Working Papers of the Finnish Forest Research Institute 111
http://www.metla.fi/julkaisut/workingpapers/2009/mwp111.htm
ISBN 978-951-40-2151-0 (PDF)
ISSN 1795-150X

The Multi-source National Forest Inventory of Finland – methods and results 2005

Erkki Tomppo, Markus Haakana, Matti Katila, Kai Mäkisara and Jouni Peräsaari

METLA www.metla.fi

Working Papers of the Finnish Forest Research Institute publishes preliminary research results and conference proceedings.

The papers published in the series are not peer-reviewed.

The papers are published in pdf format on the Internet only.

http://www.metla.fi/julkaisut/workingpapers/
ISSN 1795-150X

Office
Post Box 18
FI-01301 Vantaa, Finland
tel. +358 10 2111
fax +358 10 211 2101
e-mail julkaisutoimitus@metla.fi

Publisher
Finnish Forest Research Institute
Post Box 18
FI-01301 Vantaa, Finland
tel. +358 10 2111
fax +358 10 211 2101
e-mail info@metla.fi
http://www.metla.fi/

Authors

Tomppo, Erkki, Haakana, Markus, Katila, Matti, Mäkisara, Kai & Peräsaari, Jouni

Title

The Multi-source National Forest Inventory of Finland – methods and results 2005

Year	Pages	ISBN	ISSN
2009	277	978-951-40-2151-0 (PDF)	1795-150X

Unit / Research programme / Projects

Vantaa Research Unit / National Forest Inventory Designs and Methods 3428

Accepted by

Pasi Puttonen, Director of Research, 2 February 2009

Abstract

This article presents the results of the Finnish multi-source forest inventory based on satellite images, digital map data and field data. The main purpose of the article is to make multi-source forest inventory results available for the foresters, and all the readers interested in forests and forest environment. The field data originates from the 10th Finnish National Forest Inventory from years 2004 and 2005 and the updated field data of the 9th National Forest Inventory (NFI9) from years 1996–2003. The basic features of the employed improved k-NN method are described, as well as the new updating method briefly. The data from the field sample from the NFI9 were updated to the mid of year 2005 identifying the major changes on the field plots with the help of satellite images. Tree level growth models were applied to the remained trees to add the volume increments between field measurement and 2005.

The results are presented by forestry centres and within forestry centres by municipalities. The estimates are given, for example for land areas, areas of tree species dominance, age and development classes and often separately for forests available for wood supply. The mean volume and total volume estimates are given in many different ways: mean volumes by tree species and by timber assortments for forest land, poorly productive forest land and their combination and also for forests available for wood supply, as well as by age and development classes. The biomass estimates are given by tree species groups in young thinning stands in which first commercial thinning was proposed for the first 5-year-period, separately for stem and bark and branches and foliage. The biomass estimates of mature forests are presented separately for branches, foliage and stem residuals, and stumps and large roots by tree species groups. All biomass estimates are given separately for land available for wood supply. The presented results can be employed in calculating forest resource estimates also for regions larger than municipalities. The error sources for the estimates as well as the reliability of the estimates are briefly discussed in Section 3.4.3 and Chapter 5.

Keywords

multi-source forest inventory, national forest inventory, satellite images, genetic algorithm, k-nearest neighbours, small-area estimation, stratification, statistical calibration, change detection, growth models.

Available at

http://www.metla.fi/julkaisut/workingpapers/2009/mwp111.htm

Replaces

Is replaced by

Contact information

Erkki Tomppo, The Finnish Forest Research Institute, Vantaa Research Unit, Vantaa Unit, PL 18, FI-01301 VANTAA, FINLAND. E-mail erkki.tomppo@metla.fi

Other information

Contents

List of abbreviations **5**
1 Introduction **6**
2 Materials **7**
2.1 Field data 7
2.1.1 Sampling designs 7
2.2 The employed satellite images 14
2.3 Digital map data 15
2.3.1 The use of the map data 15
2.3.2 The map data 16
2.4 Digital elevation model 19
2.5 Large area forest resource data 19
3 Methods **21**
3.1 Updating of the NFI9 field plots 21
3.1.1 Computation of the expected volumes of removal from the cuttings 21
3.1.2 Detection and updating of the field plots with regeneration cuttings 22
3.1.3 Detection and updating of the field plots with thinning 23
3.1.4 Growth models 23
3.2 Image rectification and radiometric correction of the spectral values 24
3.3 Preparation of input data sets 25
3.4 The estimation methods 27
3.4.1 Selecting estimation parameters and their values for k-NN 29
3.4.2 Area and volume estimates for small areas – correction for map errors 30
3.4.3 Assessing the errors 31
4 Results **31**
4.1 Forest resources by municipalities 31
4.1.1 Available energy wood 48
4.2 Digital thematic output maps 49
5 Discussion **54**
Acknowledgements **56**
References **56**
Appendix **60**

List of abbreviations

k-NN	k-Nearest Neighbour method
ik-NN	Improved k-NN method
NFI	National forest inventory
MS-NFI	Multi-source national forest inventory
Landsat	Land satellite
Landsat TM	Thematic Mapper, a high-resolution scanner system on board the Landsat-4 and Landsat-5 satellites
Landsat ETM+	Enhanced Thematic Mapper, sensor on-board Landsat-7
Spot XS HRV	Satellite pour l'Observation de la Terre with XS (Spot 2) or XI (Spot I) multispectral scanner, High resolution Visible (HRV) imaging systems
IRS LISS-3	Indian Remote Sensing Satellite, Linear imaging self-scanner3
NLS	National Land Survey of Finland
TOPO50	1:50 000 topographic map database
TOPO	Topographic database
PerusCD	Digital map corresponding to 1:20 000 printed maps
DEM	Digital elevation model
RMSE	Root mean square error
SE	Standard error
FRYL	Forestry land, in MS-NFI covers forest land, poorly productive forest land and unproductive forest land
FPPF	Forest and poorly productive forest land
PPF	Poorly productive forest land

1 Introduction

The First National Forest inventory (NFI) was carried out in Finland from 1921–1924. The tenth rotation began in 2004 and was completed by 2008. Based on the information from sample plots, estimates can be made for a country, or regions within a country, with a minimum size of about 300 000–500 000 hectares, depending on the parameter. The densities of plots are high enough to ensure that the resulting sampling errors are low for core variables such as areas of land classes and the volume of growing stock.

The administrative units in Finnish forestry are the forestry centres, which have forest areas between 800 000 hectares and 5 million hectares. However, the areas of the centres are divided into sub-areas, or forest management units, from which forest resource and forest status information are needed. High utilisation levels of forests and the changing practices of forestry and forest industries in the 1980's and '90's required more accurate, localised and up-to-date information than previously. The use of field data alone to respond to the increased information needs would have been an expensive alternative. Either a substantially denser field plot grid or some other type of information was required.

The development of the Finnish multi-source national forest inventory (MS-NFI) began in 1989, and the first operative results were calculated in 1990 (Tomppo 1990, 1991, 1996, 2006b). The MS-NFI was introduced during the 8th rotation of NFI (1986–1994) in the Pohjois-Savo forestry centre (see Fig. 2.1, Chapter 2). The first results for the entire country were published in 1998 (Tomppo et al. 1998b) and the second country level results in 2008 (Tomppo et al. 2008b).

For MS-NFI, methods were sought that would provide area and volume estimates, possibly broken down into subclasses, such as tree species, timber assortments, and stand-age classes. As an example of more recent information needs, the potential wood energy biomass from forests is also considered in this article due to its importance as an alternative energy source to non-renewable energy. In the optimal case, the MS-NFI method should be able to provide estimates for small areas as accurate as the field based method provides estimates at national and regional levels. Since the first implementation of this method, it has been modified continuously and new features have been added (Katila et al. 2000, Katila and Tomppo 2001, 2002). The core of the current Finnish MS-NFI method is presented in Tomppo and Halme (2004) as well as in Tomppo et al. (2008b).

A somewhat similar method, which combines field data and satellite images, has been developed and employed in a several other countries like Sweden (Nilsson 1997, Reese et al. 2002, 2003, Hagner and Olsson 2004), USA (Franco-Lopez et al. 2001, McRoberts 2006, McRoberts et al. 2002a, 2002b, 2006), Norway (Gjertsen, 2005), Austria (Koukal et al. 2005), New Zealand (Tomppo et al. 1999), China (Tomppo et al. 2001a), Germany (Diemer et al. 2000) and Italy (Maselli et al. 2005). Tomppo et al. (2008b) gives a list of references. The Swedish k-NN product is used for a multitude of purposes, as well as a basis for post-stratification to produce the official Swedish forest statistics, see also Tomppo et al. (2008b). Also, McRoberts et al. (2002a, 2002b) applied a k-NN products for post-stratified estimation. Other examples of the development work in USA are the studies by McRoberts (2006) and McRoberts et al. (2007), who presented a model-based approach to derive k-NN error estimates for a group of pixels at an arbitrary size, Finley et al. (2006) and Finley and McRoberts (2008), who presented two methods of increasing the efficiency of the k-NN search.

The progress of the Finnish NFI was changed somewhat for NFI10 (2004–2008). Measuring field plots throughout the entire country was the biggest change. One-fifth of the clusters are measured annually, except the clusters in the archipelago regions (Åland) between Finland and Sweden and the very northernmost part of the country, which was measured during one field season in 2007. The very northernmost part was not measured in NFI10 and will be measured next time during NFI11. At the same time, the inventory rotation was shortened to 5 years, or nearly half of it's previous rotation duration. Inventory progressed by regions from the fifth inventory (1964–1970) until NFI9 (1996–2003). This change makes it possible to compute the basic forest resource estimates annually for the entire country.

This change in inventory progress makes it possible to calculate also the MS-NFI results more frequently than during NFI8 and NFI9. In the NFI plans, MS-NFI results were decided to be calculated every second year for the most part of the country, and for North Finland every fourth year. The new approach to progress sets some additional challenges for the MS-NFI, e.g., field measurements from several years can and must be employed. In the first results, presented in this article, NFI10 field data come from the years 2004 and 2005. Field data from NFI9 are also employed in order to get a high enough number of field plots for training data for satellite image supported estimation. The inventory is called MS-NFI-2005 throughout this article.

The main users of the MS-NFI results, municipality level estimates and maps are the forestry authorities at forestry centres, forest industries and forest environment researchers. More details of the uses are given in Tomppo et al. (2008a and 2008b).

The objective of this article is to present the municipality-level results of the multi-source national forest inventory of Finland based on field data from NFI10 from years 2004 and 2005 and computationally updated field data from NFI9 (1996–2003). The updating method is also briefly described. Otherwise, references for methods are given.

2 Materials

2.1 Field data

2.1.1 Sampling designs

The Finnish national forest inventory is a sampling based inventory. The sample plots are arranged into clusters. An important aspect is that the cluster should represent one day's average work. The field measurements and assessments of the NFI are carried out on the field sample plots and on those forest stands that include at least a part of a field plot. The field sample plot is also a unit in the field data based estimation (Tomppo 2006a). The ninth inventory rotation was conducted during the years 1996–2003. The field work proceeded by forestry centres (Fig. 2.1).

The field sample plot both in NF9 and NFI10 has been a Bitterlich plot (angle-gauge plot) with a maximum distance (Fig. 2.2). The maximum distance detracts very little from the reliability of the estimates but decreases the amount of field work and also reduces possible errors caused by unobserved trees.

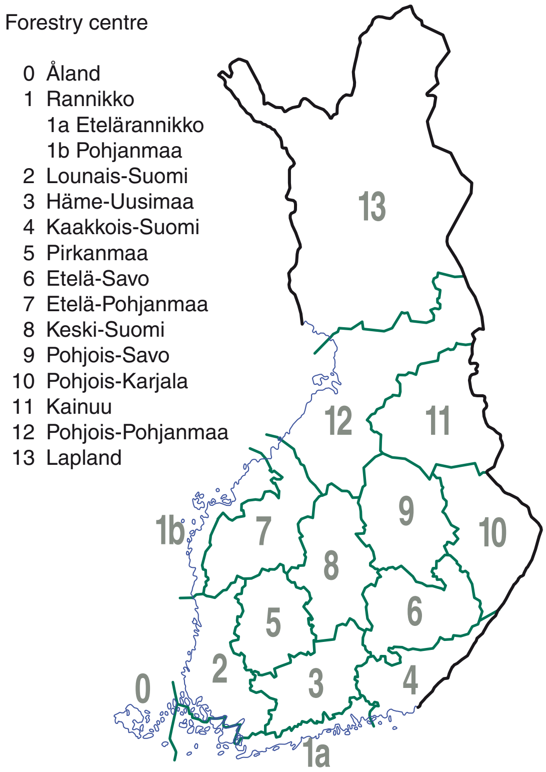


Fig. 2.1. Forestry centres 1.1.2007. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.

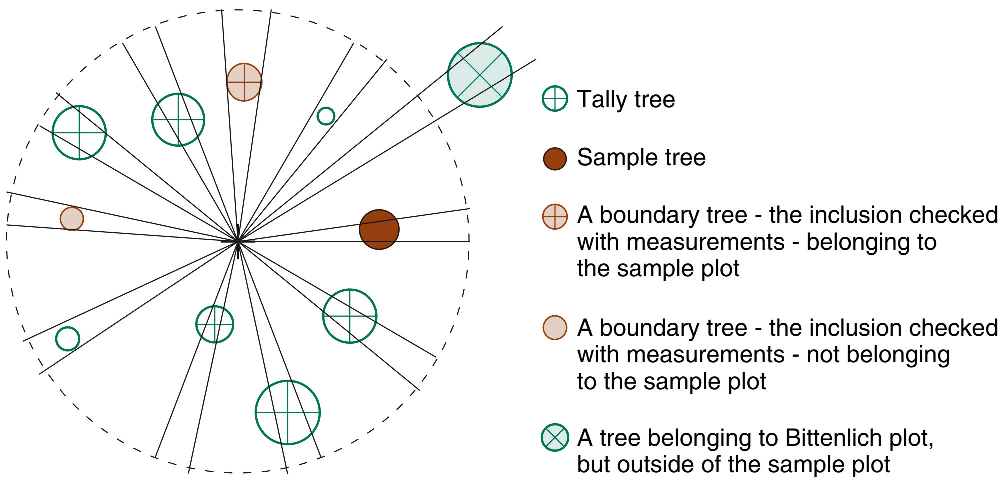
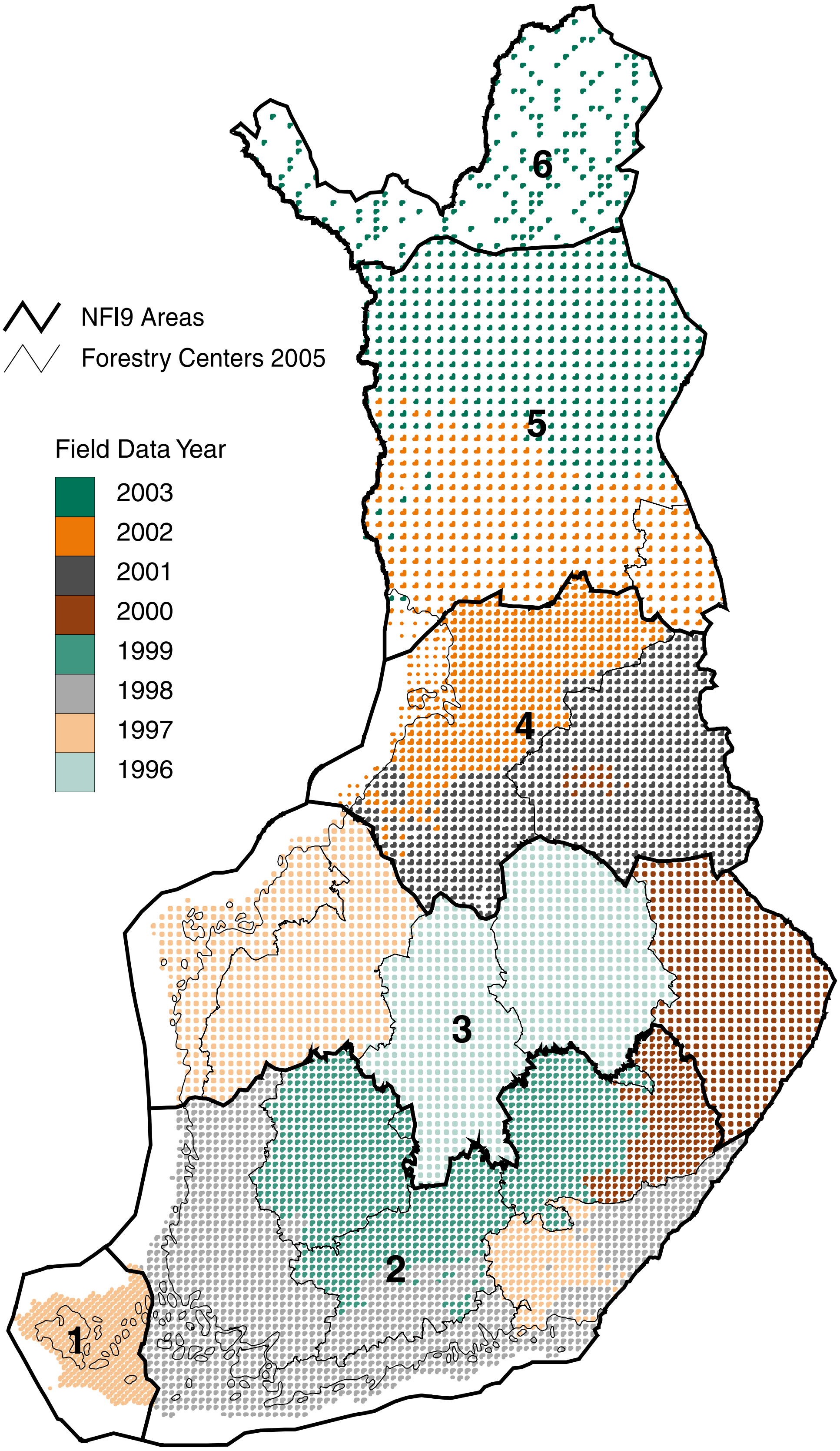


Fig. 2.2. A sample plot of NFI9 and NFI10. The maximum distance to trees to be counted is 12.52 m in South Finland (q=2) and 12.45 m in North Finland (q=1.5). Every 7th tree is measured as a sample tree, counted by crews and starting in the beginning of field season.

The sampling design was adapted to the variability of the forests. The six sampling density regions, together with field plot clusters, are shown in Fig. 2.3. The NFI9 sampling designs, cluster sizes, and distances between clusters, in the southernmost part of the country, Central Finland, North Central Finland, South Lapland, and North Lapland are shown in Fig. 2.4. The distances between clusters were 6 km by 6 km in the southernmost Finland (extra cluster lines were added to obtain high enough number of field plots in Åland, Region 1, Fig 2.3a.), 7 km by 7 km in Central Finland, 7 km by 7 km elsewhere in North Central Finland, with fewer plots than in Central Finland, and



a)

Fig. 2.3. NFI layout of clusters and the six geographic regions with different sampling densities, a) NFI9 1996-2003 and b) NFI10 2004-2005.

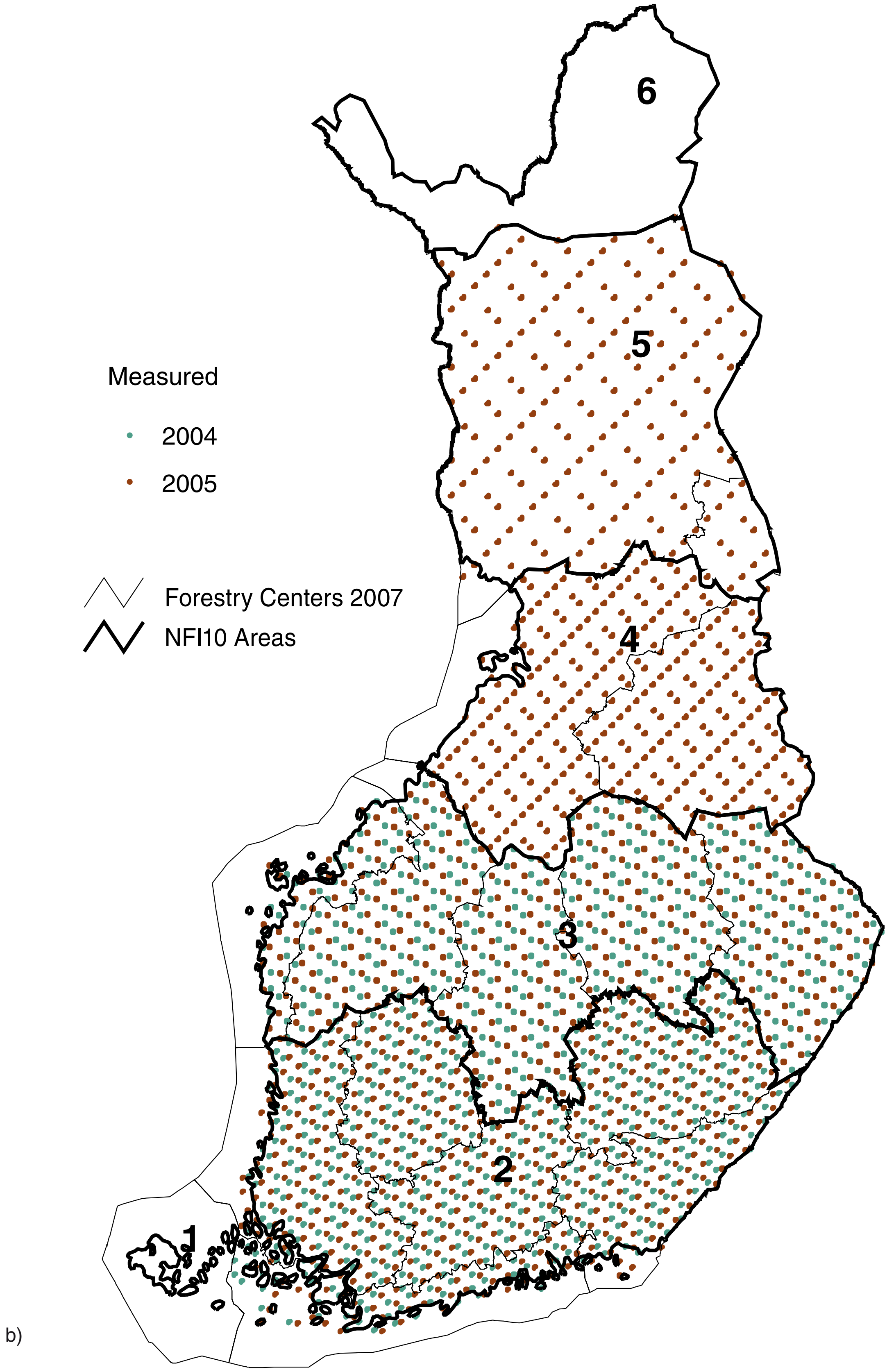


Fig. 2.3. Continued

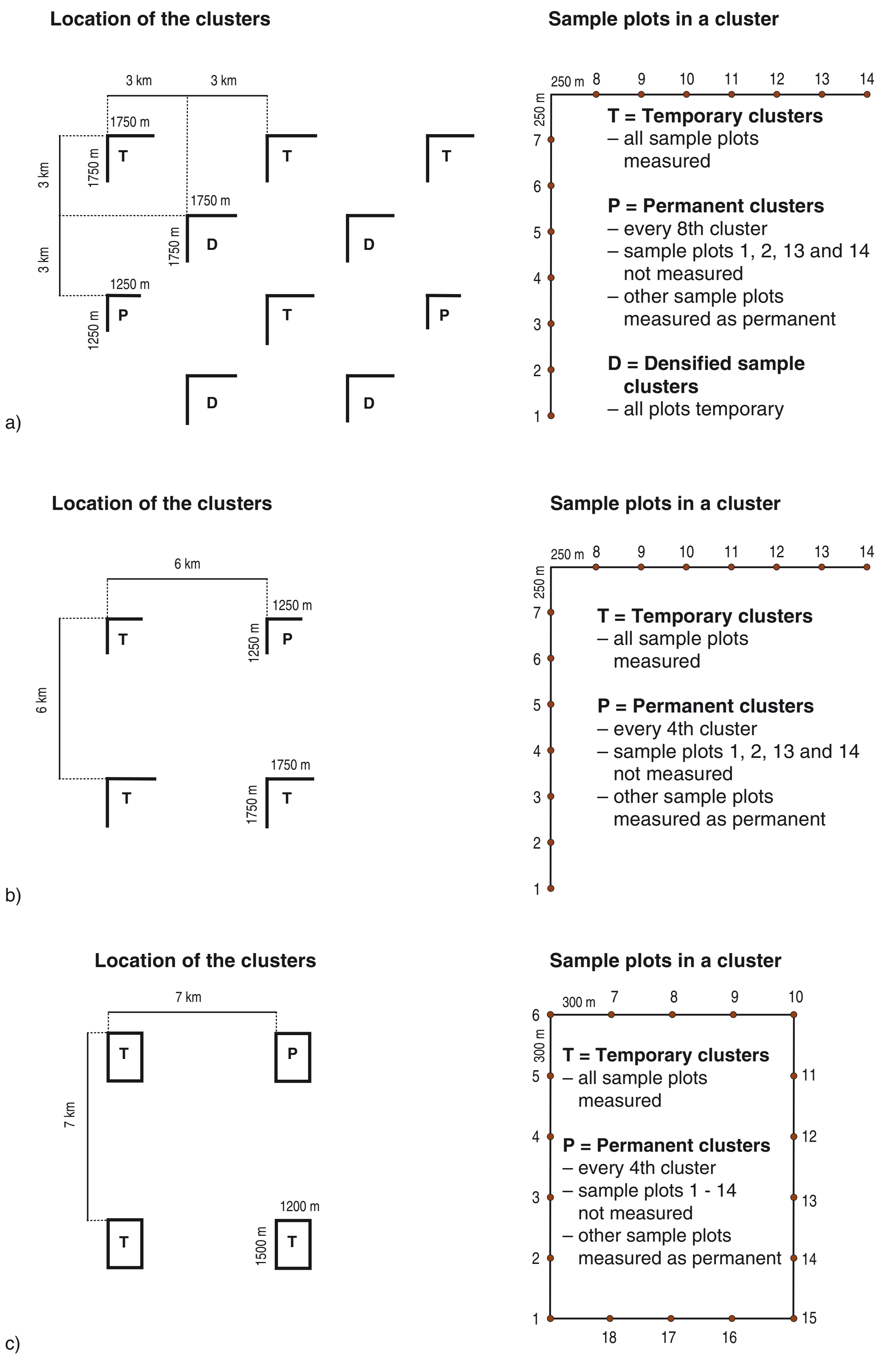


Fig. 2.4. The NFI9 sampling design in different inventory regions: (a) region 1, (b) region 2, (c) region 3, (d) region 4 (in region 5, the design is same but the distances are 10 by 10 km, (e) region 6. Stratified sampling was used in region 6.

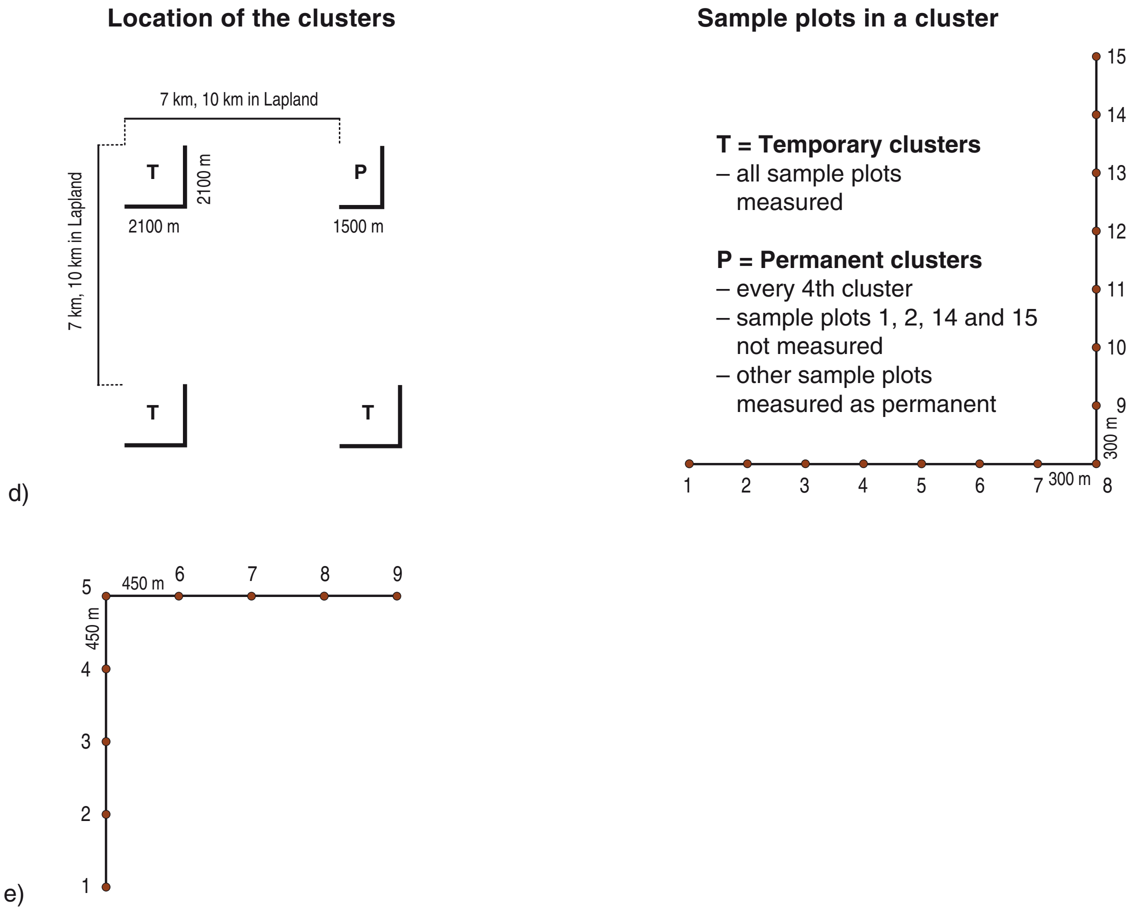


Fig. 2.4. Continued

10 km by 10 km in the municipality of Kuusamo and in South Lapland except in the area of the three northernmost municipalities.

Satellite image-based digital volume maps and sampling simulations were employed to evaluate different sampling designs. The sampling intensity was fitted to the spatial variation in forests throughout the whole country, being lower in the north than in the south (Henttonen 1991).

A similar approach was employed for NFI10 sampling design. The basic principles of the NFI10 design are similar to those of NFI9 (Tomppo 2008). The need to shift the locations of the clusters with temporary plots caused some changes. The clusters with temporary plots were shifted 1 km west and 1 km north in this rotation (Fig. 2.5). Shortening of the inventory rotation also added pressure to reduce the number of field plots slightly. New sampling simulation studies were carried out in all parts of the country for the forest area, mean volumes by tree species (m^3/ha), and total volumes (m^3) variables. The number of the field plots per cluster in different regions are shown in Fig. 2.5.

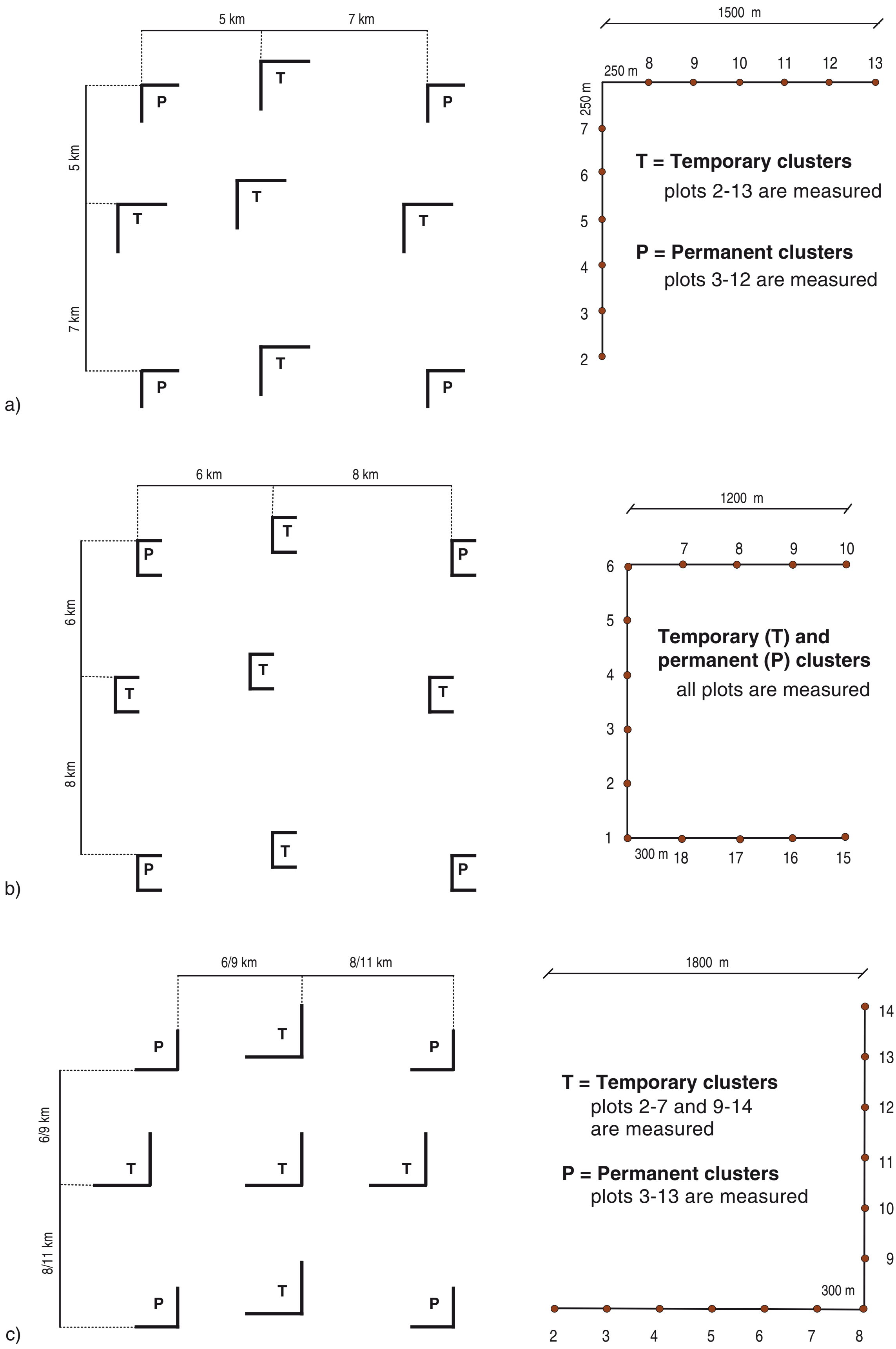


Fig. 2.5. Sampling designs of the NFI10 in different inventory regions: (a) region 2 (in region 1, the design is similar but the distances are 3 by 3 km, see Fig. 2.4), (b) region 3, (c) region 4 (in region 5, the design is same but the distances are 10 by 10 km,(d) region 6. Stratified sampling will be applied in region 6.

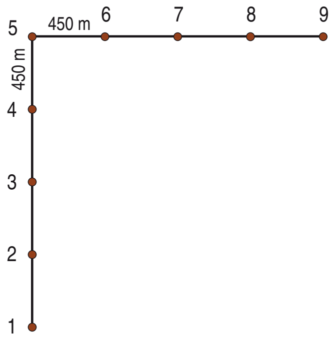
d) 

Fig. 2.5. Continued

2.2 The employed satellite images

High-resolution (about 20–30 metres pixel size) multispectral satellite images were used in this operative application. Large coverage and good availability with reasonable price were additional selection criteria. Based on these requirements, the Landsat 5 TM sensor was most suitable for this application. The IRS-P6 LISS-3 sensor provided another alternative and it was used in one area where Landsat 5 TM imagery was not available from the preferred time point.

The land area in the processed area is 20.97 million hectares, and the total area together with lakes and rivers is 23.78 million hectares. For processing, this area was covered using 13 images assembled from 24 satellite image frames (Table 2.1, Fig. 2.6). Where possible, adjacent satellite image frames from same path and date were combined to large images to simplify processing.

For MS-NFI-2005, the target year of image acquisition was 2005. However, three images (6 frames) were from 2006 and one image from 2004. A suitable imaging season in Finland is from mid-May until the end of August, with the optimal time being from early June until the end of July. The imaging season is rather short and cloud cover is frequent even in the summer. The image acquisition is, therefore, problematic and it is common that the satellite images can not be obtained from the same year as the field measurements have been carried out (Table 2.1).

Table 2.1. List of satellite images used in the MS-NFI10 2004-2005: image index number in Fig. 2.6, satellite sensor, path/row, image acquisition date and the forestry centres (Fig. 2.1) covered by the image (from largest to smallest cover).

Im-age No. (Fig. 2.6)	Sensor	Path/Row	Date	Forestry centres covered
1	Landsat 5 TM	186/16,16float	2.7.2005	10,6,4,9
2	Landsat 5 TM	186/16f	4.9.2005	10,4
3	Landsat 5 TM	187/14f,15f	9.7.2005	11,10,12,9
4	Landsat 5 TM	188/15f,16f,17f	2.9.2005	9,8,3,6,4,5,1a,11,12,10
5	Landsat 5 TM	190/14f/15f/16f	31.8.2005	12,5,8,11,7,1b,2
6	Landsat 5 TM	190/17,18	14.7.2005	2,5,1a,3,7
7	Landsat 5 TM	191/15f,16f	5.7.2005	7,12,1b,2,8,5
8	Irs P6	34/19	19.6.2005	12,11
9	Landsat 5 TM	190/14,15	17.7.2006	12,11,8,7
10	Landsat 5 TM	190/16,17,18	17.7.2006	5,8,2,7,3,1b
11	Landsat 5 TM	187/17f	4.6.2004	4,6,1a,3
12	Landsat 5 TM	188/18f	3.7.2006	1a,3,4
13	Landsat 5 TM	188/15f	14.6.2005	11,12

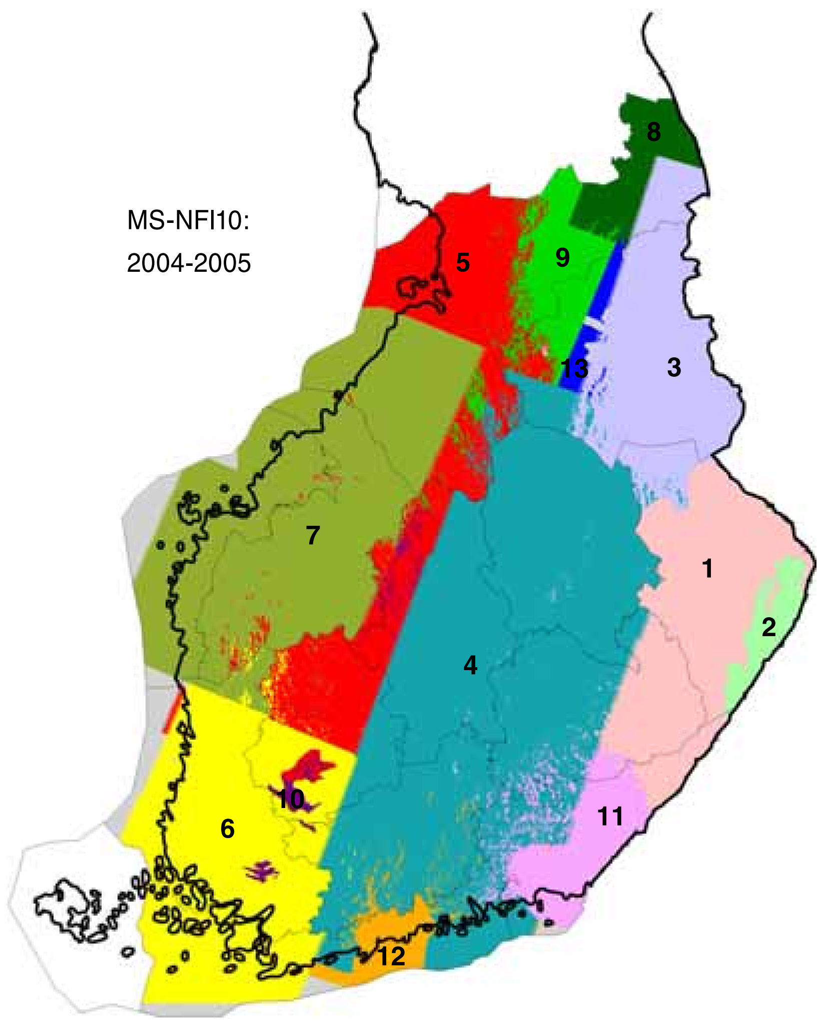


Fig. 2.6. The satellite image mosaic used to cover the Finland in MS-NFI-2005. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.

The number of satellite image frames used earlier in NFI9 (1996–2003) was 42, compared to only 24 in MS-NFI-2005. One reason for the large difference is the weather conditions combined with the satellite imaging schedule. However, there are also other reasons. The NFI9 results covered the whole country (land area 30.45 million hectares, total area with lakes and rivers 33.81 million hectares). In NFI9, the field measurements were done in only a small number of inventory regions each year and the target year of image acquisition was the year of the field measurements. This decreased the possibility of using large images assembled from several adjacent frames.

2.3 Digital map data

2.3.1 The use of the map data

Digital map data are used to reduce the errors in the estimates. The errors in both the area and total volume estimates can be reduced significantly by the multi-source method if the differentiation of forestry land from non-forestry land can be supported by digital map information in addition to satellite images (Tomppo 1996). The map information is used to separate forestry land from other land classes, such as arable land, built-up areas, roads, urban areas and single houses. The effect of possible map errors on small-area estimates is reduced by using one of two alternative statistical methods (Katila et al. 2000, Katila and Tomppo 2002, Tomppo et al. 2008b). The first is a calibration method using a confusion matrix derived from the land class distributions on the basis of field plot data and map data, and the second employs stratification of the field plots on the basis of map data (see Subsect. 3.4.2). In addition, the map data are used to stratify the forestry land area and the corresponding field plots into a mineral soil stratum and a peatland soil stratum (i.e. spruce mires, pine mires, open bogs and fens). This stratification decreases the prediction and estimation errors due to the fact that mineral soils and different types of organic soils (peatland soils) can have significantly different spectral signatures even when the growing stock is the same (e.g., Katila and Tomppo 2001). Digital map data is also used to delineate the computation units in the MS-NFI.

Almost all the map data were obtained from the National Land Survey of Finland (NLS), most often in raster format. The road data were rasterized from vector format to 25 m × 25 m pixel size. Some of the raster data were obtained in finer resolution and generalised to MS-NFI pixel size. The main principle in the rasterization and generalisation was to keep the total area covered by the map theme same as in the original data. The visual appearance of the non-forestry land classes in the MS-NFI output map was considered to be of secondary importance. The different map data were overlaid, ordered by the known accuracy of the data. The most precise map data was the final layer on the land use map. However, for the purposes of the calibration method (Subsect. 3.4.2), the overlaying was done in such a way that the formed map stratum would be as homogeneous as possible with respect to the NFI field plot based land class distribution (Katila et al. 2000). The main objective was to obtain as precise estimate as possible for the combined forest land, poorly productive forest land and unproductive forest land (denoted by forestry land) compared to the NFI field data based estimate employing the combined digital land use map (Fig. 3.2b). The accuracy of the combined land use map data was studied against the NFI field based land classes in Tomppo et al. (2008b).

2.3.2 The map data

In year 2000, the ´PerusCD´ digital map, corresponding to 1:20 000 printed maps, was obtained. The arable land and water were taken from these data. The main data source for ‘PerusCD’ was the NLS topographic database (TOPO) but in some regions -especially Lapland- map sheets were still from the 1:20 000 basic maps and dated from 1958–1990 (Table 2.2), see also Fig. 2.5 in Tomppo et al. (2008b). The topographic database has the most accurate positional data and is comparable to maps on scale 1:5 000–1:10 000 (Maanmittauslaitoksen maastotietokohteet 2005). In southern Finland, map data from the 1:50 000 topographic map database (TOPO50) were used, see Fig. 2.5 in Tomppo et al. (2008b) (Maastotietojen keruuohje 1996). TOPO50 map elements used were peatland and urban areas. All the map elements were rasterized to 25 m by 25 m pixel size.

2.3.2.1 Peatland

The spectral response of peatland differs from that of mineral soils with the same growing stock. In addition, some peatland cannot be separated from mineral soils by means of remote sensing. Therefore, digital peatland information is used in order to improve the accuracy of MS-NFI estimates (Tomppo 1996, Katila and Tomppo 2001). The site class definition is vegetation-based in the NFI: the forest stand is considered to be peatland (spruce mires, pine mires, open bogs and fens) if the organic layer covering the mineral soil is peat or if 75% of the under storey vegetation is peatland

Table 2.2 Sources and quality of the numerical map data.

Map theme(s)	Delivered by	Scale	Date of mapping	Area covered	Data source
Peatlands	NLS	1:100 000	1954-1987	northern Lapland	Topographic maps
Houses	NLS/PRC	-	1993	whole country	HR
Urban areas	Statistics Finland	-	1991	central & north Finland	HR
Roads	NLS	-	2002	whole country	Road data base
Peat production areas	Metla/NFI	raster 25m	1990-1994	central & north Finland	Digitized from satellite images
Municipality borders	NLS	1:100 000	2007	whole country	Administrational boundaries
TOPO50: urban areas, peatlands	NLS	1:50 000	1988-1995	southern Finland	Topographic maps 1:50 000
Topographic database: urban areas, peatlands, other built-up areas, open rocky soils, meadows, gardens	NLS	1:10 000 or 1:20 000	1993-2000	central & north Finland	Topographic database
PerusCD: arable land, water	NLS	1:20 000	1958-1999, mostly 1990-1999 in southern and central Finland	whole country	Base maps and topographic database

Abbreviations: Metla = Finnish Forest Research Institute, NLS = National land survey of Finland, HR = Housing register, PRC = Population Register Centre.

vegetation (Lehto and Leikola 1987). A geological definition of peatland is used for the topographic mapping: peatland is covered mainly by peat vegetation and the thickness of peat layer is over 30 cm. Thus, the peatland mask can not be used in a categorical way, but it is used to stratify the forestry land area and corresponding field plots for subsequent analysis in the estimation phase.

TOPO50 peatland data covered Southern Finland, see Fig. 2.6 in Tomppo et al. (2008b), while TOPO peatland data were obtained for the Central and Northern Finland in 2001. TOPO50 and TOPO data consisted of subclasses of open bogs, woody peatland and paludified lands. It was therefore possible to stratify the peatland in the k-NN estimation into open bogs and woody peatland (Subsect. 3.4). Paludified peatland, most often corresponding to spruce mires, were included to woody peatland. In Northern Lapland, TOPO data were not available and the peatland map data scanned from basic maps (scale 1:100 000) or GT-maps (scale 1:200 000) were used (Table 2.2) (Tomppo et al. 1998b). Thin peat layer mires (often spruce dominated) were missing from these data. In the TOPO50 and TOPO mapping process, the drained peatland, cf. corresponding main site class peatland in NFI, were not included to peatland element in Tomppo et al. (1998b). The accuracy of the scanned peatland map data and TOPO50 and TOPO peatland maps together was discussed in Tomppo et al. (2008b).

2.3.2.2 Arable land

Arable land is the third largest land class after forestry land and inland waters with an area of 2.794 million hectares in NFI9 (Tomppo et al. 2009). The area of the forestry land was 26.317 million hectares and the area of the inland watercourses 3.367 million hectares. Most of the land use changes occur between arable land and other land classes. The arable land element was taken from ´PerusCD´ digital map data (Table 2.2).

2.3.2.3 Urban areas, houses and other built-up areas

The coordinates of each house in Finland are kept in a housing register provided by the Civil Register of Finland. Buildings, when not covered by the urban or built-up area mask, were presented as squares of a size of 50 m by 50 m. These data were from the year 1993 (Tomppo et al. 1998b).

TOPO50 elements contained urban areas, or residential buildings and industrial buildings, see Fig. 2.7 in Tomppo et al. (2008b). Elsewhere, the urban area was derived from the housing register as follows: a group of houses with at least 200 inhabitants and covering a mutual distance of no more than 200 metres formed an urban area and this was combined into a connected built-up land polygon. These vector-form data were purchased from Statistics Finland and were converted into raster format after being buffered down by 75 m or 125 m from the original polygon to avoid the overestimation of urban areas (Tomppo et al. 1998b). These data originate from 1991.

TOPO map data contained a class of other built-up areas: e.g. mineral resources extraction areas, peat production areas, landfill areas, cemeteries, airfields, parks, sports and recreation areas, see Fig. 2.7 in Tomppo et al. (2008b). In addition, separate digitization from satellite images was made for these areas for MS-NFI10.

The peat production areas digitized for MS-NFI8 were used (Tomppo et al. 1998b).

2.3.2.4 Roads

The road database was based on NLS topographic database was updated in 2002. The vector format data was rasterized to 25 m by 25 m pixel size. The road class information was used to define the width of the road (13–75 m) in the rasterization process. The areal coverage of the roads was considered more important than continuity of the resulting rasterized roads (Table 2.2).

2.3.2.5 Water

The delineation of water was taken from ´PerusCD´ digital map. A conservative threshold of water applying near infrared spectral channel values was used to ‘backup’ the cases of possible missing map data or seasonal changes in the water level. An additional constraint to the above masks and water thresholds was applied in the k-NN estimation by using a digital elevation model: for the water pixels, the angle between solar illumination and the normal terrain should be near the angle defined for flat terrain, i.e., should not deviate from that angle more than a given small threshold (Tomppo et al. 1998b).

2.3.2.6 Digital boundaries of the computation units

The basic computation unit in the multi-source inventory is the municipality, of which there were 416 at the beginning of 2007. Their land areas range from around 1 000 hectares to some hundreds of thousands of hectares. Digital municipality boundaries are used to delineate the units (Tomppo 1996). The boundary information originates from NLS topographic database and was obtained in vector-form. The map data and land areas of the municipalities dating 1.1.2007 were employed to calculate the small area estimates (cf. calibration to official land areas, Subsect. 3.4.2).

2.3.2.7 Digital boundaries of protected forests

Some estimates for municipalities were calculated for forestry land available for wood supply (Subsect. 4.1, Appendix Tables). Areas of protected forests obtained from the Finnish Environment Institute were used for this purpose. The data was obtained in vector format dating 31.12.2005 and rasterized to 25 m by 25 m pixel size. The protected areas included strict nature reserves, national parks, wilderness areas, special protected areas, protected old-growth forest areas, protected herb-rich forest areas, mire conservation areas, nature reserves on private land (protected permanently or temporarily), protected areas established by the Finnish Forest and Park Service and natural habitat types preserved on the basis of Nature Conservation Act. The area of protected forests differs slightly from that employed in Nuutinen and Hirvelä (2006). The difference is caused by the fact that Nuutinen and Hirvelä (2006) includes the protected areas under nature conservation programmes as well as areas reserved for constructions. The most areas under the programmes were either realised or rejected by the date of the employed boundaries in this article wherefore the area of forests available for wood supply should be near the real one in here.

2.4 Digital elevation model

A digital elevation model is used in two ways: for stratification on the basis of elevation data and for correcting the spectral values by reference to the angle between solar illumination and the terrain normal (Subsect. 3.4) (Tomppo 1992, Tomppo et al. 2008b). Stratification in this context means using the maximum vertical distance from a pixel to its nearest neighbours. The selection of parameters for stratification and spectral correction has been studied by Katila and Tomppo (2001). The digital elevation model (DEM) employed was a raster file with a horizontal spatial resolution of 25 metres by 25 metres and with a vertical resolution of 0.1 metres (Digital Elevation Model 2007).

2.5 Large area forest resource data

The basic k-NN method was employed in NFI8. The improved ik-NN method was introduced during NFI9. The latter employs a coarse scale variation in the key forest variables to guide the selection of field plots from which the data are transferred to the pixel to be analysed. The variation is presented in the form of coarse-scale digital forest variable maps (Fig. 2.7), derived either from the current inventory data or from the data of the preceding inventory. In MS-NFI9, the field data from NFI9 was employed. The same maps were employed in calculating the MS-NFI-2005 results of the present article.

There were 81 249 field plots on land across the entire country in the NFI9 of which 67 264 were on forestry land, 62 266 on combined forest and poorly productive forest land (FPPF), and 57 457 were on forest land alone. All the plots on FPPF were used for the final large-area maps. The creation of the maps is described in Tomppo et al. (2008b).

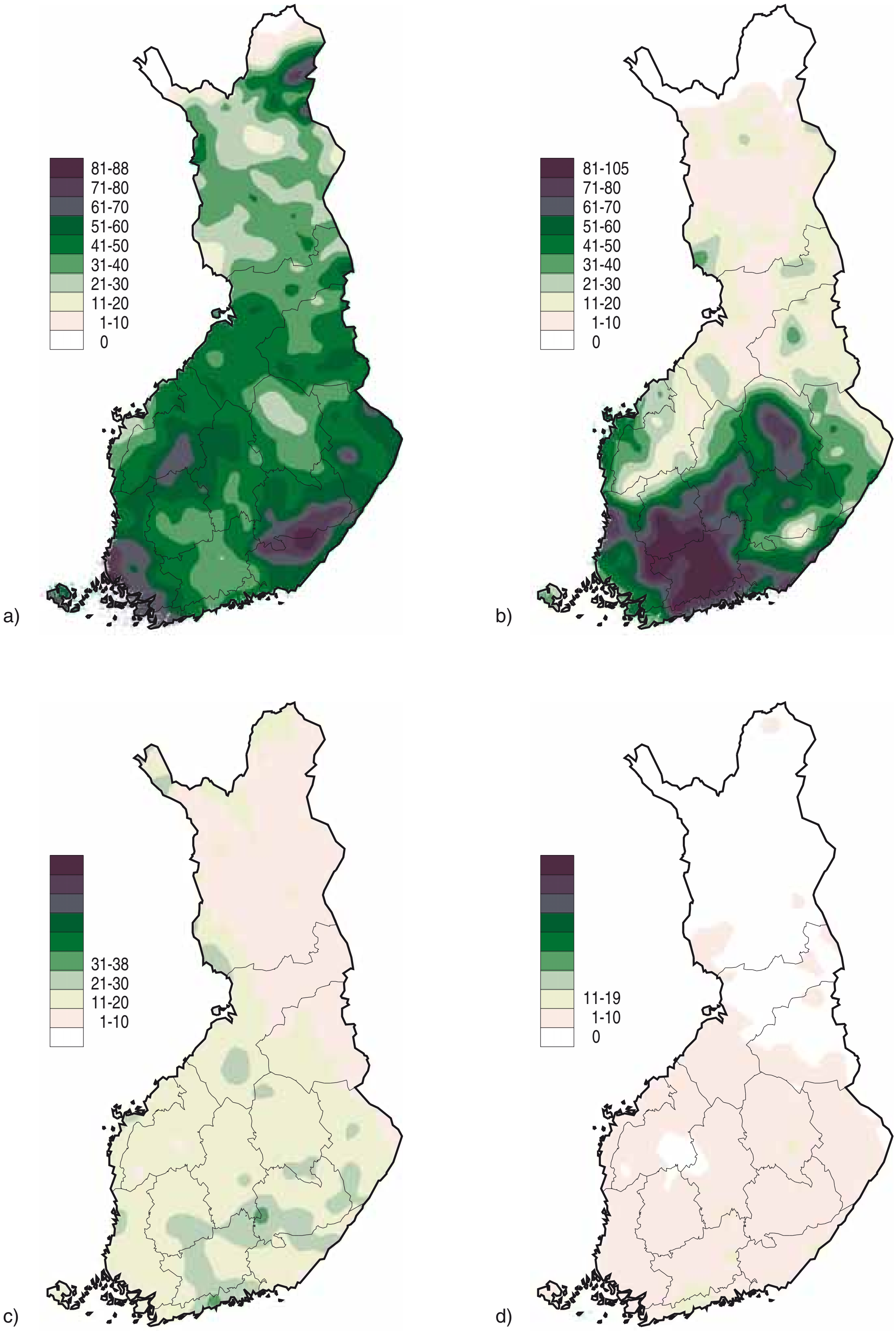


Fig. 2.7 Large scale variation of mean volumes (m^3/ha) of pine (a), spruce (b), birch (c) and other deciduous tree species (d) with boundaries of forestry centres. Digital map data: © National Land Survey of Finland, licence No. 363/MML/08.

3 Methods

3.1 Updating of the NFI9 field plots

The employed NFI9 field sample plot data was computationally updated to the date of the image data. Both image data and cutting statistics were used in the updating method. At the highest level of the method, the sample plots were divided into two groups:

1. Plots in which major changes have happened, e.g., regeneration cutting, thinning, forest damage, change in land use class, and
2. Plots where only normal growth has occurred.

The sample plots in group 1 were further divided into two subgroups:

1. Plots with regeneration cuttings, and plots having been thinned. In the first phase, the plots with likely regeneration cuttings were identified.From the remaining plots, the plots having likely been thinned were identified.
2. The rest of the plots were assigned to group 2.

Both detection of regeneration cuttings and detection of thinning use the same basic methodology with the following two components:

1. A model used to order the field plots into decreasing plausibility of regeneration cutting or thinning, and
2. A limit for the total volume of wood in cuttings in a defined region that is used to select the field plots in the order established by component 1.

The limits for cutting and household use of timber have been computed using statistics collected by Metla. Separate limits have been computed for each of tree species (pine, spruce, birch, other deciduous species) and roundwood assortments (logs, pulpwood, waste wood).

New forest variable values have been predicted for all plots. Growth models have been used to update the plots in group 2, in which no major changes had occurred. New values have been generated for group 1 based on the variables in existing plots similar to the predicted new cutting subclass (clear cut, seed tree cut, thinning) of the plots.

The updating method used in this work is based on the methods first developed in the project "Updating of the Finnish national forest inventory using remote sensing data" in years 1998–2001 (Mäkisara et al. 2001). Growth estimation is given in detail later. The method has since then been used in several cases when predicting forest variables based on current satellite data and updated sample plot data.

3.1.1 Computation of the expected volumes of removal from the cuttings

The expected volumes in both cutting classes were computed separately for each satellite image. The input data for the computations included the area covered by the satellite image, the updating interval (time between the field work and image acquisition), and cutting statistics. Several computing methods have been considered based on the statistics available from Metla (Metinfo 2007). Several different statistics and assumptions can be used to derive the cutting volume predictions.

The method used in this project was based on the statistics for private commercial cuttings for the six roundwood assortments provided by Metla annually (Metinfo 2007).

The process consisted of the following steps:

1. Extract from statistics the volumes for private commercial roundwood removals in each municipality in the region of interest for all years in the time interval of interest for each roundwood assortment.
2. Compute the ratio of total roundwood removals and private commercial roundwood removals for each forestry centre and roundwood assortment. These ratios are assumed to apply to all municipalities in the forestry centre.
3. Compute estimates for all commercial cuttings for each municipality by combining the previous two steps. The result is a prediction of total roundwood removals for each municipality.
4. Divide the cuttings in each municipality by the forest land area in the municipality to obtain total roundwood removals per unit area.
5. Compute the total expected percentage of total volume in cuttings using the field plots in the non-cloudy part of the satellite image. The representativeness (hectares) of each plot and the mean time between the field data acquisition at each plot and the date of the satellite image are used in this computation.

The cuttings per unit area were divided into regeneration cuttings and thinning using ratios between these cutting types computed for each forestry centre. These ratios have been computed using cutting statistics from years 2001 and 2002 obtained from the METINFO group (private communication).

3.1.2 Detection and updating of the field plots with regeneration cuttings

A spectral model was used to order the sample plots on the basis of likelihood of being cut. A different model was used for each satellite image. The model was in this case a linear regression model from the spectral channel values in the satellite image to the total volume, computed using the NFI10 field plots located within the satellite image. Separate models were computed for mineral soils and peatland soils. The likelihood of a regeneration cut was primarily defined as the magnitude of total volume difference between the field plot data and the prediction based on the model. The search was further constrained to plots in which the total volume in the field data exceeded a threshold (depending on the area) and the predicted volume was smaller than a threshold (depending on the area). Only plots in young and advanced thinning stands and mature stands were accepted as candidates.

The percentages of the total volumes in cuttings were transformed to absolute volumes using the field plots located within the non-cloudy part of the satellite image. The representativeness (ha) of each plot was taken into account to compensate the possibly different field plot arrangements within the satellite image area.

The tally tree data of a field plot designated as cut were modified to correspond to the new state of the plot. The regeneration cuts were divided into plots with natural and artificial regeneration (planting or sowing). The plots were assigned to the artificial and natural regeneration categories by comparing the spectral vector of a plot to the spectral average of plots of the mature stands in the NFI9 data. The average was computed based on the field data where cutting candidates have been removed. A predefined fraction of the plots nearest to the mature plot average in the spectral space were categorized as naturally regenerated. The fraction was computed separately for each forest centre using statistics on area treated with fellings from years 1995–2005.

Finally, the volume left on the plots was transformed to trees in the field data. In case of artificial regeneration, all existing trees were removed from a field plot. In case of seed-tree cut, the largest tree of a plot was left. All trees were also removed from plots with seed tree cut before the measurement.

3.1.3 Detection and updating of the field plots with thinning

The plots with likely thinning cuttings were recognized among those plots which had not been treated with regeneration cuttings. Only plots with no cuttings during past five years were accepted as candidates for thinning. The seed tree stands with seed tree removals during the update interval were also identified. The number of plots been thinned or having seed tree removals was calculated from forest statistics. The same thinning intensity was applied in the plots detected to be thinned as in the earlier project (Mäkisara et al. 2001) being 32% from the volume. Mainly the stand characteristics of NFI9 field plots were employed in allocating thinnings because they are difficult to identify using satellite images, particularly a single image. The plots with probable thinnings were selected and the strata were composed using of forest stand characteristics. The forestry centre, development class and dominant tree species group were same within each stratum. The three dominant tree species groups were pine, spruce and deciduous species. The spectral values of a satellite image and the basal area of the plot in the end of the updating period were used to order the sample plots according to the probability of being thinned within each stratum. Furthermore, the estimated number of plots to be thinned was divided into development classes. All of the pre-commercial thinnings were assumed to take place in young thinning stands (development class 4). Other thinnings were shared between advanced thinning stands and mature stands (development classes 5 and 6). The proportional shares of types of thinning in these development classes were derived from the thinning done on the plots of the satellite image coverage area, i.e., on the ground true plots, during past five years. The greatest likelihood for seed tree removals was on the seed tree stands, which had the longest time from earlier cuttings.

The tally tree data of a field plot designated as cut by thinning or seed tree removal were modified to correspond to the state of the plot after cuttings. From the plots with seed tree removal, all the trees with a diameter 200 mm or more were removed. The trees were removed from the thinning plots on the basis of a likelihood of a tree to be removed. The number of trees removed from thinning plots was calculated for each plot separately to achieve 32% of the volume on the plot which was the removal percentage for volume. By default, each tree had an equal probability to be removed, but the probabilities were adjusted for each satellite image separately in such a way that removals by roundwood assortments would be fulfilled.

3.1.4 Growth models

The NFI9 field plots were measured region-wise between 1996 and 2002 in the area which covered entire Finland, except Åland and Lapland, for which the updating was not made. The updating continued until the end of July in 2005, because most of the satellite images were acquired in that year. The length of the updating period varied from a little less than three years to nearly ten years, depending on the measurement date of the plots. The updating was carried out at the tree level, after which the updated variables were aggregated to the plot level. The updated variables were the volume of growing stock, age, basal area, mean diameter and mean height. The volumes were updated separately for tree species and timber assortments. The cutting type and cutting time were also updated for the plots with cuttings. The stand development class was updated only in case of regeneration cutting. Otherwise, the stand characteristics were assumed to remain unchanged.

Individual tree growth models of Hynynen et al. (2002) were applied to predict the growth for two following 5-year periods. The basal area growth and height growth were predicted with models and other variables like volumes were derived from these figures. The tree height is measured in NFIs only for the sample trees. For the tally trees, the heights at the beginning of the period were estimated with models developed by Mäkisara et al. (2001). In the growth models, a five-year growth period is applied (Hynynen et al. 2002) and the growths were calculated for two periods. The needed growth figures were interpolated from these predictions for each plot.

The tree volumes are estimated in a different way in the NFI than when using growth models. Therefore, all increment values were transformed into increment percentages. The updated variables were calculated by applying these increment percentages and the corresponding values of the variables on the NFI plots.

However, in the plots with growing stock less than 100 m^3/ha, the predicted values of the variables were used directly without turning growth figures into increment percentages (Mäkisara et al. 2001). All the growth predictions based on the measured tally trees, and no single trees, due to ingrowth, were generated at this phase. The natural losses were taken into account in the growth models (Hynynen et al. 2002).

Realisation of increment by growing trees did not work for the plots without any trees. A heuristic method to add trees to these stands was developed. The method copied trees to the plots in open regeneration and young seedling stands without trees from most similar plots measured in NFI10. The similarity was based on age, tree species, and development class.

3.2 Image rectification and radiometric correction of the spectral values

All images were rectified to the national uniform coordinate system. Point-type objects (e.g., small islands) were identified on both the satellite images and the basic maps and a regression model was fitted to their image coordinates and map coordinates. Second-order polynomial regression models were usually employed for this purpose. A typical number of control points has been around 50. The nearest neighbour method was applied to a re-sampling of the images to a pixel size of 25 m by 25 m. The absolute values of the residuals in the rectification models typically ranged from 0.3 pixels to 0.6 pixels.

Areas corresponding to the cloud-free parts of satellite images are used in operative applications. Forests under clouds and in cloud shadows are assumed to be similar on the average to those on the cloud-free part of the same areal unit (e.g., municipality).

The slope and aspect of the terrain locally change the illumination conditions of the surface and affect the reflectance from the ground and vegetation, as well as the radiance received by an imaging instrument. A digital elevation model was employed to remove the variation of the spectral values caused by the changes in the slope and aspect of the terrain. The details of the method are given in Tomppo (1992) and Tomppo et al. (2008b). The related parameter selection has been studied in Katila and Tomppo (2001).

3.3 Preparation of input data sets

Field measurement data, satellite images, different types of digital maps and large-area forest resource data are employed in the MS-NFI. The key input data file for the MS-NFI is the ground truth data, consisting of field plot data as well as image and map data associated to each plot. For the present results, those NFI9 sample plots in southern Finland that were incorrectly located, were removed from the ground truth data. NFI9 plots in Northern Finland and all NFI10 plots were located using GPS. Plots intersecting forestry land and other land classes (e.g., water, fields, built-up land) were rejected. About 2 to 7% of the plots were excluded from the training data on the basis of these restrictions (Katila and Tomppo 2001, Tomppo et al. 2008b).

NFI10 plots in which clear cuttings had been made between image acquisition and field measurement dates were removed from the ground truth data. Two methods were employed to remove clear-cut plots (Tomppo et al. 2008b). Plots covered by clouds or their shadows were also removed from ground truth data.

In the image analysis (Fig. 3.1), the input data sets were 1) ground truth data, i.e., one record for each centre point plot and sub-plot included in each plot part 1a) field data and 1b) satellite image data, 1c) digital map data, 1d) and other numeric feature data in text format, 2) a pre-processed satellite image, 3) a digital map of land classes and mire and open bog mask, 4) a digital elevation model and thereof derived image of the angle between terrain normal and sun illumination, 5) cloud and shadow delineation mask, 6) a large-area forest resource data and 7) a map of computation units to calculate small-area estimates (Fig. 3.2). The national uniform coordinate system was employed in the analysis. The sequential images from the same swath and all the other corresponding data sets were merged to be processed together.

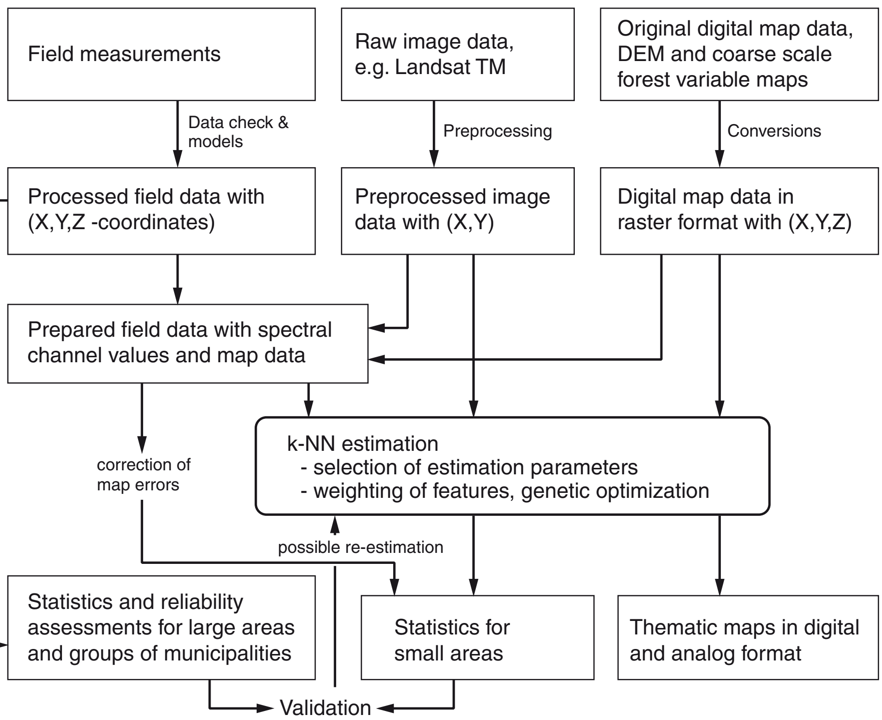


Fig. 3.1. Data flow and computational scheme for multisource NFI.

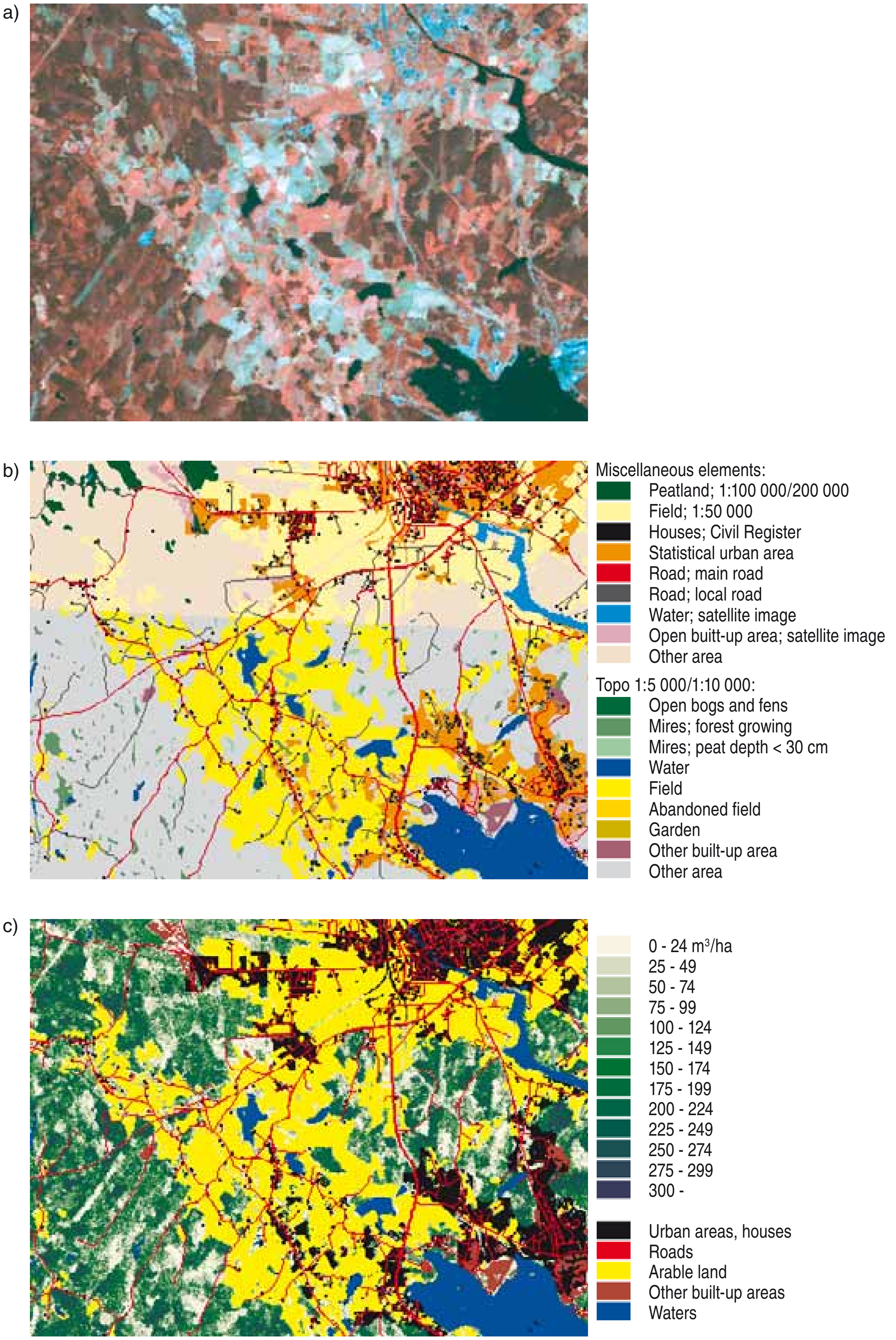


Fig. 3.2. Examples of the Landsat TM satellite image, multichannel colour composition of channels 2, 3, and 4 (a); the elements of digital land use map data, separate map data and topographic map database (b); and MS-NFI9 thematic map of total volume (m^3/ha) with other land use map data (c). Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.

The land class map was employed to distinguish combined forest land, poorly productive forest land and unproductive forest land (MS-NFI forestry land) from the other land classes. In this analysis all the area which was not peat production area, built-up land, arable land, roads or waters in the numerical map was considered forestry land. A threshold was used with satellite images to delineate those waters which were not in the numerical map, e.g. land covered by high water or fish ponds, as described in Subsect. 2.3.2.5. The numerical map data were not always up-to-date and could contain significant errors. The effect of the map errors on the estimates were corrected using two optional methods (Section 3.4.2).

3.4 The estimation methods

The non-parametric k-NN estimation has been employed in the MS-NFI calculation since operational inception in 1990. The method has been improved continuously. The details of the current method employed for the results of this article are given in Tomppo and Halme (2004) and Tomppo et al. (2008b). The basic principles are listed here. With the k-NN method, the plot weights (formula 3.4) (not equal for each plot) are computed for each plot by computation units, e.g. by municipalities (Tomppo 1996). The weights are computed for each field sample plot $i \in F$, where F is the set of field plots belonging to forestry land. These plot weights are sums of the weights which are computed for the field plots over all satellite image pixels on the forestry land mask of the computation unit. The plot weights corresponding to a single pixel (formula 3.2), in turn, are computed by a non-parametric k-NN estimation method (Tomppo 1991, 1996, Tomppo et al. 2008b). The method utilises the distance metric d, defined in the current version in the feature space of the satellite image data and coarse scale forest variables. The k nearest field plot pixels p_i, in terms of d, i.e., pixels that cover the centre of a field plot $i \in F$, are sought for each pixel p under the forestry land mask of the cloud free satellite image area. A maximum *geographical* distance is employed, if necessary, set in both horizontal and vertical directions in order to avoid selecting the nearest plots (spectrally similar plots) from a region in which the response of image variables to field variables is not equal to that of the pixel under consideration, due to, e.g., by changing atmospheric conditions or a large image frame. The feasible set of nearest neighbours for pixel p is thus

$$\{p_i \mid d_{p,p_i}^{(x,y)} \le d_{\max}^{(x,y)}, d_{p,p_i}^{z} \le d_{\max}^{z}, R(p_i) = R(p)\} \tag{3.1}$$

where $d_{p,p_i}^{(x,y)}$ is the geographical horizontal distance from pixel p to pixel p_i, d^z is the distance in the vertical direction, $d_{\max}^{(x,y)}$ and $d_{\max}^{z}$ are their maximum allowed values, and $R(p)$ is the indicator function of land class on the basis of map data (Tomppo, 1990, 1991, 1996, 2006b, Katila et al. 2000, Katila, & Tomppo, 2001, Tomppo et al. 2008b).

Denote the k nearest feasible field plots by $i_1(p),...,i_k(p)$. The weight $w_{i,p}$ of field plot i to pixel p is defined as

$$\begin{aligned} w_{i,p} &= \frac{1}{d_{p_i,p}^{t}} \Big/ \sum_{j \in \{i_1(p),\ldots,i_k(p)\}} \frac{1}{d_{p_j,p}^{t}}, \text{ if and only if } i \in \{i_1(p),\ldots,i_k(p)\} \\ &= 0 \text{ otherwise.} \end{aligned} \tag{3.2}$$

The distance weighting power t is a real number, usually $t \in [0,2]$. A small quantity, greater to zero, is added to d when $d= 0$ and $i \in \{i_1(p),...,i_k(p)\}$. The distance metric d employed in MS-NFI-2005 was

$$d_{p_j,p}^2 = \sum_{l=1}^{n_f} \omega_{l,f}{}^2 (f_{l,p_j} - f_{l,p})^2 + \sum_{l=1}^{n_g} \omega_{l,g}{}^2 (g_{l,p_j} - g_{l,p})^2 \tag{3.3}$$

where

$f_{l,p}$ is the lth normalised intensity value of the spectral band image variable, normalising done on the basis of digital elevation model, $f_{l,p_j} = f^0{}_{l,p_j} / \cos^r(\alpha)$, with α the angle between sun illumination and terrain normal, r the user given power for the cosine correction, $g_{l,p}$ the large area prediction of the lth applied forest variable, n_f the number of image variables (or features), n_g the number of coarse scale forest variables and ω_f and ω_g the weight vectors for image features and coarse scale forest variables respectively.

The values of the elements of the weight vectors ω_f and ω_g are computed by means of a genetic algorithm (Tomppo and Halme, 2004 and Tomppo et al. 2008b).

A pixel size of 1 km by 1 km is used in the coarse scale forest variable predictions $g_{l,p}$. The first phase of the improved version of k-NN, ik-NN, is to run the optimization algorithm in the applications possibly by strata, e.g., mineral soil stratum and mire and bog stratum. The estimation after that is similar to the basic k-NN estimation.

For computing forest parameter estimates for computation units, sums of field plot weights to pixels, $w_{i,p}$ are calculated by computation units, for example, by municipalities, and by map stratum h over the pixels belonging to the unit u. An example of stratum is mineral soil forestry land. The weight of plot i in stratum h to computation unit u is denoted

$$c_{i,h,u} = \sum_{p \in u_h} w_{i,p}. \tag{3.4}$$

Reduced weight sums $c_{i,h,u}^r$ are obtained from the formula (3.4), if clouds or their shadows cover a part of the area of the computation unit u. The real weight sum for plot i is obtained expanding the weight (3.4) by the ratio forestry land divided by the forestry land not covered by the clouds in each computation unit.

The weights (3.4) are computed within forestry land separately for mineral soil stratum and peatland strata. The weights are also computed for other land classes, arable land, built-up land, roads and waters using the plots falling in the corresponding stratum if the stratification based map correction method is employed (Katila and Tomppo 2002), and plots falling into forestry land map stratum if the calibration method is used (Katila et al. 2000).

After the final field plot weights to computation units ($c_{i,h,u}$) have been calculated, the ratio estimation is employed to obtain the small-area estimates (e.g., Cochran 1977). In this way, the estimation is similar to that using field plot data only (Tomppo 2006a, Tomppo et al. 2008b).

Predictions of some forest variables are written in the form of a digital map during the procedure. The land classes outside forestry land are transferred to map form predictions directly from the digital map file. Within forestry land, the variables are predicted by the weighted averages of the *k* nearest neighbours (Tomppo 1991, 1996)

A pixel-level prediction $\tilde{y}_p$, of variable Y for pixel p is defined as

$$\tilde{y}_p = \sum_{i \in I_h} w_{i,p}\, y_i \, , \tag{3.5}$$

where y_i is the value of the variable Y on plot i and I_h the set of the field plots belonging to map stratum h.

The mode or median value can be used instead of the weighted average for categorical variables. Mode has turned out to work better than median in the practical tests. The predicted variables are usually land class, site fertility class, stand age, mean diameter of stand, mean height of stand, stand basal area, volumes by tree species (pine, spruce, birch, other broad leaved trees) and by timber assortment classes. The total number of the maps was, therefore, 20 (Table 4.4).

3.4.1 Selecting estimation parameters and their values for k-NN

The basic principle of k-NN estimation is straightforward. However, practice has shown that the predictions and estimation errors depend to a large extent on the core estimation parameters of the k-NN algorithm. These are:

1. the variables employed in the distance metric, spectral bands or their transformations, possible correction for variation in illumination angle of the pixel based on elevation variation (slope, aspect), (Tomppo 1996),
2. the distance metric (Tomppo and Halme 2004),
3. the value of *k,* (Katila and Tomppo 2001),
4. the weights to be attached to the nearest neighbours, e.g., even weights or functions of the applied distance and powers (negative),
5. the variables applied in restricting the area from which the nearest neighbours are sought for a pixel, e.g., a geographical reference area (Katila and Tomppo 2001),
6. the use of additional information, e.g., large area variation of forest variables in the distance metric (Tomppo and Halme 2004),
7. the use of ancillary data in the estimation, e.g., for stratification.

The parameters and their values in MS-NFI-2005 are given in Table 3.1. The parameters are selected by image scene and the selection is documented. The criteria are the mean square error and bias of

Table 3.1. Applied ik-NN and k-NN estimation parameters employed in MS-NFI-2005.

Parameter	Choice
Variables applied in the distance metric	Illumination corrected spectral values for satellite image bands (Landsat TM 1-5, 7; IRS-P6 LISS-3 1-4) and large area forest variable estimates
Distance metric	Weighted Euclidean distance
Value of k	5-10
Weights attached to the nearest neighbours	Weights proportional to the inverse or inverse squared distance (t=1 or 2)
Restrictions for search of nearest neighbours	A maximum vertical (100 m or more in Northern Finland) and horizontal reference area(HRA) (40-120 km), since NFI9 large area forest variable maps are used to direct the NN selection, possibly with a HRA limit

pixel level predictions using leave-one-out cross validation, and particularly, the difference between areal estimates based on i) multi-source inventory and ii) on the field data based estimates and their standard errors (Tomppo et al. 2008b). The differences of the areal estimates are assessed in terms of standard error based on the field data plots (e.g., Katila and Tomppo 2002, Tomppo and Halme 2004). The values of the parameters usually vary by image depending on, e.g., imaging conditions, number of available field plots and variability of forests. The selections are not independent. A change in one parameter affects the optimal value of the other parameter. More studies are needed to optimize the values simultaneously.

3.4.2 Area and volume estimates for small areas – correction for map errors

In the multi-source estimation, numerical map data (see Sect. 2.3) are employed to decrease estimation errors. If the numerical map data would be error free, the computation unit weights (Eq. 3.4) could be calculated using pixels belonging to forestry land (according the map data) only. However, map data can be out-of-date, include location errors and does not correspond exactly to the definitions of NFI land classes. Errors can also arise during the post-processing of map data. Two methods have been developed to reduce the effect of map errors on small-area multi-source forest resource estimates: a statistical calibration method (Katila et al. 2000, Katila 2006a) and a stratified k-NN method (Katila and Tomppo 2002).

The calibration method is based on the confusion matrix between land use classes of the field sample plots and corresponding map information. The bias in the land class or other total cover estimates, obtained, e.g., from remote sensing or map data, can be corrected by means of the error probabilities expressed as a confusion matrix (Czaplewski and Catts 1992, Walsh and Burk 1993), assuming that the employed field sample are based on a statistical sampling design (Card 1982).

The applied map strata are defined in such a way that each stratum is reasonably homogeneous with respect to the ‘map errors’ and the NFI land class distribution. This enables the use of the synthetic small-area estimation method when correcting map errors (Rao 2003). The method utilises the error and land class proportions that have been estimated from a larger region. In the MS-NFI-2005 both NFI9 and NFI10 field plots were used to compute the confusion matrix. The method given in Katila et al. (2000) was used to calculate the calibrated field plot weights. The calibration typically increases the mean volume estimates and reduces the FRYL area estimates for small areas, if FRYL is overrepresented on maps. Calibration was carried out by forestry centres or groups of municipalities. However, for some calculation units, calibration did not decrease the errors of the estimates compared to the estimates based on the field data only, and was not used. Despite the rather simple idea of the calibration, it is quite laborious when implemented in the MS-NFI.

The MS-NFI employs a digital database for municipality boundaries, while the field inventory employs land and water areas from official statistics of the Finnish Land Survey (Suomen ... 2007). The area information from the latter data source is more accurate and there are slight differences between the total areas of municipalities from these two data sources. Hence, after the correction of map errors, the MS-NFI municipality land areas are calibrated to the official land areas. The calibration coefficient is straightforward $A_{U,Land\_NLS} \big/ A^*_{U,Land}$ and this ratio is assumed to also hold for forestry land and the (calibrated or stratified) weights $c_{i,U}$ are multiplied by this coefficient. For the calibrated MS-NFI, the calibrated land area $A^*_{U,Land}$ must be first estimated, see Tomppo et al. (2008b). The calibration to the official land areas is valid only for (random) deviations between the two data sets and not for the case where real and significant boundary changes between municipalities have taken place in either of the two data sources.

3.4.3 Assessing the errors

Deriving an error estimator for an arbitrary group of pixels has proven to be a challenging task. The problem can be divided into the derivation of i) an error estimator for a pixel level prediction and ii) an error estimator for a parameter for an area of interest.

Difficulties arise because:

1. errors depend on the actual value of the variable to be predicted and so pixel-level errors are spatially dependent
2. the variables measured or observed on the field plots are also spatially dependent
3. the spectral values of adjacent pixels of a satellite image are dependent due to the atmospheric properties (scattering) and imaging technique.

Furthermore, several error sources make the error estimation complex, examples of such error sources are given in Tomppo et al. (2008b).

During the data processing phase in the Finnish MS-NFI, the pixel-level root mean square error (RMSE) and the pixel level average bias are calculated using leave-one-out cross-validation using the available field plots. This is also a part of the employed genetic algorithm and the selection of the estimation parameters of k-NN and ik-NN. For a sufficiently large area consisting of a group of pixels, e.g., for areas of 200 000–300 000 ha, the MS-NFI estimates are compared to the estimates and error estimates based solely on field data. Some empirical error estimates are also available for reliability assessments (Katila 2006b, Tomppo et al. 2008a and 2008b). Standard error estimates for groups of pixels are calculated as described in Tomppo et al. (2008b). Some recent developments in error estimation, particularly in model-based error estimation are also presented in that publication, see also Kim and Tomppo (2006), McRoberts et al. (2007) and Magnussen et al. (2009).

4 Results

4.1 Forest resources by municipalities

The primary results of MS-NFI are the forest resource estimates for municipalities. With the MS-NFI method, it is possible, at least in theory, to estimate for municipalities all the parameters which are usually estimated for forestry centres using field data only. The estimates are presented in Appendix Tables 1–8 for those parameters whose estimates are considered to be sufficiently precise. Some results are given separately for land available for wood supply except in case of Åland. In the result calculation phase, the MS-NFI estimates for the groups of the municipalities and at forestry centre level (regions) were compared to the estimates and error estimates based on field data only in order to analyse possible significant errors, including biases.

The estimates can be divided into area and volume estimates. Some tables present only area estimates, some only volume estimates and some volume estimates for sub-area categories of forest land or poorly productive forest land, together with area estimates of the sub-area categories.

The estimates of forest land, poorly productive forest land and unproductive forest land (three forestry land categories) are given in Appendix Table 1a for the entire forestry land and in Appendix Table 1b for forestry land available for wood supply. The areas and proportions of forestry land of mineral soils and peatland soils are given in Appendix Tables 2a and 2b separately for three forestry land categories. Similar estimates for forestry land available for wood supply are given in Appendix Tables 2c and 2d.

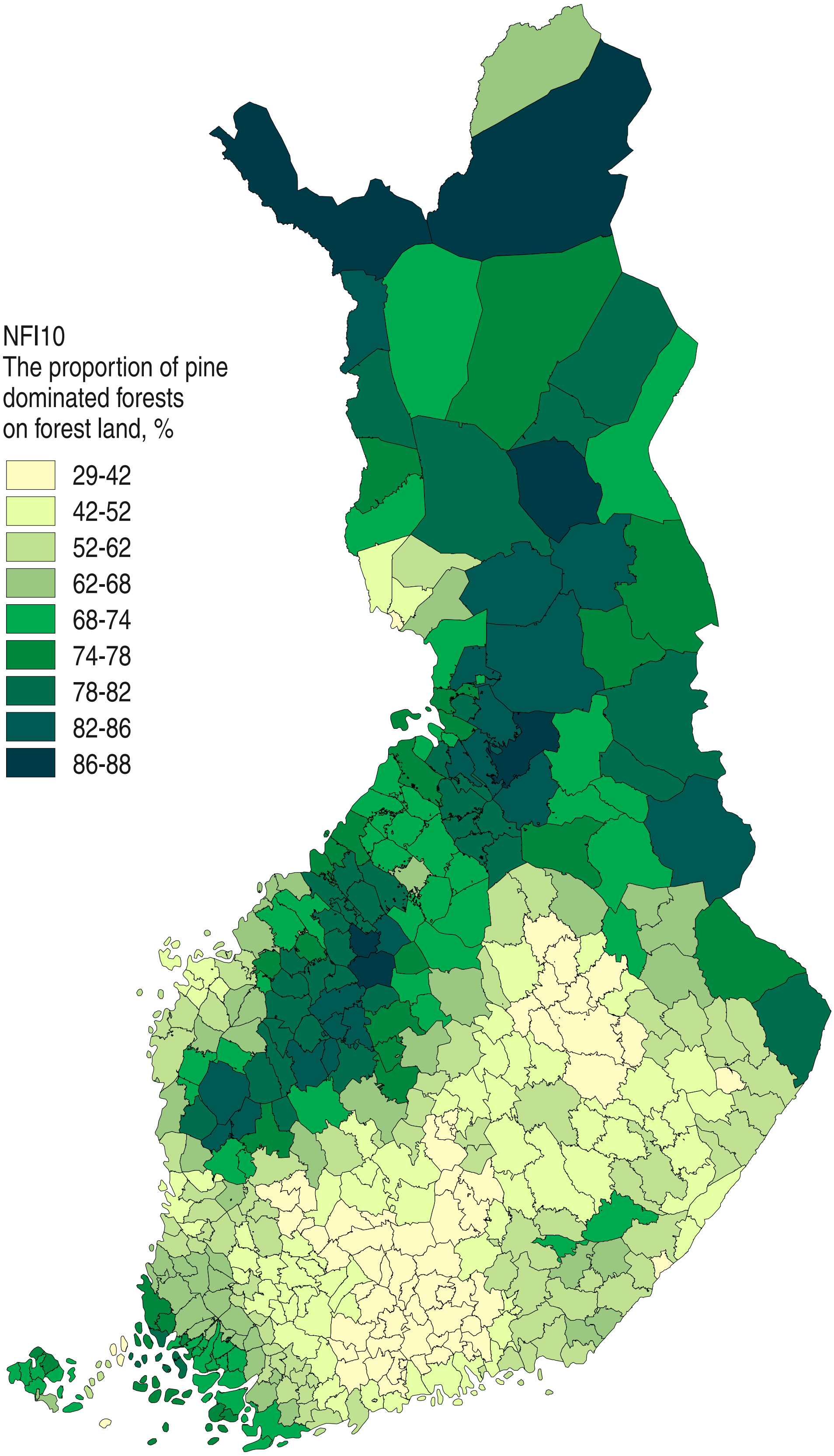
The dominant tree species by municipalities are presented in Appendix Table 3a and Fig. 4.1 for forest land and in Appendix Table 3b for poorly productive forest land. The dominant tree species is defined in the NFI for the field assessment as a stand-level variable. It is the tree species with the highest volume for the development classes from young thinning stand to mature stand, and is defined as the tree species with highest number of stems capable of development in young and advanced seedling stands. The proportion of pine dominated forests of forest land is usually highest in Northern Finland, often over 80%, with the highest proportion in the municipality of Inari, 89%. It is high also in some areas in Southern Finland in coastal regions and Central Finland in areas where Sub-xeric heath forests are common. The proportion of spruce dominated forests varies in Southern Finland between 5 and 56% and in Northern Finland between 4 and 40%, and gets its highest value, 56%, in Hauho in forestry centre Häme-Uusimaa and 55% in Pälkäne (Pirkanmaa forestry centre).

The area estimates for age classes on forest land by municipalities is presented in Appendix Table 4a and the development class estimates in Appendix Table 5a. The proportion of forest land area with an age of not more than 40 years is presented by municipalities in Fig. 4.2a and that with an age of over 120 years in Fig. 4.2b. The proportion of forest land with a stand age not more than 40 years varies by municipality in Southern Finland, from 12 to 52%. The proportional area of young forests is high in Eastern and South-Eastern Finland, and also in some municipalities in Central Finland. The range of the same proportion in Northern Finland is from 3% (in Utsjoki, in Northern Lapland) to 51% (in Kemi, in Pohjois-Pohjanmaa). One should note that the same age in Southern Finland and Northern Finland corresponds different development class of stand due to slower growth in the North.

The combined proportion of forest land area belonging to development classes open regeneration area, young seedling, advanced seedling, young thinning, seed tree or shelterwood stand are presented by municipalities in Fig. 4.3a. The proportion of the area of mature stands, ready for regeneration cuttings are presented in Fig. 4.3b.

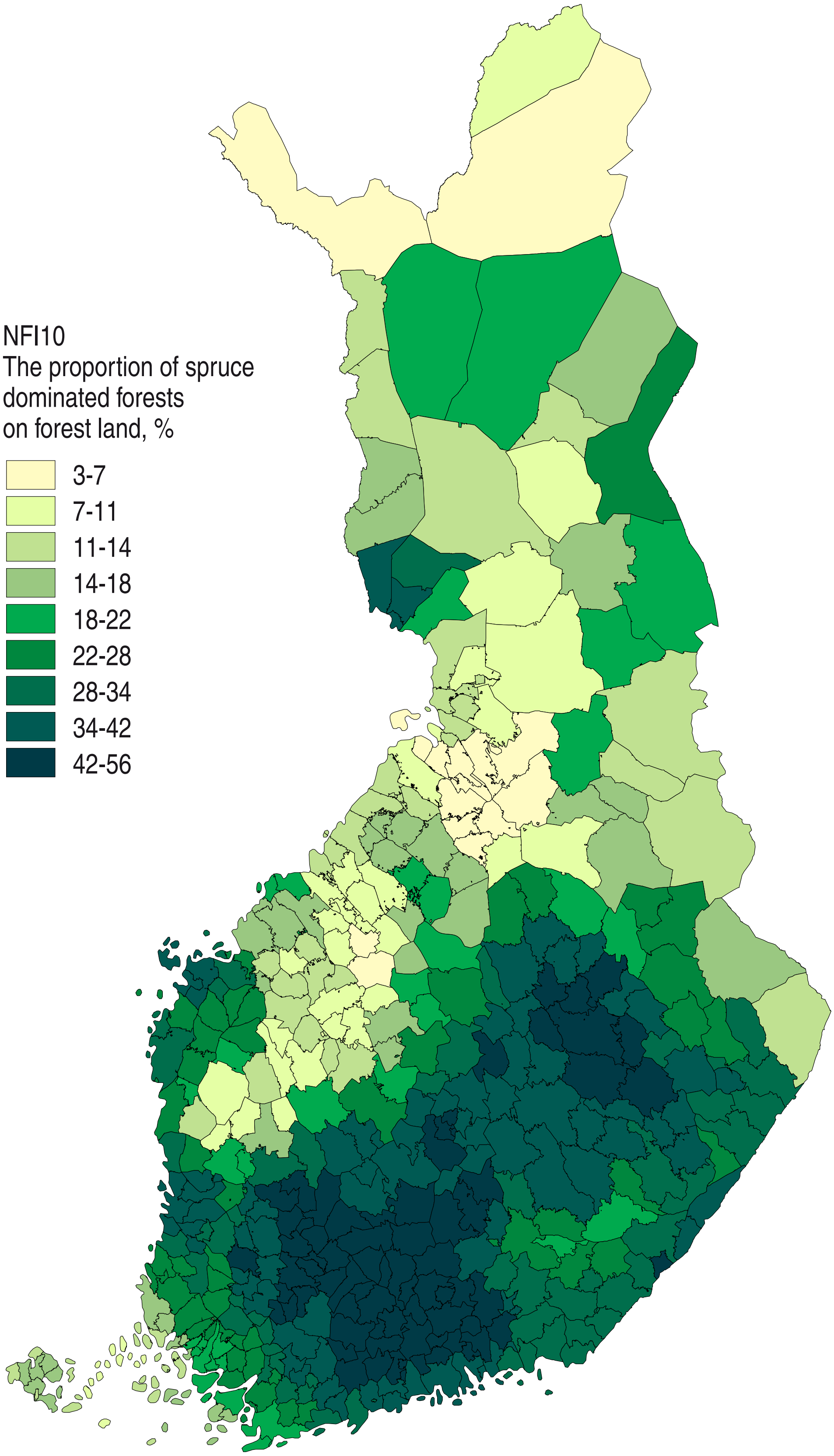
The combined proportion of young forests on forest land available for wood production, regeneration areas, seedling stands and young thinning stands is lowest in Southern Finland in some municipalities in South-Western costal region varying between 20 and 30% and is highest in northern parts of Pohjois-Savo, around 65%. Note that an ‘ideal proportion’ on the basis of some commonly used rules is 55%, 25% for seedling stands and regeneration areas and 30% for young thinning stands. The proportion of young forests on forest land available for wood supply is much higher in Northern Finland than in Southern Finland. The highest proportions vary between 75 and 85% in Lapland (Appendix Table 5c).

The mean volume and total volume estimates are given in many different ways: mean volumes by tree species and by timber assortments for combined forest land and poorly productive forest land (FPPF) (6a) likewise with total volumes (6b), see also Fig. 4.4. The similar estimates are given for FPPF available for wood production in Appendix Tables 6c and 6d. Appendix Tables 7a-d present



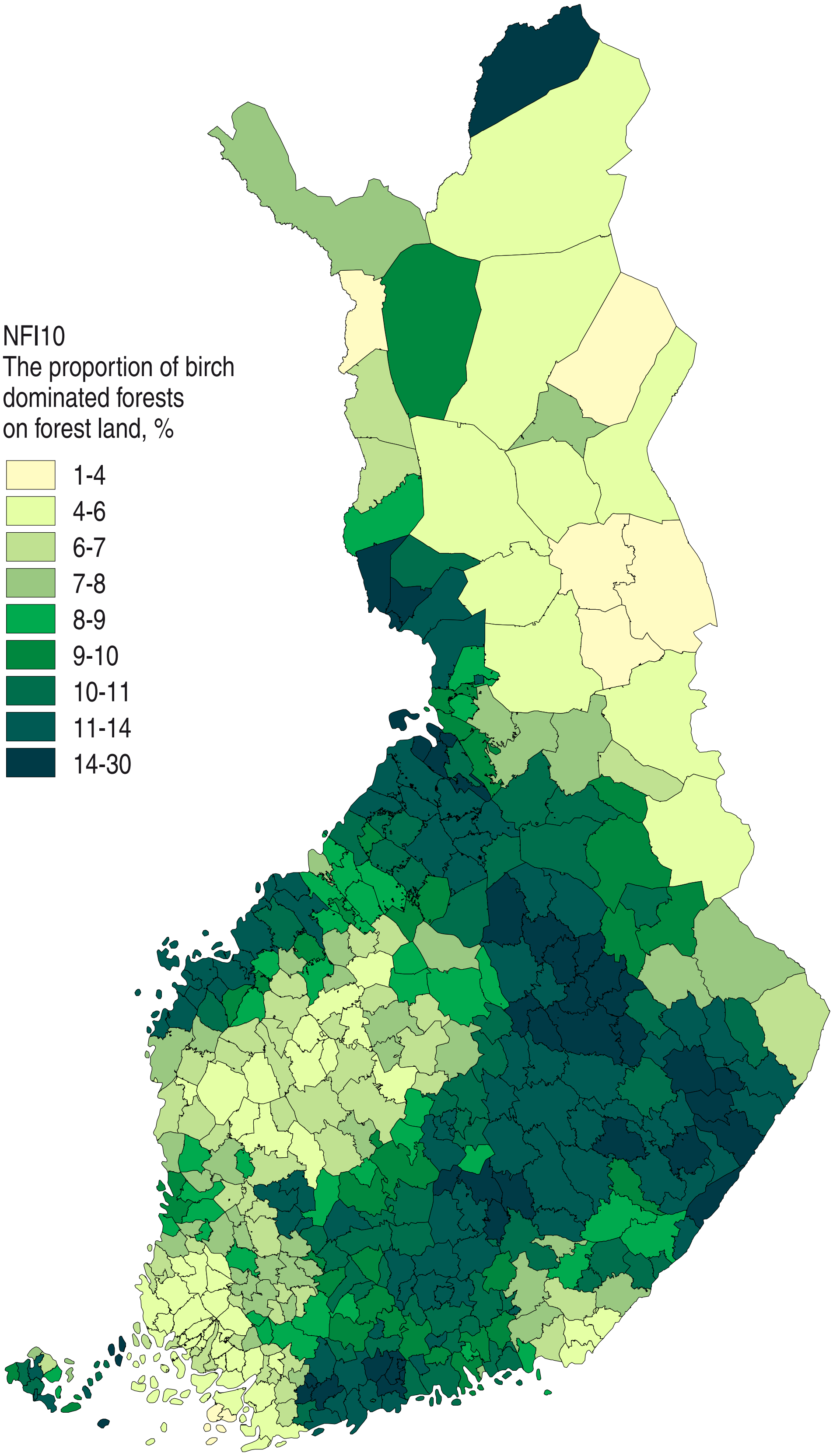
a)

Fig. 4.1. The percentage of pine (a), spruce (b), birch (c) and other deciduous species (d) dominated forests on forest land by municipalities. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



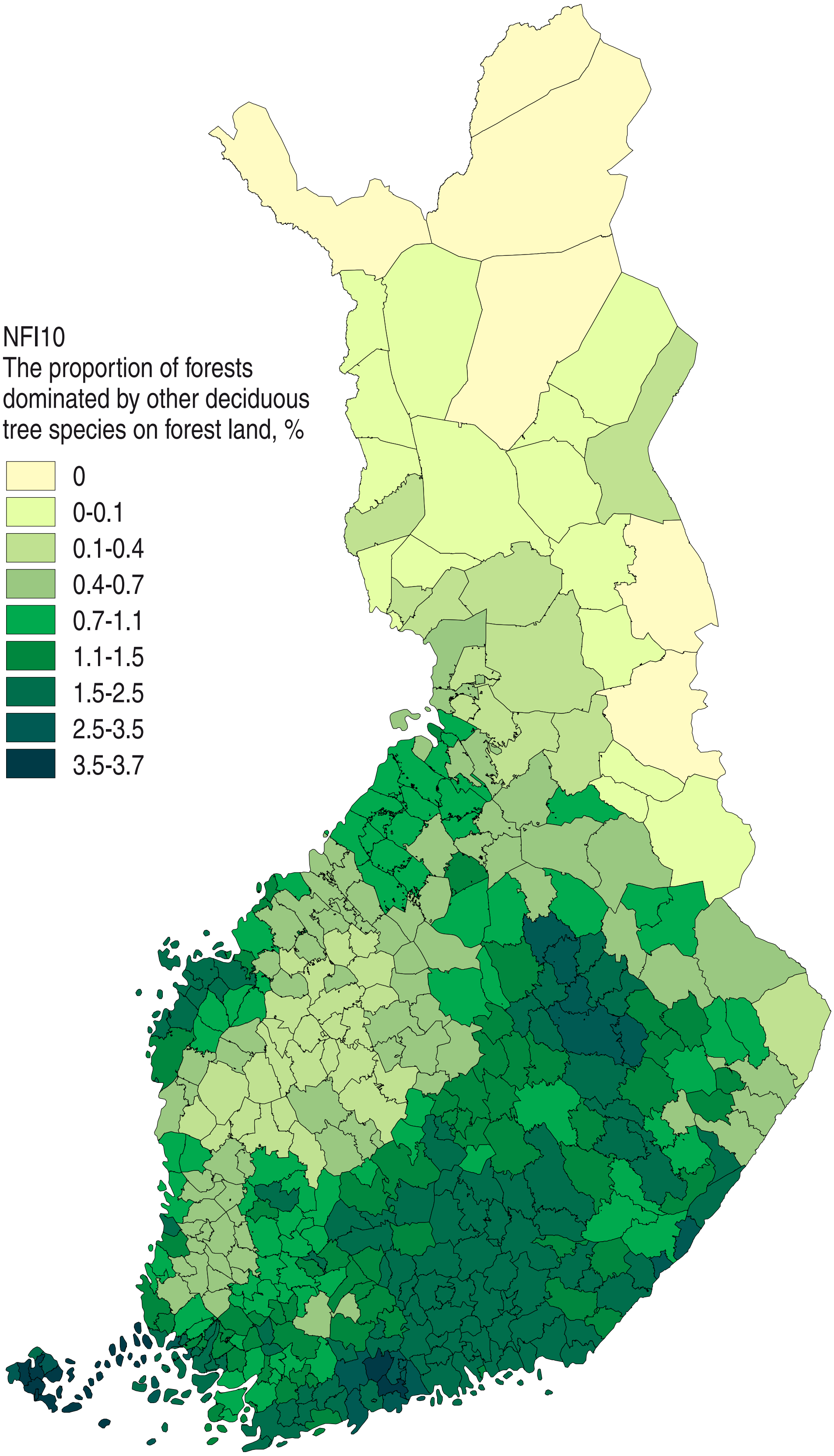
b)

Fig. 4.1. Continued



c)

Fig. 4.1. Continued



c)

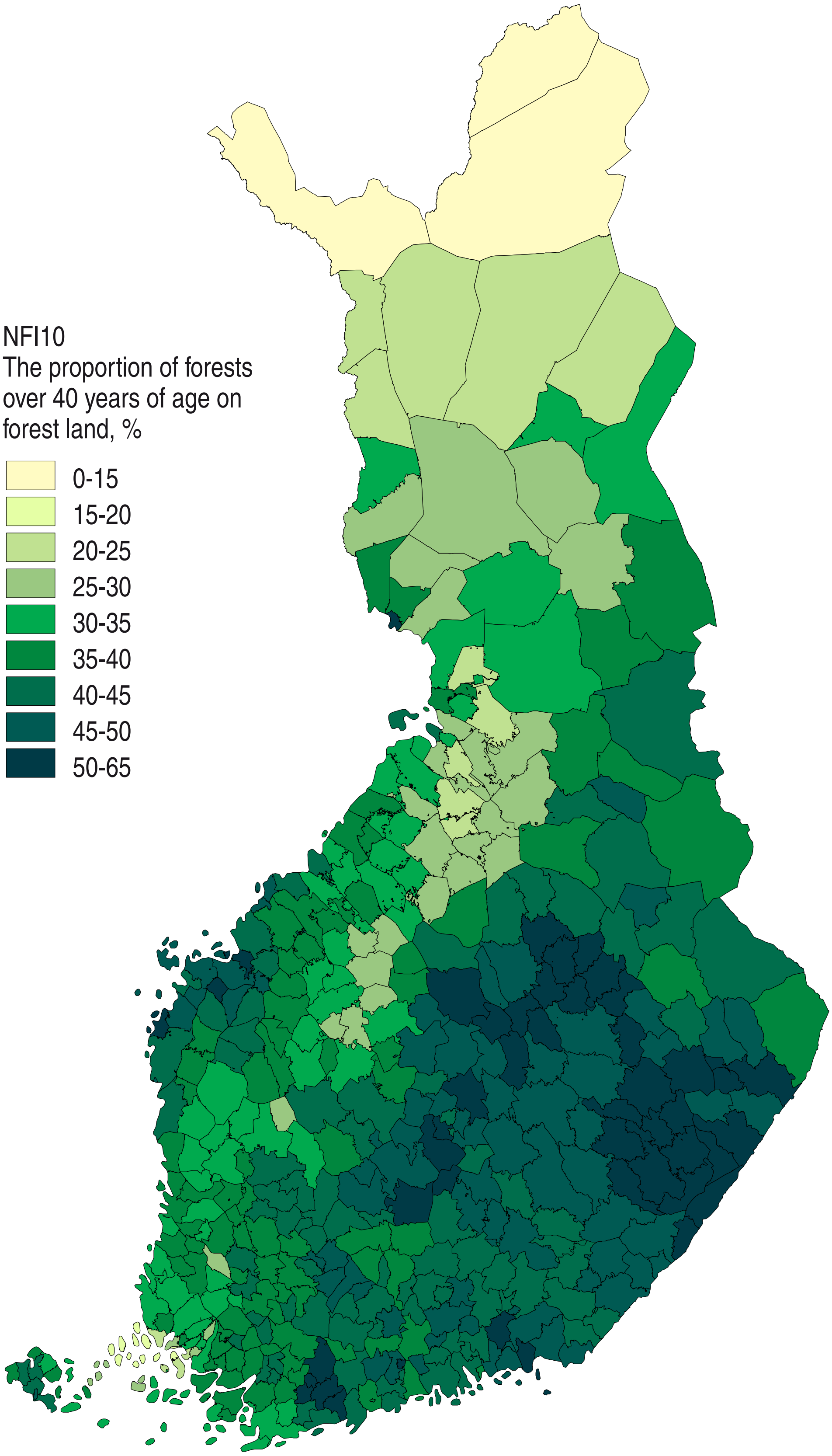
Fig. 4.1. Continued

Table 4.1. Number of municipalities in 10 m^3/ha mean volume classes on forest land and forest land area in the municipalities by forestry centres for MS-NFI-2005 (a) and MS-NFI9 (b).
a)

Forestry centre		Mean volume, m^3/ha													
		51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	141-150	151-160	161-170	171-180	Total
Åland	Number of municipalities					2	1	2		3	7	1			16
	Forest land area in the municipalities, ha					4205	874	3684		13251	39549	6231			67794
Rannikko/ Etelärannikko	Number of municipalities								1	7	3	10	3	2	26
	Forest land area in the municipalities, ha								8160	73808	51197	180765	30237	30376	374543
Rannikko/ Pohjanmaan rannikko	Number of municipalities					2	8	4		1					15
	Forest land area in the municipalities, ha					18954	308877	142343		10323					480497
Lounais-Suomi	Number of municipalities					2		2		20	29	16	1		70
	Forest land area in the municipalities, ha					50896		56721		295311	381531	158876	33212		976547
Häme-Uusimaa	Number of municipalities									1	12	18	10	3	44
	Forest land area in the municipalities, ha									4545	175378	365691	272265	105244	923123
Kaakkois-Suomi	Number of municipalities								2	7	11	3	1		24
	Forest land area in the municipalities, ha								44785	262609	355903	127096	9186		799579
Pirkanmaa	Number of municipalities						1	2	1	3	9	11		1	28
	Forest land area in the municipalities, ha						64183	119667	55773	142930	180997	311529		42625	917704
Etelä-Savo	Number of municipalities									5	10	3			18
	Forest land area in the municipalities, ha									411048	652921	128435			1192404
Etelä-Pohjanmaa	Number of municipalities				2	19	14	5							40
	Forest land area in the municipalities, ha				29685	612376	431520	180350							1253931
Keski-Suomi	Number of municipalities					1	3	6	3	11	2	1	1		28
	Forest land area in the municipalities, ha					37782	154640	397763	174810	519769	43943	38569	5754		1373030
Pohjois-Savo	Number of municipalities						4	2	10	3	3	1			23
	Forest land area in the municipalities, ha						382660	79017	500100	135038	162891	91101			1350807
Pohjois-Karjala	Number of municipalities					1	3	7	2	3					16
	Forest land area in the municipalities, ha					66855	649783	527392	106372	121722					1472124
Kainuu	Number of municipalities			3	2	4									9
	Forest land area in the municipalities, ha			699293	260668	771225									1731186
Pohjois-Pohjanmaa	Number of municipalities			8	12	15	3								38
	Forest land area in the municipalities, ha			1224942	498596	616949	122625								2463112
Lapland	Number of municipalities	10	8	3											21
	Forest land area in the municipalities, ha	2938917	1857329	216549											5012795
	Total, number of municipalities	10	8	14	16	46	37	30	19	64	86	64	16	6	416
	Total, ha	2938917	1857329	2140784	788949	2179242	2115162	1506937	890000	1990354	2044310	1408293	350654	178245	20389176

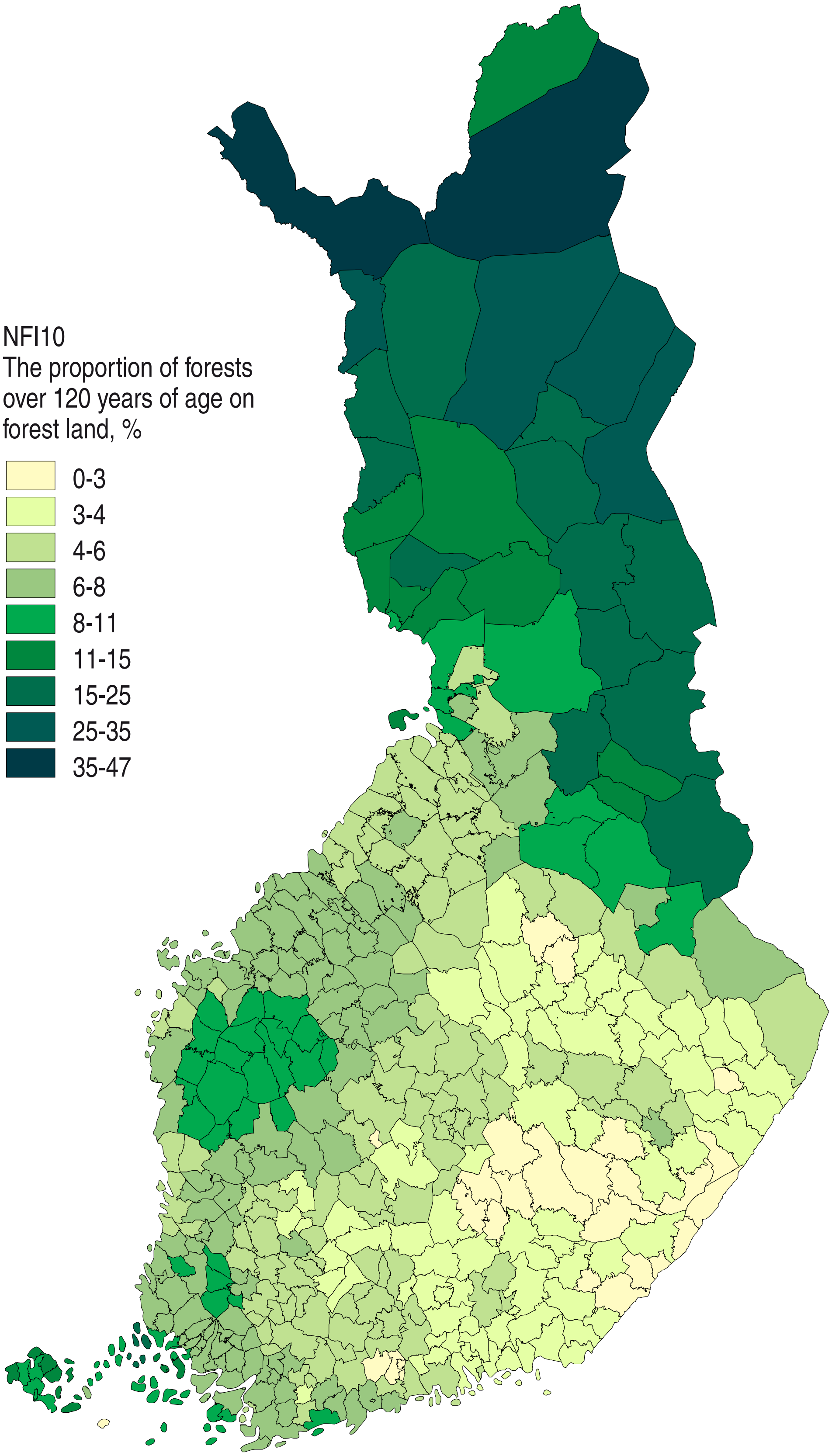
b)

Forestry centre		Mean volume, m³/ha													
		51-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	131-140	141-150	151-160	161-170	171-180	Total
Åland	Number of municipalities					2	1	2		3	7	1			16
	Forest land area in the municipalities, ha					4205	874	3684		13251	39549	6231			67794
Rannikko/ Etelärannikko	Number of municipalities								3	7	5	6	4	1	26
	Forest land area in the municipalities, ha								29736	108133	66425	78024	70933	12412	365663
Rannikko/ Pohjanmaan rannikko	Number of municipalities				1	5	5	4							15
	Forest land area in the municipalities, ha				48853	117482	141712	160530							468577
Lounais-Suomi	Number of municipalities					2		1	6	13	26	17	5		70
	Forest land area in the municipalities, ha					53373		45693	118406	136445	360245	186229	75490		975881
Häme-Uusimaa	Number of municipalities									6	7	17	11	3	44
	Forest land area in the municipalities, ha									78982	180691	369412	241981	65834	936900
Kaakkois-Suomi	Number of municipalities							1	2	8	10	3			24
	Forest land area in the municipalities, ha							2121	68442	242550	394510	77789			785412
Pirkanmaa	Number of municipalities						2	1	2	5	12	4	2		28
	Forest land area in the municipalities, ha						90127	91247	71799	159586	381231	57925	60872		912787
Etelä-Savo	Number of municipalities								5	11	2				18
	Forest land area in the municipalities, ha								394251	728834	83186				1206271
Etelä-Pohjanmaa	Number of municipalities		1	1	7	20	11								40
	Forest land area in the municipalities, ha		52997	27614	254112	569806	369720								1274249
Keski-Suomi	Number of municipalities				2	3	4	3	5	10	1				28
	Forest land area in the municipalities, ha				69858	142004	298459	219551	191866	422608	40684				1385030
Pohjois-Savo	Number of municipalities					4		6	8	4	1				23
	Forest land area in the municipalities, ha					390507		311367	338966	218475	95872				1355187
Pohjois-Karjala	Number of municipalities				1	3	2	6	4						16
	Forest land area in the municipalities, ha				68064	653265	211625	352903	196321						1482178
Kainuu	Number of municipalities			4	5										9
	Forest land area in the municipalities, ha			834574	936077										1770651
Pohjois-Pohjanmaa	Number of municipalities		2	10	19	7									38
	Forest land area in the municipalities, ha		724171	626042	805321	280919									2436453
Lapland	Number of municipalities	10	8	3											21
	Forest land area in the municipalities, ha	2944660	1853586	217624											5015870
	Total, number of municipalities	10	11	18	35	46	25	24	35	67	71	48	22	4	416
	Total, ha	2944660	2630754	1705854	2182285	2211561	1112517	1187096	1409787	2108864	1642393	775610	449276	78246	20438903



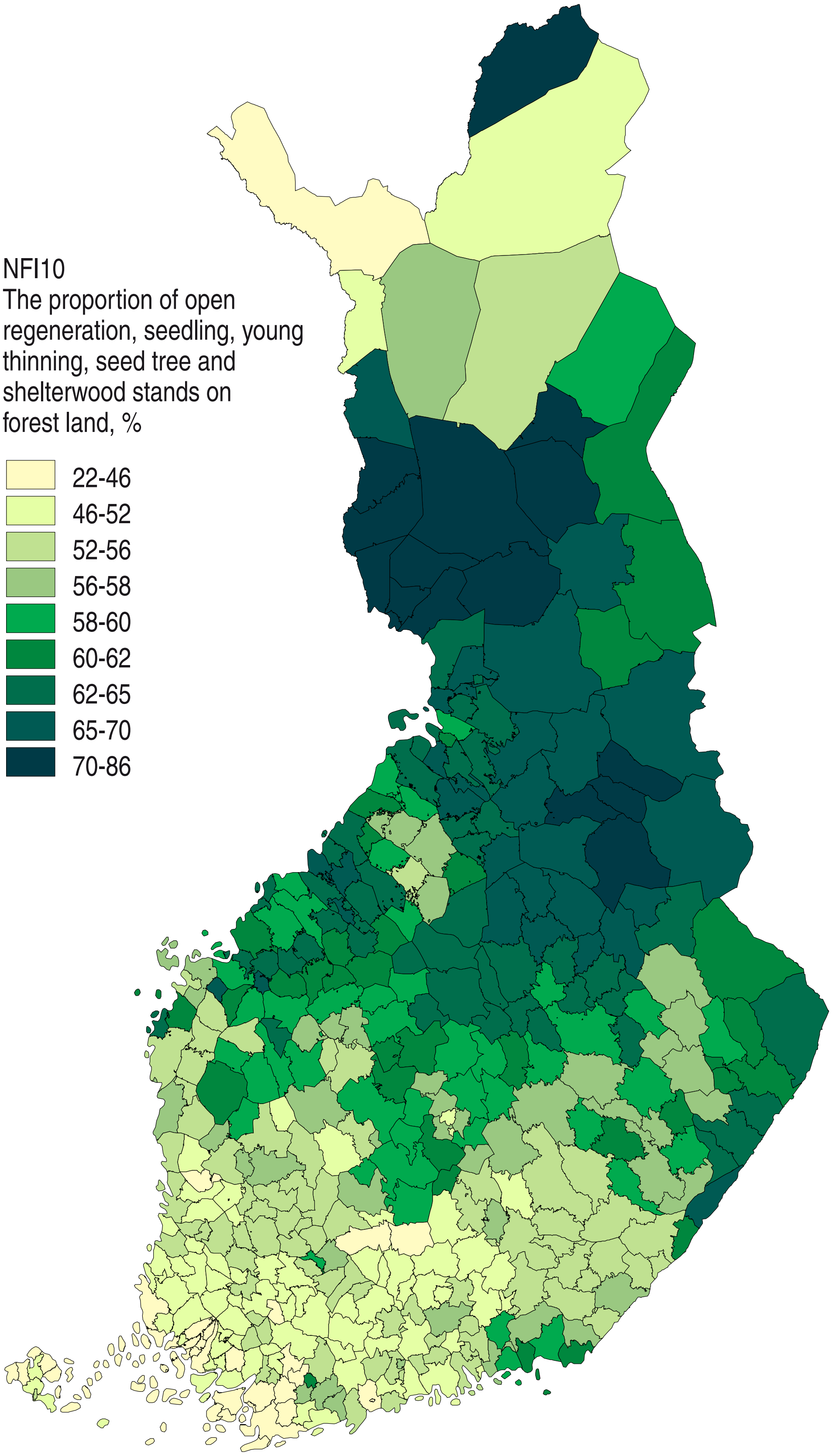
a)

Fig 4.2. The proportion of forests with age of 0–40 (a) and over 120 (b) years on forest land by municipalities. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



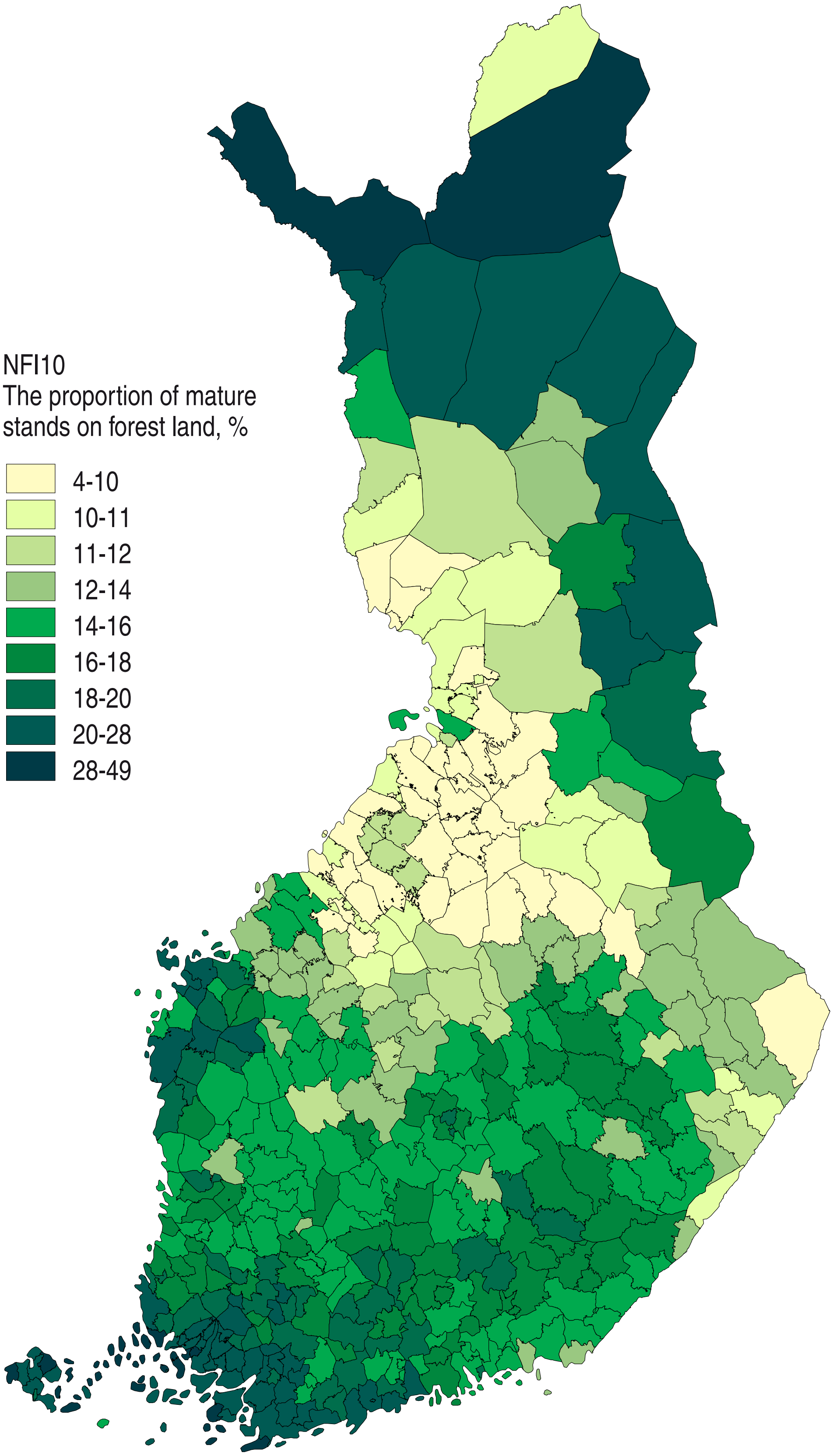
b)

Fig. 4.2. Continued



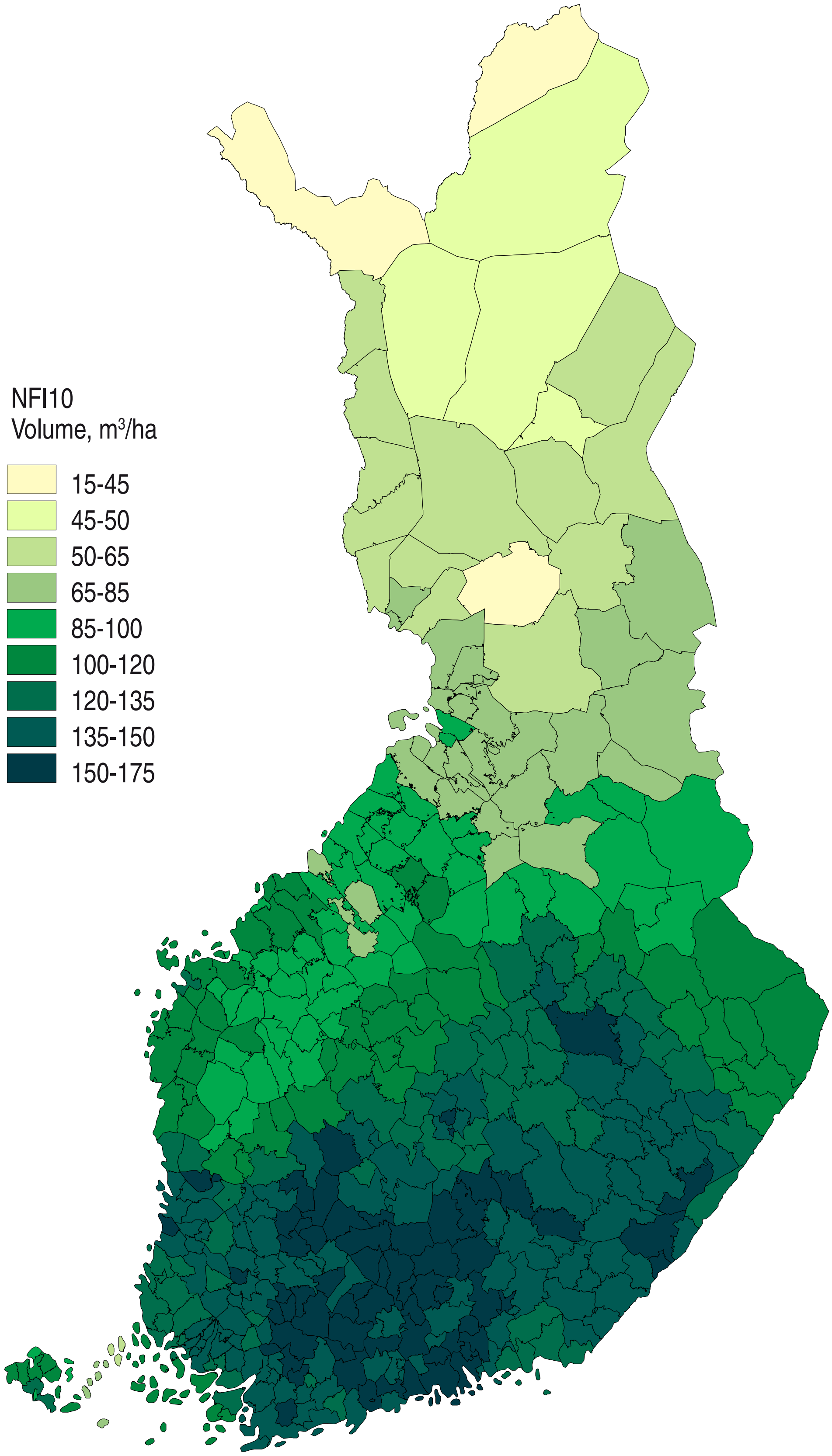
a)

Fig 4.3. The proportion of development classes of open regeneration, seedling, young thinning, seed tree, and shelterwood (a) and mature forests (b) on forest land by municipalities. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



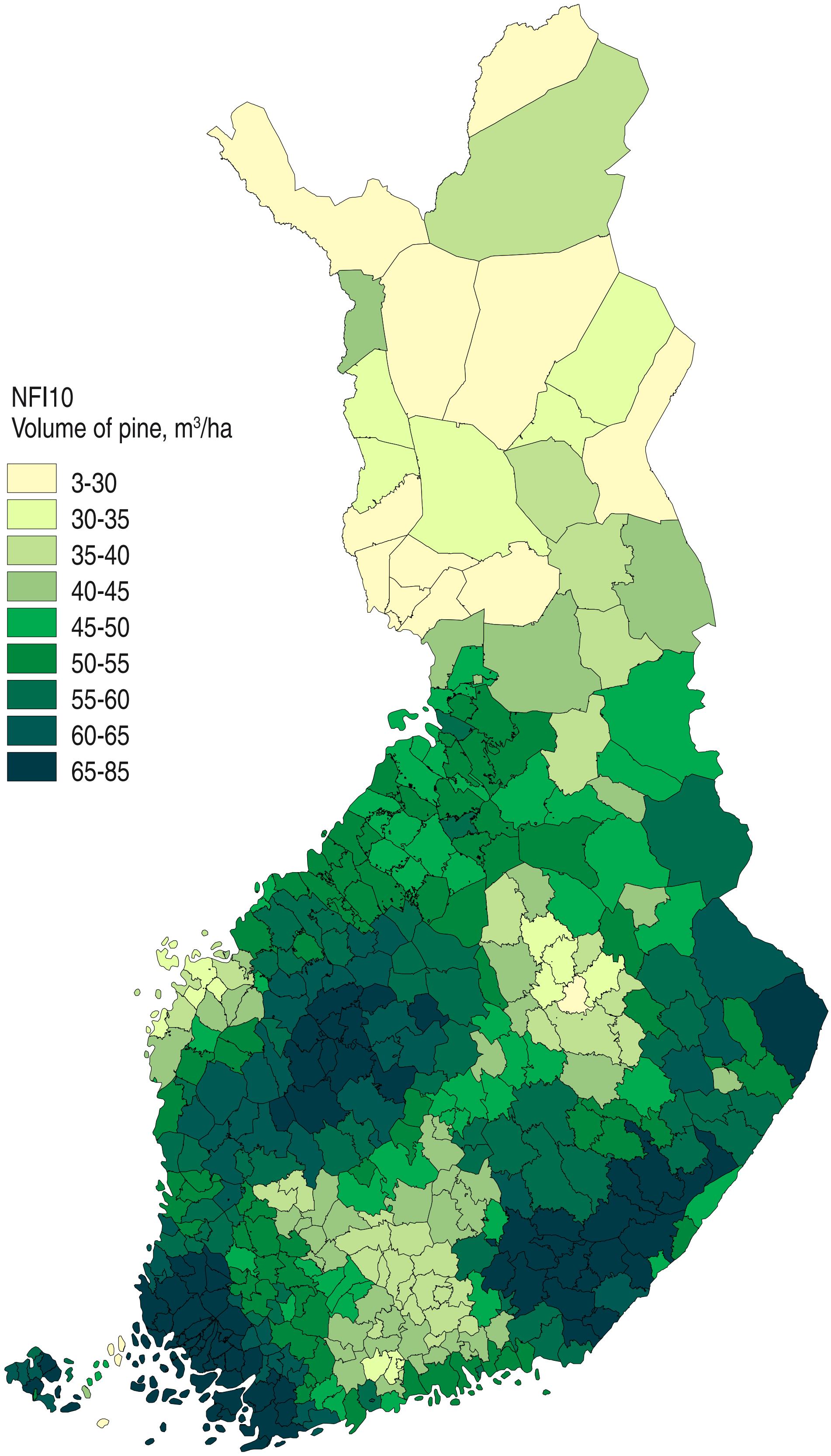
b)

Fig. 4.3. Continued



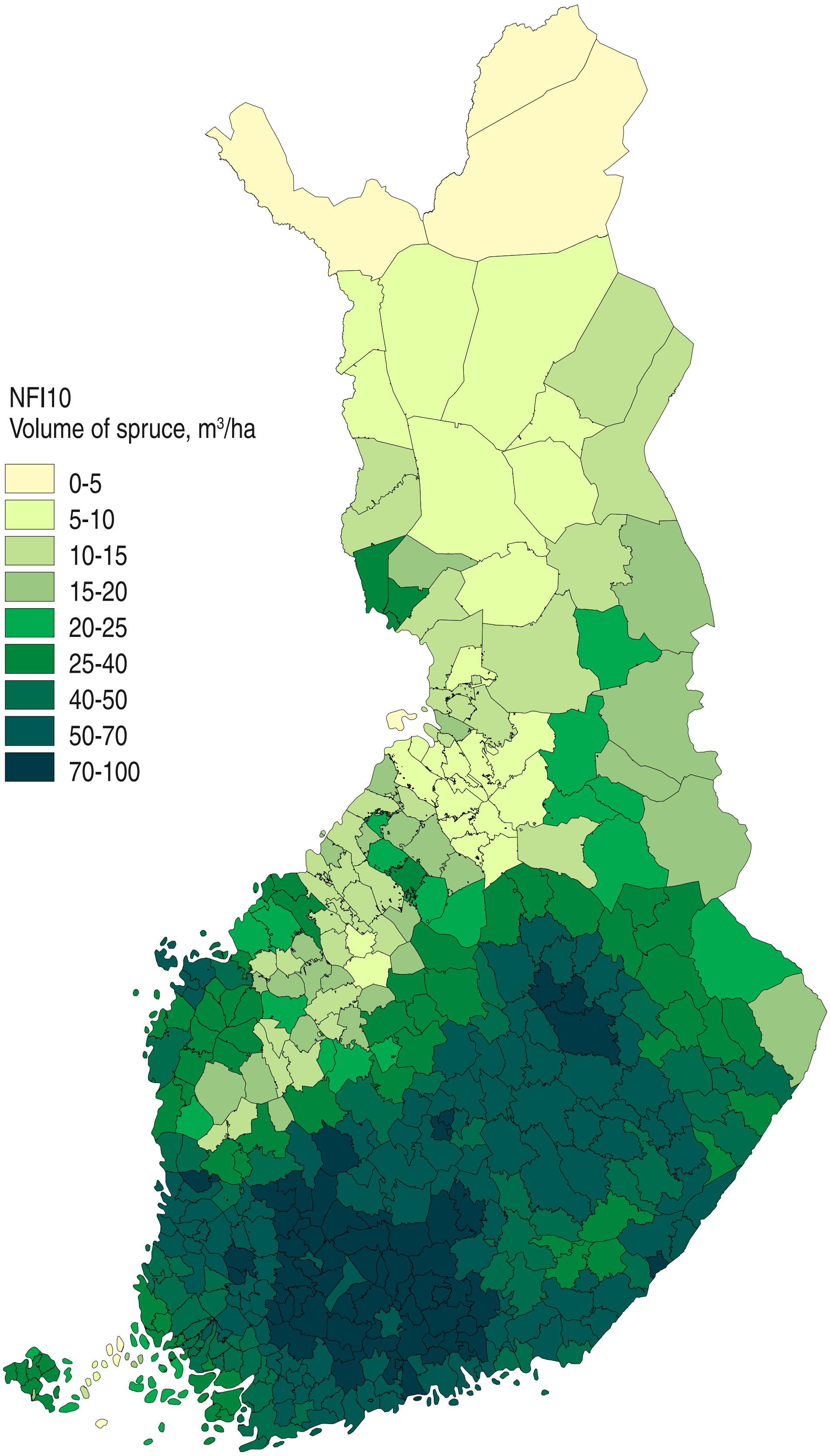
a)

Fig. 4.4. The mean volume of growing stock on forest and poorly productive forest land by municipalities: all tree species (a), pine (b), spruce (c), birch (d), other deciduous species (e). Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



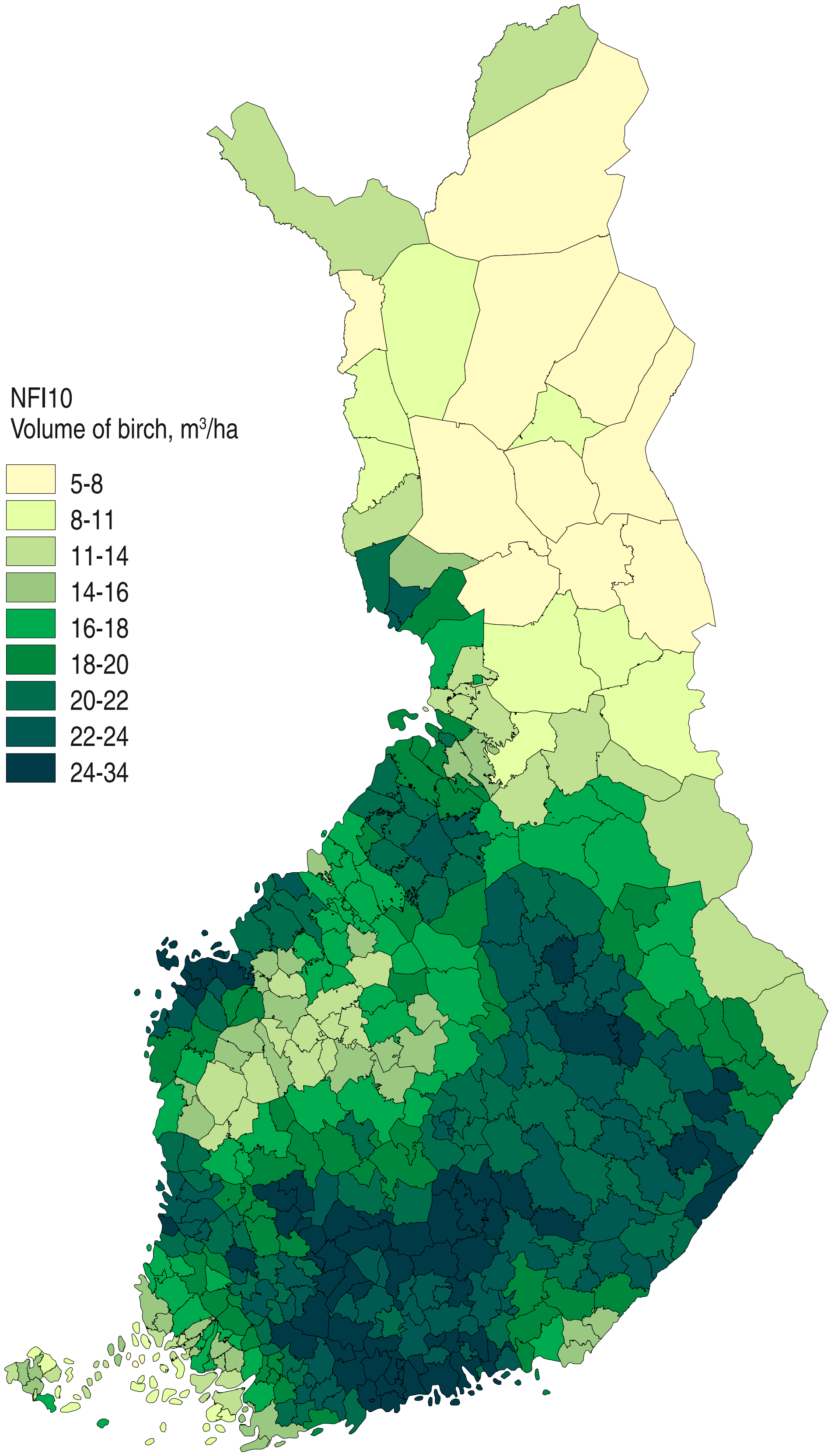
b)

Fig. 4.4. Continued



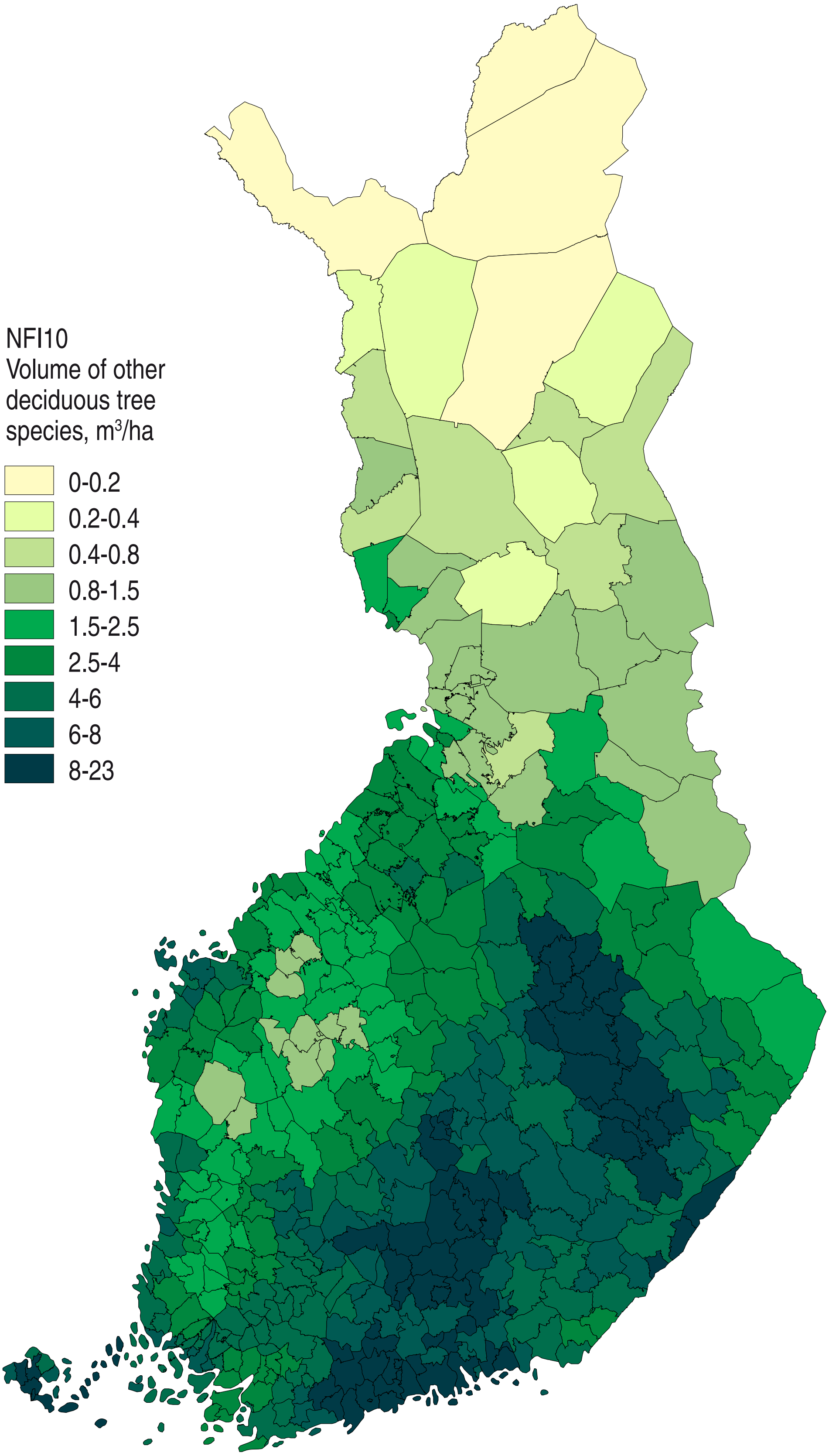
c)

Fig. 4.4. Continued



d)

Fig. 4.4. Continued



e)

Fig. 4.4. Continued

the similar estimates for forest land as Appendix Tables 6a-d for FPPF. Note that poorly productive forest land (PPF) consists either of rocky soils, fjeld forests or less fertile peatland soils, such as oligo-ombrotrophic or ombrotrophic peatland, e.g., *Sphagnum fuscum* dominated peatland. Note that the water balance of peatland soil also affects the wood production capacity and land class of peatland.

The mean volumes of the growing stock in the municipalities vary significantly between forestry centres and also within forestry centres. Table 4.1a summarises the number of municipalities by forestry centre on the basis of mean volume of growing stock on forest land in MS-NFI-2005 and Table 4.1b in MS-NFI9. The areas of forest land in the volume classes are also given. On the average, the mean volume has increased between MS-NFI9 and MS-NFI-2005.

The mean volume estimates on FPPF are given also separately for mineral soils and peatland soils, for the entire FPPF in Appendix Table 2b, and for FPPF available for wood production in Appendix Table 2d.

The mean volume estimates by age classes on forest land are given in Appendix Table 4b and the similar estimates for the forest land available for wood production in Appendix Table 4d. The corresponding mean volume estimates by development classes on forest land are given in Appendix Tables 5b and 5d.

The mean volume of the growing stock on combined forest land and poorly productive forest land by municipalities varies in Southern Finland from about 83 metres cubed per hectare (m^3/ha) in the western part of Central Finland, near the coastal region of Gulf of Bothnia, to about 170 m^3/ha in South Central Finland (170 m^3/ha in Hauho and Lammi and 175 m^3/ha in Pälkäne). The mean volume decreases towards north due to the growing conditions are vary from 15 m^3/ha in Utsjoki and 32 m^3/ha in Enontekiö (the two northernmost municipalities) to slightly over 100 m^3/ha in southern regions in Northern Finland (100 m^3/ha in Haapajärvi and Nivala).

The mean volume of spruce saw log available for wood supply by municipalities on combined forest land and poorly productive forest land is naturally highest in the region in which spruce volume is highest, that is, in southern part of Central Finland, in regions of Häme-Uusimaa, and the southern region of forestry centres Keski-Suomi and Pohjois-Savo, and is low in coastal regions and in some areas in Etelä-Pohjanmaa. The range is from 2 m^3/ha in Halsua and 4 m^3/ha in Kannus and Perho to near 50 m^3/ha in Lammi and Padasjoki.

4.1.1 Available energy wood

Biomass estimates were calculated for each field plot and plot part on FPPF in the ground truth data for energy wood estimation. Biomass estimates for tree compartments were predicted for sample trees on FPPF of NFI10 using tree level biomass models (Marklund 1988) (Table 4.2). Tree level biomass predictions were converted to kilograms per hectare (kg/ha), taking into account angle-gauge sampling (Bitterlich sampling) basal area factor, the maximum radius of the plot and sampling ratio of the sample trees. A separate MS-NFI estimation was carried out for the input images using NFI10 field plots only to obtain new weights for computation units for energy wood estimation purposes.

The biomass estimates by tree species groups in young thinning stands (development class) are presented in Appendix Table 8a with a unit of Gg (10^9 g). The biomasses of stem and bark, branches

Table 4.2. The compartments of tree biomass (Marklund, 1988).

Stem
Bark
Living branches
Foliage
Dead branches
Stump
Large roots, minimum diameter 2 cm
Small roots
Stem residual (from timber assortment class proportions and stem and bark biomass)

Table 4.3. Removal percentages applied for pre-commercial and first commercial thinnings by regions (degree days) and dominant tree species of the NFI field plot stands. Site fertility classes: Herb rich heath (HRHF), Mesic (M) and Sub-xeric (SX) forests.

Degree days	Removal % (site fertility class of the regime)		
	Pine	Spruce	Deciduous
> 1200 dd	34 (SX)	32 (M)	43 (HRHF,M)
1000-1200 dd	32 (SX)	34 (M)	43 (HRHF,M)
< 1000 dd	33 (SX)	31 (M)	28 (HRHF,M peatlands)

and foliage were calculated using those field plots on which first commercial thinning was proposed for the first 5-year-period or on which pre-commercial thinning was proposed and the treatment was already considered to be delayed in the corresponding stand. The proportion of the field plot biomass capable to be removed (28–43%) was estimated applying the regional thinning regimes of pine, spruce and birch on most common site fertility classes (Mesic forests and Sub-xeric forests) according to the dominant tree species of the field plot stand (Hyvän metsänhoidon suositukset 2006) (Table 4.3). Appendix Table 8c presents similar estimates to those in Table 8a for land available for wood supply. The aboveground biomass estimates in young thinning stands, similar to those in Table 8c, were converted to potential energy per area of forest land available for wood supply in the municipality (except in Åland forestry centre per area of all forest land), MWh/ha. These energy estimates are presented by municipalities in Fig. 4.5. 10^6 g (a ton) of biomass equals five mega-watt hours (MWh) in these calculations. The biomass estimates of mature forests are presented separately for branches, foliage and stem residuals, and stumps and large roots by tree species groups in Appendix Table 8b and for land available for wood supply in Table 8d. In practice, only spruce stumps are harvested from regeneration cutting areas. The biomass estimates for tree compartments of spruce in mature stands (Table 8d) were converted to potential energy, MWh/ha, in same way as for young thinning stands above and are presented by municipalities in Fig. 4.6.

The energy wood estimates represent energy wood potential rather than the energy wood available in practice. The practical constraints in use and harvesting of energy wood, like minimum removal and other cutting operations were not taken into account. Note also that the biomass estimates for energy wood were based on sample tree data only. For more details of deriving the biomass estimates and the reliability of the estimates, see Tomppo et al. (2008b).

4.2 Digital thematic output maps

Thematic forest maps in raster format were produced for the most important forest variables: land class, main site class, site fertility class, stand age, mean diameter and height of stands, stand basal area, and volumes by tree species and timber assortments, for four tree species or species groups (pine, spruce, two birch species combined, and other broad leaved tree species). Twelve

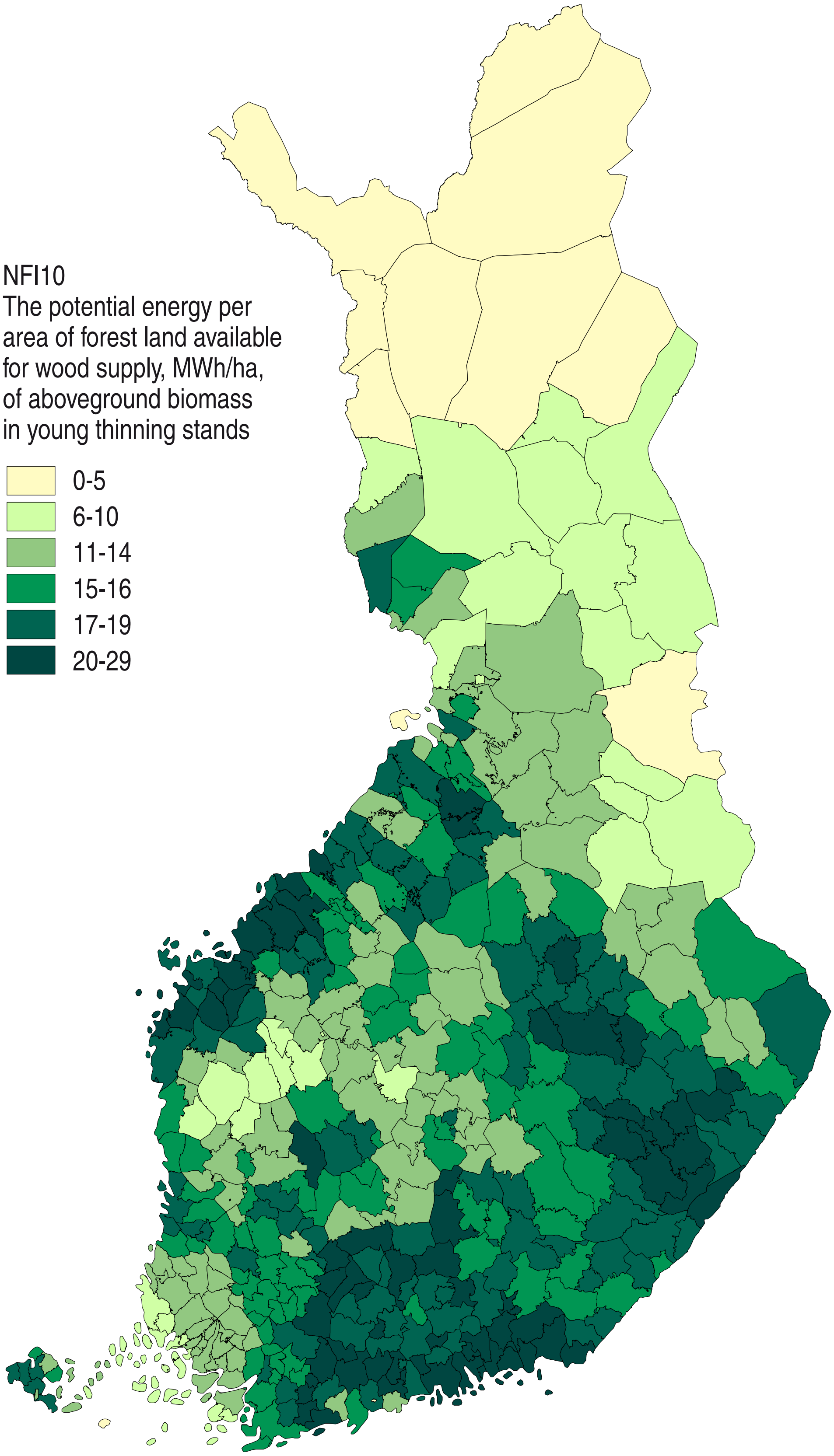
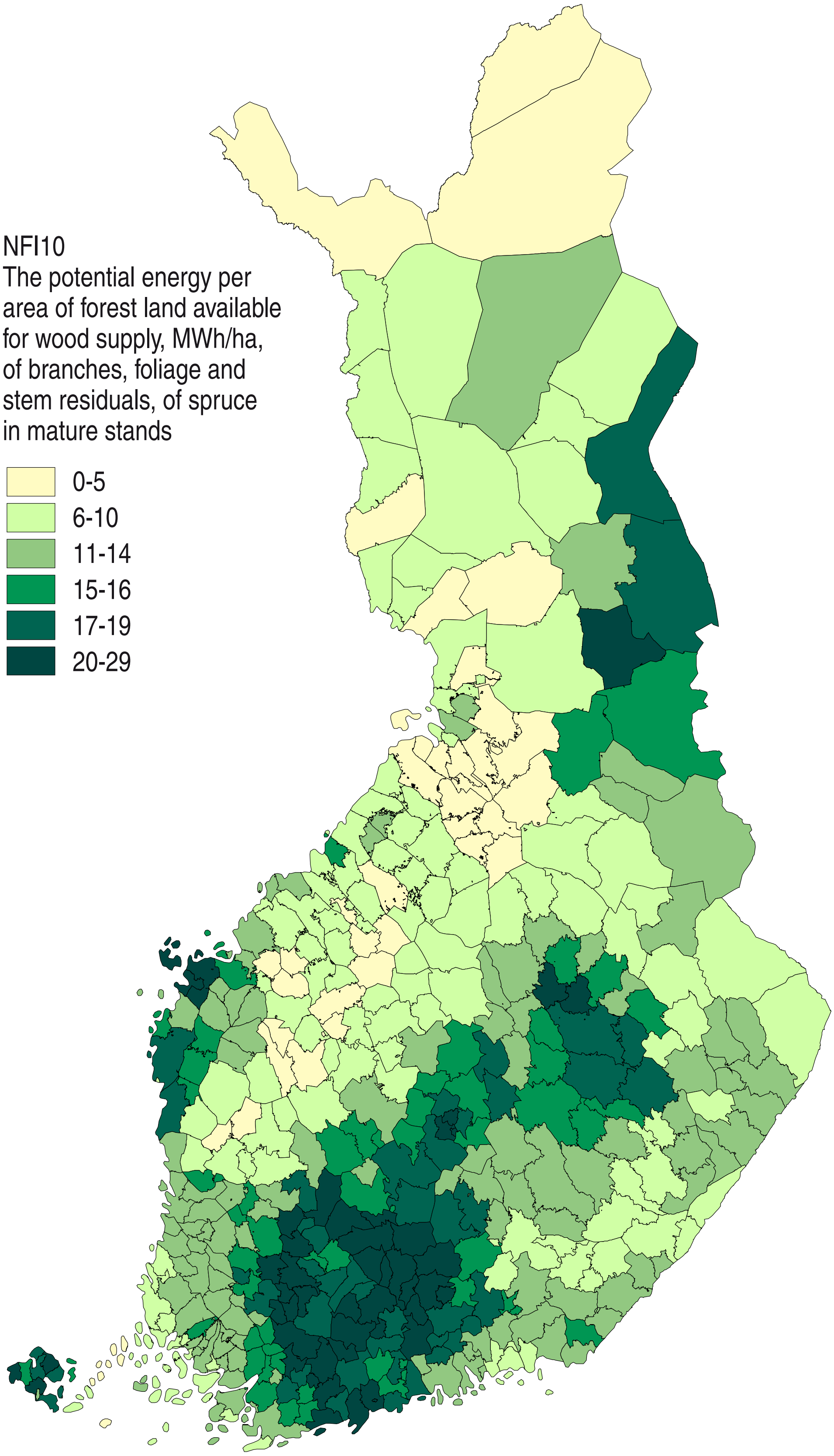
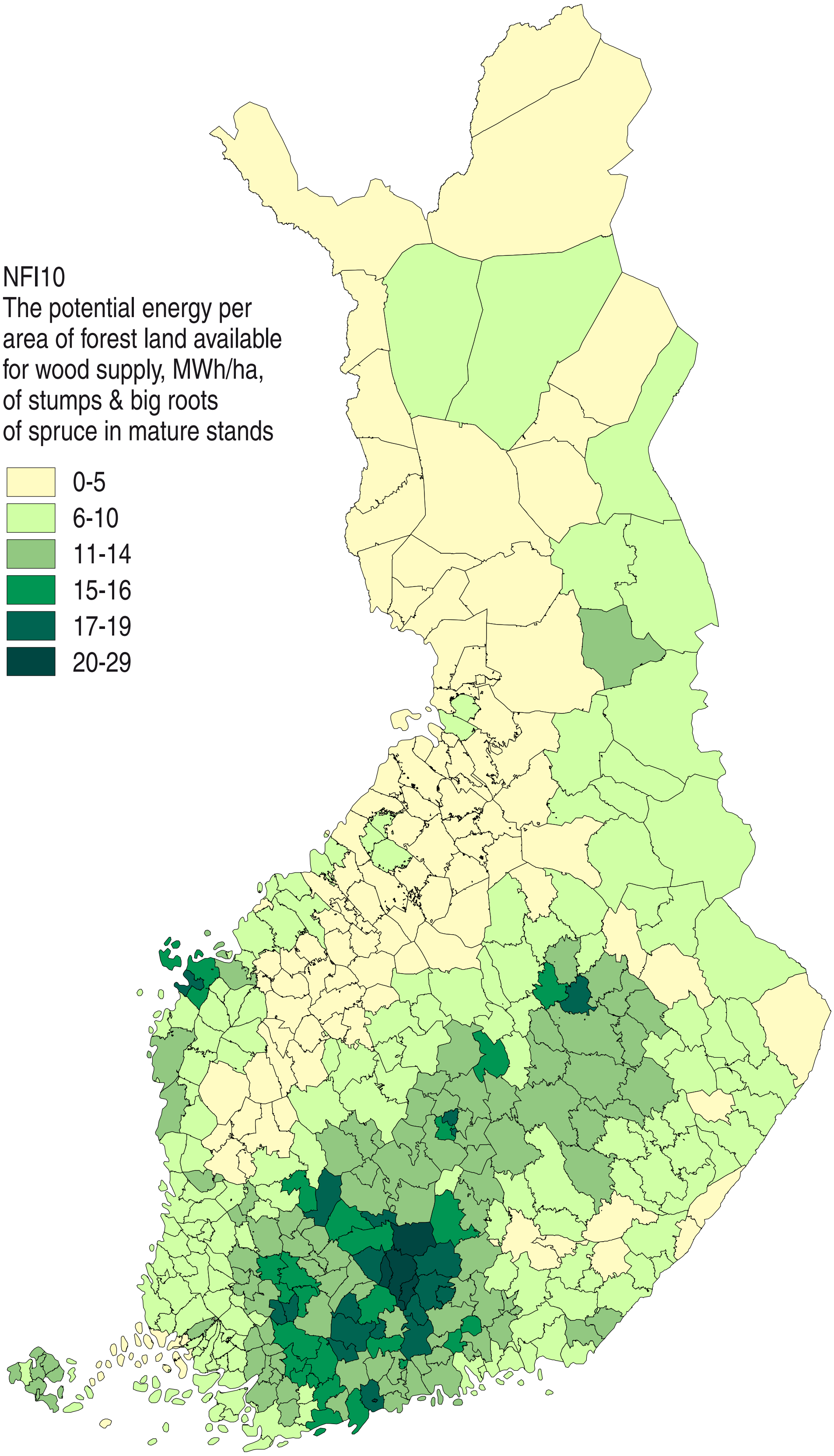


Fig 4.5. The potential energy per area of forest land, MWh/ha, of aboveground biomass in young thinning stands in forests available for wood supply by municipalities. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



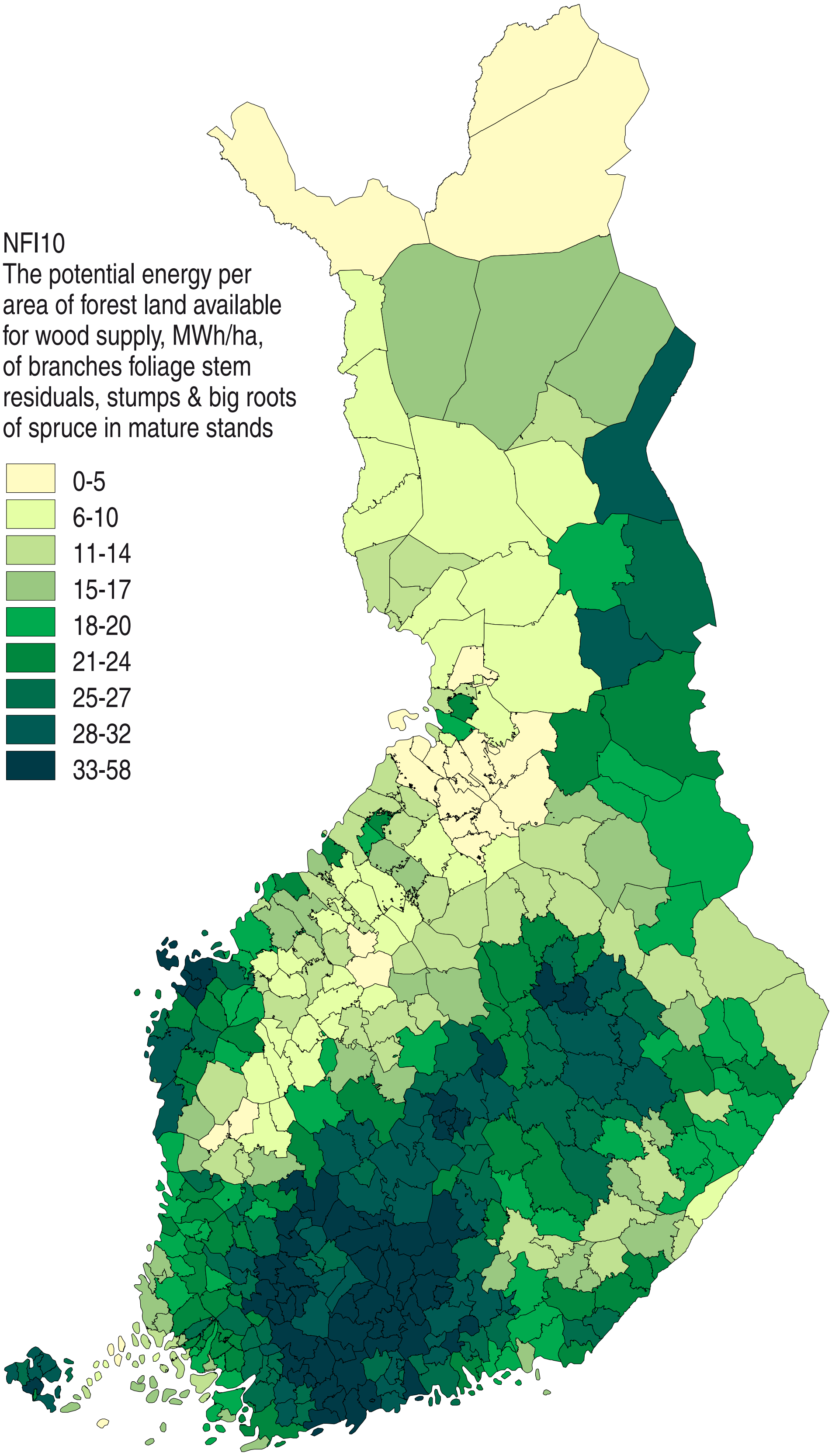
a)

Fig 4.6. The potential energy per area of forest land, MWh/ha, of branches, foliage and stem residuals (a), stumps and large roots (b), alltogether (c) of spruce in mature stands in forests available for wood supply by municipalities. Digital map data: ©National Land Survey of Finland, licence No. 363/MML/08.



b)

Fig. 4.6. Continued



c)

Fig. 4.6. Continued

Table 4.4. Twenty thematic map layers produced by MS-NFI-2005.

Thematic maps, raster format, 25 m by 25 m
20 layers - stand age, a - stand basal area, m2 - total volume, m3/ha - pine volume, m3/ha - pine log volume, m3/ha - pine pulpwood volume, m3/ha - spruce volume, m3/ha - spruce log volume, m3/ha - spruce pulpwood volume, m3/ha - birch volume, m3/ha - birch log volume, m3/ha - birch pulpwood volume, m3/ha - other broadleaved volume, m3/ha - other broadl. log volume, m3/ha - other broadl.pulpwood volume, m3/ha - mean height of stand, dm - mean diameter of stand, cm - land class (3 classes) - main site class (4 classes) - site fertility class (8 classes)

volume maps were produced for volumes of saw timber, pulp wood and total volume. The maps produced are in georeferenced raster layers of 25 m by 25 m resolution (Table 4.4), and cover Finland entirely. The data for forestry centres Lapland and Åland come from NFI9 (cf. Fig. 2.1). The non-forestry land use cover was obtained from the digital land use map data and overlaid on the satellite image data during the estimation phase. The k-NN pixel-level predictions were made for the rest of the area. An example of the total volume thematic map is given in Fig. 3.2c. The raster layers can be combined to produce new thematic maps, e.g. dominant tree species and mean volume classes by tree species dominance. More examples of digital thematic maps are given in Tomppo et al. (2008).

The sum and mean values calculated from raster layers will, in most cases, deviate from the area and volume estimates in the Appendix Tables due to the corrections for map errors. The forestry land area calculated from the maps will be greater, and the mean volume estimates smaller, than those in the Appendix Tables in most cases (cf. Subsect. 3.4.2).

5 Discussion

The development of the Finnish multi-source inventory (MS-NFI) method began in 1989 in the connection of the 8th National Forest Inventory of Finland (NFI8). The method utilises satellite images, field data of the National Forest Inventory (NFI) and digital map data. The methods and results of the first country-level MS-NFI, MS-NFI8, are presented in Tomppo et al. (1998b). The revised methods and results, MS-NFI9 methods and results, corresponding the ninth National Forest Inventory (1996–2003), were published in Tomppo et al. (2008b). The estimates obtained by MS-NFI are part of the official NFI statistics in Finland, some of which are available via the internet (Metinfo 2007).

The main purpose of the MS-NFI method is to obtain forest resource information for areas smaller than would be possible using only field data. The developed *k*-Nearest Neighbour estimation method (k-NN) meets the requirements set to the method and the results. In addition to the small area estimates, the MS-NFI provides predictions of forest variables in map form.

The method has been improved continuously and new features have been added since its first implementation. Similar development work is being carried out in several other countries.

Rapid changes in forests and in forest industries, as well as the approaching change in forest income taxation system in 2006, caused a need to update the forest resource estimates during the 9th inventory to correspond year 2002. Following that case study and the agreed NFI10 plan, the small area forest resource estimates were calculated for the region in which the NFI9 measurements were carried out in 1996–2002, using NFI9 data from years 1996–2002 and NFI10 from years 2004–2005. All the field data were updated to correspond year 2005. This article presents the MS-NFI methods and results employed to update the NFI field plot data to correspond a specific year, in this case year 2005.

The methods to update field data were based on the predicted cuttings on the field plots, carried out after the latest measurements. Satellite imagery from the new time point as well as field plot and stand level data were employed in identifying the plots from which all trees or some trees had been removed, and in predicting the removed trees. Furthermore, growth models were applied to the trees not removed after the last measurement. The developed methods are described in Section 3.1.

The updated variables were the volume of growing stock, age, basal area, mean diameter and mean height. The development class of a forest stand was updated only in the case of regeneration cutting.

When comparing the estimates with those of MS-NFI9 one should note that the updated field plot data includes updating errors. These errors also increase the prediction errors compared to the case in which field data comes from the same season as the satellite image. The pixel level prediction error is also rather high in that case. The several error sources are listed in Tomppo et al. (2008b). For this article, the estimates and pixel level predictions were validated when selecting the estimation parameters comparing MS-NFI estimates and error estimates with those based on NFI10 field data only using groups of municipalities.

The main users of the MS-NFI results, municipality level estimates and maps are the forestry authorities at forestry centres, forest industries and forest environment researchers. More details of the uses are given in Tomppo et al. (2008a and 2008b).

The most serious potential risk in the application of MS-NFI method is the availability of relevant satellite images. An individual satellite image scene should be large enough to cover high enough number of field plots, preferably several thousands, to get satisfactory ground truth data. On the other hand, the spatial resolution, a pixel size, should not be larger than about 30 metres. In addition to the problems caused by clouds, the number of the natural resource satellite with suitable specifications for forest applications is not high, particularly when one relevant instrument, Landsat 7 ETM+, suffers from malfunctioning of the scan line corrector.

This article is one in the series in which the MS-NFI estimates are calculated every second year for the greater part of the country and every fourth year for northernmost Finland. The future method development work will focus, in addition to the decrease all kinds of estimation errors, to investigate the possibilities to integrate airborne remote sensing data, e.g., lidar data, as a part of input data of MS-NFI.

Acknowledgements

The development of multi-source forest inventory was initiated by late professor Pekka Kilkki at the Finnish Forest Research Institute, Metla. The results presented in this article presume the availability of forest inventory field data, amounting about to 100 000 field plots in two inventories, NFI9 and NFI10 from 2004–2005, as well as functioning computing and data management facilities, all provided by Metla. The manuscript has been read and valuable improvements suggested by Director of Research, Professor Pasi Puttonen and Dr. Juha Heikkinen. The language was edited by Ms. Megan McCormick, and typesetting was made by Ms. Sari Elomaa. Ms. Reija Mikkola provided many types of technical assistance. We thank all the individuals whose support has made this article possible.

References

Card DH (1982) Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogramm Eng and Remote Sens 48:431-439

Cochran, WG (1977) Sampling techniques. Wiley, New York, 428 p.

Czaplewski RL & Catts GP (1992) Calibration of remotely sensed proportion or area estimates for misclassification error. Remote Sens of Environ 39:29-43

Diemer C, Lucaschewski I, Spelsberg G, Tomppo E & Pekkarinen A (2000) Integration of terrestrial forest sample plot data, map information and satellite data, An operational multisource-inventory concept. In: Ranchin T, Wald L (ed) Proceedings of the Third Conference “Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images”. Sophia Antipolis, France, January 26-28, 2000, SEE/URISCA, Nice pp. 143-150.

Digital Elevation Model. (2007) National Land Survey of Finland. http://www.maanmittauslaitos.fi/en/default.asp?id=494. Accessed 26 Mar 2007

Finley AO & McRoberts RE (2008) Efficient k-nearest neighbor searches for mult-source forest attribute mapping. Remote Sens of Environ. 112:2203–2211

Finley AO, McRoberts RE & Ek AR (2006) Applying an efficient knearest neighbor search to forest attribute imputation. For Sci 52:130−135

Franco-Lopez H, Ek AR & Bauer ME (2001) Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote Sens Environ 77:251-274

Gjertsen A (2005) Accuracy of forest mapping based on Landsat TM data and a kNN method. In H. Olsson (ed) Proceedings of ForestSat 2005. Borås, Sweden, 31 May – 3 June 2005. pp. 7-11. http://www.skogsstyrelsen.se.

Hagner O & Olsson H (2004) Normalisation of Within-Scene Optical Depth Levels in Multispectral Satellite Imagery Using National Forest Inventory Plot Data. In: Proceedings from the 24th EARSeL Symposium, Workshop on “Remote sensing of land use and land cover”, Dubrovnik, Croatia, May 28-29, 2004

Henttonen H (1991) VMI8:n Pohjois-Suomen otanta-asetelmien vertailu satelliittikuvatulkinnan avulla. Comparing of sampling designs through simulation. Finnish Forest Research Institute, 10 p. (In Finnish)

Hynynen, J., Ojansuu, R., Hökkä, H., Siipilehto, J., Salminen, H. & Haapala, P. (2002) Models for predicting stand development in MELA System. Metsäntutkimuslaitoksen tiedonantoja 835. 116 p.

Hyvän metsänhoidon suositukset (2006) Forestry Development Centre Tapio, Helsinki (In Finnish)

Katila M (2006a) Correcting map errors in forest inventory estimates for small areas. In: Kangas A Maltamo M (ed) Forest Inventory - Methodology and Applications, Managing Forest Ecosystems Vol. 10. Springer, Dodrecht, The Netherlands pp. 225-233

Katila M (2006b) Empirical errors of small area estimates from the multisource National Forest Inventory in Eastern Finland. Silva Fenn 40:729-742

Katila M & Tomppo E (2001) Selecting estimation parameters for the Finnish multi-source national forest inventory. Remote Sens. Environ 76:16-32

Katila M & Tomppo E (2002) Stratification by ancillary data in multisource forest inventories employing k-nearest neighbour estimation. Can J For Res 32(9): 1548-1561.

Katila M, Heikkinen J & Tomppo E (2000) Calibration of small-area estimates for map errors in multisource forest inventory. Can J For Res 30:1329-1339.

Kim, H.-J. & Tomppo, E. 2006. Model-based prediction error uncertainty estimation for k-nn method. Remote Sensing of Environment 104: 257-263.

Koukal T, Suppan F & Schneider W (2005) The impact of radiometric calibration on kNN predictions of forest attributes. In: Olsson H (ed) Proceedings of ForestSat 2005. Borås, Sweden, 31 May – 3 June 2005. pp. 17-21. http://www.skogsstyrelsen.se.

Lehto J & Leikola M (1987) Käytännön metsätyypit. Kirjayhtymä, Helsinki (In Finnish), 96 p.

Maanmittauslaitoksen maastotietokohteet (Feature catalog of the topographic database) (2005) National Land Survey of Finland, Helsinki, Maanmittauslaitoksen julkaisuja 97. http://www.maanmittauslaitos.fi/default.asp?id=887. (In Finnish)

Maastotietojen keruuohje (Manual of field work) (1996) Sisäinen ohje, Manuscript, National Land Survey of Finland (In Finnish)

Marklund LG (1988) Biomass functions for pine, spruce and birch in Sweden. Sveriges Lantbruksuniversitet, Uppsala, Sweden, Inst. Skogstaxering. Rapp. 45

Magnussen, S., McRoberts, R. & Tomppo, E.O. 2009. Model-based mean square error estimators for k-nearest neighbour predictions and applications using remotely sensed data for forest inventories. Remote Sensing of Environment 113: 476-488.

Maselli F, Chirici G, Bottai L, Corona P & Marchetti M (2005) Estimation of Mediterranean forest attributes by the application of k-NN procedures to multitemporal Landsat ETM+ images. International J of Remote Sens 26:3781-3796

McRoberts RE (2006) A model-based approach to estimating forest area. Remote Sens of Environmen 103:56-66

McRoberts RE & Tomppo EO, 2007. Remote sensing support for national forest inventories. Remote Sens of Environ 110: 412-419

McRoberts RE, Nelson MD & Wendt DG (2002a) Stratified estimation of forest area using satellite imagery, inventory data, and the k-Nearest Neighbors technique. Remote Sens of Environ 82:457-468

McRoberts RE, Wendt DG, Nelson MD & Hansen MH (2002b) Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. Remote Sens of Environmen 81:36-44

McRoberts RE, Holden GR, Nelson MD, Liknes GC & Gormanson DD (2006) Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the USDA Forest Service. Can J of For Res 36:2968-2980

McRoberts RE, Tomppo EO, Finley AO & Heikkinen J 2007. Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. Remote Sens of Environ 111: 466-480

Metinfo (2007) Metinfo -forest information services, Finnish Forest Research Institute. http://www.metla.fi/metinfo/index-en.htm. Accessed at 9 Mar 2007

Mäkisara K, Heikkinen J, Heiler I, Henttonen H, Katila M, Pitkänen J & Tomppo E (2001) Updating of the Finnish national forest inventory using remote sensing data (Valtakunnan metsien inventoinnin (VMI) ajantasaistus kaukokartoitusaineiston avulla), final report, Finnish Forest Research Institute.

Nilsson M (1997) Estimation of Forest Variables Using Satellite Image Data and Airborne Lidar. Ph.D. thesis, Swedish University of Agricultural Sciences, The Department of Forest Resorce Management and Geomatics, Acta Universitatis Agriculturae Sueciae. Silvestria 17, 84 p.

Nuutinen, T. & Hirvelä, H. 2006. Hakkuumahdollisuudet Suomessa valtakunnan metsien 10. inventoinnin perusteella. Metsätieteen aikakauskirja 1B/2006: 223-237.

Rao J N K (2003) Small Area Estimation. Wiley, New York, 313 p.

Reese H, Nilsson M, Sandström P & Olsson H (2002) Applications using estimates of forest parameters derived from satellite and forest inventory data. Comput and Electron in Agric 37:37-56

Reese H, Nilsson M, Granqvist Pahlén T, Hagner O, Joyce S, Tingelöf U, Egberth M & Olsson H (2003) Countrywide estimates of forest variables using satellite data and field data from the National Forest Inventory. Ambio 32:542-548

Suomen pinta-ala kunnittain (The area of Finland by municipalities) 1.1.2007 (2007) National Land Survey of Finland, Helsinki. http://www.maanmittauslaitos.fi/default.asp?id=894. (In Finnish)

Tomppo E (1990) Satellite Image-Based National Forest Inventory of Finland. The Photogramm J of Finland 12:115-120

Tomppo E (1991) Satellite Image-Based National Forest Inventory of Finland. In: Proceedings of the symposium on Global and Environmental Monitoring, Techniques and Impacts, September 17-21, 1990 Victoria, British Columbia Canada. International Archives of Photogrammetry and Remote Sens 28:419-424

Tomppo E (1992) Satellite image aided forest site fertility estimation for forest income taxation. Acta For Fenn 229, 70 p.

Tomppo E (1996) Multi-source National Forest Inventory of Finland. In: Päivinen R, Vanclay J, Miina S (ed) New Thrusts in Forest Inventory, Proceedings of the subject group S4.02-00 'Forest Resource Inventory and Monitoring' and subject group S4.12-00 'Remote Sensing Technology', vol. 1. IUFRO XX World Congress, 6-12 Aug. 1995, Tampere, Finland. European Forest Institute, Joensuu, Finland pp. 27-41

Tomppo E (2006a) The Finnish National Forest Inventory. In: Kangas A., Maltamo M (ed) Forest inventory. Methodology and applications, Managing Forest Ecosystems Vol 10. Springer, Dordrecht, The Netherlands pp. 179-194

Tomppo E (2006b) The finnish multi-source national forest inventory - small area estimation and map production. In: Kangas A., Maltamo M (ed) Forest inventory. Methodology and applications, Managing Forest Ecosystems Vol 10. Springer, Dordrecht, The Netherlands pp. 195-224

Tomppo E (2008) Tomppo, E. 2008. The Finnish National Forest Inventory. In: McRoberts, R.E., Reams, G.A., Van Deusen, P.C., and McWilliams, W.H. (eds.). Proceedings of the eighth annual forest inventory and analysis symposium. 2006 October 16-19, Monterey, CA. General Technical Report WO. Washington, DC: U.S. Department of Agriculture Forest Service.

Tomppo E & Halme M (2004) Using coarse scale forest variables as ancillary information and weighting of variables in k-nn estimation: a genetic algorithm approach. Remote Sens. Environ. 92:1-20

Tomppo E, Katila M, Moilanen J, Mäkelä H & Peräsaari J (1998b) Kunnittaiset metsävaratiedot 1990-94. Folia For. 4B/1998:619-839 (In Finnish)

Tomppo E, Goulding C & Katila M (1999) Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scand J of For Res 14:182-192

Tomppo E, Korhonen KT, Heikkinen J & Yli-Kojola H (2001a) Multisource inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. Silva Fenn 35(3):309-328

Tomppo E, Olsson H, Ståhl G, Nilsson M, Hagner O & Katila M (2008a) Combining National Forest Inventory Field Plots and Remote Sensing Data for Forest Databases. Remote Sens of Environ 10.1016/j.rse.2007.03.032

Tomppo, E., Haakana, M., Katila, M. & Peräsaari J. (2008b) Multi-Source National Forest Inventory Methods and Applications Springer. Series: Managing Forest Ecosystems, Vol. 18 2008, XIV, 374 p. 22 illus. in color., Hardcover. ISBN: 978-1-4020-8712-7.

Tomppo E, Heikkinen J, Ihalainen A, Mäkelä H, Tonteri T & Tuomainen T. (2009). The 9th Forest Inventory of Finland - methods and results. Manuscript, Finnish Forest Research Institute

Walsh TA & Burk TE (1993) Calibration of satellite classifications of land area. Remote Sens of Environ 46:281-290