Multi-Source Inventory of the Forests of the Hebei Forestry Bureau, Heilongjiang, China

Erkki Tomppo, Kari T. Korhonen, Juha Heikkinen and Hannu Yli-Kojola

Tomppo, E., Korhonen, K.T., Heikkinen, J. & Yli-Kojola, H. 2001. Multi-source inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. Silva Fennica 35(3): 309–328.

A multi-source forest inventory method is applied to the estimation of forest resources in the area of the Hebei Forest Bureau in Heilongjiang province in North-East China. A stratified systematic cluster sampling design was utilised in field measurements. The design was constructed on the basis of information from earlier stand-level inventories, aerial orthophotographs, experiences from other sampling inventories and the available budget. Sample tree volumes were estimated by means of existing models. New models were constructed and their parameters estimated for tallied tree volumes and volume increments. The estimates for the area of the Bureau were computed from field measurements, and for the areas of the forest farms estimated from field measurements and satellite images. A k-nearest neighbour method was utilised. This method employing satellite image data makes it possible to estimate all variables, particularly for smaller areas than that possible using field measurements only. The methods presented, or their modifications, could also be applied to the planning and realisation of forest inventories elsewhere in Temperate or Boreal zones. The inventory in question gave an estimate of 114 m^3/ha (the multi-source inventory 119 m^3/ha) instead of 72 m^3/ha as previously estimated from available information. Totally nineteen tree species, genera of species or tree species groups were identified (Appendix 1). The forests were relatively young, 60% of them younger than 40 years and 85% younger than 60 years.

Keywords Forest inventory, satellite images, k-nearest neighbour method, models, China

Authors' address Finnish Forest Research Institute, Unioninkatu 40 A, FIN-00170 Helsinki, Finland.
Fax +358 9 857 05717 **E-mail** erkki.tomppo@metla.fi

Received 31 October 2000 **Accepted** 5 March 2001

1 Introduction

The Hebei Forest Bureau is one of the forestry enterprises owned by the Chinese government. It is located in North-Eastern China, in the province of Heilongjiang, which is the most important wood production province of the country. The development of forestry and the forest industries of the Hebei Forest Bureau is one of key national projects. The total land area managed by the Hebei Forest Bureau is 382 401 hectares. Forests are located between longitudes 129.39E and 130.47E and latitudes 47.32N and 48.24N (east of the Xiao Xingan Ling mountains). The terrain is hilly but not mountainous. The climate is continental with very cold winters (minimum temperature lower than –30˚C) and warm summers. The average volume of the growing stock on forest land, according to previously available information, was 72 m^3/ha.

The area managed by the Bureau is divided into 19 forest farms, which are the basic production units. At present, the forest resources are mainly utilised by one sawmill located in the town of Hebei in the southernmost part of the area. Plans have been made to establish a papermill.

This study aimed to develop and test a multi-source inventory system for application in the forests of Heilongjiang. The topics addressed in this study are: planning of the sampling design, collection of data, and the computation of results. The estimates of areas, volumes and increment by computation unit (forest farms) and by reference unit (such as forest type, tree species, size class, etc.) are required for estimating the future cutting possibilities. However, the establishment of a forest management plan was not included in the scope of this study.

2 Data Collection

2.1 Selection of Sampling Design

The goal was to develop a sampling design to support a multi-source forest inventory. The forest statistics for the whole area should be computable using field plot measurements only. The field data should also serve as training data in satellite image analyses when computing the estimates by forest farms. Other limitations were:

1) the relative standard error of the estimates of the main parameters for the whole area must be $\leq 5\%$ and
2) the number of the field plots ≤ 1500,

the latter due to the available budget.

Systematic sampling has proved to be efficient in forest inventories and is always more efficient than random sampling, if the variables of interest have positive spatial autocorrelation (Matérn 1960). For practical reasons, cluster sampling must be applied, usually in such a way that measuring the field plots of one cluster is one day's work for a field crew. Sampling designs have been extensively studied in the context of national forest inventories (see e.g. Päivinen 1987, Ranneby et al. 1987, Korhonen and Maltamo 1991, Schreuder et al. 1993, Study ... 1997, Tomppo et al. 1997, Tomppo et al. 1998).

The advantage of cluster sampling is that the required time per field plot is minimised. A sparse road network also supports a cluster size of one working day, rather than, e.g., half a working day. The disadvantage of cluster sampling, compared to a completely regular layout of the plots, is a possible spatial correlation of variables between the neighbouring plots. This disadvantage can be reduced by increasing the within cluster plot distance. This also increases the time required, so a compromise is usually necessary (Ranneby et al. 1987, Tomppo et al. 1998). The sampling design of the present Chinese National Resource inventory is a pure systematic sampling without clusters. The travelling cost per field plot has proved to be very high.

Farm-level forest statistics and aerial orthophotographs were available when planning the sampling design. The statistics were based on an earlier management inventory. On the basis of the mean volumes, tree species composition, age structure and the variables to be measured, eleven field plots per cluster were estimated to be acceptable. It was not possible to estimate the covariance function of the target variables due to the lack of preliminary information. An L-shaped cluster (Fig. 2a, b) was selected in order to reduce the autocorrelation of the forest variables within

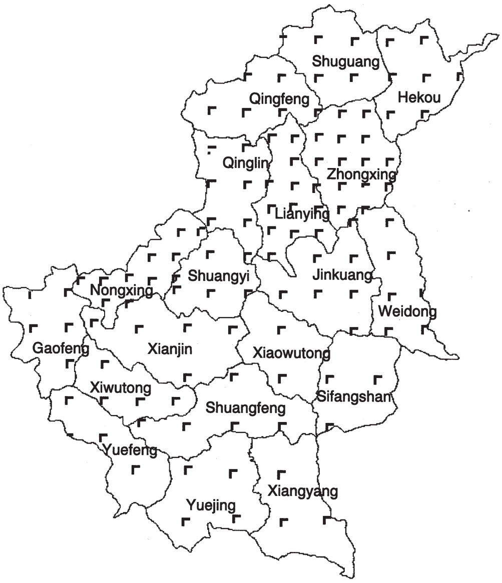

Fig. 1. Boundaries of the Hebei Forest Bureau and Forest Farms with locations of field plot clusters.

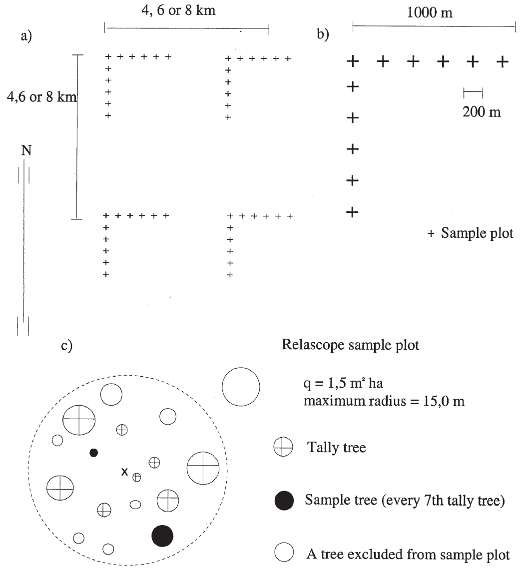

Fig. 2. Sampling design and the shape of field plot in the Hebei Forest Bureau. a) Location of clusters; b) Location of sample plots on a cluster; c) Restricted relascope sample plot.

clusters. The distance between consecutive plots in a cluster was set at 200 m. This was based on the assumption that the range of the spatial autocorrelation is not essentially larger than in Finnish and Swedish forests.

In order to arrive at a more efficient layout of the field plots, the 19 farms of the Bureau were classified into three strata according to the mean volume of the growing stock, as estimated in the earlier stand level management inventory. Different sampling intensities were then applied in each stratum. The distances between clusters of sample plots were 8 km by 8 km in stratum I (lowest mean volume); 6 km by 6 km in stratum II; and 4 km by 4 km in stratum III (highest mean volume) (Fig. 1). In this way, the above limitations could be met, the standard error of the mean volume being estimated on the basis of other inventories.

2.2 Field Data

The sample plots were located in the field with the help of forest maps and aerial orthophotographs provided by the Chinese Ministry of Forestry. They were measured as angle gauge (Bitterlich relascope) plots with a basal area factor of 1.5 and a maximum radius of 15 m (Fig. 2c). The factor 1.5 was based on the existing information concerning the mean volume. The derived mean volume estimate showed that a larger factor, e.g. 2, might have been small enough. The use of a maximum radius decreases the measuring error caused by unobserved trees belonging to the plot and reduces the amount of field work, but does not essentially increase the standard error of the estimates (Lappi 1991). Tree species and diameter at breast height were registered for each tallied tree (called tally trees). Every seventh tally tree (counted by the field crew and from the beginning of the inventory work) was measured as a sample tree. Height, bark thickness and the diameter increment of the previous five-year period were measured for the sample trees. Age and diameter increments were measured in a laboratory from increment cores taken by the field crews. One core in a random direction was taken for each sample tree as recommended by Loetsch, Zöhrer and Haller (1973).

Selected forest characteristics, such as soil type, site fertility, silvicultural measures and mean characteristics of the growing stock were recorded for the stands intersecting a sample plot. Silvicultural and harvesting measures were proposed for the coming ten-year period. A detailed description of the measurements is given in the Field Instructions (Korhonen and Yli-Kojola 1993).

A total of 1235 plots on 115 clusters were measured in autumn 1993. Due to the systematic layout, some of the clusters lay partly outside the area of the Bureau. The plots on forestry land numbered 1205. Altogether, 12 936 trees were measured, of which 1828 were sample trees.

2.3 Multi-Source Data

Estimates based on the field data were only sufficiently reliable for the area of the Forestry Bureau as a whole. For smaller areas, such as individual forest farms, the standard errors of the estimates were large, due to the sparse sampling design. Multi-source data, satellite images, and digital boundary data of the farms were applied together with the field data to decrease the farm-level standard errors. New area weights for the field sample were estimated, which made it possible to compute estimates for an arbitrary variable and for an arbitrary area. The applied method also produced thematic maps, from an arbitrary inventory variable, see Section 3.2.

2.3.1 Satellite Images

Two Landsat 5 TM images, 118-31 and 118-30, from September 9, 1993 were employed for computing the results for the forest farms. The images were received at Beijing receiving station. The images were of good quality. No atmospheric correction was needed because the images were from a same date and successive rows. Images were geometrically rectified for the applied coordinate system. This was done by modelling the relationship between the image coordinates, (*u*,*v*), and the map coordinates, (*x*,*y*), by linear regression

$$\begin{aligned} u &= a_l + b_l x + c_l y + \varepsilon_l \\ v &= a_c + b_c x + c_c y + \varepsilon_c \end{aligned} \qquad (1)$$

where *a, b* and *c* are parameters to be estimated from the data concerning objects which are clearly identifiable on both the image and the map, and where the ε's denote the model residuals. The intensity value for each pixel of the rectified image was determined by the nearest neighbour method. Resampling to a 25 m × 25 m pixel size was applied. The average absolute residuals were rather high, 50 m in east-west and 60 m in south-north directions. This indicated relatively large errors in the base map.

2.3.2 Administrative Boundaries and Land Use Data

The boundaries of the area of Hebei Forestry Bureau, as well as the boundaries of forest farms, were digitised from a 1:100 000 map. The map was provided by the Chinese Ministry of Forestry. Large agricultural areas found on the map were also digitised with a digitizer board.

The satellite images were geometrically rectified on the digitised map. Roads, gold-mining areas and parts of built areas were classified with the spectral properties of satellite images of those land use classes. The pixel-wise class probabilities, i.e., the probabilities of an arbitrary pixel *p* belonging to group *k,* were computed using discriminant functions based on the generalised quadratic distances:

$$\begin{aligned} &D_k^2(y_p) = \\ &(y_p - m_k)' S_k^{-1}(y_p - m_k) + \ln\left|S_k\right| - 2\ln(q_k) \end{aligned} \qquad (2)$$

where

y_p is the vector of intensity values at pixel *p*,

m_k is an estimate of the vector of the expected intensities in group *k*,

S_k is an estimate of the covariance matrix of the intensities in group *k*, and

q_k is the prior probability of group *k*.

The value of q_k was same for each class and was the inverse of the number of classes.

The posterior probability of observation y_p

belonging to the class *k* is, under the multivariate normal assumption, obtained from (2) by:

$$p(k \mid y_p) = \frac{e^{-0.5D_k^2(y_p)}}{\sum_{u=1}^{n} e^{-0.5D_u^2(y_p)}} \tag{3}$$

where *n* is the number of classes.

The class with the highest posterior probability was attached to the pixel *p*.

2.3.3 Water Areas

Water areas were separated from the forestry land using information from the satellite images. Thresholding of channels 4 and 7 of the Landsat TM images was employed. That is, pixel *p* was classified as water *W* by the rule:

$$p \in W \quad \text{if } f_4(p) \le k_4 \quad \text{and } f_7(p) \le k_7 \tag{4}$$

The threshold parameters k_4 and k_7 were chosen by inspecting the histograms of channels 4 and 7 and the colour composition of the satellite images displayed on a computer screen.

2.3.4 Ground Truth Data File

The multi-source estimation method needs a ground truth file consisting of field data and image data. The field data record (vector) of a field plot was utilised as such, except that tree level volumes were changed into volumes per hectare by tree species, see Section 3. It was assumed that these estimates could be used as volume estimates for the pixel corresponding to the field plot, i.e., the pixel containing the centre point of the field plot.

The geographical locations of the field plots were checked visually by displaying the locations of the plots on a colour composition of the image and utilising the forest and terrain information recorded in the field. The whole field plot cluster was shifted (with respect to the satellite image) if there was clear evidence that the cluster was incorrectly located. The coordinates of all field plots in a cluster were changed by a same amount. The purpose of this exercise was to reduce the effect of the dislocation of field plots, especially the dislocation of the starting point for line measurement. The reflectance values of the pixel containing the centre point of the plot were added to the end of the field data record of the plot.

3 Data Processing and Estimation Methods

3.1 Processing of Field Data

The inventory results for the whole Bureau area were calculated using the field data only. The field data, the digitised map, and the satellite images were employed when calculating the results for individual forest farms.

The estimation procedure based on the field data only consisted of the following steps.

1. Checking and input of data.
2. Estimation of volumes and volume increments for sample trees.
3. Estimation of volumes and volume increments for tally trees.
4. Summing area, volume and growth estimates over the sampling strata by reference unit.

3.1.1 Estimation of Volumes for Sample Trees

The volumes of the sample trees were estimated using volume functions provided by the General Bureau of Forest Industry of Heilongjiang Province. Diameter at breast height and tree height were the explanatory variables in these functions. The form of the volume functions is

$$v = a_0 d^{a_1} h^{a_2} + \varepsilon \tag{5}$$

where
v = volume of a sample tree,
d = diameter at breast height,
h = height,
a_j, $j=0,1,2$, are the (given) parameters of the model, and
$\varepsilon = N(0,\sigma^2)$ – distributed error term.

The parameter values by tree species are listed in Appendix 2. The reliability of the models was not known.

3.1.2 Estimation of Volume Increments for Sample Trees

The volume increment of the previous 5-year period was estimated for each sample tree as the difference of the present volume and the volume 5 years ago (Svensson 1988). Volume functions by tree species using the diameter as the only explanatory variable were necessary when estimating the volumes at the beginning of the previous 5-year period. These functions were estimated using the measured sample tree data. The quadratic model

$$\frac{v}{d^2} = a_0 + a_1 d + a_2 d^2 + \varepsilon \qquad (6)$$

was applied, with notation as in equation (5). The volumes were divided by the squared diameter in order to reduce heteroscedasticity (Cunia 1987). The parameter estimates by tree species are listed in Appendix 2.

Bark thickness and diameter increment of the previous 5-year period were applied when computing diameters 5 years ago. A five-year increment period was applied instead of a one-year period to reduce the effects of any annual variation of increment and measuring errors. The diameter under bark 5 years ago was calculated by subtracting the bark and diameter increment from the present above bark diameter. The above bark diameter 5 years ago was calculated assuming that the ratio *diameter inside bark/diameter outside bark* had remained unchanged for each tree during the previous 5 years (Kujala 1980). The present volume and volume 5 years ago were estimated by means of function (6) using the measured present diameter and diameter 5 years ago as respective explanatory variables. The estimate of volume increment was obtained as the difference of these two volume estimates.

3.1.3 Estimation of Volumes and Volume Increments for Tally Trees

Volume and volume increment were estimated for tally trees by means of regression analysis. The estimated sample tree volumes and model (6) (with d as the explanatory variable) were employed for estimating volumes for tally trees. The estimated volume increments of the sample trees and model (7) were utilised for estimating the volume increments for the tally trees,

$$\frac{i}{d^2} = e^{a_0 + a_1 d + a_2 \ln(d) + \varepsilon} \qquad (7)$$

where
i = volume increment of previous 5 years,
d = the present diameter,
a_j, $j = 0,1,2$, are the parameters of the model, and
$\varepsilon = N(0,\sigma^2)$ distributed error term.

Normalising by squared diameter was again applied to reduce heteroscedasticity. The parameter estimates by tree species are listed in Appendix 2.

3.1.4 Computation of Area and Volume Estimates

Area estimates of the reference units i (e.g. age-class) were obtained by

$$a_i = \sum_{j=1}^{3} \frac{n_{i,j}}{N_j} A_j \qquad (8)$$

where
$n_{i,j}$ = number of sample plots in sampling stratum j and reference unit i,
N_j = number of sample plots in sampling stratum j ($j = 1,...,3$), and
A_j = land area of sampling stratum j.

Volumes were estimated as follows. The mean volumes per hectare (m^3/ha) represented by each tally tree were first computed. For trees with a diameter less than 36.74 cm, the radius from which the tree was measured depended on the diameter of the tree (due to Bitterlich sampling). Larger trees can be considered to have been measured from circular plots of fixed radius 15 m.

(The diameter limit follows from the relascopic factor (q) of 1.5 and the maximum radius of the field plot 15 m.) The mean volume per hectare v^a represented by a tally tree of volume v and radius d was thus

$$\begin{aligned} v^a &= v\,40000\,q/(\pi d^2), \quad \text{if d} < 36.74 \\ &= v\,400/(9\pi), \quad \text{otherwise.} \end{aligned} \tag{9}$$

The mean volume increment per hectare and 5 years were obtained in a similar way. The 5 years interval was applied in order to decrease the effect of measurement error and annual growth variation on the estimates. The annual increments reported in the results section are simply the estimated 5-year increments divided by 5.

The mean volume estimates by reference units were derived from the mean volumes represented by the tally trees of that unit by taking into account the stratified sampling design:

$$\bar{v}_u^a = \sum_{j=1}^{3} w_j \frac{\sum_{k=1}^{n_{u,j}} \frac{v_{u,j,k}^a}{N_{u,j}}}{\sum_{j=1}^{3} w_j} \tag{10}$$

where

$v_{u,j,k}^a$ = mean volume (m^3/ha) represented by tally tree k in computation unit u and sampling stratum j,

$n_{u,j}$ = number of tally trees in unit u and stratum j,

$N_{u,j}$= number of centre points of field sample plots in unit u and stratum j, and

w_j = area of productive forestry land in sampling stratum j.

The volume increment estimates (m^3/ha/5a) of survivor trees were obtained in a similar way, i.e. replacing volumes v by volume increments. The total volume and total increment estimates for different reference units were obtained by respectively multiplying the mean volume and mean increment estimates by the area estimate.

The area, mean volume and increment estimators are ratio estimators and therefore biased. The bias converges asymptotically to zero as the number of observation increases and it is negligible compared to the standard error, see Section 3.3.

3.2 Estimation of Parameters by Means of Multi-Source Data

A non-parametric k-nearest-neighbour (k-nn) method was employed for estimating the values of forest parameters. The method produces both estimates for computation units and thematic maps. It is a multivariate version of the k-nn estimation method suggested by many authors (see e.g. Cover and Hart 1967 and Cover 1968). A similar 1-nn method has been earlier employed in the Finnish National Forest Inventory (NFI) by means of visual interpretation with aerial photographs (Poso 1972, Poso and Kujala 1978). Kilkki and Päivinen (1987) suggested the use of this 1-nn estimation method for forest inventories with satellite images. A distance-weighted version of the k-nn estimation method has been applied in an operative way in the Finnish NFI since 1990 (Tomppo 1996, cf. Dudani 1976). This method was adapted for the current task.

Four input files were produced for the classification. The files were:

1) ground truth file (a text file),
2) rectified satellite image file, bands 1,..,5 and 7 from the applied Landsat TM images,
3) digital land use data and computation unit file, and
4) cloud and cloud shadow mask file produced manually using the displayed satellite image.

The image processing utility written at the Finnish Forest Research Institute for the Finnish National Forest Inventory was adapted for the present files and forest parameters. The utility searches for each pixel of the image k nearest (in the feature space) pixels with known ground truth data (= pixels with sample plots). The whole ground data vectors from these k plots are attached to the pixel. The ground data vectors are weighted inversely proportionally to the squared distance in the feature space. An essential property of the process is that all inventory variables (parameters) can be estimated for the pixel at the same time. The idea of using k field plots instead of one is to reduce the random variation caused by, e.g., image noise, within-stand variation of the forest parameters, and dislocation of the field plots (compared to the image coordinates). Let us call the method a k-nearest-neighbour classification method. For the utilisation of the cloud mask file, see formula (16).

The procedure can be described more formally as follows. The Euclidean distance, $d_{p_i,p}$, is computed in the feature space from pixel p to be classified to each pixel p_i, whose ground truth is known (to pixel with sample plot i), and data from the k plots, $i_1(p),\ldots,i_k(p)$, with the shortest distances are employed in the analysis of pixel p. A maximum distance in the geographical space (usually 50 to 100 km in the horizontal direction) is set from the pixel p to the sample plots applied in order to avoid utilising sample plots from very different vegetation zones. A maximum distance is also set vertically, usually 50 to 200 m, in order to take into account the vegetation variations caused by elevation, provided that a digital terrain model is available. In this study, only the horizontal maximum distance was applied due to the absence of a reliable digital terrain model. The feasible set of nearest neighbours for pixel p is thus $\{p_i \mid d_{p_i,p}^{(x,y)} \le d_{\max}^{(x,y)}\}$ where $d_{p_i,p}^{(x,y)}$ is the geographical horizontal distance from pixel p to pixel p_i and $d_{\max}^{(x,y)}$ its maximum allowed value.

The weight of the ground data vector of plot i to pixel p is then defined by

$$w_{i,p} = \frac{\frac{1}{d_{p_i,p}^2}}{\sum_{j\in\{i_1(p),\ldots,i_k(p)\}} \frac{1}{d_{p_j,p}^2}}, \text{ if } i \in \{i_1(p),\ldots,i_k(p)\} \qquad (11)$$
$$= 0, \qquad \text{otherwise.}$$

Sums of weights, $w_{i,p}$ are calculated by computation units (for example by forest farms) in the image analysis process. The weight of plot i to computation unit U is then:

$$c_{i,u} = \sum_{p\in u} w_{i,p} \qquad (12)$$

Estimates of some (optional) forest variables are written in the form of a digital map during the procedure. The estimates are written after each processed line in the form of a multi-channel raster image. The land use classes outside forestry land are transferred directly from the digital map file. Within forestry land, the variables entered by the operator are estimated by the weighted averages of the k nearest neighbours (see Tomppo 1991, 1996)

$$\hat{m}_p = \sum_{j\in\{i_1(p),\ldots,i_k(p)\}} w_{j,p}\, m_j \qquad (13)$$

where
$\hat{m}_p$ = the multi-source estimator of the value of variable M at pixel p, and
m_j = the measured value of variable M at field plot j.

Mode value is used instead of the mean for variables of nominal or ordinal scales. Examples of the map themes are growing stock volumes by tree species, dominant tree species or tree species groups, e.g., on the basis of the tree stem volume (Fig. 3), increments by tree species, site fertility, and mean age of the growing stock.

3.2.1 Compilation of Tables and Map Production

The inventory statistics by computation units were obtained by means of digital boundary maps and the weight coefficients (12). The area of water and non-forestry land cover classes were estimated by computation units from the digital maps produced by multiplying the number of pixels classified in the land cover class by the size of the pixel:

$$A_{c,u} = \#(p \mid p \in c, p \in u)a, \qquad (14)$$

where
c = land cover class,
u = computation unit, and
a = area of one pixel.

The non-forestry land area located under a possible cloud mask can be estimated in a similar way if the land cover class estimates outside forestry land are based purely on map data. We assumed that the proportion of non-forestry land under the cloud mask was the same as that outside the cloud mask, because land cover classes were estimated from satellite images. The total non-forestry land area (14) was multiplied by *total area of a bureau / area outside cloud mask* and used as a final estimate of non-forestry land area, cf. also the formula (16).

The estimates of the areas of the forestry land reference units, by computation units, were

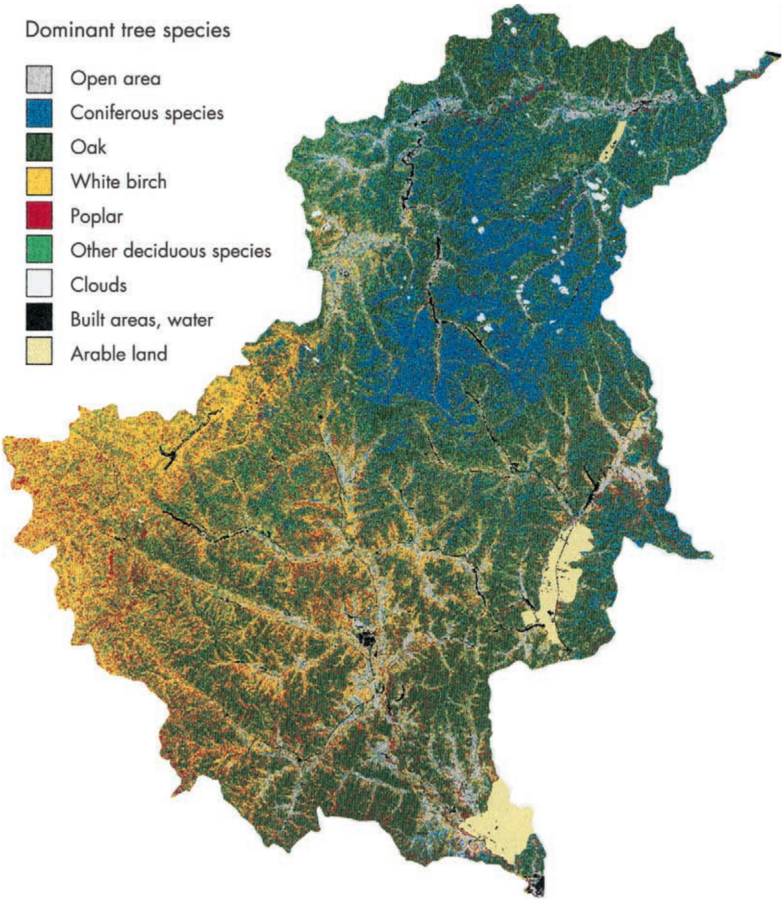

Fig. 3. An example of a thematic map, dominant tree species group.

obtained from the estimated plot weights employing the equation

$$A_{s,u} = a \sum_{i \in I_s} c_{i,u}, \tag{15}$$

where
S = forestry land reference unit (e.g. forest type),
I_S = set of sample plots belonging to the reference unit, and
u = computation unit (here a forest farm).

Reduced weight sums $c^r_{i,u}$ are obtained from the formula (12), if clouds or their shadows cover a part of the area of the computation unit u. The real weight sum for plot i is estimated by means of the formula

$$c_{i,u} = c^r_{i,u} \frac{A_{s,u}}{A^r_{s,u}} \tag{16}$$

where
$A_{s,u}$ = area of the forestry land of unit u, and
$A^r_{s,u}$ = area of the forestry land of unit u not covered by the cloud mask.

It is thus assumed that the forestry land covered by clouds per computation units is, on average, similar to the rest of the forestry land with respect to the forest variables. The proportion of non-forestry land covered by clouds is estimated as given above.

The volume estimates are computed by computation and reference units in the following way.

Mean volumes are estimated by the formula:

$$v = \frac{\sum_{i \in I_s} c_{i,u} v_{i,t}}{\sum_{i \in I_s} c_{i,u}} \tag{17}$$

where $v_{i,t}$ is the estimated volume per hectare of timber assortment (log product) t for plot i. The corresponding total volumes are obtained by replacing the denominator in formula (17) by 1. Mean and total volume increments are similarly estimated.

3.3 Estimation of Standard Errors

Estimators (8) and (10) are based on pure field data and are ratio estimators. More generally, the goal is to estimate a parameter of the form

$$M = \frac{X}{Y} \tag{18}$$

where X and Y are expectations of two random variables, x and y. Let x_i and y_i be their observed values on sample plot i, then the ratio estimator of M is

$$m = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} = \frac{\bar{x}}{\bar{y}} \tag{19}$$

where n is the number of sample plots in the inventory area. When n is large,

$$Em \approx \frac{E\bar{x}}{Y} = \frac{X}{Y} = M \tag{20}$$

that is, m is approximately unbiased. The estimation of its standard error is also complicated by the spatial autocorrelation and possible trend-like changes of the target variables. Matérn (1947, 1960) suggested the quantity $E(m - M)^2$, the error variance, as a measure of reliability of the estimator and also proposed an estimator of the error variance. Consider the cluster-wise residuals $z_r = x_r - m y_r$, where $x_r = \sum_{i \in r} x_i$ and $y_r = \sum_{i \in r} y_i$ with r a cluster of field plots i, and assume that the residuals form a realisation of a second order stationary (weakly stationary) stochastic process. The variance of the process can be estimated by means of quadratic forms $T = \sum_r \sum_s c_{rs} z_r z_s$, where $c_{rs} = c_{sr}$, $\sum_r \sum_s c_{rs} = 0$ and $\sum_r c_{rr} = 1$.

It can be shown that any estimator of this form is unbiased if the process z is spatially uncorrelated, and conservative if the process is positively correlated (Matérn 1960). The method is applied by sampling strata as follows. Within each stratum, groups g of four field plot clusters

r_3 r_4

r_1 r_2

are composed in such a way that each cluster belongs to four different groups (with necessary boundary modifications). Deviance of the cluster mean $\bar{y}_r$ from the stratum mean $\bar{y}$ is computed for each cluster j. Denote

$$z_r = (\bar{y}_r - \bar{y}) n_r \tag{21}$$

where n_r is the number of relevant sample points in cluster r. The weights c_{rs} are chosen in perhaps the most obvious way, $c_{r_1} = c_{r_4} = -c_{r_2} = -c_{r_3} = 1/2$. This choice has also been employed in the Swedish and Finnish inventories (cf. Ranneby 1981, Salminen 1985). The quadratic forms can then be expressed as $T_g = (z_{r1} - z_{r2} - z_{r3} + z_{r4})^2 / 4$, and the implied standard error estimators (the square root of the error variance estimator) for each stratum are

$$s = \frac{\sqrt{k \sum_l T_l}}{\sum_i y_i} \tag{22}$$

where
l refers to the groups of clusters in the stratum,
i refers to the relevant sample plots in the stratum, and
k indicates how many clusters each cluster group represents (in our case, k =1).

The standard error estimators for the whole Bureau area were obtained by combining the stratum-specific estimators with the usual formula for stratified sampling (e.g. Cochran 1977).

4 Examples of Forest Resource Estimates

Forest resource estimates and the standard errors of the area estimates were computed for the whole area of Hebei Forest Bureau using field measurements only. The estimates were computed by forest farms using satellite images, in addition to field measurements. The estimation parameters, the value of *k* and the maximum distance from the pixel to be analysed to the potential nearest neighbours were judged using a pixel level cross-validation technique. The chosen value of *k* was 3 and the maximum distance 30 km. Complete tables of the results are given in Tomppo et al. (1994). Only few examples are given here. Fig. 3 shows an example of a thematic map produced by means of the multi-source technique.

4.1 Land Area Estimates

The total area of the Bureau is 382 401 ha, according to the statistics provided by the Hebei Forestry Bureau. The area of forestry land was, according to the field inventory, 373 700 ha, of which 1900 ha was classified as non-productive forestry land. Thus, the area of productive forestry land was 371 800 ha (Table 1). The other land use classes and their proportions were: agricultural land 0.8%, pasture land 0.5%, built-up land 0.6%, and water 0.5%. The productive forestry land was also classified on the basis of the actual growing stock: 87.4% of the land was classified as forest land (canopy closure at least 30% or on young planted stands with a density of living trees more than 85% of reasonable planting density), 7.0% as sparse tree land, 0.9% as sapling tree land, and 4.7% as open land.

Table 1 shows the estimates obtained for the whole area when employing the field inventory

Table 1. Area of the of the forestry land subclasses by forest farms based on the multi-source inventory, and for the whole area based on the field inventory, together with the standard errors of the field inventory estimates.

Forest farm	Subclass										Total	
	Forest land		Sparse tree land		Sapling tree land		Open land		Non productive			
	1000 ha and %											
Lianying	17.0	90.7	1.0	5.2	0.1	0.6	0.5	2.8	0.1	0.7	18.7	100.0
Jinkuang	21.5	85.9	2.1	8.2	0.3	1.1	0.9	3.7	0.2	1.0	25.0	100.0
Shuguang	17.5	85.8	1.4	6.7	0.4	2.1	1.0	4.7	0.2	0.7	20.3	100.0
Zhongxing	21.5	90.3	1.2	5.1	0.3	1.1	0.7	2.9	0.2	0.7	23.8	100.0
Qinglin	14.7	80.1	1.7	9.2	0.3	1.4	1.5	8.0	0.2	1.4	18.3	100.0
Nongxing	14.7	89.9	1.2	7.1	0.0	0.2	0.3	1.9	0.1	0.9	16.3	100.0
Xianjin	23.0	87.4	1.9	7.3	0.2	0.6	0.9	3.6	0.3	1.1	26.3	100.0
Yuefeng	17.5	95.0	0.6	3.4	0.0	0.3	0.2	1.0	0.1	0.3	18.5	100.0
Gaofeng	17.6	94.7	0.7	3.8	0.0	0.0	0.1	0.7	0.2	0.9	18.6	100.0
Weidong	17.5	86.5	1.3	6.4	0.2	1.2	0.9	4.6	0.3	1.4	20.3	100.0
Xiwutong	13.9	91.7	0.8	5.1	0.1	0.4	0.3	2.1	0.1	0.7	15.2	100.0
Yuejing	20.4	83.6	2.5	10.2	0.3	1.2	0.9	3.7	0.3	1.3	24.4	100.0
Shuangfeng	20.0	82.4	2.4	9.9	0.2	1.0	1.3	5.3	0.4	1.5	24.3	100.0
Xiaowutong	14.5	84.1	1.5	8.4	0.2	1.0	0.9	5.3	0.2	1.1	17.3	100.0
Sifangshan	13.4	87.6	1.0	6.5	0.2	1.4	0.6	4.0	0.1	0.5	15.3	100.0
Xiangyang	15.5	74.8	3.3	16.1	0.4	1.8	1.2	5.6	0.3	1.6	20.7	100.0
Hekou	17.0	87.9	1.1	5.5	0.2	1.1	0.9	4.5	0.2	1.0	19.4	100.0
Shuangyi	13.3	85.6	1.4	8.8	0.1	0.4	0.6	4.1	0.2	1.2	15.6	100.0
Qingfeng	15.9	77.3	2.2	10.6	0.4	1.9	1.8	8.8	0.3	1.3	20.5	100.0
All Hebei	326.3	86.2	29.1	7.7	3.8	1.0	15.6	4.1	3.9	1.0	378.7	100.0
Field	325.1	87.0	25.9	6.9	3.9	0.9	17.4	4.7	1.9	0.5	373.7	100.0
Standard error	9.2	2.4	5.2	1.4	1.8	0.5	4.4	1.2	1.0	0.3	2.5	0.0

Table 2. Area of productive forestry land by dominant tree species groups and forest farms based on the multi-source inventory, and by dominant tree species groups based on the field inventory, together along with the standard errors of the field inventory estimates.

Forest farm	Dominant tree species groups																								Total	
	Treeless		Pines		Larch		Spruce		Fir		Maple		Other birches		White birch		Oak		Linden		Poplar		Other decid.			
											1000 ha and %															
Lianying	1.0	5.3	3.2	17.1	0.6	3.3	0.5	2.7	3.4	18.3	0.2	0.9	1.3	6.8	1.9	10.2	3.3	18.0	1.1	5.7	1.4	7.7	0.8	4.1	18.6	100.0
Jinkuang	2.0	8.2	1.4	5.7	1.2	4.9	0.2	0.8	1.7	7.0	0.1	0.4	2.4	9.8	3.5	14.3	6.8	27.5	1.6	6.6	2.1	8.5	1.6	6.3	24.7	100.0
Shuguang	1.5	7.6	1.2	6.1	0.8	4.2	0.3	1.5	2.1	10.3	0.2	1.1	1.4	6.8	2.5	12.4	6.3	31.1	1.5	7.3	1.2	5.9	1.2	5.9	20.2	100.0
Zhongxing	1.3	5.6	3.4	14.4	1.0	4.2	0.6	2.6	3.9	16.7	0.2	0.9	1.8	7.6	2.1	9.1	5.0	21.3	1.5	6.2	1.6	6.9	1.0	4.4	23.6	100.0
Qinglin	2.0	11.3	1.3	7.0	0.8	4.6	0.2	1.1	1.5	8.5	0.1	0.4	1.2	6.6	3.5	19.1	3.5	19.6	1.2	6.8	1.3	7.2	1.4	7.8	18.1	100.0
Nongxing	0.5	3.3	0.4	2.8	0.5	2.9	0.1	0.3	0.6	3.6	0.0	0.1	1.1	6.6	7.3	45.1	1.8	11.1	0.8	5.0	2.1	3.2	1.0	6.0	16.2	100.0
Xianjin	1.5	5.8	0.4	1.6	0.9	3.6	0.0	0.1	0.6	2.1	0.1	0.4	2.3	8.7	8.1	31.2	5.2	20.1	2.1	8.1	3.4	13.2	1.3	5.1	26.0	100.0
Yuefeng	0.6	3.1	0.3	1.8	0.2	1.3	0.0	0.0	0.3	1.7	0.1	0.6	1.3	6.8	6.0	32.4	4.7	25.6	1.3	7.0	3.2	17.2	0.5	2.5	18.4	100.0
Gaofeng	0.5	2.6	0.4	2.3	0.1	0.5	0.0	0.0	0.4	2.0	0.1	0.3	1.4	7.4	8.9	48.4	1.7	9.1	1.3	6.8	3.1	17.1	0.6	3.4	18.4	100.0
Weidong	2.4	12.1	1.1	5.5	1.0	4.9	0.2	1.2	1.5	7.6	0.1	0.5	1.6	8.0	2.4	12.1	5.7	28.6	0.9	4.6	1.8	9.1	1.2	5.8	20.0	100.0
Xiwutong	0.6	3.8	0.2	1.2	0.3	2.0	0.0	0.1	0.3	1.8	0.1	0.4	1.3	8.3	4.8	32.2	3.8	25.3	0.9	6.2	2.1	14.2	0.7	4.4	15.1	100.0
Yuejing	2.0	8.5	0.0	0.0	2.1	8.6	0.0	0.0	0.1	0.2	0.2	0.6	1.6	6.8	5.5	22.7	8.2	34.1	1.3	5.3	2.9	1.9	0.3	1.3	24.1	100.0
Shuangfeng	2.5	10.4	0.1	0.5	1.6	6.6	0.0	0.0	0.2	0.7	0.1	0.3	1.6	6.6	6.8	28.4	6.1	25.5	1.2	5.0	2.8	11.8	1.0	4.3	24.0	100.0
Xiaowutong	1.7	10.2	0.3	1.9	0.9	5.5	0.0	0.2	0.5	3.1	0.0	0.1	1.4	8.4	4.1	24.1	4.2	24.8	1.2	7.0	1.6	9.5	0.9	5.2	17.1	100.0
Sifangshan	1.3	8.4	0.3	2.3	1.2	7.8	0.1	0.4	0.3	2.0	0.0	0.0	1.5	9.9	2.4	15.8	5.7	37.2	0.7	4.4	1.2	7.8	0.6	3.9	15.2	100.0
Xiangyang	2.8	13.7	0.0	0.0	2.5	12.1	0.0	0.0	0.0	0.0	0.0	0.2	1.4	6.8	3.9	19.0	6.5	31.9	0.9	4.5	2.0	9.6	0.4	2.1	20.3	100.0
Hekou	1.5	7.9	1.2	6.4	0.5	2.8	0.2	1.3	2.0	10.7	0.2	1.1	1.3	6.8	2.0	10.4	6.2	32.1	1.7	8.9	1.1	5.7	1.2	6.0	19.2	100.0
Shuangyi	1.1	6.8	0.6	4.1	0.6	4.0	0.1	0.5	0.7	4.4	0.1	0.6	1.4	9.0	3.4	22.4	3.5	22.7	1.3	8.2	1.7	11.0	1.0	6.3	15.4	100.0
Qingfeng	2.4	11.8	1.5	7.6	1.4	6.8	0.3	1.7	2.1	10.3	0.1	0.5	1.4	7.0	3.1	15.4	4.1	20.1	1.7	8.4	1.0	5.0	1.1	5.3	20.3	100.0
All Hebei	29.3	7.8	17.6	4.7	18.3	4.9	2.9	0.8	22.2	5.9	1.9	0.5	28.5	7.6	82.3	22.0	92.4	24.6	24.1	6.4	37.8	10.1	17.7	4.7	374.8	100.0
Field	29.4	7.9	11.6	3.1	20.1	5.4	3.7	1.0	19.1	5.1	2.1	0.6	34.3	9.2	74.5	20.0	104.8	28.2	20.9	5.6	35.5	9.5	15.8	4.3	371.8	100.0
Standard error	5.9	1.6	2.3	0.6	5.5	1.5	1.5	0.4	3.0	0.8	0.8	0.2	5.3	1.4	7.0	1.9	8.9	2.4	3.7	1.0	4.0	1.1	2.9	0.8	2.6	0.0

alone, and for the whole area and forestry farms when using the multi-source inventory. The estimate of the area of productive forest land, based on multi-source inventory, was 374 800 ha, i.e. 0.8% higher than that based on field measurements. The relative standard error of the productive forestry land area was 0.7%. The estimates obtained with both methods for the whole area were similar. All multi-source inventory estimates were within two standard errors of the estimates of the field inventory. The proportion of forest land varied by farms from 75% to 95%, the proportion of sparse tree land from 3% to 16% and the proportion of open land from 1% to 9%. Some estimates of the growing stock for the productive forestry land area, and its annual increment, are given in the following section.

4.2 Dominance of Tree Species

The dominant tree species of a forest stand was defined as the species which has the highest stem volume or for which the silvicultural measures were intended. Nineteen tree species, genera of species (spp.) or groups of species were identified in the field measurements (Appendix 1). They were further grouped into 11 groups plus open land for multi-source inventory. The most common dominant species was oak (*Quercus* spp., mainly *Quercus mongolica*). The proportion of oak dominated stands of the area of productive forestry land was 28% (standard error 2.4%) according to the field measurements and 25% according to the multi-source inventory (Table 2).

The difference was smaller than the double standard error of the field data based estimate. The somewhat lower estimate of the multi-source inventory may have been caused by the fact that a part of the oak dominated area had been estimated to be dominated by linden, poplar or ‘other deciduous trees’. The area estimate of these species was slightly higher in the multi-source inventory than in the field inventory. The proportional estimates of the other species based on field data and multi-source data were otherwise within 1%-unit of each other. Exceptions, in addition to oak, were the group ‘other birch species’, white birch and pine. However, the total area estimates of birch species were close to each other with both methods.

There was a quite high variation in tree species dominance at the forest farm level. The proportion of the area of oak dominated forests varied from 9% to 37%, the proportion of white birch (*Betula platyphylla*) dominated forests from 9% to 48% and the proportion of pine dominated forests from 0% to 17%. Coniferous tree species dominated 15% of the productive forest area. Coniferous dominated forests were located mainly in the northern part and white birch dominated forests in the western part of the Bureau area (Fig. 3, Table 3 and Fig. 1).

4.3 Age Structure

The age of the growing stock of each sample plot was estimated by increment borings and visual estimation. The forests of Hebei are mostly young, 60% (57% according to the multi-source inventory) of the forests on productive forestry land were at most 40 years old and 85% (83%) were at most 60 years old. These figures included the treeless areas, 7.9% (7.8%) of the productive forestry land (Table 3). The estimates of the proportions of age classes obtained with field inventory and multi-source inventory were within 1%-unit of each other, except for the age classes 21–40 and 41–60 years. In the class 21–40 years, the multi-source estimate was 3.3%-units (8%) lower than the field inventory estimate. The difference was in the opposite direction in the 41–60 years age class.

The multi-source inventory showed that the proportion on old forests (older than 80 years) was higher in the Lianying and Zhongxing forest farms than in the other farms. The old pine stands which have been conserved for seed production were located on these farms. There was considerable variation in the age structure between forest farms. Some of the variation was caused by variations in tree species dominance. The pine dominated forests were, on average, older than other forests, while the larch dominated forests were younger. For larch dominated stands, the proportion of young plantations was remarkably high. The proportion of young forests was high in Xiangyang and Weidong, for example.

Of the total volume of the growing stock, 76% was in forests no more than 60 years old and 47%

Table 3. Area of productive forestry land by age classes and forest farms based on the multi-source inventory, and by age classes based on the field inventory, together with the standard errors of the field inventory estimates.

Forest farm	Age class														Total	
	Treeless		1–19		20–39		40–59		60–79		80–99		100+			
	1000 ha and %															
Lianying	1.0	5.3	1.2	6.5	4.9	26.2	4.7	25.1	2.7	14.5	2.2	11.8	2.0	10.6	18.6	100.0
Jinkuang	2.0	8.2	4.3	17.3	9.2	37.0	5.1	20.8	2.1	8.7	1.3	5.1	0.7	3.0	24.7	100.0
Shuguang	1.5	7.6	1.7	8.5	6.5	32.3	4.9	24.4	3.0	14.6	1.7	8.6	0.8	4.0	20.2	100.0
Zhongxing	1.3	5.6	2.0	8.3	6.7	28.2	5.4	22.8	3.3	14.1	2.7	11.5	2.2	9.4	23.6	100.0
Qinglin	2.0	11.3	1.9	10.8	5.9	32.8	4.9	27.0	1.6	8.8	1.1	5.8	0.7	3.6	18.1	100.0
Nongxing	0.5	3.3	1.3	8.1	7.0	43.3	6.1	37.7	0.8	5.2	0.3	1.9	0.1	0.4	16.2	100.0
Xianjin	1.5	5.8	2.5	9.8	10.6	40.7	8.5	32.7	2.2	8.5	0.5	2.1	0.1	0.4	26.0	100.0
Yuefeng	0.6	3.1	1.2	6.7	7.8	42.2	6.0	32.7	1.9	10.3	0.7	3.7	0.3	1.4	18.4	100.0
Gaofeng	0.5	2.6	0.7	3.5	7.7	41.8	7.9	43.0	1.3	7.2	0.2	0.8	0.2	1.0	18.4	100.0
Weidong	2.4	12.1	4.0	19.8	8.0	40.0	2.8	13.9	1.2	5.8	1.0	4.9	0.7	3.5	20.0	100.0
Xiwutong	0.6	3.8	0.9	6.3	6.4	42.3	5.2	34.5	1.4	9.6	0.4	2.9	0.1	0.6	15.1	100.0
Yuejing	2.0	8.5	4.2	17.6	9.5	39.4	5.7	23.9	1.6	6.7	0.9	3.8	0.0	0.1	24.1	100.0
Shuangfeng	2.5	10.4	3.9	16.3	9.7	40.5	5.9	24.6	1.4	6.0	0.5	2.1	0.0	0.1	24.0	100.0
Xiaowutong	1.7	10.2	2.7	15.9	6.6	38.6	4.4	25.5	1.2	6.7	0.4	2.4	0.1	0.6	17.1	100.0
Sifangshan	1.3	8.4	3.0	19.5	7.0	46.1	2.3	15.1	0.9	6.0	0.5	3.5	0.2	1.3	15.2	100.0
Xiangyang	2.8	13.7	5.1	25.2	7.7	37.9	3.0	14.6	1.1	5.2	0.7	3.3	0.0	0.2	20.3	100.0
Hekou	1.5	7.9	1.2	6.3	6.3	32.9	4.5	23.6	3.0	15.8	1.8	9.1	0.8	4.2	19.2	100.0
Shuangyi	1.1	6.8	2.0	13.0	5.9	38.3	4.3	27.7	1.4	9.1	0.5	3.5	0.3	1.6	15.4	100.0
Qingfeng	2.4	11.8	2.2	10.8	6.1	30.3	5.1	25.0	2.1	10.5	1.5	7.5	0.8	4.1	20.3	100.0
All Hebei	29.3	7.8	46.1	12.3	139.4	37.2	96.6	25.8	34.3	9.2	18.9	5.0	10.1	2.7	374.8	100.0
Field	29.4	7.9	44.1	11.9	150.4	40.5	90.3	24.3	32.4	8.7	18.5	5.0	6.6	1.8	371.8	100.0
Standard error	5.8	1.5	7.5	2.0	11.1	3.0	10.2	2.7	4.7	1.3	3.7	1.0	1.7	0.5	2.6	0.0

in forests no more than 40 years old. The high mean volumes of young forests also reflected the high density of forests and the need for silvicultural measures and thinning. The mean volume in the forests less than 20 years old was 47 m^3/ha and in forests aged between 20 and 40 years was 117 m^3/ha.

4.4 Growing Stock

The total volume of the growing stock on productive forest land was 42.4 million m^3 and the mean volume 114.1 m^3/ha, according to the field inventory. The corresponding figures from the multi-source inventory were 44.3 million m^3 and 118.7 m^3/ha. The somewhat higher mean volume of multi-source inventory may have been caused by the fact that a digital elevation model was not available, and so the mean volume was over-estimated on the northern slopes of hills. The mean volumes by tree species and by forestry farms are given in Table 4. The total volumes can be obtained by multiplying mean volumes by the area of the productive forestry land. Of the tree species, oak had the highest mean volume, 26.0 m^3/ha according to the field inventory and 22.8 m^3/ha according to the multi-source inventory. The multi-source estimate was 12% lower than the field inventory estimate. The multi-source estimates of pine, fir, white birch, linden, poplar and the group other deciduous trees were higher than those from the field inventory. The higher mean volume of pine may have been caused by the over-estimation on the northern slopes, while oak and other deciduous trees may have been partly mixed in the multi-source inventory.

The highest mean volumes were found in the forest farms of Lianying (170 m^3/ha), Zhongxing (157 m^3/ha) and Gaofeng (133 m^3/ha). The

Table 4. Mean volume of growing stock by tree species groups and forest farms based on the multi-source inventory, and by tree species groups based on the field inventory.

Forest farm	Tree species groups											Total
	Pines	Larch	Spruce	Fir	Maple	Other birches	White birch	Oak	Linden	Poplar	Other decid.	
						m^3/ha						
Lianying	38.6	0.9	10.2	20.7	5.1	23.3	11.4	17.9	14.4	14.1	13.0	169.6
Jinkuang	12.2	1.5	3.2	7.9	3.1	18.3	12.8	26.5	9.8	13.6	9.3	118.3
Shuguang	12.1	1.0	3.5	9.5	3.5	16.2	13.1	25.0	11.1	10.0	12.4	117.4
Zhongxing	32.3	0.8	9.1	18.5	4.7	22.1	10.3	19.6	14.2	12.4	12.6	156.7
Qinglin	16.0	1.4	4.1	9.4	3.5	16.3	18.0	19.3	11.7	12.1	11.3	123.0
Nongxing	7.7	0.6	1.5	3.4	2.8	12.6	40.6	11.2	9.5	22.8	9.4	122.1
Xianjin	5.1	0.9	1.2	2.7	3.2	14.5	26.2	19.9	11.9	20.4	9.4	115.3
Yuefeng	4.4	0.3	1.9	1.2	3.4	15.7	26.9	24.0	11.2	26.4	10.1	125.6
Gaofeng	5.7	0.1	1.8	2.2	3.8	13.7	44.6	10.2	11.3	28.4	11.6	133.4
Weidong	13.2	1.6	3.7	9.2	2.9	17.2	10.9	27.9	7.7	14.5	9.8	118.6
Xiwutong	3.6	0.5	1.4	2.0	3.6	15.0	28.4	24.5	10.9	23.1	9.8	122.8
Yuejing	0.5	0.8	0.2	0.1	2.0	13.2	15.5	31.1	8.3	17.3	5.5	94.5
Shuangfeng	1.8	0.9	0.3	0.6	2.2	13.1	21.7	23.7	7.8	17.6	6.7	96.4
Xiaowutong	7.0	1.5	1.3	3.6	2.6	15.0	19.6	23.7	9.5	15.3	7.5	106.5
Sifangshan	5.7	2.6	1.9	2.7	2.4	18.9	12.1	35.3	6.6	16.2	5.6	110.0
Xiangyang	0.3	1.2	0.0	0.0	1.1	12.3	10.5	27.9	5.7	11.8	4.0	75.0
Hekou	12.6	0.5	3.5	10.9	3.8	16.8	10.7	25.7	12.6	9.3	14.3	120.7
Shuangyi	8.5	1.1	2.2	5.1	3.5	16.1	19.6	22.2	12.0	16.9	9.9	117.0
Qingfeng	17.4	1.8	4.6	10.8	3.3	15.9	15.7	17.0	12.3	9.5	11.1	119.3
All Hebei	10.9	1.1	2.9	6.4	3.2	16.1	19.0	22.8	10.4	16.2	9.6	118.7
Field	7.7	1.3	2.7	5.6	3.1	16.6	17.5	26.0	9.4	15.7	8.3	114.1

Table 5. Mean annual increment by forest farms and dominant tree species groups based on the multi-source inventory, and by dominant tree species groups based on the field inventory.

Forest farm	Dominant tree species groups											Total
	Pines	Larch	Spruce	Fir	Maple	Other birches	White birch	Oak	Linden	Poplar	Other decid.	
						m^3/ha						
Lianying	0.39	0.03	0.22	0.63	0.13	0.29	0.21	0.41	0.27	0.28	0.40	3.28
Jinkuang	0.12	0.05	0.07	0.24	0.09	0.34	0.30	0.63	0.23	0.31	0.33	2.72
Shuguang	0.13	0.04	0.08	0.29	0.09	0.27	0.28	0.58	0.25	0.23	0.42	2.66
Zhongxing	0.33	0.03	0.20	0.56	0.12	0.30	0.20	0.45	0.26	0.26	0.40	3.11
Qinglin	0.16	0.04	0.09	0.29	0.10	0.26	0.39	0.44	0.25	0.27	0.39	2.68
Nongxing	0.08	0.02	0.04	0.12	0.09	0.23	0.90	0.28	0.25	0.52	0.35	2.89
Xianjin	0.05	0.03	0.03	0.09	0.10	0.28	0.60	0.49	0.33	0.44	0.33	2.77
Yuefeng	0.07	0.01	0.06	0.06	0.10	0.31	0.58	0.63	0.32	0.54	0.34	3.03
Gaofeng	0.08	0.00	0.06	0.10	0.13	0.24	0.93	0.25	0.31	0.55	0.40	3.06
Weidong	0.13	0.06	0.09	0.27	0.08	0.33	0.25	0.68	0.16	0.33	0.33	2.71
Xiwutong	0.05	0.02	0.05	0.08	0.11	0.29	0.63	0.61	0.30	0.47	0.33	2.95
Yuejing	0.02	0.05	0.00	0.00	0.06	0.30	0.41	0.85	0.26	0.41	0.20	2.57
Shuangfeng	0.02	0.04	0.01	0.03	0.07	0.29	0.54	0.61	0.24	0.43	0.25	2.53
Xiaowutong	0.05	0.05	0.03	0.12	0.08	0.32	0.48	0.60	0.26	0.38	0.28	2.65
Sifangshan	0.06	0.10	0.05	0.09	0.07	0.41	0.32	0.90	0.19	0.38	0.20	2.76
Xiangyang	0.01	0.06	0.00	0.00	0.03	0.32	0.32	0.76	0.20	0.34	0.17	2.20
Hekou	0.13	0.02	0.08	0.32	0.10	0.26	0.21	0.60	0.27	0.22	0.48	2.70
Shuangyi	0.08	0.03	0.05	0.16	0.10	0.29	0.45	0.53	0.30	0.38	0.34	2.73
Qingfeng	0.17	0.06	0.11	0.33	0.09	0.25	0.34	0.40	0.27	0.20	0.37	2.58
All Hebei	0.11	0.04	0.07	0.20	0.09	0.29	0.43	0.57	0.26	0.36	0.33	2.76
Field	0.09	0.05	0.07	0.18	0.09	0.32	0.40	0.66	0.24	0.36	0.29	2.74

volume of coniferous species was high on these farms. The lowest mean volume, 75 m^3/ha, was in Xiangyang. A considerable proportion of the growing stock was located in special purpose forests: 4.5% of the volume was accounted for by seed collection forests (almost totally conifers) and 4.6% by national defence forests.

4.5 Increment

The annual increment of the volume of the growing stock, based on the field inventory, was 1.01 million m^3, while the multi-source estimate was 1.03 million m^3. The mean increments by tree species and by farms with the multi-source inventory and by tree species with the field inventory are given in Table 5. The total increment can be obtained by multiplying the mean increment by the area of productive forestry land. The mean annual increment, based on the field inventory, was 2.74 m^3/ha/a, and according to the multi-source inventory, 2.76 m^3/ha/a. The two estimates were also close by tree species. The differences were similar to those of the mean volumes. The pine increment was over-estimated, probably for the same reason as the mean volume. The oak increment was under-estimated and the increments of lime and ‘other deciduous trees’ were over-estimated. The mean annual increment by forest farms varied from Xiangyangs 2.2 m^3/ha/a to Liangyings 3.3 m^3/ha/a. The increment of the drain was not included in the presented figures.

5 Discussion

The paper has presented a multi-source forest inventory method which utilises sparse field sampling and satellite images. The method has some features of the Finnish multi-source national forest inventory but also many differences. The authors believe that a similar method, or its modification, as well as the ideas in realising the inventory, can be applied when planning or realising a forest inventory in an arbitrary area in any temperate and boreal region. The sampling design was derived from existing preliminary information and experiences from other inventories. An important goal when planning the sampling design was its suitability for the utilisation of satellite images. The chosen design was not necessarily optimal even for the estimates of the basic parameters, such as mean volume of the growing stock, because of the lack of detailed preliminary knowledge of the spatial correlation structure of the forests in question. The thematic maps obtained from the multi-source inventory could be utilised for further improvement of the sampling design (e.g. Tomppo et al. 1998). Note that achieving an overall optimal design is a very complicated task, possibly an impossible one, due to the fact that different variables have different spatial correlations and thus presume different sampling designs (cf. Ranneby 1981 and Ranneby et al. 1987). The volume and increment estimates were based partly on available local volume functions and partly on new estimation models. The models could be improved if more detailed tree level measurements would be available.

The inventory can be employed for estimating regression models for various purposes. As an example, height models (estimation of height as a function of diameter) and diameter increment models were produced. These models are presented in Appendix 2.

The reliability of the inventory results was assessed using the method proposed by Matérn (1960). This method is suitable for systematic sampling, takes into account spatial correlation at different scales, and does not over-estimate the accuracy when spatial correlation is positive.

The multi-source method seemed to work in North-East Chinese conditions and to give detailed geo-referenced information concerning the forest resources and the structure of the forests, as well as land cover classes, at reasonable costs. The information can be applied to small unit (forest farm) forest management planning, including ecological planning. The farms, usually of some 10 thousand hectares, are important units in forest management planning for which farm-level information is essential. Sparse field sampling alone does not allow the estimation of parameters at the farm level.

The results show that most of the forests are dominated by species of low economic value. Further, due to the high proportion of young forests, the removals of the valuable species cannot be high in the near future. However, the growth rates of the valuable species (e.g. white birch, Korean pine and poplar) indicate that the area has great potential for wood production.

The results confirm that stand-level forest inventories often under-estimate the volume of the growing stock. The estimate of the mean volume of the growing stock was 114 m^3/ha while the old figure, based on stand-level inventories, was 72 m^3/ha. There are probably also other reasons for the large difference, e.g. small trees may not have been included in the stand-level inventories. For the whole area, the multi-source inventory gave similar results to the field inventory, also by tree species and age classes. The lack of an analytical method for the error assessment in the multi-source inventory prohibits the estimation of errors at the forest farm level. A digital terrain model was not available in this study. It would have improved the estimates of the multi-source inventory, especially on the northern sides of the hills. The volume of growing stock was over-estimated in some shadowed areas. This explains the difference in total volume between estimates based on field measurements (mean volume 114 m^3/ha) and multi-source information (119 m^3/ha).

Acknowledgements

We are grateful to the Chinese Ministry of Forestry for supporting the study. The Hebei Forest Bureau in Heilongjiang Province provided valuable assistance and all kinds of care in the planning and execution phases of the field work and in presenting and distributing the results. The Finnish Ministry of Trade and Industry and the Ministry of Agriculture and Forestry supported the study financially. The origin and the idea for the study came from Mr. Jori Pesonen, counselor of industry and the former managing director of Valmet paper machine industry, Dr. Jari Parvianen and Mr. Leo Tervo. Two anonymous referees made valuable comments on the manuscript. Dr. Ashley Selby revised the language. We wish to express our thanks to all concerned.

References

Cochran, W.G. 1977. Sampling techinques. 3rd edn. Wiley, New York.

Cover, T.M. 1968. Estimation by the Nearest Neighbor Rule. IEEE Transactions on Information Theory, Volume IT-14, Number 1: 50–55.

— & Hart, P.E. 1967. Nearest Neighbor Pattern Classification. IEEE Transactions on Information Theory, Volume IT-13: 21–27.

Cunia, T. 1987. Use of dummy variable techniques in estimation of biomass regressions. In: Estimating tree biomass regressions and their errors. p. 37–48. Proceedings of the workshop on tree biomass regression functions and their contribution to the error of forest inventory estimates. United States Department of Agriculture, Forest Service, Northeastern Forest Experiment Station.

Dudani, S.A. 1976. The Distance-Weighted k-Nearest-Neighbor Rule. IEEE Transactions on System, Man, and Cybernetics, Volume SMC-6, Number 2: 121–126.

Kilkki, P. & Päivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. In: Remote sensing aided forest inventory. Research Notes 19, University of Helsinki, Department of Forest Mensuration and Management, p. 209–215.

Korhonen, K.T. & Maltamo, M. 1990. The evaluation of forest inventory designs using correlation functions. Silva Fennica 25(2): 77–83.

— & Yli-Kojola, H. 1993. Field instructions. Inventory of Hebei. Manuscript. 15 p. Finnish Forest Research Institute.

Kujala, M. 1980. A calculation method for measuring the volume growth over bark of stemwood. Folia Forestalia 441. 8 p. (In Finnish with English summary)

Lappi, J. 1991. Sample plot size and standard error. Unpublished manuscript. Finnish Forest Research Institute.

Loetsch, F., Zöhrer, F. & Haller, K.E. 1973. Forest inventory. Volume II. BLV Verlagsgesellschaft mbH, München. 469 p.

Matérn, B. 1947. Methods of estimating the accuracy of line and sample plot surveys. Meddelanden från Statens Skogsforskningsinstitut. 36(1). (In Swedish with English summary)

— 1960. Spatial variation. Meddelanden från Statens Skogsforskningsinstitut. 49(5). 144 p. Also appeared as Lecture Notes in Statistics 36. Springer-Verlag. 1986.

Päivinen, R. 1987. A planning model for forest inventory. University of Joensuu, Publications in Sciences 11. 179 p. (In Finnish with English summary)

Poso, S. 1972. A method of combining photo and field samples in forest inventory. Communicationes Instituti Forestalis Fenniae 76(1). 133 p.

— & Kujala, M. 1978. A method for national forest inventory in Northern Finland. Communicationes Instituti Forestalis Fenniae 93(1). 54 p.

Ranneby, B. 1981. Estimation of standard errors in systematic sampling. Report 21, Swedish University of Agricultural Sciences, Section of Forest Biometry. (In Swedish with English summary)

— , Cruse, Th., Hägglund, B., Jonasson, H. & Swärd, J. 1987. Designing a new national forest survey for Sweden. Studia Forestalia Suecica 177. 29 p.

Salminen, S. 1985. The mathematic-statistical foundation of the forest inventories. Silva Fennica 19(3): 226–232. (In Finnish with English summary)

Schreuder, H.T., Gregoire, T.G., Wood, G.B. 1993. Sampling methods for multiresource forest inventory. John Wiley & Sons, Inc.

Study on European forestry information and communication systems. 1997. Reports on Forestry Inventory and Survey Systems 1.–2. European Commission, 1382 p.

Svensson, S.A., 1988. Estimation of annual stem volume increment. Swedish University of Agricultural Sciences. Department of Forest Survey. Report 46. 140 p.

Tomppo, E. 1991. Satellite image based national forest inventory of Finland. International Archives of Photogrammetry and Remote Sensing 28, Part 7–1. p. 419–424.

— 1996. Multi-source national forest inventory of Finland. 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'. Volume 1. IUFRO XX World Congress 6–12 August 1995, Tampere, Finland. EFI Proceedings No. 7. 15 p. European Forest Institute.

— , Korhonen, K.T. & Yli-Kojola, H. 1994. Inventory results of Hebei Forestry Bureau, Heilongjiang Province, China. Report. Finnish Forest Research Institute.

— , Varjo, J., Korhonen, K., Ahola, A., Ihalainen, A., Heikkinen, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S. & Tuomainen, T. 1997. Country report for Finland. In: Study on European forestry information and communication systems. Reports on Forestry Inventory and Survey Systems 1. European Commission. p. 145–226.

— , Henttonen, H., Korhonen, K.T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., Mikkelä, H., Tonteri T. & Tuomainen, T. 1998. Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968–97. Metsätieteen aikakauskirja – Folia Forestalia 2B/1998: 293–374. (In Finnish)

Total of 27 references

Appendix 1. Tree species or genera in field inventory.

1 = *Pinus koraiensis* (Korean pine)
2 = *Pinus sylvestris* (Scots pine)
3 = *Larix* sp., mainly *Larix gmelinii* (Dahurian larch)
4 = *Picea* sp. (Spruce)
5 = *Abies* sp., mainly *Abies nephrolepis* (Fir)
6 = *Pinus densiflora* (Japanese red pine)
7 = *Fraxinus mandshurica* (Manchurian ash)
8 = *Juglans mandshurica* (Manchurian walnut)
9 = *Phellodendron amurense* (Amur cork tree)
10 = *Acer mono* (Maple)
11 = *Ulmus* sp. (Elm)
12 = *Betula costata*
13 = *Betula davurica* (Dahurian birch, black birch)
14 = *Quercus* sp., mainly *Quercus mongolica* (Mongolian oak)
15 = *Tilia* sp. (Linden)
16 = *Betula platyphylla* (Manchurian birch, white birch)
17 = *Populus* sp., mainly *Populus davidiana* (Poplar)
18 = Other deciduous
19 = Other conifers

Appendix 2. Models.

Volume models for sample trees by tree species

$$v = a_0 d^{a_1} h^{a_2} + \varepsilon$$

where
v is the volume of the tree, m^3
d is the breast height diameter of the tree, cm
h is the height of the tree, m, and
ε is a $N(0,\sigma^2)$ – distributed residual.

The estimates of the parameters:

Tree species	a_0	a_1	a_2
1, 2, 6, 19	0.0000635277210	1.94354550	0.89689361
3	0.0000501687010	1.75828940	1.14966530
4,5	0.0000618598780	1.85577513	1.00705470
7–12, 15, 18	0.0000419606698	1.90945950	1.04138920
13	0.0000527864510	1.79473130	1.07126230
14	0.0000611255340	1.88100910	0.94462565
16	0.0000519351630	1.85868840	1.00389410
17	0.0000534731900	1.87789940	0.99982785

Models for volumes 5 years ago for sample trees by tree species

$$\frac{v}{d^2} = a_0 + a_1 d + a_2 d^2 + \varepsilon$$

where
v is the volume of the tree, m^3
d is the breast height diameter of the tree, cm, and
ε is a $N(0,\sigma^2)$ – distributed residual.

The estimates of the parameters:

Tree species	a_0	a_1	a_2	Maximum d(cm) for using the model
1,2,19	0.268333	0.018789	–0.00011	130
3,4,5	0.060277	0.038031	–0.00043	67
13	0.344740	0.010499	–0.00011	89
14	0.267053	0.014643	–0.00017	75
15	0.179907	0.023590	–0.00026	73
16	0.313476	0.022498	–0.00033	59
17	0.235414	0.029834	–0.00037	65
18	0.120909	0.032881	–0.00076	35
6–12	0.224845	0.020309	–0.00019	87

Volume increment models for sample trees by tree species

$$\frac{i}{100d^2} = e^{a_0 + a_1 d + a_2 \ln(d) + \varepsilon}$$

where
i is the volume increment of the tree, dm^3/5 years
d is the breast height diameter of the tree, cm, and
ε is a $N(0,\sigma^2)$ – distributed residual.

The estimates the of parameters:

Tree species	a_0	a_1	a_2
1,2,19	–7.487700	–0.033005	0.366411
3,4,5	–8.472597	–0.070551	1.076069
13	–5.135102	0.001834	–0.881438
14	–7.653503	–0.052514	0.460989
15	–7.879226	–0.064353	0.654263
16	–6.642996	–0.036436	0.023204
17	–6.745856	–0.042895	0.150808
18	–6.763920	–0.058526	0.144285
6–12	–7.443077	–0.047036	0.345476

Height models by tree species

$$\ln(h-1.3) =$$
$$a_0 + a_1 \frac{1}{d+5} + a_2 \frac{1}{(d+5)^2} + a_3 \frac{1}{(d+5)^3} + \varepsilon$$

where
d = diameter at breast height (cm)
h = estimated height (m) and
ε is a $N(0,\sigma^2)$ – distributed residual.

The estimates of the parameters:

Tree species	a_0	a_1	a_2	a_3	RMSE
1,2,19	3.3924	–8.6284	–41.34468	–	0.175236
3,4,5	3.2557	1.38862	–491.55060	2686.058929	0.252374
6–12	3.1779	–10.76361	–43.48880	–	0.266901
13	3.1112	–12.64759	–	–	0.199318
14	2.7593	–2.28497	–106.52709	–	0.258531
15	3.0081	–7.17048	–75.687691	–	0.380318
16	3.4312	–16.918131	–34.435433	–	0.215630
17	3.4250	–13.328161	–42.663523	–	0.238738
18	3.0432	–13.886421	–	–	0.268859

Diameter increment models by tree species

$$i_d = \sqrt{d^2 + i_{d^2}} - d$$

where i_{d^2} = estimate of increment of squared diameter for the future 5 year period, obtained with models:

Species 1, 2 and 19:
$$i_{d^2} = -429.346 + 51.7149\,d - 0.05002\,d^2 - 25.503845\,S_1\,d + 0.04213\,S_2\,d^2$$

Species 3,4 and 5:
$$i_{d^2} = 245.822337 + 55.167319\,d + 2494.384469\,S_1 + 2947.557241\,S_2 - 41.254773\,S_1\,d - 31.927254\,S_2\,d$$

Species 6–12:
$$i_{d^2} = -853.333592 + 37.880276\,d - 0.03389\,d^2$$

Species 13:
$$i_{d^2} = 552.272276 + 20.555934\,d - 11.075027\,S_1\,d$$

Species 14:
$$i_{d^2} = -139.094311 + 31.324879\,d - 0.006915\,d^2 - 1362.355982\,S_3 + 9.441992\,S_3\,d - 0.026362\,S_3\,d^2$$

Species 15:
$$i_{d^2} = -147.11626 + 33.963102\,d - 0.02428\,d$$

Species 16:
$$i_{d^2} = 299.68921 + 24.696109\,d$$

Species 17:
$$i_{d^2} = -5454.177778 + 44.374591\,d + 6773.346516\,S_{12} + 6915.441639\,S_{34} - 24.687306\,S_{12}\,d - 27.676095\,S_{34}\,d$$

Species 18:
$$i_{d^2} = 1394.393656 + 19.22915\,d$$

d = diameter over bark at breast height (mm)

S_1 = 1, if site class is good
0, otherwise
S_2 = 1, if site class is medium
0, otherwise
S_3 = 1, if site class is poor
0, otherwise
S_4 = 1, if site class is very poor
0, otherwise
S_{12} = 1, if site class is good or medium
0, otherwise
S_{34} = 1, if site class is poor or very poor
0, otherwise