UNITED
NATIONS

E

Economic and Social Council

Distr.
LIMITED

E/CONF.90/L.5
8 April 1997

ORIGINAL: ENGLISH

SIXTH UNITED NATIONS REGIONAL
CARTOGRAPHIC CONFERENCE
FOR THE AMERICAS
New York, 2-6 June 1997
Item 5 (e) of the provisional agenda*

REVIEW OF THE LATEST TECHNOLOGY IN CARTOGRAPHIC DATA ACQUISITION, MANIPULATION, STORAGE AND PRESENTATION, WITH SPECIAL EMPHASIS ON POTENTIAL APPLICATIONS IN DEVELOPING COUNTRIES: REMOTE SENSING IN CARTOGRAPHY

Land Cover and Forest Classification of Finland based on Landsat Thematic Mapper (TM) images and digital map data

Paper submitted by Finland**

* E/CONF.90/1.

** Prepared by Arto Vuorela, National Land Survey of Finland, Satellite Image Centre, Helsinki, Finland.

The fourth version of the Land Cover and Forest Classification is being planned in NLSF. The main goal of the project itself has been to make a digital raster database, a satellite image classification which fulfills different needs of partner organizations and other, expected users. The available ancillary information (digital map data) is used to mask the classification image and thus ease the classification task. Product has a pixel size of 25 meters, it covers the whole country and is available for all public just as any other map.

1. History

VTT (Technical Research Center of Finland) and the National Board of Waters and Environment (later Finnish Environment Institute) developed a method in 1986 for inventory of river basin characteristics (Kuittinen and Sucksdorff 1987). Basin characteristics affect the quality and quantity of water in the drainage basin. Basin characteristics are one part of Environmental Geographic Information System. Image classification for this purpose started in 1987. In autumn 1987 several organizations, including the Ministry of the Environment, the Post & Tele Radio Services, regional planning organizations and Statistics Finland were greatly interested in joining this classification project for the whole country. In early 1988 the project was started.

The first version was completed in 1991, among the first countries in the world, and the aim has been to update the data at 4–5 years interval. A new process was set up in 1992. Since 1994 the third version has been produced. The differences in processing make the change detection a bit ambiguous, but there are general reasons to improve the product (Table 1). Straightforward masking is not the only way to consider ancillary information in a classification process (Hutchinson 1982), but in Finland the good quality digital maps are usually more accurate (even when not up-to-date) than the fresh satellite image and its classification alone would be.

2. Image classification

2.1 Material

System ("bulk") corrected **Landsat Thematic Mapper (TM) images**, channels (1,) 2, 3, 4, 5 and 7 have been used for the Classification. No multitemporal image approach has been used yet.

Because of difficulties in separating some peat lands from mineral soil areas, a digitized **peat land raster mask** was used. This was 8-bit digitized mainly (95 %) from topographic maps 1:100 000 and partly (5 %) from road maps 1:200 000. A matrix camera and later also a plate scanner was used to digitize the peat elements. Corner crosses were measured and the mask rectified, then thresholded and mosaicked. The quality of this mask is not adequate for all purposes.

Table 1. Main differences between Land Cover and Forest Classification versions. See last chapter for latest work done on the process, and for future plans.

	Version 1, 1988-	Version 2, 1992-	Version 3, 1994-
Coverage	95 %, 1991	50%, 1994	at least 60%, end 1996
Training areas	Companies/own	Companies/own	National Forest Inventory (NFI) field sample plots
Peat mask	1:100 000 (1:200 000)	1:100 000 (1:200 000)	1:100 000 & 1: 50 000 (1:200 000)
Cultivated area mask	-	1:20 000-based 1:50 000	1:20 000-based & (fresher) 1:50 000
Other masks	-	Add-on buildings become available	1:50 000, peat production areas from small scale maps
Generalization	Mode filtering for segments, peat lands separately	Mode filtering by 3×3 pixels, peat lands separately	Mode filtering by 3×3 pixels, peat lands separately
Classifier	Linear discriminant	Linear discriminant	Non-parametric *k-NN* estimation (= NFI), then rule-based decisions
Training area digitization	Maps + visual screen digitization	Maps + transparency on the screen	Sample plots' *x*, *y* as image coordinates
Rectification	Rectify peat mask to TM, then classify, rectify back	First rectify TM, consider masks parallel to classifying	Automatic GCP measurements, rectify TM, consider masks parallel to the NFI production, also later

In the second classification version also **a cultivated area mask** was available to improve the classification result. New peat, water, field etc. elements are becoming available as the new topographic map 1:50 000 is digitally published (over 50 % was ready at the end of 1996).

Built-up areas can be taken from map data or from a **building register** which consists of all notable buildings in Finland. The buildings can be set to 1 or 4 pixels each, or the number of pixels of each building can represent the basal area of it, so this case is not standard, but masked add-on to the classification image. The building register does not fulfill all user requirements at the moment.

In the third version, NLSF began a closer co-operation with the Finnish Forest Research Institute (FFRI). The classification is derived from the **National Forest Inventory (NFI)** estimation results. FFRI has a wide network of sets of *x*, *y* sample plots. A lot of work is done to update this data. Satellite images have been used for NFI, but no more for deciding the class, in which phase only the NFI and the map data are used. These already have an informative meaning, so the essential information is to be filtered out. But the NFI information is relevant mainly for forest areas, and the low timber volume classes like pastures and alpine classes are hard to define from just volumes, growths etc. (Table 2). The Classification is a relevant product, since a 20-channel NFI image has less map masks included and is much more complicated to use than a single-channel classification.

Point by point training data are better than areas, because no variance occurs inside the sample. The NFI estimate for a pixel is a weighted sum of *k* spectrally nearest sample plots' properties, weights inversely proportional to spectral distances to them. See reference for more details of this non-parametric estimation method, but note missing '/' in normalization of weights (Tomppo 1990). The role of NLSF has become more like preprocessing of both TM images and map elements so that the satellite images have to be analyzed by FFRI only for the forest areas. NFI is then used to process the Land Cover product in NLSF. The overall process goes as described in APPENDIX 2.

Table 2. The National Forest Inventory data is usually a 8-bit 20-channel image, describing estimates of as many traits of forests. Not all channels are received by NLSF from FFRI.

NFI image channel original number (and received in NLSF)	Estimated forest trait (forest property)	Unit of estimate digital number
1 (1)	yield	0.1 m^3/ha/year, for peats added 100
2	needle cast/defoliation	relative 1 to 7
3 (2)	age	years
4 (3)	pine volume total	2 m^3/ha
5 (4)	pine timber volume	2 m^3/ha
6	pine other volume	2 m^3/ha
7 (5)	spruce volume total	2 m^3/ha
8 (6)	spruce timber volume	2 m^3/ha
9	spruce other volume	2 m^3/ha
10 (7)	birch volume total	2 m^3/ha
11 (8)	birch timber volume	2 m^3/ha
12	birch other volume	2 m^3/ha
13 (9)	other tree volume total	2 m^3/ha
14 (10)	other tree timber volume	2 m^3/ha
15	other tree other volume	2 m^3/ha
16 (11)	growth total	0.1 m^3/ha/year
17	growth pine	0.1 m^3/ha/year
18	growth spruce	0.1 m^3/ha/year
19	growth birch	0.1 m^3/ha/year
20	growth other	0.1 m^3/ha/year

2.2 Classification methods

Water bodies are classified parallelepiped (or feature space classified) from channels TM 4 & TM 7. The basic classification method of versions 1 and 2 was the same as the classification of river basin characteristics mentioned in the first paragraphs: a supervised method based on linear discriminant functions (Fukunaga 1990:131–138). It assumes about Gaussian distributions and gives rather similar results than the Gaussian maximum likelihood classification, but is a bit more robust. For example there will not be any class "unclassified". Programs made by VTT used image processing subroutines of DISIMP (CSIRO, Australia), which has been used by 8 to 10 organizations in Finland.

In the current process, there are several ways how the raster map database is used as an auxiliary information. Some raster database values (from the new topographic map 1: 50 000 data etc.) are converted directly into a certain value in the Land Use and Forest Classification. For example value 5 is built-up single house areas and is masked into class 30.

Some other raster database values are classified more generally, for example values 26, 27, 28, 36, 37, 38, and 73 mean different kinds of peat lands and are classified to peat land classes. It is yet possible to implement some limitations using the map raster. For example values 1, 51 and 56 are bedrock areas, and for these pixels we have limited the class to be among those with less than 102 m^3/ha timber volume, and in case this pixel has been classified as water, we assume an error to be made (steep north slope shadow) and correct this into a typical pine-dominated class.

The decisions or parametric rules to be done when deciding the class label using the NFI goes the following way: First the values to be masked are checked (from constant mask values the NFI includes or from the fresh mask image). If there is no match, some special cases are checked. If e.g. *age* < 13 years, *total yield* < 3.5 m^3/ha/year and *total growth* < 0.3 m^3/ha/year, and peat mask is not on, output class is 23, open areas. If this kind of conditions are not met, more general conditions (based on tree species proportions and timber volumes) are checked, so that some class value is set for the pixel. NFI would enable also other classifications than the one made.

The pixel by pixel classification has been generalized by mode filtering for segments (1st version) and for 3x3 pixel windows (later versions). The filterings are done for peat lands and mineral soil areas separately.

3. Products, costs

As mentioned above, the final product is a digital database, a thematic map covering the entire country with a pixel size of 25×25 m^2. See APPENDIX 1 for the class definitions. From this digital product it will be easy make different outputs depending on the needs of the customer: films, paper prints with desired classes, tapes, floppy disks, statistics covering desired areas such as counties, etc. Data is sold at 1.50 FIM/km^2 without merging the classes, so the whole country costs 400 000 FIM. One 25 μm/pixel 10" film plot will cover in full resolution an area with a maximum size of about 230×230 km^2. It is possible to process the classified imagery so that the entire country will fit on one film. A generalised product with a 200×200 m^2 pixel size and 14 classes is available.

Over 70 Landsat TM quarter images are required to get full coverage of the country. The cost of this imagery was about 600 000 FIM for version 1. Because image processing costs were over 20 000 FIM per quarter image, the total cost was 2 300 000 FIM. Later the processing costs have been a bit smaller.

4. Accuracy

About the accuracy of rectification: the mean error of the ground control point (GCP) measurement from the images and 1:20 000 (or 1: 50 000) maps has been tried to keep under 0.5 pixels (15 m). About a dozen points - very small islands - per scene have been measured. The Landsat TM system correction also got a better cubic convolution interpolation algorithm available instead of nearest neighbor. Decision about changing to use precision corrected has not been done. In Finland it is not a must. There is also a possibility to purchase raw images and implement a system that can handle the imaging geometry for rectification of these images, but this is not trivial.

Another new approach is the automatic control point measurement which is used in (Holm et al 1995). Island and lake centroids are features extracted from the satellite image to be rectified. They are matched with database features, which used to be from the version 1 classification and had the quality mentioned above, and were replaced 1996 by those based on 2 m resolution 1: 20 000 water rasters. The result of a matching is a GCP file with a few ten or hundred point pairs. Digital elevation model (DEM) could also be used in the rectification, but in the NFI production at least the grey value correction and absolute height consideration is implemented.

From version 2 on, training areas were digitized using suitably zoomed transparent copies on the screen, so that the areas were digitized looking at the image through the area map. This improved the quality of the digitization.

Some tests considering the **classification accuracies** have been made. Accuracies have been generally between 50 and 90 %, not very big on the average because of an optimistical class definition. Because of the large amount of processed images, the tests have not been made to all of them, so no accuracy estimate has been given to the user. Many of the errors are caused by segment border mixels, and as we combine the classes into main categories, the accuracies get reasonably better (Jaakkola 1994a). For the 2nd version, biggest classification error probabilities over class pairs were listed for each classification, and some gross errors could thus be eliminated. These lists can also be made available for the users (e.g. Oksanen 1996).

For the third version, a more complete and accurate estimate of the classification accuracy is available. In Table 3, a comparison of tree volume of NFI and version 1 Land Cover Classification is presented for different parts of Finland. The accuracy of NFI is very good for large areas (> 150 000 ha), but the smaller the area, the poorer the reliability of it. In NLSF some research is going on about the usability of NFI data for even a real estate tree stand value estimation (MML 1994). Areas less than a few ten (20–50) ha are usually no more reliably estimated for that purpose. NFI has a tendency to average the estimates a bit in cases where there is very small or very big tree stand volume. We must say, that the reliability of the version 3 classification for the small tree stand volume classes cannot be very good, because we had so many such classes in the previous versions and they are still included, and the original satellite images are no more used. But if in the standard NFI data there would be a channel estimating the development class of the forest, this would yield better results.

Table 3. Comparison of mean timber volume (m^3/ha) of the Classification (version 1 !) and National Forest Inventory. Mean error of the NFI total timber volume for each Regional Forestry Board is also shown. The four last are based on the 7th NFI and the others on the 8th.

Regional Forestry Board	Classification volume (m^3/ha)	NFI volume	Updates from 1.1.1990.	Mean error of NFI volume
Helsinki	112	126		3.3
Lounais-Suomi	110	121		2.5
Uusimaa-Häme	141	146		2.8
Itä-Häme	147	138		2.7
Etelä-Karjala	126	118		2.5
Satakunta	116	109		2.5
Pirkka-Häme	116	125		2.7
Etelä-Savo	134	129		2.0
Kainuu	68	59	(63)	1.7
Pohjois-Pohjanmaa	80	51	(59)	1.2
Koillis-Suomi	72	40	(40)	1.3
Lappi	52	38	(40)	0.8

Choosing the preprocessing methods has not been very simple. And of course there are some errors in the digital map input data. We can only say that the quality of them is rather good already and is also being improved stage by stage.

5. Usage

In NLSF a digital elevation model (DEM) which covers the entire country has been produced from basic maps where the contour interval is 5 (somewhere 2.5) meters. Land use classes (eg. forest covered areas) together with DEM and buildings are necessary to locate the optimum places to construct the mobile telephone base stations so that the areas of poor reception will be minimized.

The Finnish Environment Institute continues the work with the Environmental Geographic Information System. Classified satellite images are useful for example in predicting spring time floods in large areas and pollution distribution. Regional planning organizations will also use classification products in their field work. The statistics were calculated for each municipality and published for each county (Statistics Finland 1994). Product is used as a background or as a digital map base with vector data, for example in GPS-based vehicle navigation systems (Koskelo & Tolkki 1993). The Forest Research Institute will in turn use the classification result (= refreshed masks, in practise) to improve the results of National Forest Inventory.

For small scale thematic map production, further generalization of the classification data is needed. An iterative method, where the result segment size is increased and where the areas are merged into hierargically most similar neighbor areas, seems to be most promising (Jaakkola 1994). Generalisation for CORINE Land Cover project runs on ARC/INFO GRID environment. During 1996 also the classification process was transferred into GRID (Määttä 1997). The map element situation of each 40 × 30 km map sheet is analysed from a database. Attribute information (field work year, image date) can be set for each pixel.

6. Future and discussion

During 1997 the 4th version of the product is planned in NLSF. Aaro Mikkola works on this subject. International compatibility, for example with neighbour countries, has much emphasis, and the CORINE Land Cover will be made from this product. Yet there is a plan to maintain a multi-level definition of classes that still suits the Finnish needs for the whole country. The Finnish product has so far been more detailed for forests than in the international standards.

GIS data market has developed. User requirements have actually been born and can now be considered better. Users are concerned about land use/land cover (LU/LC) correspondance, and usage of e.g. open areas is hard to define. Big-scale map users like municipalities want data especially from built-up areas and ought to also contribute to data management. Nature of the classification would remain the same, although the input material would change. But these users are a somewhat disjointed group. Many want several attribute data for a pixel, e.g. more forest parameters, nature conservation area data and other vector data based information.

References

Fukunaga, K.: Introduction to Statistical Pattern Recognition. Second Edition, 1990 Academic Press, Inc.

Mikael Holm, Eija Parmes, Kaj Andersson & Arto Vuorela: A nationwide automatic satellite image registration system. *SPIE AeroSense'95 Conference on Integrating Photogrammetric Techniques with Scene Analysis & Machine Vision II*, Orlando, Florida, USA, 17.–21. April 1995, SPIE Vol. 2486-18, pp. 156–167.

Hutchinson, Charles F.: Techniques for Combining Landsat and Ancillary data for Digital Classification Improvement. *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 1, January 1982, pp. 123–130

Jaakkola, O.: FINNISH CORINE LAND COVER - A FEASIBILITY STUDY OF AUTOMATIC GENERALIZATION AND DATA QUALITY ASSESSMENT. *Reports of the Finnish Geodetic Institute* 94:4 (1994).

Koskelo, I. & Tolkki, S.: Seeing the Trees through the Forest: GPS Streamlines Finnish Logging Operation. *GPS World* Vol. 4, Number 10, October 1993, pp. 20-27, 66

Kuittinen, Risto and Sucksdorff, Yrjö: INVENTORY OF RIVER BASIN CHARACTERISTICS IN FINNISH CONDITIONS USING SATELLITE IMAGERY. *Aqua Fennica* 17,2 1987 pp. 97-114

MML: Metsä- ja kiinteistöverotuksen arviointimenetelmien kehittämishanke (Haukipudas). Maanmittauslaitos – VARVI-työryhmän raportti 1994 (in Finnish)

Määttä, Paavo: Maankäyttö- ja puustotulkinnan prosessin siirto ARC/INFO GRID-ympäristöön. Unpublished report, National Land Survey 1997 (in Finnish)

Oksanen, Juha: Reliability of Land Cover and Forest Classification in Åland Main Island. Unpublished Laudatur journey final report at Helsinki University / Geography 1996 (in Finnish)

Statistics Finland: Environment Statistics. Official Statistics of Finland (SVT)/Environment 1994:3

Tomppo, Erkki: Satellite Image-Based National Forest Inventory of Finland. *The Photogrammetric Journal of Finland* Vol. 12 No. 1 1990 pp. 115-120. The Finnish Society of Photogrammetry and Remote Sensing & Institute of Photogrammetry and Remote Sensing, Helsinki University of Technology. See also http://www.metla.fi/projects/vmi/nfi.html.

Arto Vuorela: Satellite Image Based Land Cover and Forest Classification of Finland. Risto Kuittinen, editor, Proceedings of the Finnish-Russian Seminar on Remote Sensing in Helsinki 29 August – 1 September 1994. Reports of the Finnish Geodetic Institute 97:2 (1997), pp. 41–51.

APPENDIX 1

Soili Karppinen and Arto Vuorela
National Land Survey of Finland, 20 Dec. 1996

Land Use and Forest Classification of Finland: Used class codes

(1. Clouds, cloud shadows)
(4. Cloud shadows)
7. Water bodies (classified)
8. Rushes etc. wettish open lands
9. Water bodies (1:50 000)
10. Shoreline, rivers, drains 1:50 000
11. Filling-in water bodies, flood areas (1:50 000)
12. Cultivated area (classified)
13. 1:20 000-based cultivated areas
14. 1:50 000-based cultivated areas (more up-to-date than 13)
19. Clear cutting areas (new), mineral soil
20. Clear cutting areas (old), mineral soil
21. Clear cutting areas (new), peat land
22. Clear cutting areas (old), peat land
23. Open lands. Sand (quarry, rock, other bare soil; classified), mineral soil
24. Open lands. (only certain rather permanent out of 1:50 000 & databases)
(27. Add-on buildings from national building register)
29. Built-up areas / urban (1:50 000)
30. Built-up areas / single house (1:50 000)
31. Industrial areas (1:50 000), (beacons)
37. Peat production areas (1:50 000 or NLSF databases)
38. Peat production areas (digitized from VAPO maps)
39. Peat production areas (classified)
40. Open peat lands (oligotrophic, classified in version 1)
41. Open peat lands (minerotrophic, classified in version 1)
42. Open peat lands (1:50 000)

45. Spruce-dominated peat lands	< 52 m^3/ha
46. "	52-101 m^3/ha
47. "	102-151 m^3/ha
48. "	152-201 m^3/ha
49. "	> 201 m^3/ha
50. Deciduous-dominated peat lands	< 52 m^3/ha
51. "	52-101 m^3/ha
52. "	102-151 m^3/ha
53. "	152-201 m^3/ha
54. "	> 201 m^3/ha
55. Deciduous-coniferous mixed peat lands	< 52 m^3/ha
56. "	52-101 m^3/ha
57. "	102-151 m^3/ha
58. "	152-201 m^3/ha
59. "	> 201 m^3/ha

Continues

Continued

66. Pine-dominated peat lands	< 12 m^3/ha
67. "	12-51 m^3/ha
68. "	52-101 m^3/ha
69. "	102-151 m^3/ha
70. "	152-201 m^3/ha
71. "	> 201 m^3/ha
77. Pine-dominated mineral soil forests	< 52 m^3/ha
78. "	52-101 m^3/ha
79. "	102-151 m^3/ha
80. "	152-201 m^3/ha
81. "	> 201 m^3/ha
88. Spruce-dominated mineral soil forests	< 52 m^3/ha
89. "	52-101 m^3/ha
90. "	102-151 m^3/ha
91. "	152-201 m^3/ha
92. "	> 201 m^3/ha
99. Deciduous-dominated mineral soil forests	< 52 m^3/ha
100. "	52-101 m^3/ha
101. "	102-151 m^3/ha
102. "	152-201 m^3/ha
103. "	> 201 m^3/ha
(104. Oak forests)	
(105. Little-leaf linden/Small-leaved basswood forests)	
(106. European aspen forests)	
110. Pine-Spruce mixed mineral soil forests	< 52 m^3/ha
111. "	52-101 m^3/ha
112. "	102-151 m^3/ha
113. "	152-201 m^3/ha
114. "	> 201 m^3/ha
115. Coniferous-deciduous mixed mineral soil forests	< 52 m^3/ha
116. "	52-101 m^3/ha
117. "	102-151 m^3/ha
118. "	152-201 m^3/ha
119. "	> 201 m^3/ha
132. Conifer plantations without coppice, mineral soil	
133. Conifer plantations with coppice or coppice thickets, mineral soil	
134. Alpine birch forests	
135. Sparsely stocked areas close to timber line, or coppice thickets	
136. Bare mountaintops	
142. Conifer plantations without coppice, peat land	
143. Conifer plantations with coppice or coppice thickets, peat land	
(252. Roads / NLSF database & 1: 50 000, railroads.) See class 24.	
255. or other values: missing data, background value, clouds etc.	

At least these classes are no more produced in the 3rd version: 110-114, 134-136. Class is put in parentheses, when it is not much used or if it does not include in the standard product.

/

Appendix 2: Production process of Land Cover and Forest Classification process, using National Forest Inventory. A. Vuorela, National Land Survey of Finland 1997