

This is an author produced version of a paper published in *Computers and Electronics in Agriculture*. This paper has been peer-reviewed. It does not include the journal pagination.

Citation for the published paper:

Reese, H., Nilsson, M., Sandström, P. and Olsson, H. (2002)
Applications using estimates of forest parameters derived from satellite
and forest inventory data. *Computers and Electronics in Agriculture*.
37:1, 37-55. ISSN 0168-1699.

[http://dx.doi.org/10.1016/S0168-1699\(02\)00118-7](http://dx.doi.org/10.1016/S0168-1699(02)00118-7)

Access to the published version may require journal subscription.
Published with permission from: Elsevier Science.



Epsilon Open Archive <http://epsilon.slu.se>

Applications using estimates of forest parameters derived from satellite and forest inventory data

Heather Reese¹, Mats Nilsson, Per Sandström and Håkan Olsson
Remote Sensing Laboratory
Department of Forest Resource Management and Geomatics
Swedish University of Agricultural Sciences
901-83 Umeå, Sweden

¹ Correspondence to: Heather.Reese@resgeom.slu.se

Abstract

From the combination of optical satellite data, digital map data, and forest inventory plot data, continuous estimates have been made for several forest parameters (wood volume, age and biomass). Five different project areas within Sweden are presented which have utilized these estimates for a range of applications. The method for estimating the forest parameters was a "k-Nearest Neighbor" algorithm, which used a weighted mean value of k spectrally similar reference plots. Reference data were obtained from the Swedish National Forest Inventory. The output was continuous estimates at the pixel level for each of the variables estimated. Validation results show that accuracy of the estimates for all parameters was low at the pixel level (e.g., for total wood volume RMSE ranged from 58-80%), with a tendency toward the mean, and an underestimation of higher values while overestimating lower values. However, when the accuracy of the estimates is assessed over larger areas, the errors are lower, with best results being 10% RMSE over a 100 ha aggregation, and 17% RMSE over a 19 ha aggregation. Applications presented in this paper include moose and bird habitat studies, county level planning activities, use as input information to prognostic programs, and computation of statistics on timber volume within drainage basins and smaller land holdings. This paper provides a background on the k NN method and gives examples of how end users are currently applying satellite-produced estimation data such as these.

Key words: applications, Landsat TM, SPOT, forest inventory, continuous estimates, kNN

1 Introduction

Approximately 65% of Sweden's land area is forested (National Atlas of Sweden, 1990). Satellite data, due to their synoptic and repeated coverage, can be useful in providing information with which to map these large areas of forest (Iverson et al. 1989; Leckie 1990; Innes and Koch 1998). At the Swedish University of Agricultural Sciences (SLU), one method being used to interpret satellite data into information about the forest has been the k-Nearest Neighbor (kNN) algorithm. The output from kNN can be continuous estimates of specified forest parameters, such as total wood volume, wood volume by tree species, stand age, and above-ground tree biomass. These continuous data can be categorized and combined in various ways, suiting different uses of the data. The estimated forest parameter data are currently being used in forestry, planning, and ecologically related studies.

The kNN algorithm has been described in earlier studies (Tomppo 1990; Tomppo 1993) and is also referred to as the Reference Sample Plot method (Kilkki and Päivinen 1987; Muinonen and Tokola 1990; Tokola et al. 1996). Finland has used satellite data and the kNN algorithm to generate forest parameter estimates for operational use in their forestry program since 1990. Trotter et al. (1997) applied kNN in New Zealand's coniferous forest areas.

Variants of kNN are currently being used in the U.S. for forest estimations in the Oregon University/U.S. Forest Service's "Coastal Landscape Analysis and Modeling Study" (CLAMS) and the NASA-funded Upper Midwest Regional Earth Science Application Center (RESAC) (A. Ek, pers. comm.). Several studies have also applied this method to Swedish forest conditions (Nilsson 1997; Fazakas et al. 1999; Holmgren et al. 2000; Nilsson and Sandström 2002). In several of these studies, it had been found that pixel level accuracies of the results were low, but when accuracy was assessed over larger areas, more acceptable accuracy was reached.

Many studies have investigated the relationship between wood volume and satellite spectral reflectance. SPOT and Landsat data have a primarily negative correlation with wood volume (Trotter et al. 1997), except for the near-infrared band, in which studies show various correlations, (see Spanner et al. 1990; Franklin 1986; Ripple et al. 1991; Danson and Curran 1993). The correlation between spectral data and wood volume tends to be stronger for younger stands than older stands (Franklin 1986; Horler and Ahern 1986; Peterson and Nilson 1993). Studies have investigated the influence of shadow on the correlation between spectral reflectance and forest properties, and varied in their findings on which bands were most important and why (see Horler and Ahern 1986; Ripple et al. 1991; Ardö 1992; Cohen and Spies 1992; Trotter et al. 1997). Brockhaus and Khorram (1992) found the mid-infrared bands important for determining wood volume, age class and basal area due to their correlation to moisture and forest density.

Other studies examining the relationship between volume and reflectance data at the pixel level show a weak relationship (low r^2 values) between wood volume and spectral reflectance (Danson 1987), yet a high significance (Franklin 1986; Trotter et al. 1997). This can result in wood volume estimates that can have low accuracy at the pixel level (Tomppo 1993; Tokola et al. 1996; Nilsson 1997) yet are more accurate for larger areas such as stands or a landscape level (Franklin 1986; Poso et al. 1987; Muinonen and Tokola 1990; Hagner 1990; Ahern et al. 1991). It is also of significance that once wood volumes increase and the canopy closes,

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

the relationship between spectral reflectance and wood volume is asymptotic (Franklin 1986; Spanner et al. 1990; Danson and Curran 1993; Ekstrand 1994), and relationships for these higher volumes become less accurate.

In this paper, we describe a number of projects carried out at SLU that have used the *k*NN algorithm to estimate different forest parameters. Descriptions of the data used for these studies are given, which include Landsat Thematic Mapper (TM) and Satellite Probatoire d'Observation de la Terre (SPOT) satellite data, and field data from the Swedish National Forest Inventory (NFI). The *k*NN algorithm is described, as is an explanation of the accuracy of the forest parameter estimates. We then illustrate uses of the forest parameter estimation data by describing several applications that include planning activities, descriptive statistics, and habitat suitability mapping.

2 Study areas

Five study areas are described in this paper (Figure 1). All test areas are dominated by managed boreal forest with Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* L. Karst), Hairy or Downy birch (*Betula pubescens* Ehrh.), and Silver or Warty birch (*B. pendula* Roth.). The first area, near the town of Älvsbyn, lies furthest north (approx. 65° 38' N, 21° 00' E). The elevation ranges from 0 to 530m above sea level, and the soil is primarily till. The second area, the coastal part of Västerbotten county (approx. 64° 30' N, 18° 30' E), has elevation ranging from 0 to approximately 600 m above sea level, and till soil predominates with some small areas of sand and glaciofluvial soils. The other study areas, further to the south, are Dalarna county (approx. 61° 00' N, 14° 30' E), Gävleborg county (approx. 61° 30' N, 16° 00' E), and an area referred to as Uppland (approx. 60° 00' N, 17° 00' E). Dalarna county has elevations ranging from 0 to approximately 1200 m above sea level in the western mountain region, with soils predominated by till, and pockets of sand and glaciofluvial soils, peat and fine-grained sediments. Gävleborg county has similar soils and elevations ranging from 0 to approximately 700 m above sea level. Uppland's elevation ranges from 0 to approximately 100 m above sea level, and has a mixture of till, peat, fine-grained sediments and clayey till, and sand and glaciofluvial soils.

3 Data

Three different sources of data were used to make forest parameter estimates in each of the projects: satellite data, digital map data, and forest inventory plots. Single date optical satellite data were processed individually (i.e., unmosaicked). The Älvsbyn study consisted of two adjacent SPOT-3 scenes; for Västerbotten one Landsat-5 TM scene was used; in Dalarna two Landsat-5 TM scenes were processed; Gävleborg used three Landsat-5 TM scenes; and, in the Uppsala area a quarter Landsat-5 TM scene was used. The data were geometrically precision-corrected to the Swedish National Grid (RT90) to within ½ pixel error, and resampled to 25m x 25m pixels for TM and 20m x 20m for SPOT, using cubic convolution (done by Satellus Corporation, Kiruna, Sweden). TM bands 1-5 and 7, and SPOT XS1-3 were used in the estimations. The image data were not radiometrically nor atmospherically corrected due to time constraints; all images were nearly cloud-free (0 to 10% cloud cover).

Digital map data from either 1:50,000 scale topographic maps or 1:100,000 scale vegetation maps were used to create a forest mask, as only "productive forestland" areas were estimated. Productive forestland was defined as land that was suitable for wood production with an

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

annual mean production potential of at least 1 m³/ha (Anon. 2000). Non-forestland areas were excluded from the analysis.

Field data were used to construct relationships between the measured forest parameters and the satellite data. The field data were taken from the Swedish NFI that is an annual national field sample of forest parameters. It is based on a systematic grid across the country consisting of “tracts” or squares on which circular sample plots are located (Ranneby et al. 1987). The tracts vary in size according to the defined region of Sweden to which they belong (Anon. 2000). Both permanent and temporary plots exist, the difference being in size of both tract and plot, and that permanent plots are re-measured periodically. Permanent plots were 10 meters in radius, and located on tracts that varied in side length from 300 to 1200 meters, with eight plots to the tract. Temporary plots were seven meters in radius, and were located on tracts that varied in side length from 300 to 1800 meters, with 12 plots to the tract (Hägglund 1983). The plot coordinates in the NFI database were derived from one of three sources: nominal coordinates, digitized locations recorded by NFI field crews on 1:10,000 scale orthophoto field maps, or Global Positioning System (GPS) coordinates (introduced to the NFI in 1996). Positional accuracy for the NFI plots has been estimated on a nation-wide level, giving average errors for nominal coordinates of 67 meters Root Mean Square Error (RMSE), while digitised locations had 33.5 meters RMSE (Hagner et al. 2002). GPS coordinates for all permanent plots were differentially corrected, giving a positional accuracy from 5 (Nilsson and Sandström 2002) to 6.5 meters RMSE (Hagner et al. 2002).

4 Methods

The *k*NN algorithm, used in this study to calculate the estimates of selected forest parameters was implemented using an in-house C program. The algorithm required the input of forest inventory plots and digital satellite data.

4.1 Forest inventory dataset preparation

For each individual satellite scene to be estimated, a subset of NFI plots was extracted. Plots were chosen over the scene’s full geographic extent from a minimum of four years and, in this paper, up to a nine-year time span, in order to give a large number of reference plots. A bigger time span was especially necessary for the project involving SPOT data, as the SPOT scene covered a smaller geographic area (60 km x 60 km) than a TM scene (185 km x 185 km). Holmström et al. (2002) has noted that use of reference plots from up to a 25-year time span (if plots with thinnings and clear-cuttings were excluded from the reference material) produced acceptable results. Variables such as volume, biomass, age, ground moisture, site quality and field layer were extracted for the plot data subset. A growth function (standard deviation calculated to be about 10 per cent) projecting 5-year growth (Söderberg 1986) was applied to stem volume measurements to reconcile the differences between the image acquisition date and the date the NFI plots were surveyed. Each of the plot data subsets also included the satellite scene’s spectral values for each plot location. This was done by geographic overlay of the plot dataset coordinates on the image data and extraction of each image band’s spectral values for that location using nearest neighbor. Figure 2 shows the extent of a satellite scene and the location of the NFI plots used in the estimation.

Because many of the NFI plots were measured previous to the date of image acquisition, there was a likelihood that changes in the vegetation had occurred (exclusive to normal growth), and therefore some data points should not be used in the estimation. Errors in the geographic registration between the NFI plots and the image may also exist. In an effort to

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

eliminate potentially erroneous data from being used in the estimation, outliers were analysed using regression. The regression model included several variables from the NFI database in order to explain the digital numbers from each band. In most cases, approximately 15 to 20% of the plots with the highest residuals were subsequently removed from the plot dataset. In examining field data used for the estimations, approximately 15% of the plots can have been affected by clearcutting, misregistration between plot location and satellite data, or atmospheric conditions. After outlier removal, datasets to be used in the estimations contained an average of 400 NFI plots for a SPOT scene and 1,250 NFI plots for a Landsat TM scene, ranging from a four to nine years time span. Table 2 presents some summary statistics describing the NFI data used for estimation and the resulting estimations for three satellite scenes.

4.2 The k NN algorithm

Using the NFI plot dataset with outliers removed, parameters such as total wood volume, wood volume for the different tree species, biomass and age were calculated for every pixel within the forest mask. The estimated forest parameter value (\hat{v}) for each pixel (p) was calculated as a weighted mean value of the reference plots' (j) forest parameter (v_j) of the k nearest samples in spectral space (Equation 1). Weights (w) assigned to each of the k samples were proportional to the inverse squared Euclidean distance (d) between the pixel to be estimated and the reference plot (Equation 2).

$$\hat{v}_p = \sum_{j=1}^k w_{j,p} * v_{j,p}$$

where (1)

$$w_{j,p} = \frac{1}{d_{j,p}^2} / \sum_{i=1}^k \frac{1}{d_{i,p}^2}$$

$$d_{1,p} \leq d_{2,p} \leq \dots \leq d_{k,p}$$

(2)

and

A number of studies have proposed different numbers for the optimal number of nearest neighbors (k) (Mouer and Stage 1995; Tokola et al. 1996; Nilsson 1997; Trotter et al. 1997; Nilsson and Sandström 2002). Hagner et al. (2002) contended that an optimal number for k was dependent on the data and goals of the estimation. Nilsson and Sandström compared the use of either $k=5$ or $k=10$, and found $k=10$ gave slightly higher accuracy at the pixel level than $k=5$, and both $k=5$ or 10 produced the same accuracy when assessed over aggregated areas. In most of the projects described in this paper, $k=5$ was used. The reason for using $k=5$ was due to the finding that five nearest neighbors produces nearly equal accuracy results, while not over-generalizing the data. The higher the number k , the more the pixel-level results will average towards the mean (Nilsson 1997).

5 Results

The accuracy assessments from each of the projects described in this paper gave similar and consistent results, showing accuracy of estimates on a pixel level to be poor, yet improved when assessed over aggregated areas. For every parameter estimated on the pixel level, there was a tendency toward the mean value of the parameter with an average low bias, with

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

underestimation of higher values and overestimation of lower values. Measures of RMSE and bias were calculated according to Equations 3 and 4, respectively.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

$$\text{Bias} = \hat{y}_i - \bar{y}_i \quad (4)$$

where \hat{y} was the estimated parameter, and y was the field-measured parameter.

The pixel level results for each of the projects are given in Table 3. Some of the projects used different types of evaluation data. Dalarna county's accuracy was assessed at the pixel level using a total of 400 GPS-located NFI plots (Reese and Nilsson 1999); Älvsbyn and Gävleborg used cross-validation to assess accuracy (Reese and Nilsson 2000); Västerbotten used cross-validation for the whole scene, and in a subset of the study area, used 627 field-measured plots; and, the Uppland study used cross-validation for the whole scene, and 453 field-measured plots for a subset of the study area (Fazakas et al. 1999). An example of the distribution of pixel-level errors by volume class from the Dalarna data is given in Table 4. In Gävleborg county, proportions of pine, spruce and broad-leaved trees were assessed at the pixel-level using cross-validation, with the results being a relative RMSE of 73%, 88%, and 163%, respectively.

Estimates for total volume have also been assessed over larger spatial areas. To assess accuracy over these areas, aggregation of the estimates was not done physically, but instead took the mean value of the estimates within the aggregated area and compared this to the mean value in the NFI data for the same area. Therefore, the RMSE was a reflection of how well the estimates represented the forest parameters over a certain area. This was done because the original intended use of *k*NN estimations was to estimate timber volumes over larger areas. In Västerbotten, total wood volume estimates were assessed over 14.9 ha areas, giving an overall RMSE of 38.9%, and at a 350 ha level assessment, overall RMSE was 12% (Nilsson and Sandström 2002). When the estimates for the Uppland study area were assessed over 50 ha areas, the RMSE was 20%, and at 100 ha areas, total wood volume had approximately 10% RMSE while biomass was slightly higher (Fazakas et al. 1999). Figure 3 shows the effect of assessing accuracy of the total wood volume estimates over varying size areas.

6 Discussion

The results from the five study areas were quite similar. Pixel level estimates had a high RMSE and a low bias. The errors tended to be more accurate toward the mean value, with underestimations of higher values and overestimations of lower values. Assessment of larger areas of the data showed a decrease in error, ranging from 39 to 17% at a stand level and down to 10% for a 100 ha assessment. As a comparison to traditional methods, the accuracy of 1:30,000 scale photo-interpreted wood volume was assessed as having 13.5% RMSE on a stand level (Ståhl 1992).

The cross-validation results for the species proportions from the Gävleborg project show that improvement of the estimates for broad-leaved species volume is needed. The broad-leaved proportion RMSE was quite large (163.3%), and this could be attributed to several reasons. One reason could have been that the mean value in the NFI data for broad-leaved volume was

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

only 18.5 m³/ha. Also, there were far fewer NFI reference plots with broad-leaved forest used for the estimation. This may have contributed to a lower accuracy for these less represented classes. In a Finnish study, Tokola et al. (1996) assessed the accuracy of volume by species based on estimates of wood volume made with the Reference Sample Plot method. Using TM data in one area, they found the best estimates for broad-leaved species to have a standard error of 157.5% (Tokola et al., p. 234). As with the Gävleborg project, the errors in estimating broad-leaved volume were notably high.

Other ways of improving the estimates have been tried at SLU. In Älvsbyn, an illumination correction using a cosine with c-correction (Teillet et al. 1982) was necessary due to a low sun angle in one of the images (25°, see Table 1). The correction improved the results by 9 m³/ha, or 7% of the NFI mean value of 115m³/ha (Reese and Nilsson 2000). In Dalarna, Gävleborg and Västerbotten, a geographic distance limitation, which limited the geographic distance between the pixel to be estimated and the location of the reference plots, improved the results (Nilsson and Sandström 2002). Holmgren et al. (2000) showed that estimation accuracy could be improved for total wood volume at the stand level when ancillary data such as mean tree height were incorporated (from 36% RMSE to 17% RMSE at 19 ha area assessment). Holmström and Fransson (2002) combined SPOT-4 and CARABAS-SAR data, and using the *k*NN algorithm they derived an RMSE of 22% for total wood volume at the stand level. Work is also being done in which artificial neural networks and atmospheric correction will likely improve *k*NN estimate results (Hagner et al. 2002).

7 Applications

The *k*NN estimation data represent large area databases that could be used for multiple applications. Having continuous data, as opposed to thematic, allows the user to decide how to divide the data into various classes. The users have been informed about the error associated with pixel level estimates, and therefore have been using the *k*NN data to look at stand level characteristics. Specific examples are given below of how the estimates are being used in both forest planning and ecological applications.

7.1 Planning applications

Agencies such as county administrative boards or county forestry boards needed more spatially detailed data about the forests over large areas, such as the data produced from the *k*NN estimations. In Dalarna and Gävleborg, the county boards are currently combining the forest parameter estimation data (total wood volume, wood volume by species, and age) with other data layers in a GIS. Their aim is to identify areas within their counties that might be purchased and added to nature reserve areas for protection of biological diversity. Feedback from the users indicated that the data have pointed them towards ecologically important areas that they might not otherwise have investigated without the use of these data. However, they were disappointed with the discrimination of broad-leaved proportions (Olle Kellner, pers. comm.). In Älvsbyn, the forestry board uses the estimation data to compare to and supplement other existing forest information databases, giving them a better idea of the distribution of their forest.

Investigation of the quantity of forest bio-fuels was conducted using estimation data from both Älvsbyn and Västerbotten. Using the *k*NN algorithm, with *k*=1, the weights were used as input to a non-spatial prognostic program called HUGIN. This program was used to predict the consequences of management practices on the forest, given existing legal restrictions and

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

regulations. The available standing biomass was estimated, and projected 30 and 50 years into the future using HUGIN (Bååth et al. 2002; Hallsby et al. 2000).

7.2 Habitat studies

The k NN data from Västerbotten were used in a study of potential relationships between forest parameter estimates and moose (*Alces alces*) forage habitat (Sandström and Edenius 1999). Estimates of species composition and wood volume were used in conjunction with data from GPS-collared moose and daily homeranges (adaptive kernel method). GPS locations were available from four “stationary” (winter and summer range the same) and five “migratory” (winter range geographically separate from summer range) moose. Analysis was carried out on 114 daily homeranges. Habitat composition within daily homeranges was evaluated against k NN-derived estimates of tree species composition to assess whether there were differences in use between different categories of animals (stationary vs. migratory and moose cows with vs. without calves). Initial analysis showed that all the moose preferred areas of younger forests and relatively low wood volume. Further work will include estimating seasonal resource selection probability functions for moose based habitat information from daily homeranges and thereby producing Habitat Suitability Maps (HSM) for a larger area.

Study of available bird habitat was carried out using k NN estimates from Dalarna. Because species composition and vertical structure of the forest can be important for bird habitat, information about tree species composition, standing wood volume, basal area and age were analysed. The model-species used was the long-tailed tit (*Aegithalos caudatus*). Thematic maps were produced for the occurrence and distribution of the key variables in the landscape and a habitat suitability map was created. The long-tailed tit prefers habitat that is predominantly mature broad-leaved trees, in stands greater than 2 ha in size. The analysis revealed only one such suitable patch of habitat in the study area, and this patch was already a protected area (Nilsson and Löfstrand 1999).

Another project is investigating the use of k NN estimates for mapping reindeer grazing habitat. It is a cooperative project between the Västerbotten County Board, Västerbotten County Forestry Board, Norwegian Institute for Nature Research (NINA), SLU and two organizations of Sami reindeer herders in Västerbotten. One Landsat TM and three SPOT-4 scenes are being used to create k NN estimates of forest parameters such as wood volume, age, crown closure and species proportion, as well as different types of vegetation important to reindeer grazing (e.g., lichen-rich coniferous forest). Fieldwork has been completed for the year 2000, with more scheduled through 2002, with resulting field data to be used as groundtruth. With the k NN estimates, the co-operators hope to create new maps of potential reindeer grazing areas (Sandström and Pahlén 2000).

7.3 Regional and local statistics

With regional databases such as the k NN estimates, statistics for different forest parameters can be calculated at varying scales. Two examples have been carried out to demonstrate possibilities for using k NN estimate data in compiling statistics about the forest, both on a watershed basis and also for approximations of statistics about forest estate holdings. This can be useful, as the current methods in Sweden for calculating information about the potential standing wood volume of forest are made using much less spatially dense data.

Volume estimates for an 116,000 ha drainage basin of the Sävar river were extracted from the Västerbotten estimation data (Figures 4 and 5). It was found that 74% of the land area was

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

forested and within the forested area, total wood volume for the watershed was calculated to be 7,964,300 m³. Distribution by area of volume classes showed that the largest area was covered by 0-25 m³/ha pixels (>20,000 ha), with a median of 100-124 m³/ha (13,000 ha).

Statistics on a smaller scale may also be useful. To demonstrate this, a 106 ha forest estate (the Sävars-Brattby estate, shown in the inset to Figure 4) within the Sävar watershed was randomly selected to compare the results from the *k*NN estimations to statistics compiled from a subjectively measured forest management plan from 1994 (error approximately 10%). The total wood volumes compared quite well, with a total of 10,425 m³ from the forest management plan, and 11,080 m³ from the *k*NN estimates. The proportion of coniferous to broad-leaved trees compared well, with 86% coniferous and 14% broad-leaved from the *k*NN estimations, and 88% coniferous and 12% broad-leaved from the forest management plan. The proportions of the coniferous species (pine to spruce) was 50% pine, 36% spruce for the *k*NN estimates, and 61% pine, 27% spruce for the management plan (Reese, Nilsson and Olsson 1999). Due to the errors of the *k*NN estimates on small areas, the size of the forest estate should be kept in mind while using the *k*NN data.

8 Conclusion

Using optical satellite data, extensive field data from the Swedish National Forest Inventory, and digital map data used as a forest mask, the *k*-Nearest Neighbor algorithm has been used to produce continuous estimates of various forest parameters over large areas. Forest variables such as total wood volume, wood volume by species, age and biomass have been estimated. These continuous data can be combined in a number of ways, such as mapping species proportion, or identifying young pine areas. Users of the data can tailor the classes of the estimated forest parameters to fit their own study needs, making them useful both for ecological and forestry purposes. The examples given in this paper, such as use in planning, in prognostic programs, creating habitat suitability maps, and descriptive statistics at regional and local scales, shows the variety of present and potential uses of the estimation data. However, users must be well informed of the accuracy level of these data when applying the data; pixel level accuracy has been shown to be quite poor on average (ranging from 58-80% RMSE for total wood volume). Data have been therefore been assessed at stand level, and accuracies have been in a range of 39% to 17% RMSE, with the lowest error being 10% at a 100 ha area assessment. In consideration of this, the estimates should be used at an aggregated level, and not at the pixel level. Errors in estimation of broad-leaved volume are higher than for the coniferous species, and should be a consideration in their use. Future efforts will aim at improving these estimates. As users begin to apply these data, it is increasingly important for the data producers to communicate the accuracy of the data, and therefore the potential uses. The studies given in this paper are examples of how the data can be used on the more aggregated level.

Acknowledgements

This paper stemmed from a conference poster made for the Second International Conference on Geospatial Information in Agriculture and Forestry, Lake Buena Vista, Florida, 10-12 January 2000. The reviewers are thanked for taking time in making helpful comments that improved this paper. The work has been funded by the following sources: the Swedish research program Remote Sensing for the Environment (RESE), financed by the Swedish Foundation for Strategic Environmental Research (MISTRA); the Swedish National Space Board; the County Administration Boards of both Dalarna and Gävleborg; and the County

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Forestry Board of Norrbotten. Ronny Lofstrand and Tina Pahlén are also acknowledged for their work on the habitat studies.

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

References

- Ahern, F. J., Erdle, T., Maclean, D.A., and Kneppeck, I. D. 1991. A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements. *International Journal of Remote Sensing*, 12:387-400.
- Anonymous. 2000. Instruktion för Fältarbetet vid Risksskogstaxeringen. Swedish University of Agricultural Sciences, Department of forest resource management and geomatics, Umeå, Sweden.
- Ardö, J. 1992. Volume quantification of coniferous forest compartments using spectral radiance recorded by Landsat Thematic Mapper. *International Journal of Remote Sensing*, 13:1779-1786.
- Brockhaus, J.A. and Khorram, S. 1992. A comparison of SPOT and Landsat-TM data for use in conducting inventories of forest resources. *International Journal of Remote Sensing*, 13:3035-3043.
- Bååth, H., Gällersprång, A., Hallsby, G., Lundström, A., Löfgren, P., Nilsson, M., and Ståhl, G. 2002. Remote sensing, field survey, and long term forecasting: an efficient combination for local assessments of forest fuels. *Biomass and Bioenergy*, 22(3):145-157.
- Cohen, W.B., and Spies, T.A. 1992. Estimating Structural attributes of Douglas-Fir/Western Hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, 41:1-17.
- Danson, F.M. 1987. Preliminary evaluation of the relationships between SPOT-1 HRV data and forest stand parameters. *International Journal of Remote Sensing*, 8:1571-1575.
- Danson, F. M. and Curran, P.J. 1993. Factors affecting the remotely sensed response of coniferous forest plantations. *International Journal of Remote Sensing*, 43:55-65.
- Ekstrand, S. 1994. Assessment of Forest Damage with Landsat TM: Correction for varying forest stand characteristics. *Remote Sensing of Environment*, 47:291-302.
- Fazakas, Z., Nilsson, M., and Olsson, H. 1999. Regional forest biomass and wood volume estimation using satellite data and ancillary data. *Agricultural and Forest Meteorology* 98-99: 417-425. Sp. Iss. SI Dec. 31, 1999.
- Franklin, J. 1986. Thematic Mapper analysis of coniferous forest structure and composition. *International Journal of Remote Sensing*, 7:1287-1301.
- Hägglund, B. 1983. En ny svensk riksskogstaxering. Swedish University of Agricultural Sciences, Department of Forest Survey. 74pp.
- Hagner, O. 1990. Computer aided forest stand delineation and inventory based on satellite remote sensing. In: SNS/IUFRO workshop, The Usability of Remote Sensing for Forest Inventory and Planning, Umeå, Sweden. p.94-105.
- Hagner, O., Nilsson, M., Joyce, S., and Olsson, H. 2002. K-nearest neighbour methods and artificial neural networks for estimation of forest parameters from satellite imagery. Manuscript.
- Hallsby, G., Bååth, H., Gällersprång, A., Löfgren, P., Lundström, A., Nilsson, M., and Ståhl, G.,. 2000. Metodik för skattning av lokala skogsbränsleresurser. Working Paper 65:2000, Swedish University of Agricultural Sciences. 12p. + Appendix
- Holmgren, J., Joyce, S., Nilsson, M., and Olsson, H. 2000. Estimating stem volume and basal area in forest compartments by combining satellite image data with field data. *Scandinavian Journal of Forest Research* 15:103-111.
- Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.
- Journal Home Page:
http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

- Holmström, H., Nilsson, M., and Ståhl, G. 2002. Forecasted reference sample plot data in estimations of stem volume using satellite spectral data and the *k*NN method. *International Journal of Remote Sensing* 23(9):1757-1774
- Holmström, H. and Fransson, J.E.S. Combining remotely sensed optical and radar data in *k*NN estimation of forest parameters. *Forest Science*, in press.
- Horler, D.N.H, and Ahern, F.J. 1986. Forestry information content of Thematic Mapper data. *International Journal of Remote Sensing*, 7:405-428.
- Innes, J.L. and Koch, B. 1998. Forest biodiversity and its assessment by remote sensing. *Global Ecology and Biogeography Letters* 7:397-419.
- Iverson, L.R., Graham, R.L. and Cook, E.A. 1989. Applications of satellite remote sensing to forested ecosystems. *Landscape Ecology* Vol. 3, 2:131-143.
- Killki, P. and Päivinen, R. 1987. Reference sample plots to combine field measurements and satellite data in forest inventory. *In: Remote Sensing-Aided Forest Inventory. Seminars organised by SNS and Taksaattoriklubi, Hyytiälä, Finland*, pp. 209-212.
- Leckie, D.G. 1990. Advances in remote sensing technologies for forest surveys and management. *Canadian Journal of Remote Sensing*, 20:464-483.
- Mouer, M., and Stage, A.R. 1995. Most similar neighbor – an improved sampling inference procedure for natural resource planning. *Forest Science* Vol. 41, Iss. 2, pp. 337–359.
- Muinonen, E. and Tokola, T. 1990. An application of remote sensing for communal forest inventory. *In: The Usability of Remote Sensing for Forest Inventory and Planning. SNS/IUFRO workshop, Department of forest resource management and geomatics, Swedish University of Agricultural Sciences, Umeå, Sweden*, pp. 35-42.
- National Atlas of Sweden. 1990. National Atlas of Sweden: The Forests. Nilsson, N-E. (ed.). SNA, Italy.
- Nilsson, M. 1997. Estimation of forest variables using satellite image data and airborne Lidar. *Acta Universitatis Agriculturae Suecia. Silvestria* 17. Swedish University of Agricultural Sciences. Department of forest resource management and geomatics, Umeå, Sweden.
- Nilsson, M., and Löfstrand, R. 1999. Creating and using forest parameter estimates using the *k*NN algorithm. Abstract from the November 11-13, 1999 RESE Annual meeting, Jukkasjärvi, Sweden.
- Nilsson, M. and Sandström, P. 2002. Evaluation of a non-parametric method for estimation of wood volume and biomass in boreal forests by combining remote sensing and field data. Accepted for publication in *International Journal of Remote Sensing*.
- Peterson, U. and Nilson, T. 1993. Successional reflectance trajectories in northern temperate forests. *International Journal of Remote Sensing*, 14:609-613.
- Poso, S., Paananen, R. and Simila, M. 1987. Forest inventory by compartments using satellite imagery. *Silva Fennica*, 21, pp. 69-94.
- Ranneby, B., Cruse, T., Hägglund, B., Jonasson, H., and Swärd, J. 1987. Designing a new national forest survey for Sweden. *Studia Forestalia Suecica*, No. 177. Swedish University of Agricultural Sciences, Faculty of Forestry, Uppsala, Sweden.
- Reese, H. and Nilsson, M. 1999. Using Landsat TM and NFI data to estimate wood volume, tree biomass and stand age in Dalarna. Working Paper 53:1999, Swedish University of Agricultural Sciences. 19p. + Appendices.

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

- Reese, H. and Nilsson, M. 2000. Wood volume estimations for Älvsbyn Kommun using SPOT satellite data and NFI plots. Working Paper, Swedish University of Agricultural Sciences. 20p. + Appendices.
- Reese, H., Nilsson M., and Olsson, H. 1999. Applications of *k*NN estimation data. Abstract from the November 11-13, 1999 RESE Annual meeting, Jukkasjärvi, Sweden.
- Ripple, W.J., Wang, S., Isaacson, D.L., and Paine, D.P. 1991. A preliminary comparison of Landsat Thematic Mapper and SPOT-1 HRV multispectral data for estimating coniferous forest volume. *International Journal of Remote Sensing*, 12:1971-1977.
- Sandström, P. and Edenius, L. 1999. Estimation of occurrence and distribution of different habitat parameters from satellite images, field observations, and ancillary data. Abstract from the November 11-13, 1999 RESE Annual meeting, Jukkasjärvi, Sweden.
- Sandström, P. and Pahlén, T. 2000 Remote sensing: A useful tool for reindeer management. Abstract from the November 15-17, 2000 RESE Annual meeting. Hasselludden, Sweden.
- Spanner, M.A., Pierce, L.L., Peterson, D.L., and Running, S.W. 1990. Remote sensing of temperate coniferous forest leaf area index: The influence of canopy closure, understorey vegetation and background reflectance. *International Journal of Remote Sensing*, 11:95-111.
- Söderberg, U. 1986. Functions for forecasting of timber yields – Increment and form height for individual trees of native species in Sweden. (In Swedish with English summary pages 126 – 135) Report No. 14. Swedish University of Agricultural Sciences. Department of Forest Mensuration and Management.
- Teillet, P.M., Guindon, B., and Goodenough, D.G. 1982. On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing* 8:84-106.
- Tokola, T. Pitkänen, J., Partinen, S., and Muinonen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing*, 17:2333-2351.
- Tomppo, E. 1990. Designing a satellite image-aided national forest survey in Finland. *In: The Usability of Remote Sensing for Forest Inventory and Planning*. SNS/IUFRO workshop, Department of forest resource management and geomatics, Swedish University of Agricultural Sciences, Umeå, Sweden, pp. 43-47.
- Tomppo, E. 1993. Multi-source National Forest Inventory of Finland. *In: Proceedings of Ilvessalo Symposium on National Forest Inventories*, August 17-21, Finland. pp. 52-59.
- Trotter, C.M., Dymond, J.R., and Goulding, C.J. 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. *International Journal of Remote Sensing*, 18 (10): 2209-2223.

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Figure Captions

Figure 1. Project area locations.

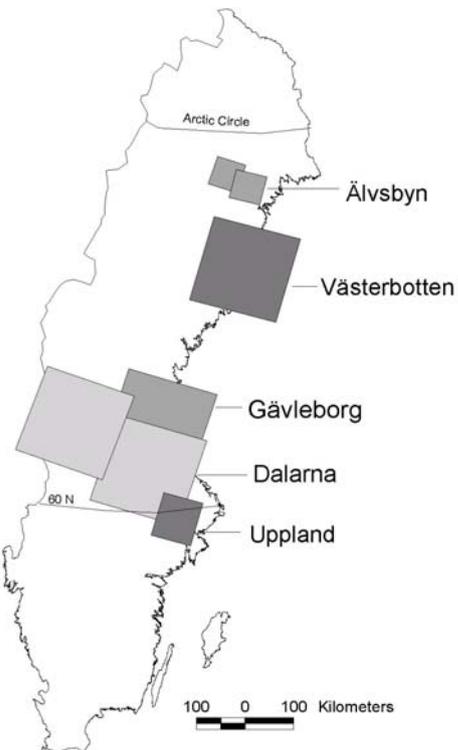
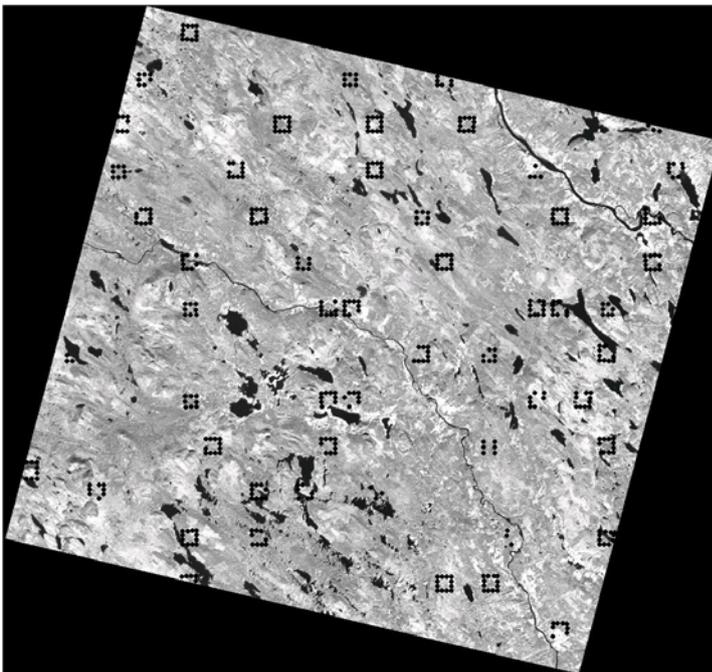


Figure 2. The extent of a SPOT satellite scene with NFI plots used for estimation.



Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Figure 3. The effect of assessing accuracy over larger areas on wood volume estimates (Uppland from Fazakas et al. 1999; Västerbotten from Nilsson and Sandström 2002).

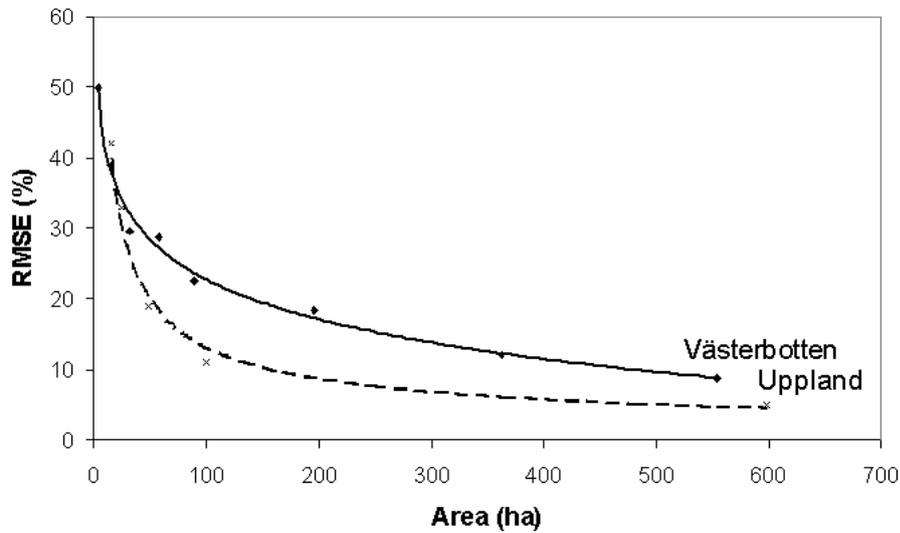
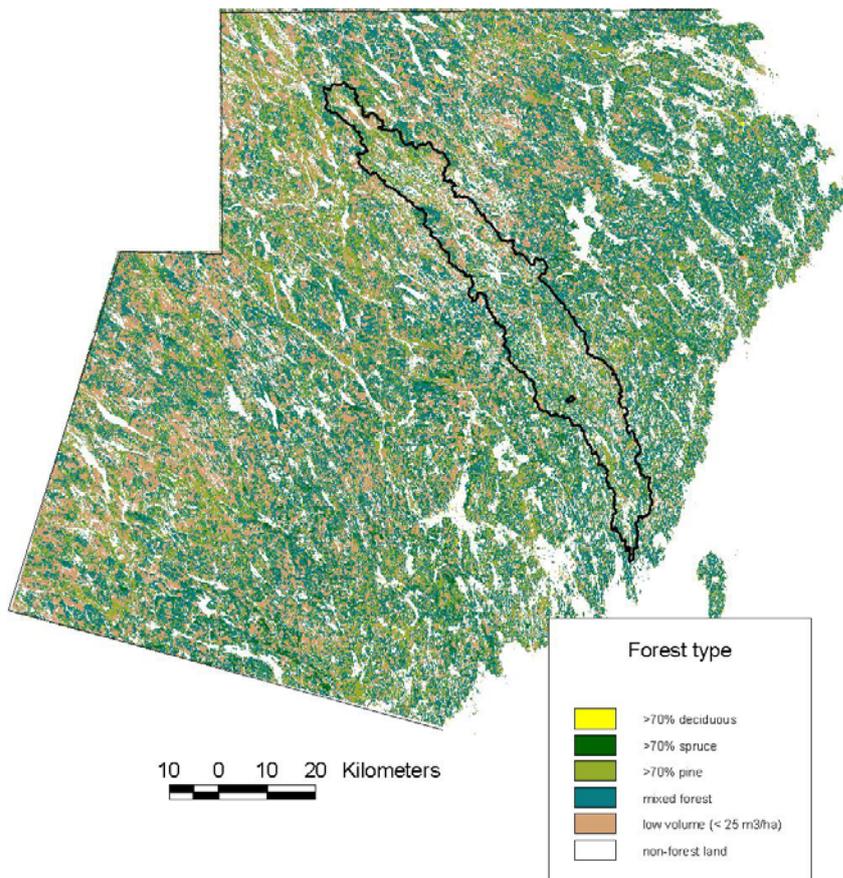


Figure 4. Forest types estimated from the Västerbotten data. The Sävar watershed is outlined.

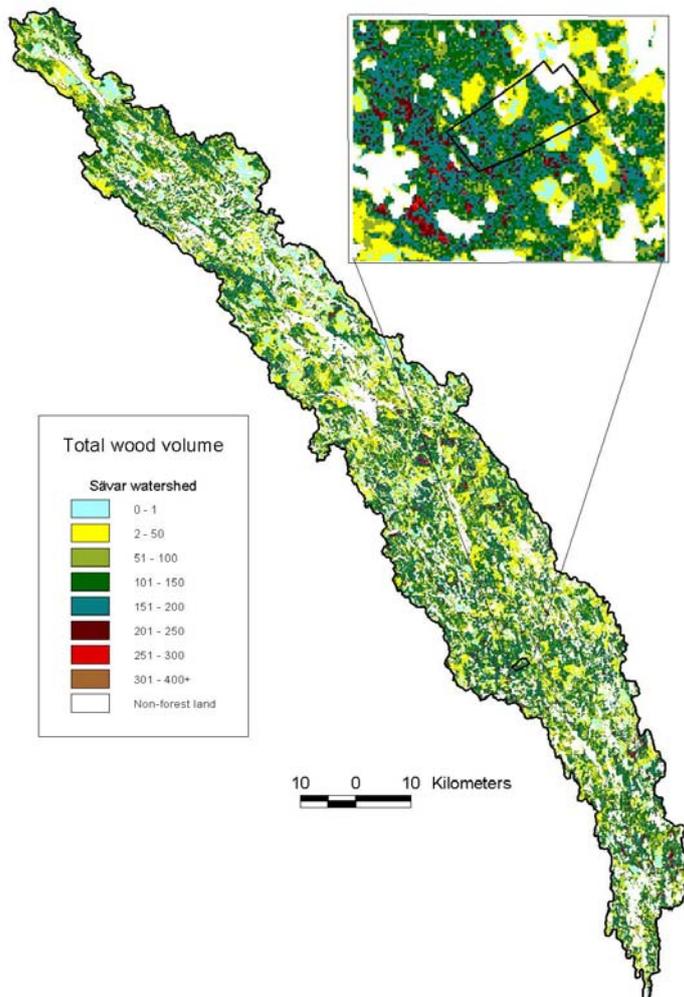


Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Figure 5. Total wood volume estimated for the Sävar watershed. Inset of the Sävars-Brattby forest estate.



Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Table 1. Satellite data properties for each project.

Project	Sensor	Path/row	Acquisition date	Acquisition Angle	Sun angle
Västerbotten	Landsat 5	193/015	July 5, 1994	---	46°
Gävleborg	Landsat 5	194/017	Aug. 8, 1997	---	39°
	Landsat 5	194/017*	Aug. 8, 1997	---	39°
	Landsat 5	196/017	June 27, 1995	---	48°
Dalarna	Landsat 5	194/017	Aug. 8, 1997	---	39°
	Landsat 5	196/017	June 27, 1995	---	48°
Uppland	Landsat 5	192/018†	Aug. 2, 1995	---	44°
Älvsbyn	SPOT 3	053/213	Aug. 22, 1996	-12.16°	35°
	SPOT 3	054/214	Sept. 17, 1996	- 1.04°	25°

* Scene shifted to south † Quarter TM scene

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Table 2. Examples of dynamic range in the NFI plots from three projects and the dynamic range of pixel-level estimates of total wood volume.

Project	Volume 0-24 m3/ha (%)	Volume 25-99 m3/ha (%)	Volume 100-200 m3/ha (%)	Volume 200-300 m3/ha (%)	Volume 300+ m3/ha (%)	Years of NFI plots	No. NFI plots	Range total wood volume m3/ha	Mean value total wood volume
Gävleborg NFI	15	22	28	19	16	1990- 1998	2760	0-768	164
Gävleborg estimates*	19	14	33	26	8			0-644	151
Dalarna NFI	27	30	30	10	3	1992- 1995	680	0-330	72
Dalarna estimates*	32	30	30	7	1			0-330	113
Älvsbyn NFI	31	35	28	3	1	1990- 1998	283	0-405	77
Älvsbyn estimates*	23	43	33	1	1			0-323	75

*Gävleborg scene = 194/017, Dalarna scene = 196/017, Älvsbyn scene = 053/213

Table 3. Pixel-level results (RMSE given as a percentage relative to the reference plot mean value).

Project	Path/row	RMSE estim. total wood vol.(%)	Estim. bias (m3/ha)	NFI mean total wood vol. (m3/ha)	RMSE estim. biomass (%)	NFI mean biomass (tons/ha)	RMSE estim. Age (yrs)	NFI mean Age (yrs)
Västerbotten	193/015	59%	-0.1	120.3	53%	66.0	---	---
Gävleborg	194/017	70%	-0.6	153.2	63%	82.3	50%	62.3
	194/017	66%	1.4	165.1	61%	83.7	49%	58.9
	196/017	71%	-1.2	99.5	66%	57.3	55%	76.8
Dalarna	194/017	58%	12.0	187.0	69%	95.0	57%	60.0
	196/017	75%	6.0	100.0	77%	64.0	60%	78.0
Uppland	192/018	74%	10.0	139.1	79%	81.5	---	---
Älvsbyn	053/213	80%	0.6	77.0	---	---	---	---
	054/214	69%	2.0	115.0	---	---	53%	69.0

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description

Table 4. RMSE and bias for total wood volume estimates from the Dalarna project.

	Eastern TM Scene (194/017)			Western TM scene (196/017)		
	0-99 m ³ /ha	100-299 m ³ /ha	300+ m ³ /ha	0-99 m ³ /ha	100-299 m ³ /ha	300+ m ³ /ha
RMSE (m ³ /ha)	79	91	199	43	67	244
Bias (m ³ /ha)	54	26	-158	32	-34	-228
n	65	90	24	134	53	6

Full Citation: Reese, H., Nilsson, M., Sandström, P., and Olsson, H. 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture* 37(1):37-55.

Journal Home Page:

http://www.elsevier.com/wps/find/journaldescription.cws_home/503304/description#description