pine volumes. The explanation is that saw-timber generally occurs also within relascope plots occupied by other tree species than pine (primarily Norway spruce). Another explanation reads that the plots including saw-timber generally are greater than is the average size of the relascope plots. The phenomenon becomes apparent in sub-areas where there are plenty of spruce-dominated forests on hill-top sites. However, the tracts catch the variation so effectively that the variation of the tract values is relatively lower with pine than with saw-timber.

The volume of the total growing stock has been assessed from aerial photographs and it has also been used as a stratum variable in forming the super cells. This partly explains the similarity of characteristics $\bar{x}$ and $\bar{x}_p$ (mean volume on land and forest land) with respect to both error components. No attempt has been made to divide the volume of the total growing stock into constituents from aerial photographs, which explains the relatively higher errors of pine volumes compared with the errors of $\bar{x}$ and $\bar{x}_p$. Actually the dominant tree species has been assessed from aerial photographs and also used to some degree in the stratification of the photo sample. The effect of this information on the error of pine volume is probably fairly small. The utility of this stratum variable is greatest in the estimation of the proportions of the pine- and spruce-dominated forests.

The relative error from the interpretation increases with decreasing proportion to be estimated. Differences in the accuracy of the photo-interpretation cause only slight deviations from this general rule. Concerning proportions, the behaviour of the error components is fairly similar with a few exceptions (see characteristics 1 and 3 in *Savukoski* and *Salla*). The inconsistencies can be explained by systematic changes occurring within great sub-areas. The change is not the same for different characteristics. A systematic change increases the error of the photo sample but does not necessarily increase the error from the interpretation.

It seems to be possible to predict the error from the interpretation fairly well by the error of the photo sample. In the present case, it happens that the error components are of the same order of magnitude. With another sampling design and photo material the ratio between the errors will naturally be different. So an increase of the field sampling fraction or a better accuracy in the photo-interpretation results in a smaller error from the interpretation with respect of the error of the photo sample. The ratio between the error components must be determined for each case separately. As discussed above, however, noticeable deviations from the general trend can occur due to area-specific and characteristic-specific features.

453. Total sampling error

The estimator for a mean volume or a proportion can be expressed as a sum of two terms (Cochran 1977, p. 329):

$$e = \sum_{h=1}^{L} w_h e_h = \sum_{h=1}^{L} w_h e'_h + \sum_{h=1}^{L} w_h (e_h - e'_h)$$

where e_h refers to the field sample and e'_h to the photo sample. Assuming that the two terms are independent of each others the variance of e is a sum of the variances of the terms. The variance of the first term is estimated as the error of the photo sample and that of the second term as the error from the interpretation. Consequently, the total sampling error is approximated by the formula

$$(33) \quad v(e) = v_p(e) + v_I(e)$$

The values of the error components expressed as relative standard error have been presented in Tables 19 and 22. These are transformed into absolute variances, summed up and the results are presented as relative standard error in Table 23. Two estimates are given, to find the range within which the true error is probably located. The limits have been obtained by using the upper and the lower limits of the error of the photo sample in the calculations.

The variance of each estimated quantity (area or total volume) in *Koillis-Suomi* has been estimated by summing the corresponding variances in the sub-areas (p. 24). The relative standard error of the total volume on forest land in any sub-area is approximated as

$$s_{V_p}\% \simeq \sqrt{(s_{A_p}\%)^2 + (s_{\bar{x}_p}\%)^2}$$

where $s_{Ap}\%$ equals to $s_p\%$. Let A_k refer to the forest land area and $\bar{x}_k$ to the mean vol-

Table 23. Total sampling error expressed as relative standard error (%) by sub-area and characteristic. Estimation of the error of the photo sample: 1 = on the basis of the variation of the tract values only and 2 = the variation both between the tracts and within the tracts included.

Taulukko 23. Yhdistetty otantavirhe ilmaistuna suhteellisena keskivirheenä (%) eri muuttujilla osa-alueittain. Ilmakuvanäytteen virheen arvioimistapa: 1 = lohkoarvojen vaihtelun perusteella ja 2 = sekä lohkojen välisen että sisäisen vaihtelun perusteella.

Characteristic[1]) *Tunnusluku*[1])		Sub-area — *Osa-alue* Pelkos.	Posio	Kemij.	Kuusamo	Salla	Savukoski	Koillis-Suomi
		%						
1	1	9,66	4,62	4,72	3,38	3,80	2,53	1,61
	2	7,49	4,07	4,07	3,03	2,83	2,03	1,32
2	„	3,44	1,85	1,96	1,82	1,91	1,45	0,79
		3,04	1,69	1,70	1,64	1,68	1,22	0,69
3	„	6,46	3,45	3,54	2,61	3,16	2,38	1,31
		5,06	3,13	3,18	2,36	2,55	2,05	1,12
4	„	19,48	14,02	22,74	11,43	9,23	8,66	4,87
		18,59	12,29	20,70	10,67	7,82	7,59	4,31
5	„	9,05	7,97	7,70	6,48	5,78	5,38	2,77
		8,53	6,71	6,72	5,65	4,94	4,57	2,39
6	„	8,64	7,82	7,37	6,15	6,00	5,37	2,73
		8,44	6,71	6,55	5,47	5,11	4,63	2,38
7	„	14,83	11,61	9,29	9,52	9,65	8,18	4,13
		14,25	10,73	8,20	9,00	9,09	7,60	3,83
8	„	17,99	12,03	13,59	9,71	10,28	8,45	4,55
		17,21	10,61	11,85	8,67	9,47	7,62	4,09

[1]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4 (s. 27), alaviitta 1*

ume on forest land in sub-area k. The estimator for the relative standard error of the mean volume on forest land in *Koillis-Suomi* reads (Svensson 1983, p. 26): $s_{\bar{x}}\% =$

$$\frac{\sqrt{\sum_{k=1}^{n}\left\{A_k^2\bar{x}_k^2\left[(s_{A_k}\%)^2+(s_{\bar{x}_k}\%)^2\right]+A_k^2(s_{A_k}\%)^2\left[\bar{\bar{x}}^2-2\bar{\bar{x}}\bar{x}_k\right]\right\}}}{\sum_{k=1}^{n} A_k\bar{x}_k}$$

where n refers to the number of the sub-areas.

The above estimator takes into account that forest land areas A_k are estimates. A fairly good estimator for $s_{\bar{x}}$ % reads

$$(34) \qquad s_{\bar{x}}\% = \frac{\sqrt{\sum_k A_k^2\bar{x}_k^2 * (s_{\bar{x}_k}\%)^2}}{\sum_k A_k\bar{x}_k}$$

According to Svensson (1984, written comment) formula 34 gives almost the same outcome as the more complex and statistically correct estimator.

The sampling error of the total volume on forest land in *Koillis-Suomi* can be approached in two ways. Firstly, the variance of V_p is a sum of the corresponding variances in the sub-areas:

$$v(V_p) = \sum_{k=1}^{n} v(V_{p_k})$$

and the relative standard error is calculated as

$$s_{V_p}\% = 100 * \frac{\sqrt{v(V_p)}}{V_p}$$

where $V_p = \sum_k V_{p_k}$. Practically the same result is obtained by combining the relative standard error of the mean volume on forest land in *Koillis-Suomi* (formula 34) and the corresonding error characteristic of the total forest land area.

Table 24. The division of the total variance of the estimates into components error of the photo sample (v_P) and error due to the interpretation (v_I).
Taulukko 24. Estimaattien varianssien jakaantuminen ilmakuvanäytteen virheeseen (v_p) ja tulkintavirheeseen (v_I).

Sub-area *Osa-alue*		Characteristic[1]) — *Tunnusluku*[1]) 1	2	3	4	5	6	7	8
Pelkosenniemi	v_P	67	42	66	24	29	20	23	25
	v_I	33	58	34	76	71	80	77	75
Posio	”	48	45	41	43	53	49	30	47
		52	55	59	57	47	51	70	53
Kemijärvi	”	50	50	40	33	44	39	40	46
		50	50	60	67	56	61	60	54
Kuusamo	”	47	43	45	28	44	40	23	42
		53	57	55	72	56	60	77	58
Salla	”	74	44	62	49	50	48	25	32
		26	56	38	51	50	52	75	68
Savukoski	”	65	57	54	42	49	44	28	39
		35	43	46	58	51	56	72	61
Koillis-Suomi	”	62	47	53	40	47	43	29	40
		38	53	47	60	53	57	71	60

[1]) See Table 4 (p. 27), footnote 1 — *Ks. taulukko 4 (s. 27), alaviite 1*

Table 25. The estimated area of forest land (A_P) and the estimated total volumes on forest land (V_P), as well as on land (V) and the corresponding relative standard errors by sub-area.
Taulukko 25. Metsämaan pinta-alan (A_P), metsämaan puuston tilavuuden (V_p) ja kokonaispuuston tilavuuden (V) arviot sekä vastaavat suhteelliset keskivirheet osa-alueittain.

Sub-area *Osa-alue*	Characteristic — *Tunnusluku* A_P sq.km.	s_{A_P} %	V_P 1000 m³	s_{V_P} %	V 1000 m³	s_V %
Pelkosenniemi	1 152	6,46	4 928	10,79	5 404	9,05
Posio	2 148	3,45	11 339	8,55	12 024	7,97
Kemijärvi	2 458	3,54	11 044	8,18	11 619	7,70
Kuusamo	3 361	2,61	18 348	6,68	19 337	6,48
Salla	3 754	3,16	16 565	6,78	18 083	5,78
Savukoski	4 381	2,38	20 573	5,87	21 987	5,38
Koillis-Suomi	17 254	1,31	82 796	3,03	88 454	2,77

The total mean square error in the sample consists of the error from the interpretation v_I and of the error of the photo sample v_P as presented in Table 24. It is clear that the error from the interpretation is relatively greater the poorer is the accuracy of the photo-interpretation. So the share in the MSE of the proportion of mineral soils (characteristic 1) in *Koillis-Suomi* is only 38 % because the interpretation of sub-class is accurate. Forest land (3) is well discernible on aerial photographs, as well. The other extreme concerns the characteristics mean volume of pine and saw-timber on land (7 and 8), as well as proportion of mature stands (4). As mentioned earlier, the variables in question have not been assessed from aerial photographs. As to the other characteristics, the two error components occupy about an equal part of the total MSE.

The area of forest land and the total volumes on land as well as on forest land and the corresponding relative standard errors by sub-area are in Table 25. The error characteristics refer to the higher estimated errors presented in Table 23. The lower limits of s_{Ap}, s_{Vp} and s_V in *Koillis-Suomi* are 1,12 %, 2,63 % and 2,39 %. The total volumes do not include usable dead trees the volume of which is 4,9 milj. cu.m. on forest land and 5,7 milj. cu.m. in all. The amount of this tree material is about 6 % of the total growing stock. The error characteristics have been calculated on the basis of the total tree stock (= growing stock + usable dead trees). However, it can be assumed that they are also applicable to the growing stock.

The total forest land area and volumes in *Koillis-Suomi* have been calculated by summing the corresponding quantities in the sub-areas. The results can also be calculated after the samples of the sub-areas have been put together. The outcome is not quite the same because the latter approach implies that the area per one photo point is the same in the whole study area. The two estimates and the relative difference between them, by characteristic, are:

Characteristic[1]			Estimate[2] 1	Estimate[2] 2	Difference[3] ‰
1		%	59,69	59,69	—
2		,,	84,83	84,84	0,1
3	(A_p)	,,	67,21	67,19	0,3
4		,,	19,43	19,42	0,5
5	(V)	m^3/ha	36,72	36,66	1,6
6		,,	50,80	50,73	1,4
7		,,	18,74	18,71	1,6
8		,,	10,52	10,48	3,8

[1]) See Table 4 (p. 27), footnote 1
[2]) 1 = Sum of the quantities in the sub-areas
2 = Calculated from the whole sample of *Koillis-Suomi*
[3]) Per mille of estimate 2

As to the total volume on forest land (V_p) the difference is 1, 8 ‰.

The differences are very small compared with the estimated sampling error and they are not of practical importance. For the sake of simplicity, approach 2 is applied to the official calculations. The estimation of the error from the interpretation must, however, be made by sub-area, because the division into the super cells is different in the sub-areas. The error of the photo sample in *Koillis-Suomi* can also be estimated from the total sample. Expressed as relative standard error, the two approaches give the following estimates for the error component in question (notation as above):

Characteristic		Estimate 1	Estimate 2
		%	
1		1,26	1,48
2		0,54	0,55
3	(A_p)	0,95	0,98
4		3,10	3,58
5	(V)	1,91	1,97
6		1,80	1,86
7		2,22	2,64
8		2,88	3,17

Approach 2 gives slightly greater estimates, which is natural, because the total sample in *Koillis-Suomi* catches a greater variation than the samples in the sub-areas on average. It was earlier stated that the ratio between the error components is fairly stable. Thus, assuming that the ratio is know, one can approximate the total mean square error in any set of sub-areas simply by the error of the photo sample calculated from the corresponding total sample.

One goal of the inventory is that the relative standard error of the estimated volume of the total growing stock must not exceed 4 % in the main reference areas. This goal has been achieved in *Koillis-Suomi*. The 'overprecision' present is mainly due to the demand that the main results must be usable at the sub-area level.

The relative standard error of the total tree stock volume in the sub-areas varies from about 5—6 % in Savukoski to 9—10 % in Pelkosenniemi (Table 23, p. 65). The applicability of the inventory results does not depend solely on the estimated statistical precision. The concept of accuracy also comprises other sources of error such as measurement errors. It is accuracy not statistical precision that is the decisive factor in the evaluation of the applicability of the results.

The importance of random measurement errors, like sampling error, decreases with an increasing size of sample. The number of the measurements in a large-area forest inventory is large enough to make this error negligible at least in the main reference areas. Systematic measurement errors are independent of sample size, which makes them the main additional error in these areas. Measurement errors can be quantified only by carefully repeated measurements in a part of the sample. A study of this kind carried out in Sweden suggested that the total volume in the whole country was underestimated by about 1—2 % in the inventory period of 1973—1977 (Daamen 1980, p. 132). A bias of this magnitude may not be impossible in the results of the Finnish National Forest Inventory.

A bias increases the confidence interval at a given probability level. The relative increases with different amounts of bias in the estimated total volume are:

Sub-area	Amount of bias, % 0,5	1,0	1,5	2,0
	Increase of confidence interval, %			
Pelkosenniemi	0,1	0,7	1,5	2,6
Savukoski	0,4	2,0	4,4	7,8
Koillis-Suomi	1,9	7,4	15,5	26,4

A bias of a fixed size increases the confidence interval to a relatively greater amount the smaller is the sampling error.

It has not been objectively verified whether there are systematic measurement errors or not in the National Forest Survey in Lapland. The above discussion is, therefore, merely of a theoretical nature. Biases are also possible due to the systematic sampling. Systematic errors are likely to arise if the pattern of the population variation fits the pattern of the sampling. Such an accident cannot be proved without special studies of the population variation.

The estimation technique used (ratio estimate) introduces a bias into the results. The importance of this error decreases with increasing sample size (number of sampling units). It has been proved that the bias due to the ratio estimates is insignificant in the new field sample (p. 29). The relative standard error of the mean tract size gives the maximum possible ratio between $|B|$ and s_e (Cochran 1977, p. 162). The maximum ratio in the photo sample by sub-area is:

Sub-area	Characteristic $\bar{x}_D$	others
Pelkosenniemi	0,087	0,084
Posio	0,068	0,064
Kemijärvi	0,057	0,054
Kuusamo	0,052	0,049
Salla	0,055	0,046
Savukoski	0,040	0,035
Koillis-Suomi	0,018	0,015

All ratios are below the critical value of 0,1 and clearly less than the corresponding ratios in the new field sample. Consequently, the bias due to the ratio estimates is negligible in the photo sample.

The unit volumes of the measured trees have been estimated by volume functions developed for the purposes of the National Forest Survey in Finland (Laasasenaho 1982). The statistical precision of an estimated unit volume depends on tree species and tree size the average relative standard deviation being about 4 %. Trees for the volume and growth estimation have been measured 300—1000 st. per sub-area. Accordingly, the randomlike error component in question can be regarded as negligible. It cannot be stated with certainty whether or not the volume functions are unbiased under Lappish conditions.

Taking into account all the error components discussed, it can be concluded that the applicability of the main results is *sampling error limited* in the sub-areas and probably also in the whole study area. This means that the sampling error is the main component of the total mean square error (Maxim and Harrington 1983, p. 1431). This does not necessarily imply that the most efficient way to increase the accuracy is to decrease the sampling error.

The estimated mean square errors in *Koillis-Suomi* are at a level which is usual in the forestry board districts of Finland (see Uusitalo 1983, p. 6). The figures in Table 26 concern the seven northernmost forestry board districts (Uusitalo ibid., p. 50). The new inventory in *Koillis-Suomi* excepted, the s_e %-values have been taken from Kuusela and Salminen 1983, p. 77—78, Salminen 1973, p. 24—25 and Poso and Kujala 1977, p. 46. The most important comparison concerns the error characteristics of *Koillis-Suomi* in the two successive inventories in 1976 and 1982—83. The comparison reveals how well the new inventory design works compared with the old one. A short efficiency analysis will be carried out in the next chapter. The error characteristics only are discussed here without any cost comparison.

Concerning *Koillis-Suomi* and the two last inventories, the results presented above are not directly comparable. This is partly due to the different approaches in the estimation of the sampling error. The main reason for the inconsistency, however, is that the vari-

Table 26. The relative standard error (s_e %) of the estimated proportion of forest land (p), mean volume on forest land ($\bar{x}_p$) and total volume on forest land (V_p) in the northernmost forestry board districts.

Taulukko 26. Metsämaan osuuden (p) ja metsämaan puuston keskitilavuuden ($\bar{x}_p$) sekä kokonaistilavuuden (V_p) arvion suhteellinen keskivirhe (s_e %) pohjoisimmissa piirimetsälautakunnissa.

Forestry board district *Piirimetsälautakunta*	Inventory *Inventointi*	Land area *Maa-ala*	P	$\bar{x}_p$ s_e %	V_p
Koillis-Suomi	1982—83	25 670	1,3	2,7	3,0
„	1976	„	1,4	2,1[1])	—[2])
„	1969	„	1,7	3,5	3,9
Lappi	1974—76[3])	44 480	1,3	1,7[1])	—[2])
„	1969—70[4])	73 050	1,7	—[2])	2,9
Kainuu	1982[5])	21 760	0,9	2,6	2,8[6])
„	1968	„	1,2	2,9	3,1
Pohjois-Pohjanmaa	1982—83[5])	24 130	1,5	2,1	2,6[6])
„	1968	„	1,4	3,1	3,0
Keski-Pohjanmaa	1981—82	10 940	2,7	2,4	3,6
Pohjois-Karjala	1979—80	17 600	1,2	1,9	2,2
Pohjois-Savo	1980	16 720	1,1	1,6	1,9

1) Concerns mean volume on forested area — *Koskee keskitilavuutta metsä- ja kitumaalla*
2) Unavailable — *Ei saatavilla*
3) The three northernmost municipalities excluded — *Kolme pohjoisinta kuntaa poislukien*
4) The three northernmost municipalities included — *Kolme pohjoisinta kuntaa mukaanlukien*
5) Preliminary estimates — *Alustavia arvioita*
6) Estimated by formula
Arvioitu kaavalla $s_{V_p}\% \simeq \sqrt{(s_{A_p}\%)^2 + (s_{\bar{x}_p}\%)^2}$

ances of the older inventory have been corrected due to the systematic location of the photo sample. The correction coefficient applied is as high as 0,5 (Poso and Kujala 1977, p. 45—46). It can be concluded that to obtain consistent error characteristics from the new inventory, the corresponding reduction should be made from the value of the g-function on one hand and from the error of the photo sample (v_p) on the other. The effect of the systematic field sample is somewhat obscure in connection with two-phase sampling. The proportion of the systematic field sample in the inventory of 1976 was about 40 %. It seems, however, that correction due to the partly systematic field sampling have not been made, although it should be, if applied, some 20 %.

The estimate of the accuracy of the mean volume in *Koillis-Suomi* in 1976 concerns the forested area (forest land plus scrub land), which makes the comparison still more difficult. However, it has been noticed in this investigation that the relative accuracy of the mean volumes is fairly independent of whether the mean volume estimates concern forest land or total land area (cf. characteristics 5 and 6 in Table 23, p. 65). So it is assumed here that the relative accuracy of the mean volume estimate in *Koillis-Suomi* in 1976 is the same on forest land as on the forested area. The corresponding s_e %-value of V_p would then be 2,52.

The following error characteristics of the new inventory in *Koillis-Suomi* have been obtained by applying a correction coefficient of 0,5 to the error of the photo sample (v_p) as well as to the value of the g-function. The approximation gives following s_e %-values:

Inventory	Characteristic A_p	$\bar{x}_p$	V_p
1982—83	1,1	2,2	2,5
1976	1,4	2,1	2,5

As to the volumes, the error characteristics are now fairly similar. The accuracy of the estimated proportion of forest land, however, is clearly better in the new inventory. The difference between the estimates cannot be wholly explained by the larger samples in the new inventory. Possible explanations include better accuracy of photo-interpretation, more efficient stratification, differences in the estimation of the accuracy etc. Any reduction due to the systematic field sample has not been done in the derivation of the above figures.

Systematic field sampling causes problems in the extension of the field data to the photo sample. One must use groups (cells in the new inventory) of varying size which lessens the accuracy compared to the grouping method. The results indicate that the harmful effects of the systematic sampling are relatively greater concerning volumes than proportions. This is because it is difficult to construct homogenous photo strata with respect to volumes.

Summarizing, it can be stated that the new inventory in *Koillis-Suomi* probably gives a more accurate estimate of the forest land area than the earlier one. This is mainly due to the increased sampling intensity of the photo sample. The accuracy of the estimated mean volumes is at about the same level in the two successive inventories. The above conclusions are, of course, valid at the sub-area level, too. In *Posio,* for example, the adjusted error characteristics are (Poso and Kujala 1977, p. 46):

Inventory	Characteristic		
	A_p	$\bar{x}_p$	V_p
1982—83	2,7	6,1	6,6
1976	3,5	5,4	6,4

As can be seen, there are deviations from the general rule. The difference between the error characteristics of $\bar{x}_p$ can be explained by area-specific features. In one or several subareas there must occur a deviation to the other direction.

The land areas of *Kainuu* and *Pohjois-Pohjanmaa* are of the same order of magnitude as the land area of *Koillis-Suomi.* This enables a direct comparison of the presented error characteristics to be made. Despite of the different inventory methods, the accuracy of the main results is about the same in these forestry board districts. This is supporting evidence of the correctness of the sampling intensity used in *Koillis-Suomi.*

5. POTENTIAL USES OF THE SAMPLES

51. Original samples

511. Sub-area or a set of sub-areas as target area

The applicability of the original samples depends on the size of the target area and on the characteristic of interest. This item has been touched on earlier in this investigation in connection with the statistical precision of the different samples. Estimation is possible on the basis of both the bare field sample and the combined field and photo sample. Table 27 illustrates the degree of accuracy of the different samples.

The characteristics for comparison have been chosen to represent different levels of photo-interpretational accuracy. In this connection, the accuracy of the photo-interpretation refers to both direct and indirect correlations between the interpreted and field measured data. It can be clearly seen that the photo sample decreases the mean square error the more the better is the accuracy of the photo-interpretation. As to the eight main characteristics dealt with in this investigation, the reduction from the new field sample to the combined field and photo sample is, on average, 60—70 %. The corresponding reduction from the total field sample (the new and the remeasured field samples together) is 45—60 %. So the results calculated on the basis of the field samples are really of a preliminary nature in connection with the two-phase sampling in question.

In the forest conditions of Lapland it is difficult to find a variable which is totally independent of the stratum variables. The photo sample tends to increase the accuracy

Table 27. The relative standard error in the new field sample (1), as well as the ratio between the mean square errors in the total field sample and in the new field sample (2), and the ratio between the mean square errors in the combined field and photo sample and in the new field sample (3) by characteristic and sub-area.

Taulukko 27. Suhteellinen keskivirhe uudessa maastonäytteessä (1) ja keskineliövirheiden suhde verrattaessa koko maastonäytettä uuteen maastonäytteeseen (2) sekä vastaava suhdeluku verrattaessa yhdistettyä ilmakuva- ja maastonäytettä uuteen maastonäytteeseen (3) osa-alueittain eräillä tunnusluvuilla.

Sub-area Osa-alue	Characteristic[1]) — *Tunnusluku*[1])											
	p			P_m			$\bar{x}$			$\bar{x}_{pine}$		
	1	2	3	1	2	3	1	2	3	1	2	3
	%											
Pelkosenniemi	13,14	69	24	21,50	96	82	11,01	118	68	24,87	92	36
Posio	4,89	74	50	27,37	91	26	14,33	87	31	16,20	74	51
Kemijärvi	5,34	84	44	40,17	49	32	10,72	80	52	14,09	81	43
Kuusamo	5,57	60	22	16,61	77	47	12,04	71	29	16,50	72	33
Salla	5,67	81	31	18,64	76	25	11,37	75	26	16,64	81	34
Savukoski	4,58	88	27	13,10	76	44	8,06	72	45	11,49	76	51
Koillis-Suomi	2,41	74	30	8,40	75	34	4,75	77	34	6,42	77	41

[1]) p = Proportion of forest land — *Metsämaan osuus*
P_m = Proportion of mature stands — *Uudistuskypsien metsien osuus*
$\bar{x}$ = Mean volume on land — *Puuston keskitilavuus maalla*
$\bar{x}_{pine}$ = Mean volume of pine on land — *Männyn keskitilavuus maalla*

of the characteristics even though the corresponding variable has not been assessed from aerial photographs. Let site type serve as an example. This variable is determined on mineral soils and on peatland in land use classes 1—3. The classification is made on the basis of the field layer vegetation which is supposed to indicate the fertility of the site. The estimated proportion of site type 3 ('damp' sites) varies in the combined field and photo sample about from 30 % to 40 %. The error characteristics (notation as in Table 27) and the structure of the mean square error in the combined field and photo sample are:

Sub-area	Characteristic				
	1	2	3	v_p[1])	v_I[2])
	s_e %	%		Per cent of 3	
Pelkosenniemi	16,26	91	40	28	72
Posio	13,85	82	42	23	77
Kemijärvi	17,37	75	40	20	80
Kuusamo	6,92	68	67	26	74
Salla	9,79	79	38	22	78
Savukoski	9,54	68	47	19	81
Koillis-Suomi	4,84	72	42	23	77

1) Error of the photo sample
2) Error due to the interpretation

It appears that the photo sample decreases the mean square error of the estimated proportion of site type 3 by 58 % from the new field sample and by 42 % from the total field sample. The reduction is not significantly smaller than with characteristic $\bar{x}_{pine}$ although dominant tree species has been assessed from aerial photographs. As can be expected, the error originating from the interpretation, in this case, contributes the major part of the total mean square error.

The indirect correlations become effective by the detailed stratifications used in the extension of the field data to the photo sample. Site type is connected with several variables used in the stratification. On mineral soils these are land use class, dominant tree species and stocks. Drainage status is an especially effective indicator on peatland because drainage is carried out only on the better half of the site type distribution. The indirect correlations have the favourable consequence that the photo sample rarely impairs the accuracy of the results from the level attained by the field sample.

512. Target area other than sub-area or a set of sub-areas

5121. New combination of the samples

For a target area of sufficient size, say at least 1800 sq.km., one can consider a separate extension of the field data to the photo sample. The procedure is quite the same as in the sub-areas, with the exception that the external photo and field points are first excluded from the original samples. Ownership categories, such as private forests and state-owned forests, make typical target areas within the sub-areas. Commercial forests without nature conservation areas also make an interesting target area.

An independent combination for the nature conservation areas in itself can also be considered.

A land area of 1800 sq.km. includes, on

Table 28. The number of the photo points in some domain areas by sub-area.
Taulukko 28. Ilmakuvapisteiden lukumäärä eräissä ositteissa osa-alueittain.

Target area *Kohdelaue*	Sub-area — *Osa-alue* Pelkosen-niemi	Posio	Kemijärvi	Kuusamo	Salla	Savukoski	Koillis-Suomi
Whole sub-area *Koko osa-alue*	1720	2721	3137	4356	4998	5748	22680
State-owned forests *Valtion metsät*	679	923	948	448	2096	5040	10134
Private forests *Yksityismetsät*	984	1741	2122	3806	2864	679	12196
Commercial forests[1]) *Talousmetsät*[1])	1654	2660	3121	4187	4747	4516	20885
Nature con. areas[2]) *Luonnonsuojelualueet*[2])	66	61	16	169	251	1232	1795

1) Including hill-top sites and protection forests — *Lakimetsät ja suojametsät mukaan lukien*
2) All forests outside forestry due to nature conservation — *Varsinaiset luonnonsuojelualueet*

average, about 1500 photo points and 200 field points. The numbers of the photo points on land within the different target areas by sub-area are presented in Table 28. A separate combination is quite possible for the commercial forests in all sub-areas. The private forests and the state-owned forests can also be dealt with separately in several sub-areas. The latter division is desirable because the forests in the two ownership categories in question generally differ in many respects.

It is also possible to combine two or more sub-areas to get a sufficient amount of photo and field points to a certain target area. This solution is the only one possible with the nature conservation areas in *Koillis-Suomi.* A problem in this context is that the transfer of the field data can be several hundred kilometers which tends to increase the variation within the super cells. This is partly avoided by using the sub-area as a stratum variable in the largest super cells.

A new combination of the samples provides results the accuracy of which can be estimated independently of the original combination(s). This approach is the most simple and accurate one where the target area is distinctly smaller than the original sub-area. The other alternative is to try to predict the accuracy by the sample size. Several assumptions must then be made, the validity of which are more or less questionable.

A decrease of the sample tends to increase both the error of the photo sample and the error from the interpretation. Assuming a similar stratification, a similar distribution into the super cells and a similar variation by super cell as in the original sub-area, the error due to the interpretation can be estimated without bias by the number of the photo points and by the corresponding error at the sub-area level. In practice, however, the three qualifications are rarely fulfilled at the same time. A smaller sample in a target area generally leads to a different (coarser) stratification than in the corresponding sub-area. The stratifications are generally different if a target area includes photo points from several sub-areas. Even where the stratifications used are similar, a different distribution into the super cells is typical, for example, of the ownership categories within a sub-area. Only the variation within a given super cell can be assumed to be fairly independent of the domain area.

The error of the photo sample depends on the number of the tracts and on the number of the photo points per tract. The areal distribution of a target area has a strong effect on this error component. The item will be dealt with more thoroughly in the next section. As with the error originating from the interpretation, one can conclude that the best estimate of the error of the photo sample is obtained by the actual characteristics of the recombined samples.

5122. Original combination of the samples

The basic idea of the extension of the field data to the photo sample is to permit calculations which are independent of the original field sample. After the primary combination of the field and the photo samples by sub-area, target areas within the original sub-areas usually include external field information. This causes a bias to the estimates, here called *the bias due to the extension.* This bias arises because the expected values of the estimates do not equal the true values in the target area.

Of course, the bare field samples can also be used for the estimation in the target areas. Then the bias due to the extension is avoided but the small size of the samples make the estimates very inaccurate. The whole purpose of the photo sample is to improve the accuracy in small target areas. A comparison of the mean square errors in the field samples and in the combined field and photo sample reveals that the field samples give estimates clearly inferior to those obtained from the combined sample even at the sub-area level (p. 71). In addition, one must keep in mind that ratio estimates are always biased irrespective of the kind of the sample.

Using the original combined field and photo sample for estimation in a target area gives rise to three kinds of errors. *The error of the photo sample* can be estimated from the variation of the tract values in the target area. This approach takes into account the area-specific features in the best possible way provided that the target area is large enough for this purpose. The nature of the areal distribution affects the increase of the error of the photo sample with decreasing size of a target area. The effect of a random distribution can be studied by omitting randomly an increasing part of the photo plots from the original combined sample. Denoting the mean square error of the tract values in the total sample by 100 gave following relative errors in *Kuusamo:*

Characteristic[1])	Sample size				
	Total	80 %	60 %	40 %	20 %
1	100	106	130	164	280
2	100	97	97	141	204
3	100	98	102	147	254
4	100	115	136	182	246
5	100	101	115	126	185
6	100	101	115	121	173
7	100	108	121	130	210
8	100	103	124	136	197

[1]) See Table 4, footnote 1, p. 27

As a rule of thumb, the mean square error of the tract values becomes 2,0—2,5-fold when the random sub-sample is decreased to one fifth of the total sample. This happens because the number of the tracts stands almost unchanged. The simulation reveals that the number of the photo plots per tract can be decreased 30—40 % without a significant increase of the error of the photo sample. As such, the result is an evidence for a sufficient sampling intensity within the photo tracts at the sub-area level. The relative MSE of the photo sample in a random sub-sample of 20 % by sub-area and characteristic is:

Sub-area	Characteristic[1])							
	1	2	3	4	5	6	7	8
Pelkosenniemi	174	127	128	361	218	620	296	228
Posio	226	282	228	251	173	170	372	234
Kemijärvi	185	223	231	193	344	347	399	215
Kuusamo	280	204	254	246	185	173	210	197
Salla	153	181	195	227	291	318	424	320
Savukoski	207	273	333	152	175	145	276	190

[1]) As above

As can be expected, there are considerable deviations from the general trend. The deviations are mainly due to area-specific and characteristic-specific features. The effects of random accidents become apparent in characteristics 2 and 3 in the sub-sample of 80 % and 60 % in *Kuusamo.*

The results of the simulation are not, as such, applicable to true target areas within the sub-areas. A true target area rarely has a random areal distribution. In addition, the population parameters of a true target area usually differ from those of the whole sub-area. The former factor affects the number of the tracts in the sub-sample. Both factors affect the variation of the tract values. So it is desirable to estimate the error of the photo sample separately for each target area of interest. Some examples from *Kuusamo* are now presented. The target areas in question are:

0. Total sub-area (100 %)
1. All private forests (87 %)
2. Private forests without common forests (70 %)
3. Common forests (17 %)
4. A random sub-sample of 20 %

The sample sizes and the estimated values of the three main characteristics by target area are in Table 29. The number of the tracts in the random sub-sample of 20 % is almost the same as in the total sample. The number of the tracts decreases clearly with decreasing area of the true sub-are. However, the decrease of the tracts is smaller than that of the plots. The error of the photo sample, expressed as relative standard error in the different cases, is (cf. Table 19, p. 57):

Target area	Characteristic		
	p	$\bar{x}$	$\bar{x}_p$
		%	
0	1,75	4,32	3,91
1	1,72	3,86	3,47
2	1,84	3,83	3,53
3	3,85	5,80	6,16
4	2,74	5,90	7,24

The relative errors concerning volumes are smaller in the common forests than in the random sub-sample of 20 %. This occurs despite the fact that there are less points and especially tracts in the common forests. It can be concluded that the variation of the tract values is relatively smaller in the common forests than in the other forests of *Kuusamo.* The example indicates that the population parameters are not without importance in the estimation of the error of the photo sample.

One form utilizing the combined field and photo samples is to study the forest resources in the different parts of the sub-areas. The sub-areas can be mechanically divided into halves and quarters for that purpose. For sake of demonstration, the error of the photo sample is studied, in this investigation, by dividing each sub-area into parts, as follows:

1. NW = north-western part
2. NE = north-eastern „
3. SW = south-western „
4. SE = south-eastern „

Table 29. Some characteristics concerning the target areas in *Kuusamo*.
Table 29. Kuusamon sisäisiä osa-alueita koskevia tunnuslukuja.

Characteristic Tunnusluku	Target area[1]) — *Kohdealue*[1]) 0	1	2	3	4
Number of — *Lukumäärä:*					
— tracts on land *maalohkot*	74	73	63	29	73
— tracts on forest land *metsämaalohkot*	72	72	63	28	71
— points on land *pisteet maalla*	4 066	3 523	2 829	694	820
— points on forest land *pisteet metsämaalla*	2 730	2 378	1 884	494	545
Proportion of forest land, p (%) *Metsämaan osuus*	67,14	67,50	66,60	71,18	68,29
Mean volume on land, $\bar{x}$[2]) *Keskitilavuus maalla*	40,62	39,62	35,44	56,68	40,45
Mean volume on forest land, $\bar{x}_p$[2]) *Keskitilavuus metsämaalla*	56,96	55,66	50,31	76,06	56,33

[1]) The target areas are defined in the text — *Kohdealueet määritellään tekstissä*
[2]) Cu.m. per hectare including bark and usable dead trees — *m³/ha kuorineen mukaan lukien käyttökelpoiset kuolleet puut*

5. N = northern part
6. S = southern ,,
7. W = western ,,
8. E = eastern ,,

To make the trends more readily perceivable the mean value of the error of the photo sample at the different area levels (1—4 = 25 % and 5—8 = 50 %) are calculated. Denoting the mean square error of the photo sample in the whole sub-area as 100 results in relative mean square error by sub-area, characteristic and area level as follows:

Sub-area and characteristic		Area level 50 %	25 %
Pelkosenniemi	p	174	282
	$\bar{x}_p$	223	510
	$\bar{x}$	197	473
Posio	,,	188	350
		189	289
		197	367
Kemijärvi	,,	180	309
		199	412
		192	385
Kuusamo	,,	205	420
		200	410
		202	413
Salla	,,	203	430
		194	345
		190	332
Savukoski	,,	203	410
		186	342
		188	361

The results are consistent in that the ratio between the observed mean square errors at the different area levels is approximately the inverse of the ratio of the areas.

The examples discussed above range from randomly scattered target areas to compact parts of sub-areas. The extreme cases do not necessarily represent the minimum and the maximum error of the photo sample as indicated by the case of the common forests in *Kuusamo*. This is partly due to the differences of the population parameters in the true target areas. The population parameters being the same, the interrelations of the variations between and within the tracts affect the final outcome. A more thorough analysis of this item falls, however, outside the scope of this investigation.

The estimation of *the error due to the interpretation* in a true target area requires knowledge of the distribution of the photo points into the original super cells. The following discussion concerns the case that all photo points are furnished with the mean values $\bar{x}_h$ and p_h of the photo strata. As to the estimation of p (proportion of forest land) and $\bar{x}$ (mean volume on land), the estimator for the error due to the interpretation reads (as to the notation see p. 60).

$$(35) \quad v_I(e) = \sum_{h=1}^{L} \frac{w_h s_h'^2}{n'} \left(\frac{1}{f_h} - 1\right) + \sum_{h=1}^{L} \frac{s_h'^2 (n - n_h)}{n m^2}$$

where n and m refer to the numbers of the plots (plot = cluster or single point, see p. 62) in the photo sample and in the field sample falling in the target area, $f_h = m_h/n_h$ and n' = total number of the photo plots in the original sub-area. The first term on the right approximates the expected difference of the mean square errors in the field sample and in the photo sample falling in the target area. The value of the latter term on the right (the effect of the g-function) increases with decreasing size of target area because m is decreased. The increase is justified considering that the estimation of the stratum weights is the more unprecise the smaller is the target area. As to the mean volume on land ($\bar{x}_p$), the corresponding estimator reads

(36)

$$v_I(\bar{x}_p) = \frac{1}{p^2} * \sum_{h=1}^{L} \frac{p_h w_h s_h'^2}{n'} \left(\frac{1}{f_h} - 1\right) + \sum_{h=1}^{L} \frac{s_h'^2 (n - n_h)}{nm^2}$$

where p and p_h refer to the estimated proportions of forest land in the target area. Assuming that $E(s'^2_h)$ and $E(f_h)$ are constant by super cell, the best estimates for each s'^2_h and f_h are obtained from the total original combined sample independent of the target area in question.

Concerning the case where the photo points are furnished with the mean values by photo stratum, the two error components and the total error expressed as relative standard error in *Kuusamo* by characteristic and target area are:

Target[1]) area	Characteristic[2]) p 1	p 2	p 3	$\bar{x}$ 1	$\bar{x}$ 2	$\bar{x}$ 3	$\bar{x}_p$ 1	$\bar{x}_p$ 2	$\bar{x}_p$ 3
					%				
0	1,75	1,92	2,60	4,32	4,88	6,52	3,91	4,80	6,19
1	1,72	1,94	2,60	3,86	5,02	6,33	3,47	4,90	6,00
2	1,84	2,06	2,76	3,83	5,31	6,55	3,53	5,18	6,24
3	3,85	3,71	5,35	5,80	6,86	8,98	6,16	6,17	8,72
4	2,74	3,04	4,09	5,90	7,01	9,16	7,24	6,28	9,58

[1]) As above

[2]) $1 = 100 * \sqrt{v_p(e)}/e$

$2 = 100 * \sqrt{v_I(e)}/e$

$3 = 100 * \sqrt{v_p(e) + v_I(e)}/e$

The error due to the interpretation in *Kuusamo* (sub-area 0) is now not quite the same as presented earlier in this investigation (Table 22, p. 63). The differences are due to the different approach in the estimation of this error component. The disagreements are, however, insignificant. The values of $v_I(e)$ in target area 4 (random sub-sample of 20 %) are not the same as in the whole sub-area, although the w'_hs are supposed to be equal. This is because the value of the g-function is 25-fold in target area 4. A comparison of these target areas reveals that the estimation of the stratum weights have a strong effect on the accuracy of the results in small target areas.

A separate combination of the field sample and the photo sample is possible in target areas 1 and 2 in *Kuusamo*. The error characteristics are then as follows (notation as above):

Target area	Characteristic p 1	p 2	p 3	$\bar{x}$ 1	$\bar{x}$ 2	$\bar{x}$ 3	$\bar{x}_p$ 1	$\bar{x}_p$ 2	$\bar{x}_p$ 3
					%				
1	1,66	2,06	2,64	3,78	5,18	6,41	3,41	5,06	6,10
2	1,98	2,39	3,10	3,87	6,45	7,52	3,56	6,38	7,30

The stratifications used in the separate combinations are almost the same as in the original combination of the samples. As can be expected, the error of the photo sample in the target areas is about the same as that estimated on the basis of the original combination. The error from the interpretation, on the other hand, is now greater, which is mainly due to the decreased field sample. The magnitude of this error component in a new combination can also be predicted on the basis of the original combination. This is accomplished by using n instead of n' in the estimators (formulae 35 and 36). The procedure implies similar stratifications in the combinations. Estimated in this way, the error from the interpretation in target areas 1 and 2 is (expressed as relative standard error):

Target area	Characteristic p	$\bar{x}$	$\bar{x}_p$
1	2,09	5,41	5,27
2	2,45	6,32	6,14

The values are not very different from those estimated above from the new combinations. The differences are due to the fact that the estimated values of s'^2_h and f_h are not the same in the different combinations.

The practical realization of the extension of the field data is achieved by randomly distributing the true field plot data to the photo points within the super cells. So the expected variation of the values of the photo points in a target area is the same as in the corresponding sub-area by super cell. The variance of the stratum mean values is greater in target areas than in sub-areas. Consequently, both sub-components of the error due to the interpretation increase by decreasing size of target area. In this respect the case is, in fact, the same as with separate combination of the samples.

Using the original combination of the samples for the estimation in a target area leads to *bias due to the extension* which lessens the usability of the results. In this respect there are differences between the characteristcis. The outcome depends primarily on the stratum variables used in the extension of the field information. In principle, bias due to the extension is avoided if the corresponding variable has been used as stratum variable *and* the classification used in the photo-interpretation is unambiguous with the classification made in the field. Consequently, the following characteristics are estimated without bias due to the extension:

— land use class distribution
— division into mineral soils and peatland
— division into drained and undrained sites
— mean volume of the total growing stock

The estimated proportions of the treatment classes can be biased because each photo class includes more than one treatment class. Other characteristics possibly associated with this error component are:

— proportions of the forests dominated by pine, spruce or deciduous trees
— division of the mean volume into tree species and timber assortments
— site type and taxation class distribution
— all characteristics concerning silvicultural measures
— stand quality distribution

As to the eight main characteristics, the statistical accuracy of which has been presented in this investigation (Table 23, p. 65), the possibility of *bias due to the extension* is as follows:

Characteristic	Possibility of bias No	Yes
Proportions:		
Mineral soils	x	
Forested area	x	
Forest land	x	
Mature stands		x
Mean volumes:		
Growing stock on land	x	
Growing stock on forest land	x	
Pine on land		x
Saw-timber on land		x

In the case that a bias due to the extension can occur, its magnitude is determined by the difference between the true values of a characteristic in the target area and in the sub-area. To verify and quantify a difference is a difficult task in itself. It can be attempted with the aid of the field sample or by the combined field and photo samples using separate combinations. The result is, however, very inaccurate due to the high sampling error associated with the estimation of differences. A subjective idea of the matter formed by practical experience can, in this context, be better than an objective estimate. So a practical forestry professional can, with good reason, claim, for example, that the proportion of treeless cutting areas in the forests of a certain ownership category is higher or lower than in the whole sub-area on average.

The effect of a bias on the accuracy of an estimate can be quantified if the absolute bias is known. The total mean square error is calculated as

$$\mathrm{MSE} = v(e) + B^2$$

where $v(e)$ refers to the sampling error expressed as absolute variance or mean square error and B to the absolute bias. Let e_1 and e_2 refer to two estimates the sampling errors of which are known ($v(e_1) < v(e_2)$) and suppose that e_1 is associated with an extra bias of magnitude B. It is then worth utilizing estimate e_1 if

$$v(e_1) + B^2 < v(e_2) \rightarrow B < \sqrt{v(e_2) - v(e_1)}.$$

This ~~can~~ also be expressed by the ratio of B to $\sqrt{v(e_1)}$. e_1 is preferable if $B/\sqrt{v(e_1)}$ is smaller than critical ratio R which is calculated as

$$(37)\quad R = \sqrt{\frac{v(e_2) - v(e_1)}{v(e_1)}}$$

The critical ratio can be calculated for any pair of estimates if only the variances or mean square errors are known. In Table 30 are given the absolute mean square errors for some characteristics in certain sample alternatives in some target areas of *Kuusamo*. The three characteristics chosen for the comparison are those main characteristics which can be associated with bias due to the extension in sample 1. The stratification made in target area S for alternative 2 is somewhat coarser than in the other target areas. The error characteristics of target areas 1, 2 and S in alternative 1 represent the hypothetical case that all photo points are furnished with stratum mean values. As to the absolute mean square error, alternative 1 is in several cases inferior to alternative 2 in target areas 1 and 2, even though bias due to the extension is not considered. This happens because the population variation of the characteristics in question is greater in the whole sub-area than in the target areas. Alternatives 1 and 2 give relative error in target areas 1 and 2 as follows:

Characteristic	Target area 1 alt. 1	Target area 1 alt. 2	Target area 2 alt. 1	Target area 2 alt. 2
		%[1])		
p_m	11,80	11,68	13,12	15,17
$\bar{x}_{pine}$	9,70	9,87	10,19	12,12
$\bar{x}_{s-t}$	9,55	9,42	10,15	11,41

[1]) 100 * $\sqrt{MSE}$ /e

A separate combination of the samples is highly recommendable in target area 1. It is also preferable in target area 2 if one suspects a noticeable bias due to the extension. A comparison of alternatives 3 and 1 gives the following critical ratios (formula 37):

Characteristic	Kuusamo	Target area 1	Target area 2	Target area S
p_m	1,00	1,13	0,73	1,19
$\bar{x}_{pine}$	1,34	1,51	1,47	2,16
$\bar{x}_{s-t}$	1,20	1,52	1,11	2,14

Let the meaning of the critical ratios be illustrated by an example. The value of p_m in target area 1 estimated from sample 1 is 16,55 % and $\sqrt{v(p_m)}$ = 1,95 %. The maximum allowable bias of sample 1 is 2,20 % (= R * 1,95 %). The field sample gives a more accurate estimate for p_m if the actual bias exceeds the allowable maximum in sample 1.

The usability of the original combination of the samples decreases with decreasing size of target area for two reasons, even in the case that the photo points are furnished with the mean values by photo stratum. First, bias due to the extension tends to be the greater the smaller is the target area in relation to the total sub-area. Secondly, the estimation of the stratum weights is inaccurate in small target areas, which increases the error from the interpretation. Bias due to the extension is avoided by making a separate combination of the samples with a coarse stratification of the photo sample. The variation within the super cells is greater than in the original combination which tends to increase the error due to the interpretation. The outcome is not quite sure in advance, however, because the estimation of the stratum weights is more accurate with a smaller number of super sells. The following example concerns the common private forests of Kuusamo. The sample alternatives are

A = Original combination of the samples at the sub-area level (photo points with mean values)
B = Hypothetical separate combination of the samples with the same stratification as above
C = Separate combination of the samples with a coarse stratification of the photo sample

The error due to the interpretation in alternative B is estimated by using n instead of n' in the estimators (formulae 35 and 36). The number of the super cells is 27 in alternatives A and B and 9 in alternative C. The error characteristics in the different cases are as follows:

Characteristic		Sample alternative A	Sample alternative B	Sample alternative C
p_m	1[1])	9,03	9,03	6,03
	2[2])	11,39	19,76	16,63
	3[3])	14,53	21,73	17,69
$\bar{x}_{pine}$	”	7,47	7,47	8,79
		10,66	19,11	16,07
		13,02	20,52	18,32
$\bar{x}_{s-t}$	”	8,25	8,25	8,11
		8,71	15,75	11,71
		12,00	17,78	14,25

[1]) 100 * $\sqrt{v_p(e)}$ /e
[2]) 100 * $\sqrt{v_I(e)}$ /e
[3]) 100 * $\sqrt{v_p(e) + v_I(e)}$ /e

Table 30. The absolute mean square error of the estimated proportion of mature stands (p_m), mean volume of pine on land ($\bar{x}_{pine}$) and mean volume of sawtimber ($\bar{x}_{s-t}$) in some sample alternatives by target area in *Kuusamo*.

Taulukko 30. Uudistuskypsien metsien osuuden (p_m), männyn keskitilavuuden ($\bar{x}_{pine}$) ja tukin keskitilavuuden ($\bar{x}_{s-t}$) arvion absoluuttinen keskineliövirhe eräissä näytevaihtoehdoissa Kuusamon osa-alueissa.

Target area¹) *Kohdealue*¹)	Sample alternative²) — *Otosvaihtoehto*²) 1			2 Characteristic — *Tunnusluku*			3		
	p_m ³)	$\bar{x}_{pine}$	$\bar{x}_{s-t}$	p_m ³)	$\bar{x}_{pine}$	$\bar{x}_{s-t}$	p_m ³)	$\bar{x}_{pine}$	$\bar{x}_{s-t}$
Kuusamo	3,80	4,07	2,04				7,63	11,41	4,95
1	3,81	4,26	1,83	3,59	4,54	1,88	8,64	13,91	6,03
2	3,20	3,60	1,83	2,93	4,39	1,61	4,93	11,42	4,09
S	4,94	4,93	2,13	9,98	10,50	5,67	11,95	27,97	11,90

¹) 1 = Private forests — *Yksityismetsät*
2 = Private forests without common private forests — *Yksityismetsät ilman yhteismetsää*
S = The southern part of Kuusamo — *Kuusamon eteläpuolisko*
²) 1 = Original combination of the samples made for sub-area *Kuusamo* — *Alkuperäinen Kuusamon yhdistetty ilmakuva- ja maastonäyte*
2 = Separate combination of the samples made for the target area in question — *Osa-alueen ilmakuva- ja maastonäyte yhdistettynä uudestaan*
3 = Bare field sample — *Pelkkä maastonäyte*
³) Multiplied by 10 000 — *Kerrottuna 10 000:lla*

The small size of the target area in question does not permit the estimation of the error characteristics in the field sample accurately enough. It appears that the error due to the interpretation is smaller with the coarser stratification. So the decrease of the value of the g-function more than compensates the effect of the increased variation within the super cells. Alternative A is clearly better than alternatives B and C assuming that considerable bias due to the extension does not occur. This qualification is, in principle, fulfilled with the main characteristics other than the three discussed above.

52. Additional sampling

521. Increase of the photo sample

5211. Effect on the error of the photo sample

The error of the photo sample can be decreased by increasing the number of the photo plots. The additional photo plots can be located within either the existing or new photo tracts. The former alternative decreases the variation of the tract values. It has been contented earlier in this investigation that photo plots can be randomly omitted at the sub-area level some 20—30 % without a marked increase of the error of the photo sample (p. 74). Increasing the sampling intensity within the photo tracts decreases this error component very slowly at the sub-area level. A more rapid decrease is attained by establishing new photo tracts. Assuming that the variation of the tract values stands unchanged, the percentage decrease of the mean square error of the tract values follows a rule

$$D_{MSE}\% = \frac{100\,p}{100+p}$$

where p refers to the increase of the number of the tracts in per cent. The present photo sample includes 2/3 of the possible photo tracts. So the number of the photo tracts can be increased 50 % at the most. The corresponding decrease of MSE of the tract values is 33 %.

One possibility to increase the photo sample is to broaden the interpretation within the present photo tracts in the west-east direction. The present breadth of the photo strip is 5 km, which is mainly determined by the technical details of the photo material. A broader strip often leads to the lack of stereo coverage on either side of the tract. This is not, however, the critical point because the additional photo plots can be located on either side of the present tracts. The magnitude of the increase must, however, be the same over the whole study area. It can be concluded that this kind of additional sampling decreases the variation of the tract values more than is the case when the additional plots are located within the present 5 km photo strip.

The single photo tracts between the continuous photo strips (Figure 5, p. 16) also make it possible to interpret more plots on the present photo material in the south-north direction. This is because each aerial photograph covers 13 km longitudinally. The minimum possible addition in the south-north direction is 4—5 km per single photo tract. The possible increase is 10 km when the single photo tract is located on two aerial photographs in the south-north direction. So establishment of new photo tracts is to some degree possible even with the present photo material.

A marked increase of the photo tracts requires the acquisition of new aerial photographs between the present single photo tracts. A photo coverage of 100 % is attained by one new stereo pair per 384 sq. km., on average. The corresponding additional cost of the photo material is about 300 Fmk. Assuming that the sampling intensity within the tracts is not changed, the increase in the number of the photo tracts by 50 % affects the costs by sub-area as follows:

Sub-area	Cost[1]) Fmk 1	2
Pelkosenniemi	1400	1900
Posio	2400	3200
Kemijärvi	2800	3700
Kuusamo	3900	5200
Salla	4500	6000
Savukoski	5000	6700

[1]) 1 = Cost of the additional photo material
2 = Cost of the additional interpretation

The mean square error of the tract values is decreased by 33 % with the above arrangement. A further decrease can be achieved by increasing the sampling intensity within the photo tracts. As stated before, however, this measure affects the MSE much less than an increase of the tracts at the sub-area level. The number of the photo plots in the present tracts should be doubled to obtain the same decrease of the error of the photo sample as achieved by increasing the photo tracts by 50 %. The former alternative means the interpretation of 160 additional photo plots per 384 sq.km. instead of 80 plots in the latter alternative. The cost of the additional photo material is saved in the former alternative but the total cost is greater than in the latter alternative because the cost of the additional interpretation is doubled. In other words, increasing the photo tracts is a more efficient way to decrease the MSE of the tract values than increasing the sampling intensity within the present photo tracts. The difference would be even smaller, or perhaps vanish altogether, if the new plots were partly located outside the present photo tracts in the west-east direction.

The above consideration justifies a ranking of the alternatives in increasing the photo sample at the sub-area level as follows:

A. Increased sampling of the photo tracts
B. Increased sampling within the photo tracts
   B1. The new plots are partly located outside the original photo strips (broadening)
   B2. No broadening

It is possible that alternative B1 is almost as efficient as alternative A. A greater decrease than 33 % of the MSE of the tract values is possible by alternative B alone or by a combination of A and B.

The above discussion concerns an increased photo sample at the sub-area level. The situation is somewhat different when the increase of the photo sample concerns only a small target area within a sub-area or a set of sub-areas. Suppose a random target area of 20 % which occurs within almost every original photo tract. An increase of the photo tracts decreases the error of the photo sample in the target area in the same relation as in the whole sub-area or set of sub-areas. Suppose that an increase of the sampling intensity within the present photo tracts (alternative B) does in the same way. It has been noticed that doubling the sampling intensity within the tracts decreases the mean square error of the tract values by some 33 %. The absolute cost of this measure in the random target area of 20 % is one fifth of the corresponding cost in the whole sub-area or set of sub-areas. About the same relative decrease of the error of the photo sample is achieved by increasing the number of the photo tracts by 50 %. The cost of alternative A is about the same or little smaller than the cost of alternative B for the whole sub-area or set of sub-areas. But alternative A is more expensive than alternative B in any random-like small target area when the increased sampling concerns only the target area. This is because the cost of the additional photo material is the same at both area levels. The area-dependent difference in the efficiencies of alternatives A and B becomes smaller with increasing relative size of target area.

Concerning true target areas, the order of the alternatives of increasing the photo sample cannot be determined in advance. The choice depends on the location of the target area with respect to the present photo sample. A situation can arise where the target area is located mainly outside the present tract net and a considerable increase of the photo sample in the target area is possible only by increased coverage of the photo tracts. The opposite is, of course, just as possible. Consequently, the alternatives must be considered separately for each case.

An increased sampling of the photo plots always implies a recombination of the field sample and the photo sample. There are two possibilities for combining the samples where the photo sample has been increased only in a target area within a sub-area or a set of sub-areas. The target area being great enough the samples can be recombined independent of the sub-area(s). For small target areas the recombination must be done at the sub-area level. The resulting characteristics of the target area are then associated with bias due to the extension. It is, however, more important to remember that the resulting characteristics in the whole sub-area or a set of sub-areas also become biased. This is because the increased photo sample in the target area also affects the estimates outside the target area when the recombination of the samples is carried out at the sub-area level. The bias in question is avoided only where the distribution of the photo plots into the super cells is the same at the both area levels. It seems reasonable to use this kind of combined data for the estimation only in the target area where the photo sample was increased.

5212. Effect on the error due to the interpretation

The error due to the interpretation can be approximated as (cf. p. 62).

$$(38)\quad v(e)_I = \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h}\left(1-\frac{m_h}{n_h}\right) + \sum_{h=1}^{L} \frac{s_h^2(n-n_h)}{n} * \frac{1}{m^2}$$

where n refers to the number of the photo plots and m to the number of the field plots. It is readily seen that an increase of the photo plots increases the value of the first term on the right because the value of m_h/n_h ($= f_h$) is decreased. The value of the second term on the right remains unchanged because $E[(n-n_h)/n]$ is independent of n. An increase of the photo plots from n to n' causes an absolute increase in the error due to the interpretation

$$(39)\quad I_{v(e)_I} = \frac{m}{n}\left(1-\frac{n}{n'}\right) * \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h}$$

The maximum increase approaches

$$I_{v(e)_{I\max}} = f * \sum_{h=1}^{L} \frac{w_h^2 s_h^2}{m_h}$$

when n/n' approaches zero. The value of m/n in the field and photo sample of Lapland is 13,33 %. The relative increase of the value of the first term on the right (RI) is as follows:

	Increase of photo plots %	RI %
	20	2,6
	40	4,4
$A_{max.}$		
	60	5,8
	80	6,8
$B_{eq.}$	100	7,7
	150	9,2
	200	10,3
	300	11,5
	400	12,3
	.	.
	.	.
	.	.
	∞	15,4

A and B above refer to the alternative means of increasing the photo sample discussed in the previous section. The value of the second term on the right in formula 38 takes 5—10 % of the error due to the interpretation at the sub-area level. So the above percentages only slightly overestimate the increase of the total error component at the sub-area level. Alternative A improves still further compared with B when the effect of an increased photo sample on the error due to the interpretation is taken into account. The reason is that the number of the new photo plots is smaller in alternative A.

An increase of the photo tracts by 50 % increases the error due to the interpretation by some 5 % ($A_{max.}$). Relatively the same decrease of the MSE of the tract values (33 %) is achieved by doubling the sampling intensity within the present photo tracts, which increases the error from the interpretation by some 7—8 % ($B_{eq.}$). The two main error components are about equal at the sub-area level. So the total mean square error of the estimates is decreased by 13—14 % at the sub-area level.

The total effect of an increased photo sample on accuracy depends on the structure of the total mean square error of the estimates. Where the photo points are furnished with the stratum mean values, the relation between the two main error components seems to stand unchanged when results are calculated for target areas from the original combination of the samples. However, the structure of the error due to the interpretation is changed in the way that the error arising from the g-function takes the greater proportion the smaller is the target area. So the relative increase of this main error component caused by an increased photo sample decreases with decreasing size of target area when the samples are combined at the sub-area level. In this case the efficiency of increasing the photo sample increases with decreasing size of target area.

The total error due to the interpretation becomes greater when the samples are combined separately for a target area. The increase of this error component is greater than the increase of the error of the photo sample. It also happens that the relative contribution of the g-function to the total error due to the interpretation decreases. The decrease is the greater the coarser is the stratification. Both factors decrease the efficiency of increasing the photo sample in a target area when the samples are combined separately for the target area in question. It is not impossible to imagine a situation where the absolute increase of the error from the interpretation is greater than the absolute decrease of the error of the photo sample.

5213. The order of the possible applications

In conclusion, the possible applications of increasing the photo sample can be ranked with respect of the relative efficiency as follows:

1. An increased photo sample in a target area. The plots are located according to alternative B1 or A (p. 80). The samples are combined at the sub-area level.
2. An increased photo sample in the whole sub-area. The plots are located according to A or a combination of A and B1.
3. As case 1, but the samples are combined separately for the target area.

A realization of case 2 implies a realization of case 1. Case 1 alone generally implies a greater relative increase of the photo sample than for case 2. Case 3 is not reasonable in small target areas. The limit is determined by the ratio between the main error components and the structure of the error from the interpretation.

522. Increase of the field sample

The error due to the interpretation can be decreased by increasing the field sample. An increase of the field sample affects both subcomponents as can be seen from formula 38. However, the relative decrease is greater with the g-function. The change of the total error component can be studied by simulation. To see the effect, the number of the field plots in the original sub-areas was both gradually decreased and increased and the error due to the interpretation was estimated at each sample size level. It appears that the relative change is almost the same for the eight main characteristics. The relative mean square error from the interpretation by sub-area and sample size is:

Sub-area	Sample size[1])						
	0,5	0,75	1,00	1,25	1,50	1,75	2,00
Pelkosenniemi	234	142	100	76	60	49	40
Posio	237	143	100	75	59	47	39
Kemijärvi	227	141	100	76	61	50	42
Kuusamo	224	140	100	77	61	50	42
Salla	231	141	100	76	60	49	41
Savukoski	227	141	100	76	61	50	42

[1]) In relation to the original field sample

The pattern of the relative change is independent of sub-area. Halving the field sample more than doubles the error due to the interpretation. On the other hand, the error in question is decreased by some 60 % when the field sample is doubled.

It is now possible to compare the efficiencies of increasing the field sample and the photo sample at the sub-area level. Let the goal be to decrease the total mean square error by 14 %. The cheapest alternative of increasing the photo sample costs 700 Fmk per a unit area of 384 sq.km. The same decrease of the total MSE requires 30 % more field plots which costs 1100 Fmk per unit area. Consequently, increasing the photo sample is more efficient in this case. However, the relative efficiency of field sampling increases when the required decrease of the MSE is increased.

The error of the photo sample contributes to the total MSE less in target areas within a sub-area than it does at the sub-area level. So the relative efficiency of increasing the field sample is greater in target areas. This happens although a certain relative increase of the photo sample is more effective in target areas within a sub-area than at the sub-area level. In small target areas it seems reasonable to consider an increase of both samples. The best solution is very area-specific, why the construction of a general decision model is not attempted in this investigation.

6. GENERAL CONSIDERATIONS

The two-phase systematic sampling design studied in this investigation is a sequel to experiments in the use of aerial photographs in forest inventories in Finland. The series of investigations reaches back to the latter half of the 1960's and includes small-scale as well as large-scale practical applications and theoretical considerations (Poso 1969 and 1972, Poso and Kujala 1971 and 1977, Nyrhinen 1976). The present study was commenced in 1982 in *Kuusamo* (Mattila 1983).

The latest innovations in the inventory method used in Lapland include the totally systematic field sample, the establishment and remeasurement of semi-permanent field plots, the requirement of improved usability of the inventory results at the sub-area level and the partial adoption of the multi-resource inventory idea. Perhaps the greatest progress, however, concerns practical data management. The computer-based system developed in the course of the present study allows rapid recombination of the field sample and the photo sample for whichever reference area independent of the original sub-areas.

The performance of an inventory design can be expressed in a general functional form as (Aldred 1971, p. 15)

$$P = f\,(X_1 \ldots X_i \ldots X_m;\ Y_1 \ldots Y_j \ldots Y_n)$$

where X refers to the control variables and Y to the environmental or "given" variables. The goal in planning a forest inventory is to maximize the value of P, for example the accuracy of the results at a certain cost level. The "planning environment" in Lapland strongly affects the sampling design. The prefixed variables are:

— characteristics of the study area
— inventory budget
— precision goal
— photo material
— kind of plot and sampling unit
— type of sampling (two-phase systematic)

The only control variables are, in fact, those concerning sampling intensity and location of the samples in both sampling phases.

It seems rational to consider a substitution of satellite imagery for the present photo material provided that the spatial resolution is improved to some 10 m. According to the preliminary information, the spatial resolution of the material produced by satellite SPOT to be launched in 1985 will be 10—20 meters (Hardy 1981, p. 30, Astermo 1984). Satellite imagery can be utilized either in analog form as conventional aerial photographs or in digital form which permits automatic interpretation on the basis of the spectral characteristics. The latter alternative generally implies a representation of the information also in analog form for field sampling.

The relascope technique is no doubt superior in bare field sampling in the estimation of the characteristics of the tree stock. For the purpose of the assessment of volumes from small-scale aerial photographs, however, the relascope plot with factor 2 used in Lapland is too small a unit. The problem can be reduced by using a smaller factor or by using a greater fixed-sized plot. The latter alternative, at least would lessen the efficiency of the field sampling from the present level.

The plot used for the inventory in Lapland is a cluster of two relascope plots 40 meters apart. Using satellite imagery with automatic interpretation would probably lead to a single relascope plot or a different cluster of relascope plots than the present one. It is worth noting that with automatic interpretation, the relascope plot is indisputably the best alternative for acquiring the field information. In the case that satellite pictures are used as conventional aerial photographs, the disagreements between the interpreted volumes and the corresponding measured volumes are greater than with usual photographs because the spatial resolution is poorer.

Is two-phase sampling with the present photo material efficient? The question is now discussed, concerning characteristics area of forest land (A) and total volume of the growing stock (V). The ratio of the mean square error of the new field sample to the mean square error of the combined field and photo sample by sub-area is:

Sub-area	*Characteristic* A	V
Pelkosenniemi	1,97	4,27
Posio	4,06	3,25
Kemijärvi	1,83	3,01
Kuusamo	4,39	3,04
Salla	2,61	3,33
Savukoski	4,12	3,44
Koillis-Suomi	3,41	3,25

The mean square errors of the field sample have bee approximated on the basis of the generalizations described in page 28. Despite this, considerable deviations from the general trend with A still occur.

The mean square errors of the new field sample are, on average, more than 3-fold compared with the mean square error of the combined field and photo sample. Consequently, the number of the tracts in the new field sample should be increased by some 220—240 % to reach the same accuracy as that attained by the photo sample, assuming that the variation between the tract values is not changed. Practically almost every tract should be measured as a field tract, with the present sampling intensity within the tracts.

The cost of the new field sample in *Koillis-Suomi* was 260 000 Fmk and the cost of the photo material and the additional training needed 80 000 Fmk. The cost of the photointerpretation need not be included because the task could be executed by the permanent personal. The total field and photo cost was 340 000 Fmk, which is only 40 % of the field cost required in the case that the same accuracy had been acquired by only increasing the field fracts. So with *the present field tract* of Lapland, the two-phase sampling design is a highly efficient solution.

A true alternative for the national forest inventory in Lapland is to use the same field-based method as in South Finland. An efficiency comparison with the two-phase sampling is difficult because the variation of the tract values with the new tracts, under the conditions of *Koillis-Suomi,* is not known. The field tract used in South Finland is measured in one day and all units of 8 km * 8 km are sampled. The cost in *Koillis-Suomi* would have been 640 000 Fmk which is almost double the corresponding cost of the two-phase sampling design. The statistical precision would, however, have probably been better.

A trial derivation of the variability of the tract values with the tracts of South Finland in the conditions of Lapland is possible by certain error characteristics of the forestry board districts of *Kainuu, Pohjois-Pohjanmaa* and *Koillis-Suomi.* At the same time, the different sizes of the sub-areas and the differences in the population parameters must be taken into account. The valid data concerning the proportion of forest land and mean volume on forest land are (as to the references see p. 68):

Sub-area	Land area sq.km	P	CV_P %	$\overline{X}_P$ m³/ha	$CV_{\overline{X}P}$ %
Kainuu	21760	76,9	0,9	66,1	2,6
Pohjois-Pohjanmaa	24130	66,1	1,5	57,8	2,1
Koillis-Suomi	25670				
1982—83		67,2	1,0	47,9	2,1
1976		69,5	1,4	46,3	2,1
1969		60,6	1,7	51,0	3,5

The tract used in *Koillis-Suomi* in 1969 was L-formed and included 26 relascope plots for the volume estimation along a line of 5100 m (Kuusela and Salminen 1969, p. 10). One field tract represented an area of 121 sq.km (Salminen 1973, p. 6). The corresponding area of tracts in the latest inventory in *Kainuu* and *Pohjois-Pohjanmaa* (1982—83), as well as for tracts in South Finland, is 64 sq.km.

The characteristics concerning mean volume on forest land in *Pohjois-Pohjanmaa* and *Koillis-Suomi* are readily compared with each other. It has been noticed that the relative variation of the tract values of the volumes increases slightly with increasing volume. Taking into account the greater mean volume and the smaller land area in *Pohjois-Pohjanmaa* one can conclude that the location of the field tracts of the size employed in South Finland could be more sparse in *Koillis-Suomi* than in *Pohjois-Pohjanmaa* to reach the same level of accuracy in these forestry board districts. The sampling intensity of the tracts cannot be decreased by

7 408500946V

50 % as indicated by the last row above. A decrease of 25 % would give an inventory cost of 480 000 Fmk which still is 140 000 Fmk or 40 % greater than the comparable actual inventory cost in *Koillis-Suomi.*

Taking into account the natural conditions, the error characteristics of the estimated proportion of forest land in *Kainuu* and *Koillis-Suomi* are more comparable than the corresponding ones of *Pohjois-Pohjanmaa* and *Koillis-Suomi*. The relatively low CV_P%-value in *Kainuu* is mainly due to the high P-value there. As to the estimation of p, one can conclude that the relative variation of the tract values would be clearly greater in *Koillis-Suomi* than it is in *Kainuu.* So there is no sure evidence that the decrease of the sampling intensity could be greater than that approximated above with the volumes.

The above conclusions concern the present L-formed field tract of South Finland (Fig. 9, p. 18) which probably is not the most efficient one in the forest conditions of Lapland. The best outcome implies a tract including fewer plots with greater intervals between the plots (cf. Hägglund 1983, p. 15). The form of the tract and especially the interval between the tracts can also be different in the best solution. It is not possible to predict, without far-reaching studies, to what extent the best sampling design would be better measured by the inventory cost. However, it is unlikely that the best field sampling design is better than the present two-phase sampling design in *Koillis-Suomi.*

The cost of photos for *Koillis-Suomi* in 1982—83 was 35—40 % greater than it would have been with the sampling design used in 1976. This is due to the additional intermediary photo tracts in the new inventory (Fig. 5, p. 16). Some 20 % more field plots were measured in 1982—83 than in 1976, which probably increased the field cost by an equivalent amount. Consequently, the total inventory cost with the old two-phase sampling design would have been some 275 000 Fmk which is 19 % less than the actual inventory cost in 1982—83. The difference can be regarded as the 'price' of the new inventory goals.

The cost comparison made above does not include the remeasurement of the old semi-permanent fiels plots. This omission is justified because these plots have not been used as a part of the two-phase sample. The following discussion, also, only concerns the two-phase samples.

Was there any rationality in the additional monetary inputs to the inventory in *Koillis-Suomi* in 1982—83? The estimated accuracy of the main results serves as one evaluation criteria. It seems that the increased field and photo sample has increased the accuracy of the estimated proportions significantly (p. 69). This is essential in the forest conditions of Lapland where site characteristics are relatively more important than in South Finland. It also seems that the accuracy of the volume estimates has not decreased, despite the use of the systematic field sample. It is possible that the distribution of the total volume into tree species and assortments has been estimated more accurately by the increased field sample.

Another evaluation criteria concerns the estimation of characteristics other than the conventional forest characteristics. In general, an increased and decentralized field sample favours multi-resource inventories. A systematic field sample is important for items that cannot be linked to the photo sample. The survey of the reindeer winter ranges integrated in the NFI in Lapland profits by the increased sampling in both sampling phases.

One reason for the increased field sample was to compensate for the harmful effects of the systematic field sample on the extension of the field data to the photo sample. However, the systematic field sample has favourable consequences with respect to the application of the SPR-technique (p. 35). The gain from this detail can be realized on the next inventory occasion.

A further aspect must be mentioned in this connection. The field sample being systematic, the interpretation of the photo sample can be done resiliently both before and/or after the measurement of the field sample. The location of the field plots must, however, be done in the field using the same aerial photographs as used in the photo-interpretation. Consequently, the photos for the field tracts must be acquired before the field work although the interpretation is done in the following winter season. This was the arrangement in *Koillis-Suomi* in 1982—83 as well as in the southern part of *Lappi* in 1984—85. In fact, it was impossible to interpret the whole photo sample before the measurement of the field sample for two reasons. Firstly, the work was done by the permanent personal and therefore had to be extended over a longer period of time. The other reason was that all aerial photographs needed for the interpretation could not be supplied before the field work. The latter fact is decisive in the sense that the last inventory would not have been possible within the normal timetable by using the original grouping method.

The above-mentioned consequence reveals an important feature of the two-phase sampling designs with systematic sampling in the second phase. The feature is great *flexibility in time* which in the case of the NFI carried out in 1982—85 in Lapland saved both time and money.

Interpretation after field work introduces a potential risk of bias in the interpretation. The field plots must not be interpreted by the field cruiser because he possibly remembers the reality of the field and interprets accordingly, rather than by the view on the photograph. This has been taken into account in the realization of the interpretation in Lapland.

Systematic sampling causes difficulties in the estimation of the sampling error on the basis of the observations as pointed out by Matérn and Ranneby (1983, p. 4). The estimation of error with special reference to forest inventories, as well as the theory of topographic variation and spatial variation in general, has been studied by Matérn (1947, 1960). Using the random sampling estimators, as is the case in this investigation, generally leads to an overestimation of the sampling error. This must be kept in mind in an evaluation of the error characteristics.

Another inconvenience with systematic sampling is that a periodicity in the study area can in one way or other fit in with the systematic location of the samples resulting in a bias in the estimates. On the basis of discussions by Milne (1959) and Matérn (1960), however, the risk of such an accident in a natural forest can be considered to be negligible.

The combination of the field sample and the photo sample was made at the sub-area level without taking into account the geographical distance between the plots. The maximum transfer of the field data can then be more than 100 km which tends to increase the bias due to the extension (p. 73).

One solution would be to use some kind of 'locative' variable as a stratum variable. This can only be done in the largest sub-areas without increasing the variation within the super cells.

Theoretically, the best way of extending the field data within a super cell is to give to each photo point a weighted mean of all field points in the same super cell. The method is called 'kriging' when the weights of the field points are chosen in an optimum way with respect of the distances and the correlation-functions (Matérn and Ranneby 1983, p. 6). The procedure in question would minimize the bias due to the extension. However, an application of the method to practice is a complex task.

The method used to extend the field data to the photo sample is inefficient with respect to the estimation of the characteristics in the target areas. This is because the expected mean square of the photo sample is the same in target areas as in the corresponding sub-areas by super cell. The increase of the error due to the interpretation is tempered by furnishing each photo point with the mean of the field points by super cell. An immediate consequence of this is that the calculational flexibility of the inventory method is lost. A better alternative is to form a group of field points for each photo point to approximate the weighted means of the photo points and use the extended data with the corresponding weights in the calculations. In its most sophisticated form, the method approaches 'kriging' which requires knowledge of the correlation functions.

A goal of this investigation was to determine the future study and development needs of the forest inventory method in question. Some items brought to prominence by the experiences discussed in this paper are listed below.

1. The use of satellite information as a complement to or a substitution for the present photo material in the first phase of sampling.
2. The kind of field plot. A change is required, especially where satellite imagery is employed (cf. Poso et al. 1984, p. 283).
3. The construction of a field tract which is efficient both in bare field sampling and in two-phase sampling.
4. The extension of the field data to the photo sample in an efficient way which preserves the calculational flexibility of the grouping method.
5. The development of a system for checking changes by the photo materials acquired on successive inventory occasions.

7. SUMMARY

This paper deals with the two-phase forest inventory method used in the 7th National Forest Inventory (NFI) in Finnish Lapland in 1982—85. The actual study area is the *Koillis-Suomi* forestry board district including six sub-areas of 1 800—6 400 sq. km. The total land area is 25 700 sq.km. Two-phase sampling was utilized in this area for the first time in the 6th NFI in 1976.

The goals set for the present inventory included achieving a systematic field sample, semi-permanent field plots, usability of the inventory results in the sub-areas and integration of multiresource aspects to the NFI. The precision goal being that the relative standard error of the estimated total volume in the forestry board district must not exceed 4 %.

The study area was divided into 8 km * 8 km squares a part of which was sampled as photo tracts. For the photo-interpretation small-scale (1 : 50 000) black and white panchromatic material was used. A part of the photo tracts was sampled in the field in the second phase of the sampling. The field plots were located systematically independent of the interpreted data. The selection of the photo tracts and the field tracts was also systematic.

The samples have been increased and decentralized compared with the previous inventory to guarantee the usability of the inventory results in the sub-areas. The inventory cost would have been some 20 % smaller with the previous sampling design. The cost comparison does not include remeasurement of the old semi-permanent field plots which were established in the previous inventory.

The field sample alone allows a preliminary estimation which in itself is one implicit goal of the new inventory. The following error characteristics concern the new field sample excluding the remeasured semi-permanent field plots. The accuracy of the estimated total volume in *Koillis-Suomi* expressed as relative standard error is 4—5 %. The corresponding figures in the largest *(Savukoski)* and smallest *(Pelkosenniemi)* sub-area are 9—10 % and 17—19 %. As to the estimated proportion of forest land, the error characteristics are 2—3 %, 4—5 % and 8—9 % correspondingly.

The location of the old semi-permanent field plots was also systematic. So, after remeasurement, these plots can be used for the preliminary estimation with the new field plots. The resulting error characteristics are then some 10—20 % lower than the values given above.

Change analysis is the main function of the semi-permanent field plots. Using the permanent tracts as units, the correlation of the total volume between the two successive measurements varied from 0,650 to 0,986 in the sub-areas and was 0,854 in the whole study area. The standard error of the observed difference is about the same as that estimated on the basis of the total combined field and photo samples of the two inventories.

The temporary field plots of the previous inventory have been selected to the sample with varying probabilities, so they cannot now be utilized according to the SPR-principles. In this sense the situation will be better on the next inventory occasion because the whole present field sample has been selected with an equal sampling intensity. The number and the location of the new semi-permanent field plots within a field tract is the same as in the old field sample.

The remeasured field plots are not used as a part of the two-phase sample. This is because their exact location should be manually transferred from the old photographs to the new ones for the interpretation. The task is laborous and it introduces a risk of bias to the interpretation.

The photo sample was interpreted both before and after the field work. This is possible because the location of the field plots

does not depend on the interpreted data. By this arrangement the task could be executed by the permanent personal which saves costs. The arrangement was also necessary because all aerial photographs could not be acquired before the field work.

The most important variables assessed are *land use class, sub-class, treatment class, dominant tree species, drainage status* and *volume*. The accuracy of the assessment of the nominal scale variables from photos was studied by a K-statistic. The agreement is very good with sub-class (mineral soils/ peatland), good with land use class and drainage status and relatively poor with treatment class and dominant tree species. The correlation between the assessed volumes and the field volumes varies from 0,516 in *Pelkosenniemi* to 0,679 in *Posio*. The results concerning the accuracy of the photo-interpretation are normal with respect to the photo material in question.

The stratum variables for the stratification of the photo sample were chosen on the basis of the photo-interpretational accuracy and the importance of the corresponding characteristic. The power of the stratum variables is different in the different main photo strata. The goal is to form homogenous strata with respect of the interpreted variables. Homogenous strata are required because the field data is randomly extended to the photo sample by photo strata. All photo plots obtain formally complete field data which makes the method flexible when calculating the results.

The effect of the stratification on certain error characteristics was studied by simulation. It was noticed that a fairly detailed stratification is recommendable. However, an error component arising from the estimation of the stratum weights tends to increase with increasing number of strata.

A computer programme was constructed for the extension of the field data to the photo sample. In addition to the samples, the programme requires a definition of the photo strata as input. The output includes the combined sample, the main results and the error characteristics of the main results. The programme allows a rapid recombination of the samples for any reference area. The output sample can be used as such in the general calculation routines of the NFI.

On average, the photo sample decreases the mean square errors of the estimates to one third. The sampling errors of the total volume (V) and the forest land area (A) expressed as relative standard error by subarea are:

Sub-area	Land area sq.km	Characteristic A	V
		CV-%	
Pelkosenniemi	1826	6,5	9,1
Posio	3096	3,5	8,0
Kemijärvi	3578	3,5	7,7
Kuusamo	5006	2,6	6,5
Salla	5745	3,2	5,8
Savukoski	6423	2,4	5,4
Koillis-Suomi	25674	1,3	2,8

The above estimates have been obtained by random sampling estimators, so the true sampling errors in *Koillis-Suomi* are probably between 1,0—1,2 % (A) and 2,1—2,5 % (V).

The present sampling design probably gives more accurate estimates for proportions than is the case with the old design. On the other hand, it seems that the accuracy of the volume estimates has not decreased from the previous inventory. The systematic field sample tends to decrease the accuracy of the volume estimates because it is difficult to construct homogeneous photo strata with respect to volumes.

The total sampling error is divided into two main components which are called *error of the photo sample and error due to the interpretation*. The former one is estimated on the basis of the variation of the tract values and it is the only source of sampling error in the case that all photo points have been measured in the field. The latter component is affected by the variation within the photo strata and the size of both the field sample and the photo sample. It includes a sub-component the value of which refers to the error arising from the estimation of the stratum weights. The two main components take about equal parts of the total sampling error at the sub-area level. Estimation of the stratum weights contributes some 5—10 % to the error due to the interpretation.

The potential uses of the samples at different area levels are dependent both on the magnitude of the total sampling error and its structure. The two main categories of the potential uses concern i) the possibilities with the original samples, and ii) the possibilities of additional sampling.

Results for target areas within sub-areas can be calculated on the basis of the original combination of the samples at the sub-area level. The estimates are then associated with *bias due to the extension*. This can be avoided by recombining the samples for each target area separately.

The error of the photo sample can be decreased by increasing the photo sample. At the same time, however, the error due to the interpretation is increased. The increase is *relatively* smaller than the decrease, so the final outcome depends on the relation between the main error components.

An increase of the field sample decreases the error due to the interpretation. The relative decrease of the error component is greater than the relative increase of the field sample. This is because the effect of an additional field sample on the error arising from the estimation of the stratum weights is inversely exponential.

A comparison carried out at the sub-area level revealed that a decrease of the total mean square error of the estimates by 14 % can be achieved more efficiently by increasing the photo sample than by increasing the field sample. The efficiency difference decreases when the required decrease of the error is increased. The precedence of the alternatives in increasing the samples is very area-specific. Even an increase of the photo sample can be made in several ways. The choice is partly dependent on the area level.

The present case study revealed some study needs concerning the two-phase forest inventory method in Lapland. One must be ready for an adoption of satellite imagery as a source of information in the first phase of the sampling. The present field plot and field tract are probably not optimum. The extension of the field data to the photo sample is inefficient as it is now. Knowledge of correlation — functions is needed to determine the optimum sampling design. Such studies made in Sweden are partly applicable to Finnish Lapland.

REFERENCES

Aldred, A. H. 1971. A planning model for forest inventories based on combined photo and ground sampling methods. Dissertation. School of Forestry, University of California, Berkeley. 192 pp.

Arman, V. 1969. Riksskogstaxering åren 1958—1967. Results from the National Forest Survey in 1958—1967. Royal College of Forestry, Department of Forest Survey. Research Notes 13: 1—74. Stockholm.

Astermo, S. 1984. Den svenska SPOT-satsningen. Sammandrag. [The Swedish contribution to the SPOT-satellite. Summary]. A discourse of a remote sensing seminar 15.—16. 11. 1984 in Espoo, Finland.

Bengtsson, G. 1978. The Swedish National Forest Survey. Paper presented at IUFRO-meeting S4.02 and S4.04, June 1978, Bucharest, Romania.

Bitterlich, W. 1948. Die Winkelzählprobe. [Relascope plot]. Allg. Forst- u. Holzwirtschaftl. Ztg. 59 (1/2) : 4—5.

Burk, T. E. & Ek, A. R. 1982. A procedure for combining independent sample estimates. Minnesota Forestry Research Notes 280: 1—3.

Cochran, W. G. 1977. Sampling techniques. Third edition. John Wiley & Sons. 428 pp.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. Vol. 20 (1) : 37—46.

Congalton, G. R. & Mead, A. R. 1983. A quantitative method to test for consistency and correctness in photointerpretation Photogrammetric Engineering and Remote Sensing. Vol. 49 (1) : 69—74.

Daamen, W. P. 1980. Kontrolltaxeringen åren 1973—1977. Resultat från en kontroll av datainsamlingen vid riksskogstaxeringen. [Control survey in 1973—1977. Results of a control study conserning data gathering in the National Forest Survey]. Swedish University of Agricultural Sciences, Department of Forest Survey. Report 27: 1—189. Umeå.

Ek, A. R. 1983. Use of forest growth models in management inventories. University of Helsinki, Department of Forest Mensuration and Management. Research Notes 17: 129—136.

Eriksson, A. & Janz, K. 1975. Riksskogstaxeringen åren 1968—1972. [National Forest Survey in 1968—1972]. Royal College of Forestry, Department of Forest Survey. Research Notes 21: 1—165. Stockholm.

Eriksson, L., Ingelög, T. & Kardell, L. 1979. Blåbär, lingon och hallon. Förekomst och bärproduktion i Sverige 1974—1977. [Occurrence and productivity of bilberry, lingonberry and raspberry in Sweden in 1974—1977]. Swedish University of Agricultural Sciences, Section of environmental Forestry. Report 16: 1—124.

Frayer, W. E. 1978. Stratification in double sampling — 'The easy way out may sometimes be the best way'. Colorado State University, Department of Wood & Forest Science. Resource Inventory Notes No. BLM 10: 1—3. Fort Collins.

Green, A., Jonasson, H., Talts, J. & Wallin, M. 1981. Landsat data för riksskogstaxering. [Landsat data for the National Forest Survey]. Lantmäteriet. Tekniska Skrifter 1981 (1) : 1—36. Gävle.

Hardy, J. R. 1981. Data collection by remote sensing for land resources survey. In: Townshend, J. (ed.) 1981. Terrain analysis and remote sensing. London. George Allen & Unwin. pp. 16—37.

Husch, B. 1971. Planning a forest inventory. FAO Forestry and Forest Products Studies no. 17. 120 pp.

— Miller, C. & Beers, T. 1982. Forest Mensuration. Third edition. John Wiley & Sons. 402 pp.

Hägglund, B. 1983. En ny svensk riksskogstaxering. Sveriges lantbruksuniversitet, Institutionen för skogstaxering. [A new national forest survey in Sweden. Swedish University of Agricultural Sciences, Department of Forest Survey]. Umeå. 74 pp.

Ilmakuvatulkinnan ohjeet, VMI 7 Lappi. 1982. Moniste. [Instruction for the photo-interpretation in the 7th National Forest Inventory in Lapland. Stencil]. The Finnish Forest Research Institute, Department of Forest Inventory. Helsinki — Rovaniemi. 4 pp.

Ilvessalo, Y. 1923. Tutkimuksia yksityismetsien tilasta Hämeen läänin keskiosissa. Referat: Untersuchungen über den Zustand der Privatwälder in den mittleren Teilen des Läns Tavastehus. Acta For. Fenn. 26 (2) : 1—137 (15).

— 1957. Suomen metsät metsänhoitolautakuntien toiminta-alueittain. Valtakunnan metsien inventoinnin tuloksia. Summary: The forests of Finland by forestry board districts. Results of the National Forest Inventory. Commun. Inst. For. Fenn. 47 (3) : 1—128.

— 1962. IV valtakunnan metsien inventointi: 1. Maan eteläpuoliskon vesistöalueryhmät. Summary: The Fourth National Forest Inventory. 1. Southern water system areas. Commun. Inst. For. Fenn. 56 (1) : 1—56.

Instruks for proveflatetaksering i 13 fylker syd for Saltfjellet 1982—. 1982. Norsk Institutt for Skogforskning, Avdelning for Landsskogtaksering. [Instruction for a plot survey in 13 municipalities south of Saltfjellet from 1982 on. The Norwegian Forest Research Institute, Division of National Forest Survey]. Stencil 30 pp. Ås.

Jonasson, H. 1982. Provtaxering 1981. [Experimental survey in 1981]. Swedish University of Agricultural Sciences, Department of Forest Survey. NYTAX-report 9: 1—25. Umeå.
Kilpatrick, D. J. 1981. Optimum allocation in stratified sampling of forest populations on successive occasions. Forest Science 27(4): 730—738.
Kuusela, K. 1960. Volume and increment calculation of a sample plot determined with the relascope. Selostus: Kuution ja kasvun laskenta relaskoopilla määritetyllä koealalla. Acta For. Fenn. 71 (6): 1—19.
— 1966. A basal area-mean tree method in forest inventory. Seloste: Pohjapinta-alakeskipuumenetelmä metsäninventoinnissa. Commun. Inst. For. Fenn. 61(2): 1—32.
— 1979. Sampling of tree stock by angle gauge in proportion to tree characteristics. Seloste: Puuston arviointi kulmamittarilla suhteessa puun tunnuksiin. Commun. Inst. For. Fenn. 95(7): 1—16.
— 1982. Peräpohjolan hakkuusuunnitteen arvioimisen perusteet. Metsäntutkimuslaitoksen tiedonantoja 105: 2—6. [The grounds of the allowable cut in Peräpohjola, North Finland. The Finnish Forest Research Institute. Research Notes 105: 2—6]. Rovaniemi.
— 1984. Euroopan metsävarat ja niiden käyttö. [The forest resources of Europe and their usage]. SITRA. B75: 1—75. Helsinki.
— & Poso, S. 1970. Satellite pictures in the estimation of the growing stock over extensive areas. 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. Commun. Inst. For. Fenn. 83 (4): 1—31.
— & Salminen, S. 1969. The 5th national forest inventory in Finland. General design, instructions for field work and data processing. Commun. Inst. For. Fenn. 69(4): 1—72.
— & Salminen, S. 1978. Koillis-Suomen metsävarat vuonna 1976 ja Lapin metsävarat vuosina 1970 ja 1974—76. Forest resources in the forestry board districts of Koillis-Suomi in 1976 and Lappi in 1976 and 1974—76. Folia For. 337: 1—35.
— & Salminen, S. 1983. Metsävarat Etelä-Suomen kuuden pohjoisimman piirimetsälautakunnan alueella 1979—1982 sekä koko Etelä-Suomessa 1977—1982. Summary: Forest resources in the six northernmost Forestry Board Districts of South Finland, 1979—1982, and in the whole of South Finland, 1977—1982. Folia For. 568: 1—79.
— & Salovaara, A. 1971. Kainuun, Pohjois-Pohjanmaan, Koillis-Suomen ja Lapin metsävarat vuosina 1969—70. Forest resources in the forestry board districts of Kainuu, Pohjois-Pohjanmaa, Koillis-Suomi and Lappi in 1969—70. Folia For. 110: 1—49.
Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce and birch. Seloste: Männyn, kuusen ja koivun runkokäyrä- ja tilavuusyhtälöt. Commun. Inst. For. Fenn. 108: 1—74.
Langsaeter, A. 1926. Om beregning av middelfeilen ved regelmessige linjetakseringer. (Über die Berechnung des Mittelfehlers des Resultates einer regelmässigen Linientaxierung). Meddelelser fra det Norske Skogforsøksvesen 2(7): 5—47.
Lappi-Seppälä, M. 1924. Linja-arvioimisesta ja sen tarkkuudesta. Über die Linientaxierung und deren Genauigkeit. Commun. Inst. For. Fenn. 7(3): 1—54.
Lindeberg, J. W. 1923. Über die Berechnung des Mittelfehlers des Resultates einer Linientaxierung. Acta For. Fenn. 25(5): 1—22.
Loetsch, F. & Haller, K. E. 1973. Forest Inventory. Volume 1. Second edition. München. 436 pp.
Marshall, P. L. & Nautiyal, J. C. 1980. Optimal sampling size in multivariate forest inventories: a programming procedure. Canadian Journal of Forest Research 10(4): 579—585.
Matern, B. 1947. Metoder att uppskatta noggrannheten vid linje- och provytetaxering. Summary: Methods of estimating the accuracy of line and sample plot surveys. Meddelanden från statens skogsforskningsinstitut 36(1): 1—136.
— 1960. Spatial variation. Meddelanden från statens skogsforskningsinstitut 49(5): 1—144.
— & Ranneby, B. 1983. Variational structure in forests. Implications for sampling. University of Helsinki, Department of Forest Mensuration and Management. Research Notes 17: 1—10.
Mattila, E. 1979. Kangasmaiden luppometsien ominaisuuksia Suomen poronhoitoalueella 1976—1978. Characteristics of the mineral soil forests with arboreal lichens (Alectoria, Bryoria and Usnea spp.) in the Finnish reindeer management area, 1976—1978. Folia For. 417: 1—39.
— 1981. Survey of reindeer winter ranges as a part of the Finnish National Forest Inventory in 1976—1978. Seloste: Porojen talvilaitumien arviointi osana valtakunnan metsien inventointia Suomessa 1976—1978. Commun. Inst. For. Fenn. 99(6): 1—74.
— 1983. Pohjois-Suomessa suoritettavat inventointitutkimukset. Metsäntutkimuslaitoksen tiedonantoja 105: 116—136. [Inventory studies in North Finland. The Finnish Forest Research Institute. Research Notes 105: 116—136]. Rovaniemi.
— & Kujala, M. 1980. Utsjoen, Inarin ja Enontekiön metsävarat 1978. Summary: Forest resources of Utsjoki, Inari and Enontekiö, North Finland, in 1978. Folia For. 436: 1—21.
Maxim, L. D. & Harrington, L. 1982. To mix or match: On choosing matched samples in comparative aerial surveys. Photogrammetric Engineering and Remote Sensing 48(12): 1863—1867.
Meier, P. 1953. Variance of a weighted mean. Biometrics 9: 59—73.
Milne, A. 1959. The centric systematic area-sample treated as a random sample. Biometrics 15: 270—297.
Mitchell, B. & Bare, B. 1981. A separable goal programming approach to optimizing multivariate sampling designs for forest inventory. Forest science 27(1): 147—162.
Nersten, S. 1981. Takstmetoder i Norge. Norges landbrukshøgskole, Institutt for skogtakssasjon. [Forest survey methods i Norway. Agricultural University of Norway, Institute of Forest Mensuration and Management]. 43 pp.
Netto, Péllico, S. 1981. National forest inventory in Brazil. XVII IUFRO World Congress in Japan. Proceedings, Division 4: 213—226.

Neyman, J. 1938. Contribution to the theory of sampling human populations. Jour. Amer. Stat. Assoc. 33: 101—116.
Nyrhinen, T. 1976. Kaksivaiheisen metsän inventoinnin koe Lounais-Suomessa. A test of two-step forest inventory in South-West Finland. Folia For. 285: 1—24.
Nyyssönen, A. 1954. Kuutiomäärän arvioiminen relaskoopin avulla. Summary: Estimation of stand volume by means of the relascope. Commun. Inst. For. Fenn. 44(6): 1—31.
— 1955. On the estimation of the growing stock from aerial photographs. Seloste: Puuston arvioimisesta ilmakuvien avulla. Commun. Inst. For. Fenn. 46(1): 1—57.
— 1967. Remeasured sample plots in forest inventory. Medd. Norske Skogforsøksvesen 84: 189—220.
— Poso, S. & Keil, C. 1968. The use of aerial photographs in the estimation of some forest characteristics. Seloste: Ilmakuvien käyttö eräiden metsän tunnusten arvioimisessa. Acta For. Fenn. 84(4): 1—33.
— Roiko-Jokela, P. & Kilkki, P. 1971. Studies on improvement of the efficiency of systematic sampling in forest inventory. Seloste: Systemaattiseen otantaan perustuvan metsän inventoinnin tehokkuudesta. Acta For. Fenn. 116: 1—26.
Ohlsson, S. 1975. Riksskogstaxeringen. Ändamål, utformning och användningsområden. The National Forest Survey. Objectives, design and applications. Royal College of Forestry, Department of Forest Survey. Research Notes 20: 1—66 + app. 30 pp.
Omyle, S. A. & Kozak, A. 1982. Estimators for successive forest sampling with partial replacement. Canadian Journal of Forest Research 12: 753—760.
— & Williams, D. H. 1982. Optimum allocation by dynamic programming for sampling on successive occasions with partial replacement of units. Canadian Journal of Forest Research 12: 264—269.
Parry, B. T., Vaux, H. J. & Dennis, N. 1983. Changing conceptions of sustained-yield policy on the national forests. Journal of Forestry 81(3): 150—154.
Poso, S. 1969. Ilmakuvat metsätilojen inventoinnissa. [Aerial photographs in an inventory of two forest estates]. Metsä ja Puu 7/8: 28—29. Helsinki.
— 1972. A method of combining photo and field samples in forest inventory. Seloste: Ilmakuva- ja maasto-otokseen perustuva metsän inventointimenetelmä. Commun. Inst. For. Fenn. 76(1): 1—133.
— Häme, T. & Paananen, R. 1984. A method for estimating the stand characteristics of a forest compartment using satellite imagery. Seloste: Menetelmä kuvioittaisten metsikkötunnusten estimointiin satelliittikuvia käyttäen. Silva Fenn. 18(3): 261—292.
— Keil, C. & Honkanen, T. 1968. Combarisons of film-scale combinations in examining some stand characteristics from aerial photographs. Seloste: Eri filmi-mittakaavayhdistelmät eräiden metsikkötunnusten ilmakuvatulkinnassa. Folia For. 53: 1—22.
— & Kujala, M. 1971. Ryhmitetty ilmakuva- ja maasto-otanta Inarin, Utsjoen ja Enontekiön metsien inventoinnissa. Summary: A two-phase forest inventory method based on photo and field plot sampling as applied in northernmost Lapland. Folia For. 132: 1—40.
— & Kujala, M. 1977. A method for national forest inventory in northern Finland. Seloste: Menetelmä valtakunnan metsien inventointiin Pohjois-Suomessa. Commun. Inst. For. Fenn. 93(1): 1—54.
Päivinen, R. & Yli-Kojola, H. 1983. Pysyvät koealat Etelä-Suomen valtakunnan metsien inventoinnissa. Käsikirjoitus. [Permanent plots in the National Forest Inventory in South Finland. Manuscript]. 30 pp. Helsinki.
Ranneby, B. 1979. Model studies of tract sizes in forest survey. In: Frayer, W. E. 1980. Forest resource inventories. Volume I. Colorado State University, Dep. For. & Wood Sci. Fort Collins. pp. 289—297.
— 1981a. Den topografiska variationen inom olika områden. En redovisning av skattade korrelationsfunktioner. [The topographic variation in the different regions. A review of the estimated correlation functions]. Swedish University of Agricultural Sciences, Department of Forest Survey. NYTAX-report 5: 1—54. Umeå.
— 1981b. Medelfelsformler till skattningar baserade på material från den 5:e riksskogstaxeringen. [Estimation of standard errors in systematic sampling]. Swedish University of Agricultural Sciences, Section of Forest Biometry. Report 21: 1—19 + app. 4 pp.
Rühling, Å. & Skärby, L. 1979. Landsomfattande kartering av regionala tungmetallhalter i mossa. [A countrywide survey of the regional heavy metal content in mosses]. Statens Naturvårdsverk. Rapport SNV PM 1191.
Saastamoinen, O. 1982. Economics of Multiple-Use Forestry in the Saariselkä Forest and Fell area. Seloste: Metsien moninaiskäytön ekonomia Saariselän metsä- ja tunturialueella. Commun. Inst. For. Fenn. 104: 1—102.
Salminen, S. 1973. Tulosten luotettavuus ja karttatulostus valtakunnan metsien V inventoinnissa. Summary: Reliability of the results from the fifth national forest inventory and a presentation of an output — mapping technique. Commun. Inst. For. Fenn. 78(6): 1—64.
— 1981. Vuosien 1971—75 valtakunnallisia metsävaratietoja karttamuodossa. Summary: A cartographic presentation of forest resources in Finland 1971—75. Folia For. 483: 1—42.
Saukkola, P. 1982. Satelliittikuviin perustuva puuston inventointi. [Timber inventory based on satellite imagery]. Technical Research Centre of Finland, Research Reports 85: 1—79 + app. 19 pp. Espoo.
Schmid-Haas, P. 1983. Information on drain by forest inventories. University of Helsinki, Department of Forest Mensuration and Management. Research Notes 17: 50—61.
Shain, W. A. & Rudolph, V. J. 1965. Continuous forest inventory in American forest management. Quat. Bull. Mich. Agric. Exp. Sta. 47: 387—429.
Siegel, S. 1956. Nonparametric statistics for the behavioral sciences. McGran-Hill Kogakusha Ltd. 312 pp.
Skogsdata 83. Aktuella uppgifter om de svenska skogarna från riksskogstaxeringen. [Actual data of the swedish forests based on the National

Forest Survey]. Swedish University of Agricultural Sciences, Department of Forest Survey. 55 pp. Umeå.
Svensson, S. 1980. Riksskogstaxeringen 1973—77. Skogstillstånd, tillväxt och avverkning. The Swedish National Forest Survey 1973—77. State of forests, growth and annual cut. Swedish University of Agricultural Sciences, Department of Forest Survey. Report 30: 1—167.
— 1983. Medelfel i riksskogstaxeringens skattningar 1973—82. Standard error of estimates based on the National Forest Survey 1973—82. Swedish University of Agricultural Sciences, Department of Forest Survey. Report 34: 1—98.
Taksering av Norges skoger. Landsskogtakseringen 50 år 1919—1969. 1970. [Survey of the norwegian forests. National Forest Survey 50 years 1919—1969]. 210 pp. Oslo.
Taksering av Norges skoger utført av Landsskogtakseringen. Sammendrag for hele landet. 1933. [The Norwegian National Forest Survey. Summary for the whole country]. 123 pp. Oslo.
Tiihonen, P. 1966. IV valtakunnan metsien inventointi. 3. Maan pohjoispuoliskon vesistöjen ja metsänhoitolautakuntien alueryhmät. Summary: Fourth National Forest Inventory in Finland. 3. Northern water system areas and forestry board districts. Commun. Inst. For. Fenn. 60 (6): 1—35.
Townshend, J. R. G. 1981. Image analysis and interpretation for land resources survey. In: Townshend, J. R. G. (ed.) 1981. Terrain analysis and Remote Sensing. London. George Allen & Unwin. pp. 59—108.
U.S.D.A. Forest Service 1982. Forest Service Resource Inventory: An overview. Forest Resources Economics Research Staff. Washington D.C. 22 pp.
Uusitalo, M. (ed.) 1984. Metsätilastollinen vuosikirja 1983. Yearbook of forest statistics 1983. Folia For. 590. Official Statistics of Finland XVII A: 15: 1—224.
Ware, K. & Cunia, T. 1962. Continuous Forest Inventory with partial Replacement of samples. Forest Science Monograph 3: 1—40.
Vuokila, Y. 1959. Relaskooppimenetelmän tarkkuudesta puuston arvioinnissa. Summary: On the accuracy of the relascope method of cruising. Commun. Inst. For. Fenn. 51 (4): 1—62.
Yang, Y. C. 1981. Forest Resource Inventory Method in Taiwan. XVII IUFRO World Congress in Japan. Proceedings, Division 4: 291—297.
Zöhrer, F. 1980. Forstinventur. Ein Leitfaden für studium und Praxis. Paul Parey. 207 pp.

Total of 98 references

SELOSTE

Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa

Sovellutus

Kaksivaiheista ilmakuva- ja maasto-otantaa on käytetty valtakunnan metsien inventoinnissa Lapin ja Koillis-Suomen piirimetsälautakuntien alueella vuodesta 1970 lähtien. Tässä tutkielmassa tarkastellaan menetelmän viimeisintä sovellutusta valtakunnan metsien 7. inventoinnissa Lapissa 1982—83. Tutkimusalueena on Koillis-Suomen piirimetsälautakunta, joka käsittää kuusi osa-aluetta (= kuntaa). Maapinta-ala tutkimusalueella on 25 700 km^2 vaihdellen osa-alueissa 1800 km^2:sta 6400 km^2:iin. Edellinen inventointi tutkimusalueella suoritettiin vuonna 1976.

Uusia kohtia inventoinnin tavoitteenasetannassa ovat kokonaan systemaattinen maastonäyte, edellisessä inventoinnissa puolipysyviksi perustettujen maastokoealojen uusintamittaus, inventointitulosten käyttökelpoisuus kuntatasolla ja metsän moninaiskäyttöön liittyvien erillisselvitysten integrointi osaksi inventointirutiinia. Tarkkuustavoitteena on se, että piirimetsälautakunnan elävän puuston runkotilavuuden arvion suhteellinen keskivirhe ei saa olla suurempi kuin 4 %.

Otantajärjestely

Tutkimusalue jaettiin 8 km * 8 km ruutuihin, joista osa (2/3) valittiin tulkintalohkoina mukaan ilmakuvanäytteeseen. Tulkinta tehdään pienimittakaavaisilta (1 : 50 000) mustavalkoisilta ilmakuvilta. Osa (3/8) tulkintalohkoista mitataan myös maastolohkoina. Maastokoealat sijoitetaan systemaattisesti, ts. ilmakuvatulkinta ei vaikuta maastonäytteen valintaan. Tulkintalohkojen ja maastolohkojen valinta myöskin on systemaattinen. Näytteen rakenne on esitetty kuvissa 4—8 sivuilla 15—18.

Sekä ilmakuva- että maastonäytettä on lisätty ja hajautettu edellisestä inventoinnista. Pääsyynä tähän on ollut tulosten käyttökelpoisuutta kuntatasolla koskeva tavoite. Osatekijänä on myös systemaattinen maastonäyte. Inventointikustannus vanhalla näytteen rakenteella olisi ollut noin 20 % nykyistä pienempi. Kustannusvertailussa ei ole mukana vanhojen puolipysyvien maastokoealojen mittaamisesta aiheutuva lisäkustannus.

Maastonäytteen antamien tulosten tarkkuus

Systemaattinen maastonäyte mahdollistaa harhattoman estimoinnin ilmakuvanäytteestä riippumattomasti, mikä sinänsä on eräs uuden inventoinnin implisiittinen tavoite. Seuraavat virhearviot koskevat vain uutta maastonäytettä. Piirimetsälautakunnan puuston runkotilavuuden arvion virhe ilmaistuna suhteellisena keskivirheenä on 4—5 %. Kyseinen virhe suurimmassa osa-alueessa (Savukoski) on 9—10 % ja pienimmässä osa-alueessa (Pelkosenniemi) 17—19 %. Metsämaan pinta-alan arvion virheet ovat vastaavasti 2—3 %, 4—5% ja 8—9 % (s. 27—29).

Vanhojen puolipysyvien koealojen sijainti myöskin on systemaattinen. Näin ollen niitä voidaan uusintamittauksen jälkeen käyttää yhdessä uuden maastonäytteen kanssa. Ko. lisänäyte alentaa yllä esitettyjä virhearvioita noin 10—20 %.

Puolipysyvien maastokoealojen merkitys

Maastokoealojen uusintamittaus mahdollistaa tehokkaan muutosanalyysin. Muutoksen rakenteen lisäksi muutosten syistä saadaan selkeä käsitys jo yhden inventointijakson sisällä. Tutkimuksessa arvioitiin pysyvien lohkojen lohkoarvojen (puuston keskitilavuus lohkolla) kahden mittauskerran välinen korrelaatio. Osa-alueissa korrelaatio vaihtelee välillä 0,650—0,986 ollen koko tutkimusalueella 0,854 (s. 33). Puolipysyvien koealojen perusteella saadun puuston tilavuuden muutosarvion keskivirhe ei ole paljoakaan suurempi kuin koko yhdistetyn ilmakuva- ja maastonäytteen perusteella arvioidun muutoksen keskivirhe.

Edellisen inventoinnin kertakoealoja ei voida nyt käyttää hyväksi SPR-tekniikan (Sampling with Partial Replacement) mukaisesti, koska nämä koealat on valittu maastonäytteeseen vaihtelevin otantatiheyksin. Tätä haittaa ei ole seuraavassa inventoinnissa, koska otantatiheys on sama koko nykyisessä maastonäytteessä. Uusien puolipysyvien koealojen lukumäärä ja sijainti maastolohkolla on sama kuin edellisessä inventoinnissa.

Uudelleen mitattuja puolipysyviä maastokoealoja ei voida käyttää kaksivaiheisen ilmakuva- ja maastonäytteen osana. Tulkintaa varten koealat tulisi paikallistaa ja merkitä käsityönä vanhoilta ilmakuvilta uusille ilmakuville. Tehtävä on työläs ja aiheuttaa systemaattisen virheen vaaran tulkinnassa.

Ilmakuvanäytteen tulkinta

Ilmakuvatulkintaa Koillis-Suomen piirimetsälautakunnan alueella tehtiin ennen maastotöitä ja tulkinta saatettiin loppuun maastotöitä seuraavana talvena. Tämä on mahdollista, koska maastonäytteen valinta ei riipu tulkituista koealatiedoista. Järjestely mahdollisti tulkinnan suorittamisen oman vakinaisen henkilökunnan voimin, mikä on taloudellisesti edullista. Myöhemmin myös osoittautui, että ilmakuvien toimitukseen liittyvien yksityiskohtien vuoksi tulkinnan suorittaminen kokonaan ennen maastotöitä ei olisi ollut edes mahdollista.

Tärkeimmät tulkitut muuttujat ovat *maaluokka, alaryhmä, kehitysluokka, vallitseva puulaji, ojitustilanne* ja *kuutiomäärä*. Luokitteluasteikon muuttujien osalta tulkinnan onnistumista tutkittiin erään K-tilastosuureen avulla. Tulkinnan vastaavuus oli paras alaryhmällä (jako kankaisiin ja soihin), hieman heikompi ojitustilanteen ja maaluokan osalta sekä melko huono vallitsevalla puulajilla ja kehitysluokalla (taulukko 10, s. 40). Korrelaatio tulkitun ja maastossa mitatun kuutiomäärän välillä vaihteli 0,516:sta Pelkosenniemellä 0,679:ään Posiolla (taulukko 12, s. 47). Ilmakuvamateriaali huomioon otettuna tulkintatarkkuus on tavanomainen.

Ilmakuvanäytteen osittaminen

Luokittelumuuttujat valittiin tulkinnan vastaavuuden ja metsikkötunnusten tärkeyden perusteella. Ositus tehtiin eri tavalla ilmakuvanäytteen eri pääositteissa. Osituksen tavoitteena on muodostaa tulkintatietojen suhteen homogeenisia ilmakuvakoealajoukkoja, joiden sisällä maastotiedot laajennetaan satunnaisesti koko ilmakuvanäytteeseen. Ositteiden saaminen suhteellisen homogeenisiksi myös tulkitun kuutiomäärän suhteen edellyttää yksityiskohtaista stratifiointia.

Osituksen vaikutusta eräisiin virhekomponentteihin tutkittiin simuloimalla. Osoittautui, että osa-aluetasolla on suotavaa käyttää suhteellisen yksityiskohtaista luokitusta. Ositteiden painojen arvioinnista aiheutuva virhe pyrkii kuitenkin lisääntymään ositteiden määrän kasvaessa.

Maastotiedon laajentaminen

Tarkoitusta varten kehitettiin tietokonesysteemi tutkimuksen alkuvaiheessa. Systeemin input käsittää kohdealueen ilmakuvanäytteen ja maastonäytteen erillisinä sekä ilmakuvanäytteen osittamista ohjaavan parametritiedoston. Systeemin output sisältää yhdistetyn näytteen sekä kohdealueen päätulokset ja vastaavat virhetunnusluvut. Näytteiden teknillinen yhdistäminen ei enää ole ko. inventointimenetelmän minimitekijä.

Yhdistetyn näytteen antamien tulosten tarkkuus

Ilmakuvanäytteen avulla estimaattien keskineliövirhe pienenee noin kolmanteen osaan siitä, mitä se on maastonäytteen perusteella lasketuissa tuloksissa. Metsämaan pinta-alan (A) ja puuston runkotilavuuden (V) arvion virhe ilmaistuna suhteellisena keskivirheenä osa-alueittain on (taulukko 23, s. 65):

Osa-alue	Maa-ala km²	Characteristic A	V
		CV-%	
Pelkosenniemi	1826	6,5	9,1
Posio	3096	3,5	8,0
Kemijärvi	3578	3,5	7,7
Kuusamo	5006	2,6	6,5
Salla	5745	3,2	5,8
Savukoski	6423	2,4	5,4
Koillis-Suomi	25674	1,3	2,8

Virheet on arvioitu satunnaisotannan kaavoilla, minkä vuoksi ne todennäköisesti ovat yliarvioita. Näytteen systemaattisuus huomioon otettuna Koillis-Suomen todelliset virhetunnusluvut lienevät välillä 1,0—1,2 % (A) ja 2,1—2,5 % (V).

Uusi inventointi ilmeisesti antaa tarkemman arvion pinta-alajakaumista kuin edellinen inventointi. Toisaalta on mahdollista, että puuston runkotilavuuden arvion tarkkuus ei ole huonontunut. Systemaattinen maastonäyte huonontaa tilavuusestimaattien tarkkuutta, koska homogeenisten ilmakuvaositteiden muodostaminen tulkitun kuutiomäärän suhteen ei aina ole mahdollista. Maastonäytettä lisättiin juuri tämän vuoksi.

Otantavirheen rakenne

Otantavirhe jaetaan tutkimuksessa kahteen pääkomponenttiin, joista käytetään nimitystä *ilmakuvanäytteen virhe* ja *tulkintavirhe*. Edellistä estimoidaan lohkoarvojen vaihtelun perusteella ja se on ainoa virhelähde pelkässä maastonäytteessä. Tulkintavirheeseen vaikuttaa kuvaositteiden sisäinen vaihtelu sekä näytteiden koko. Ositteiden painojen arvioinnista aiheutuva virhe sisältyy tulkintavirheeseen. Osa-aluetasolla ilmakuvanäytteen virhe ja tulkintavirhe ovat lähes yhtäsuuret. Ositepainojen estimoinnista koituva virhe on 5—10 % tulkintavirheestä.

Näytteiden käyttömahdollisuuksia

Otantavirheen suuruus ja rakenne eri osa-aluetasoilla vaikuttavat näytteiden käyttömahdollisuuksiin. Operointi alkuperäisillä näytteillä on yksi tie näytteiden hyödyntämisessä, lisäotannan tarjoamat mahdollisuudet toinen.

Osa-alueiden sisäisille kohdealueille voidaan joustavasti laskea tuloksia alkuperäisistä yhdistetyistä näytteistä. Näin menetellen kohdealueiden tuloksiin sisältyy *maastotiedon laajennuksesta aiheutuva harha*. Tämä harha vältetään yhdistämällä maastonäyte ja ilmakuvanäyte kullakin kohdealueella erikseen.

Ilmakuvanäytteen virhettä voidaan alentaa tulkitsemalla lisää ilmakuvakoealoja. Samalla kuitenkin tulkintavirhe kasvaa, tosin *suhteellisesti* hitaammin kuin ilmakuvanäytteen virhe pienenee. Lopullinen vaikutus otantavirheeseen riippuu näiden virhekomponenttien keskinäisestä suuruussuhteesta.

Maastonäytteen lisääminen alentaa tulkintavirhettä. Ko. keskineliövirhekomponentin aleneminen on suhteellisesti nopeampaa kuin näytteen lisäys. Tämä johtuu ositteiden painojen arvioinnista aiheutuvan virhevarianssin nopeasta alenemisesta.

Osa-aluetasolla suoritettu vertailu osoitti, että estimaattien keskineliövirheen alentaminen noin 14 prosentilla on tehokkaampaa lisäämällä ilmakuvanäytettä kuin lisäämällä maastonäytettä. Tehokkuusero kuitenkin pienenee, kun vaadittu tarkkuuslisä kasvaa. Näytteiden lisäämisvaihtoehtojen edullisuusjärjestys on hyvin tapauskohtainen. Myös ilmakuvanäytteen lisäämisvaihtoehtojen edullisuus on harkittava kussakin tapauksessa erikseen.

Tutkimustarpeita

Tämän tutkielman perusteella voidaan osoittaa eräitä lisätutkimusta ja jatkokehittämistä vaativia kohtia nykyisessä Lapin kaksivaiheisen metsäninventoinnin menetelmässä. Tulevaisuudessa on oltava valmius siirtyä tavanomaisista ilmakuvista satelliittikuvainformaation käyttöön ensimmäisessä otantavaiheessa. Maastokoeala ja maastolohko kaipaavat kehittämistä optimaaliseen suuntaan. Maastotiedon laajentamisessa ilmakuvanäytteeseen jää nykyisellään suuri osa menetelmän potentiaalisesta hyödystä osa-alueen sisäisillä kohdealueilla käyttämättä. Sekä koealan ja lohkon että laajennusmenetelmän optimointi edellyttävät korrelaatiofunktioiden selvittämistä. Alan tutkimustulokset Ruotsissa ovat osittain käyttökelpoisia myös Lapin olosuhteissa.

Mattila, E. 1985. The combined use of systematic field and photo samples in large-scale forest inventory in North Finland. Seloste: Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa. Commun. Inst. For. Fenn. 131: 1—97.

ODC 585+524.41+524.6+905.2+(480.99)
ISBN 951-40-0702-6
ISSN 0358-9609

MATTILA, E. 1985. The combined use of systematic field and photo samples in a large-scale forest inventory in North Finland. Seloste: Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa. Commun. Inst. For. Fenn. 131: 1—97.

The study concerns the inventory method used in the 7th National Forest Inventory in Finnish Lapland in 1982—84. The two-phase sample consisted up a systematic photo sample, a systematic part of which was measured in the field. The systematic plots of the old (1976) field sample were remeasured for change analysis. A part of the new field sample was established as semi-permanent, for a remeasurement in the future.

Author's address: The Finnish Forest Research Institute, Rovaniemi Forest Research Station, Eteläranta 55, SF-96300 Rovaniemi, Finland.

ODC 585+524.41+524.6+905.2+(480.99)
ISBN 951-40-0702-6
ISSN 0358-9609

MATTILA, E. 1985. The combined use of systematic field and photo samples in a large-scale forest inventory in North Finland. Seloste: Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa. Commun. Inst. For. Fenn. 131: 1—97.

The study concerns the inventory method used in the 7th National Forest Inventory in Finnish Lapland in 1982—84. The two-phase sample consisted up a systematic photo sample, a systematic part of which was measured in the field. The systematic plots of the old (1976) field sample were remeasured for change analysis. A part of the new field sample was established as semi-permanent, for a remeasurement in the future.

Author's address: The Finnish Forest Research Institute, Rovaniemi Forest Research Station, Eteläranta 55, SF-96300 Rovaniemi, Finland.

Tilaan kortin kääntöpuolelle merkitsemäni julkaisut (julkaisun numero mainittava).

Please, send me the following publications (put number of the publication on the back of the card).

Nimi
Name ______________________

____

____

Osoite
Address ______________________

____

____

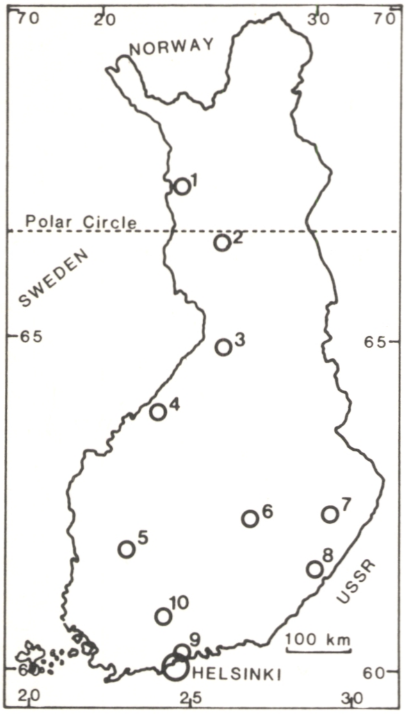
____

Metsäntutkimuslaitos
Kirjasto/*Library*
Unioninkatu 40 A
SF-00170 Helsinki 17
FINLAND

Folia Forestalia

Communicationes Instituti Forestalis Fenniae

Huomautuksia
Remarks



THE FINNISH FOREST RESEARCH INSTITUTE

DEPARTMENTS (Helsinki)

Administration Office
Information Office
Experimental Forest Office
Dept. of Soil Science
Dept. of Peatland Forestry
Dept. of Silviculture
Dept. of Forest Genetics
Dept. of Forest Protection
Dept. of Forest Technology
Dept. of Forest Inventory and Yield
Dept. of Forest Economics
Dept. of Mathematics

RESEARCH STATIONS

1 Kolari
2 Rovaniemi
3 Muhos
4 Kannus
5 Parkano
6 Suonenjoki
7 Joensuu
8 Punkaharju
9 Ruotsinkylä
10 Ojajoki

FACTS ABOUT FINLAND

Total land area: 304 642 km^2 of which 60—70 per cent is forest land.

Mean temperature, °C:	Helsinki	Joensuu	Rovaniemi
January	-6,8	-10,2	-11,0
July	17,1	17,1	15,3
annual	4,4	2,9	0,8
Thermal winter (mean temp. < 0°C):	20.11.—4.4.	5.11.—10.4.	18.10.—21.4.

Most common tree species: Pinus sylvestris, Picea abies, Betula pendula, Betula pubescens

Communicationes Instituti Forestalis Fenniae

128 Saarenmaa, H. Within-tree population dynamics models for integrated management of *Tomicus piniperda* (Coleoptera, Scolytidae). Seloste: Pystynävertäjän lisääntymiskauden populaatiodynamiikkamallit tuhojen integroitua hallintaa varten.

129 Sepponen, P. The ecological classification of sorted forest soils of varying genesis in northern Finland. Seloste: Syntytavaltaan erilaisten lajittuneiden kangasmetsämaiden ekologinen luokittelu Pohjois-Suomessa.

130 Uusvaara, O. The quality and value of sawn goods from plantation-grown Scots pine. Seloste: Viljelymänniköistä saadun sahatavaran laatu ja arvo.

131 Mattila, E. The combined use of systematic field and photo samples in a large-scale forest inventory in North Finland. Seloste: Systemaattisen ilmakuva- ja maastonäytteen yhteiskäyttö laajan metsäalueen inventoinnissa Pohjois-Suomessa.

ISBN 951-40-0702-6
ISSN 0358-9609