SATELLITE IMAGE-BASED NATIONAL FOREST INVENTORY OF FINLAND *

Erkki Tomppo
The Finnish Forest Research Institute
Unioninkatu 40 A
SF-00170 Helsinki, Finland

1. INTRODUCTION

The National Forest Survey (NFI) data of Finland has been based until today on field measurements only. Systematic cluster sampling has been applied in the last four inventories in Southern Finland. It has been considered generally that the area for which the results can be computed reliably enough is about 150 000 ha. The inventory of the whole country takes about 10 years. This implies that a part of the data concerning the amount of growing stock and the level of growth is out of date all the time. Other drawbacks of the recent inventory are: 1) cutting and mortality can not be estimated reliably enough by means of the temporary sample plots; 2) diseases and damages can not be detected continuously; 3) time variation of the growth can not be estimated reliably in different regions of the country; and 4) results are not in an easy form to use by other organizations.

Using ground measurements only to remove these drawbacks and to get localized information is too expensive. Therefore, The Forest Research Institute has started to develop a new inventory system. The method exploits satellite image data and digital map data as well as other geographical data, for example, meteorological data as well as ground measurements.

2. DATA

2.1. Satellite image data

The possible image data are at this moment Landsat TM and Spot images (and MOS-1 images). Of these, TM will be used in the first phase. The reasons are: 1) One TM scene covers a larger area (about $180km \times 180km$) than a Spot image ($60km \times 60km$). This is of importance at the beginning, when a reliable image calibration system is not available. This makes it also more possible to obtain cloudfree images covering the whole country. 2) The price of TM data is about 0.01 Fmk/ha compared with 0.025 Fmk/ha of Spot -data. However, according to our hypothesis, the best results would be obtained if both TM and Spot imagery could be used because of the better spatial resolution of Spot. This approach will be used after the computer capacity of The Institute becomes large enough. The satellite images will be rectified with a pixel size of $25m \times 25m$ to the base map coordinate system.

* To appear in: Proceedings of ISPRS, COMISSION VII, Mid-Term Symposium Global and Environmental Monitoring, Techniques and Impacts, September 17-21, 1990 Victoria, British Columbia Canada

2.2. Digitized map data

Digital map data will be used for improving the accuracy of estimates and to separate forest and non-forest land from each other. The spectral response of peatlands differs from that of the mineral soils with the same growing stock according to our earlier studies. Further, some peatlands can not be separated from mineral soils. Therefore, digital peatland information will be used in order to improve the accuracy of estimates. These digital data are provided by the National Board of Survey. The peatland data, digitized from a generalization of a base map with the scale of 1 : 100 000, are now available for the whole country.

Agricultural areas and roads will be digitized from base maps having a scale of 1 : 50 000 (if the preliminary tests show these data to be useful). A combination of a numerical interpretation and digital map information will be used in classifying agricultural areas, because the map data are not necessarily up to date.

Urban areas can be obtained from the house register provided by the statistical centre of Finland. The coordinates of each house in Finland are known. A digital urban area mask can be produced from this information.

Water areas could be obtained from base maps but they can be obtained relatively reliably also from satellite images.

Some administrative information such as community boundaries and, in the future, boundaries of forest holdings will also be used in the digital form in order to differentatiate computation units.

2.3. Ground truth data

The recent NFI data can serve as ground truth data. In the future sampling design, the needs of satellite image interpretation must be taken into account more thoroughly.

Such problems as 1) the sampling intensity of permanent and temporary plots, 2) the rotation period of inventory, and 3) the size and the form of the sample plot and of the cluster will have to be addressed. Choices must be made, for example, between one relatively small or several even smaller plots in the same stand. (The satellite image information is essentially standwise information rather than sample plot information.)

3. IMAGE INTERPRETATION

The image analysis consists of preprocessing of the image (removal of noise, striping, etc.), choice of features, classification, and postprocessing (generalization). All inventory variables should be estimated. One possibility is to use some non-parametric method such as nearest neighbor classification (see Appendix 1). This method can also be referred to as *a 'nearest neighbor fuzzy classification'*. Each interpretation class consists of one sample plot (one or several pixels depending on what kind of features will be used). Each pixel to be classified will be shared between several classes according to Appendix 1. The shares are inversely proportional to the squared feature space distances $d^2_{(i),p}$. Our preliminary tests have shown that this method works quite well. The estimation of some changes such as thinning may require other methods, cf. Thomas (1990). The spatial information of the image can be taken into account in the feature choice and/or in the postprocessing. Segmentation techniques or Gibbsian random field modelling, for example, are possible postprocessing methods (see Besag (1986) and Tomppo (1987,1989)). Bias and the accuracy of the estimates must of course be kept under control.

4. DATA MANAGEMENT

Special attention must be paid to data management; total single date coverage of the country with TM imagery involves about five Gbytes of data. The output of the image analysis will be compatible with the the forest management planning system (MELA) used by the Finnish Forest Research Institute. This system allows simulation of the development of the forest between two ground measurement (and image analysis) timepoints. In addition, themes can be presented in a raster form and the maps can be produced at different scales.

The system allows the use of digital map data of other organizations, for example the National Board of Survey and the National Board of Forestry. In the future, the output will be formatted to be suitable for the databases of those organizations.

5. PRELIMINARY RESULTS

The above methodology and input data have been applied in estimating NFI -variables in a few areas in eastern Finland. Sum characteristics of ordinary standwise data, measured for forest management planning purposes, are available as comparison material. These data (referred here as FBD -data) are based on visual ground estimation and information from false colour aerial photographs. The data are measured by the local Forestry Board District and only non-company private forests are included. All the above mentioned digital map data are so far available only from the area of two communes, Tohmajarvi and Värtsilä, the total area of non-company private land being 52 000 hectares.

The digital communal boundaries were applied in order to restrict the test area. The boundaries of lands of two forest companies were digitized in order to remove non-private forest areas from the test site, because the forest characteristics of company-ovned land were not known.

Table 1 shows the satellite image-based estimates (NFI) and the estimates of Forestry Board District (FBD) for some mean characteristics. (Note that the total number of NFI -variables is about 200.)

Table 1. The NFI and FBD estimates of some mean characteristics in an area of 52000 hectares.

Mean tree stem volume (m^3/ha)	NFI	FBD
- of pine	47.4	37.7
- of spruce	43.0	56.8
- of broad leave trees	18.7	13.7
All tree species	109.1	108.2
Development classes (%)		
- open area	2.4	3.2
- seed tree or shelterwood stand	1.0	0.9
- small seedling stand	9.2	10.3
- advanced seedling stand	17.6	21.0
- young thinning stand	32.6	24.4
- advanced thinning stand	25.0	24.2
- mature stand	12.3	13.1
- seedling stand with seed trees	-	1.6
- low yielding stand	-	1.9

The estimates of the total mean tree stem volume are about the same with both methods. Some differences occur however in tree species proportions. The percentages of different development classes in both data are fairly close to each other except the groups 'advanced seedling stand' and 'young thinning stand'. One reason, in addition to the possible confusion, may be that the definitional boundary between these groups is not same in different inventory systems. One should also remember that FBD-estimates are based on visual assesment.

However, this preliminary test shows that the satellite image-aided inventory method may work fairly well at the communal level. Differences in tree species proportions are rather high. The reasons must be traced in the future. The testing will continue with other variables and other data. If systematic errors occur, the applicability of new features and perhaps new classifiction techniques will be tested. The final conclusions can be derived after thoroughly-measured test data are available.

6. THE ADVANTAGES OF THE METHOD

The satellite image-aided NFI method is expected to make it possible to:

1) estimate all the variables of the national forest survey for each point of the country; The variables include the properties of site type, soil, growing stock, mortality, cutting, and damages and diseases. It is anticipated that some improvement in estimates of all variables can be expected compared with the current situation. The estimates of variables concerning a single pixel may be unreliable. However, the communal level estimates, and in the case of some variables possibly also the forest holding and stand estimates, are expected to be applicable.

2) keep the data up-to-date in the whole country; Clouds prevent obtaining imagery

which covers the whole country each year. Tree and stand models, and simulation, can be applied to update the data between two image acquisition dates. Other remote sensing data such as radar data or different kinds of airborne data can be utilized later.

3) keep the geographically localized information in a digital form and to transfer it easily into the databases of the users; and

4) estimate the time and spatial variation of variables more reliably than before. The new ground sampling design, which includes permanent sample plots, will support this task.

References

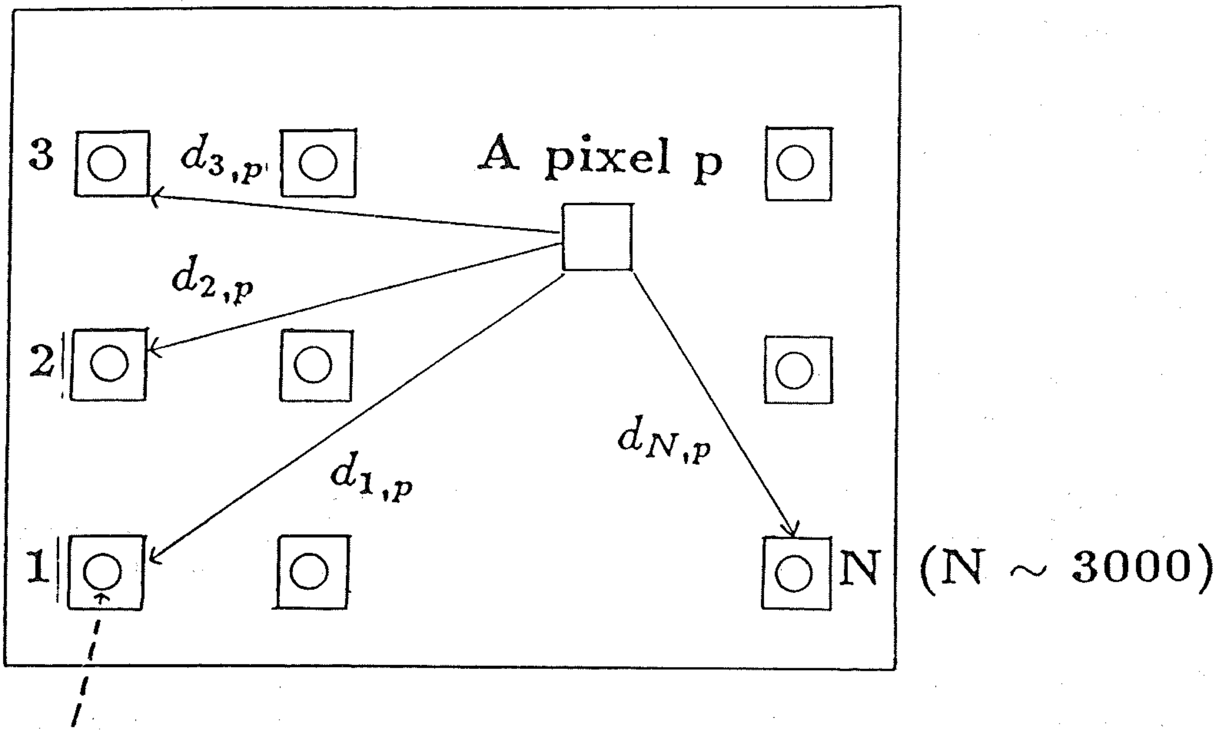
1. Besag, J.E. 1986. On the statistical analysis of dirty pictures (with Discussion). J.R.Statist.Soc. B, 48, 259-302.

2. Tomppo, E. 1987. An application of a segmentation method to the forest stand delineation and estimation of stand variates from satellite images. Proceedings of the 5th Scandinavian Conference on Image Analysis. Stockholm, June 1987. pp. 253-260.

3. Tomppo, E. 1989. Comparisons of some classification methods in satellite image aided forest tax class estimation. In: SCIA '89 Proceedings of The 6th Scandinavian Conference on Image Analysis. Oulu, Finland, June 19 - 22, 1989. Vol 1. 150-157.

4. Thomas, R. W. 1990. Some Thoughts on the Development of Remote Sensing-Aided National Forest Inventory in Sweden. To appear in: Proceedings of the SNS/IUFRO Remote Sensing and Forest Inventory Workshop in Umeå, Sweden, February 26 to 28, 1990.

Appendix I: Nearest neighbor fuzzy classification method



NFI -sample plots with geographic coordinates and with the corresponding pixels $i, i = 1, ..., N$

$d_{i,p}$ is the Euclidean distance in the feature space from the pixel p to the pixel i (sample plot i).

Take $d_{(1),p}, ..., d_{(n),p}, (d_{(1),p} \leq .... \leq d_{(n),p})$, $n \sim 5 - 10$.

Define $w_{i,p} = \frac{1}{d^2_{(i),p}} \sum_{i=1}^{n} \frac{1}{d^2_{(i),p}}$.

Define the estimate of the variable M

$$\hat{M}_p = \sum_{j=1}^{n} w_{(j),p} \cdot M_{(j),p},$$

where $M_{(j),p}$, $j = 1, ..., n$, are the values of the variable M in the n closest pixels to the pixel p.