

## ELEVATION ACCURACY OF LASER SCANNING-DERIVED DIGITAL TERRAIN AND TARGET MODELS IN FOREST ENVIRONMENT

J. Hyyppä<sup>1</sup>, U. Pyysalo<sup>2</sup>, H.Hyyppä<sup>2</sup> and A. Samberg<sup>2</sup>

1. Finnish Geodetic Institute, Department of Photogrammetry and Remote Sensing, P.O. Box 15, FIN-02431 Masala, Finland, Tel: +358-9-29555305, Fax: +358-9-29555200, Email: [juha.hyyppa@fgi.fi](mailto:juha.hyyppa@fgi.fi)
2. Helsinki University of Technology, FIN-02015 HUT, Finland

### ABSTRACT

This paper evaluates and discusses the accuracy of laser scanner in DTM (digital terrain model) generation and digital 3-D height model generation in forested areas. High-pulse-rate laser scanners are capable of detecting single trees in a boreal forest zone, since a significant amount of the laser pulses reflect directly from the ground without any interaction with the canopy. This allows for a detailed investigation of forest areas and the creation of a 3-dimensional tree height model. Special emphasis is laid on the optimisation of the selection of ground hits used for the creation of the DTM of future high-pulse-rate laser scanners. A novel DTM algorithm is depicted in detail. In order to develop the algorithm, five phases were created: 1) Calculation of the original reference surface, 2) removal of the vegetation from the reference surface, 3) classification of the original cloud of points, 4) calculation of the DTM based on the classified ground hits, and 5) interpolation of the missing points. A standard error of 15 cm was obtained for flat forest areas and the error increased with increasing terrain slope to the value of approximately 40 cm at the slope of 40 %. The average standard error for forest area was about 22 cm. The laser-derived DTM of the forest road deviated from the true height by 8.5 cm only. An optimum performance for the DTM generation was obtained by averaging the ground hits which were located, at the maximum, 60 cm above the minimum terrain values. It was also shown that tree heights of individual trees in the dominating storey can be obtained with less than 1 m standard error.

### INTRODUCTION

The topographic measurement of the land surface of the Earth has been a standard task of photogrammetry. However, the automatic assessment of large-scale, high-quality topographic data in forested areas has been a difficult task to perform for a long time. Photogrammetric measurements suffered from very dark shadows on the ground and from difficulties in finding the same ground point in two aerial images (1). Laser scanning is a relatively new method for direct measurement of digital terrain models (DTM) from the air particularly suited for forest areas, as long as the laser pulses are capable of at least partly penetrating the vegetation in order to reach the ground. Since the penetration capability of signals at optical and infrared wavelengths is basically zero through any object, the penetration requires gaps within the canopy layer. The requirement for DTM generation by laser scanning is a reasonable amount of laser pulses reflecting directly from the ground without any interaction with the canopy. These hits are evaluated as ground hits and the DTM is reconstructed using them. As to the reflection of a laser pulse on vegetation, there are several possibilities. The laser pulse might directly reflect on the canopy or reach the ground undisturbed. Typically, part of the energy is dissipated in the vegetation and partly reflected at intermediate levels on its way through the vegetation, giving several distinguishable return signals. The laser scanners typically record the distances for the first and last echoes neglecting the intermediate returns. Previously, the pulse rate of the laser scanners was rather limited resulting in a low number of ground hits obtained. The typical assumption in DTM calculation is that the minimum heights correspond to terrain heights and the DTM is finalised by interpolations.

Today, the most advanced laser scanners can produce up to almost 100,000 samples per second with two pulse modes (first and last pulse) allowing for a more advanced optimisation and selection of ground hits for the creation of DTM, and not just looking for the minimum elevations. Previous studies have shown that by using the last pulse mode there exist some reflections from beneath the terrain surface (partly due to multiple reflections). Thus, the use of minimum filters and last pulse mode data will result in a false interpretation of surface height. Additionally, the use of minimum values might result in a systematic shift in the DTM (typically a couple of tens of cm below the surface). However, despite the inaccuracy compared to the optimal system, the use of minimum filters especially with varying spatial window size is an economic way of calculating the DTM.

The airborne laser scanner may offer huge opportunities for rapidly estimating tree height, timber volume, and forest biomass over extensive forest areas (2). Previously, laser systems applied for forest studies were profiling sensors capable of data collection merely along the flight track, such as in Nelson et al. (3), or the pulse rate of the laser scanner limited the capability of detecting individual trees (4). However, the situation changes when the number of pulses transmitted by the laser scanner increases. In the boreal forest zone and in many forest areas, there exist gaps between the forest crowns. For example in Finland, roughly speaking more than 30 % of the first pulse data reflects directly from the ground without any interaction with the canopy. By increasing the number of pulses, it is possible to have samples from each individual tree and also from the gaps between the trees. Basically this means that several laser pulses can be recorded per m<sup>2</sup>. This allows for a detailed investigation of forest areas and the creation of 3-dimensional tree height map. The tree height map can be calculated from the digital terrain and crown models, both obtained with the laser scanner data. By analysing the 3-dimensional tree height model using image vision methods, it is possible to locate individual trees, estimate individual tree heights, crown area and derive stem diameter, number of stems, basal area, and stem volume using those data. The use of aerial photographs helps in the discrimination of tree species and improves the accuracy obtained.

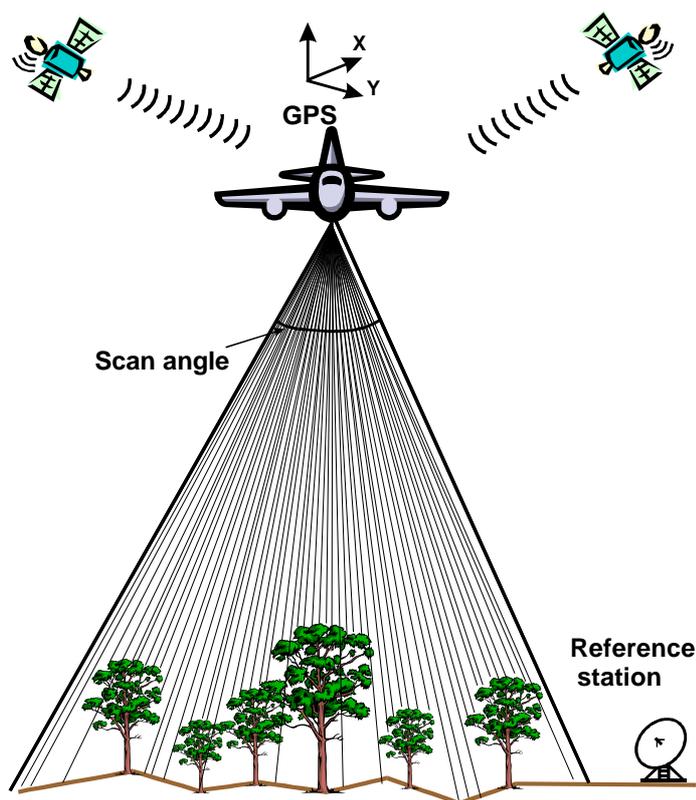
This paper evaluates and discusses the elevation accuracy of laser scanner in DTM and 3-D tree height model generation in forested areas.

## METHODS

### Laser scanning measurements

Laser scanning is based on distance measurements and precise orientation of these measurements between a sensor (the position of which is well known) and a reflecting object (the position to be defined). By knowing the sensor position, the distance and the incidence angle of each measurement, one can easily calculate the co-ordinates of the reflecting object. The scanning mechanism sweeps the laser beam across the flight line providing coverage across the flight track. Along the track coverage is provided by the aircraft's motion. By using sensitive and noise-suppressing kinematic DGPS receivers, the position of the sensor can be measured with an accuracy of about 0.1 m. Further, a corresponding reference station must be placed within or close to the surface area. Sensor's orientation is obtained with an accuracy better than 0.2 mrad. This results in an accuracy better than 1 m in x- and y-directions from the altitude of 800 m (see 5). With respect to the reference plane, each measured point can then be characterised by three co-ordinates (x, y, z). The detailed measurement principle of laser scanning over forested areas is depicted in Figure 1.

Concerning the DTM generation, measurement density, use of first/last pulse modes, and incidence angle require careful validation. High measurement density is required for a high spatial resolution and a sufficient number of ground hits. The steep incidence angle helps in obtaining a sufficient number of ground hits. Test flights (5) have shown that at incidence angles of more than 10° off-nadir, the amount of shadowed areas heavily increases, i.e., the number of measured ground hits decreases and gaps in the DTM occur more frequently. The experiments in central Europe between 1988 and 1992 gave penetration rates to the ground ranging between 24 and 29 % for coniferous and 22 and 25 % for deciduous trees during summer time (6). The latter rate can be increased sig-



nificantly in winter times. Thus, last pulse mode measurements during winter time has been increasingly used for the collection of terrain elevation data especially in wooded areas.

There are few airborne laser scanners available on the market today. Main providers are TopoSys, Optech, and Saab Survey Systems. The TopoSys laser scanner was selected for the study due to its high measurement density and steep incidence angle. The performance of the TopoSys-1 laser is depicted in Table 1.

**Figure 1: The measurement principle of a laser scanner in forested area.**

The laser scanner campaign was carried out early in September 1998. The TopoSys-1 laser scanner was installed in the local aircraft. Three DGPS receivers were employed

to record the carrying platform position: one on board the aircraft, and two as ground reference GPS stations (the first as basic receiver, the second for backup). A test site named Kalkkinen was selected in southern Finland, 130 km north of Helsinki. This rather typical boreal forest area was selected for the study in order to maximize the availability and adequacy of good field inventory data and remote sensing data. An area of 100 hectares (2 km by 0.5) was measured with the TopoSys-1 laser scanner. The test site is dominated by minor hills, otherwise it is flat and situated about 110 m above the sea level. The main tree species are Norway spruce and Scots pine, while the mean stand size is 1.2 hectares. The test site was flown with the laser scanner from an altitude of 400 m. The first pulse mode was applied in the analysis.

**Table 1: TopoSys-1 laser scanner performance parameters.**

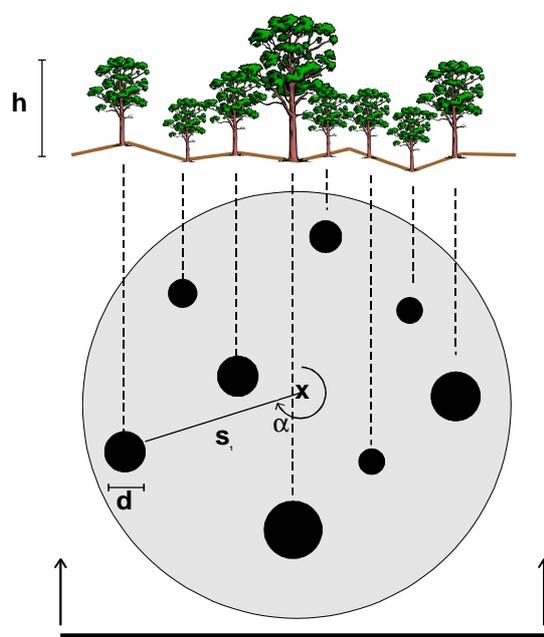
Parameter	Performance(s)
Sensor	pulse-modulated, TopoSys-1
Laser pulse frequency	83 000 Hz
Scan frequency	630 Hz
Field of View	$\pm 7.1$ degrees
Measurement density	8...10 per m <sup>2</sup> at 400 m
The number of shots per scan	128 parallel shots (one of which is the reference)
Swath width at 800 m	About 100 m
Position accuracy	x, y < 1.0 m
Elevation accuracy (WGS84)	z < 0.15 m
Laser classification	class 1 by EN 60825 (eye-safe)

**Field-measured data for verification**

In Kalkkinen test site, a small forest area corresponding to about 1.4 hectare in size was selected for detailed studies of DTM. The field data was obtained using a tachymeter for 750 points. The tachymeter measurements were transformed from the local coordinate system to the Finnish coordi-

nate system using GPS measurements of the Finnish Road Administration. In such a small area, the effect of the Geoid can be neglected. 340,000 individual laser-derived points were obtained from the test site resulting in an average measurement density of 24 measurements per  $m^2$ .

A systematic sample plot network with 100-m spacing was designed for the test site of 2 km by 0.5. The location of the centre of each sample plot was determined with a standard deviation of better than 1 m using advanced GPS/GLONASS system by Finnish Road Administration. From each plot (Figure 2), the basal area with stratification by species, diameter, and tree species of each tree,



height and age (of at least 3 trees of every species and stratum) was measured. From every 5<sup>th</sup> plot, the location, diameter at breast height (1.3 m), and height of every tree were recorded. The location of every tree was measured as a reference to the centre of the sample plot. Distance and angle deviation from compass north were recorded. Of these data, 89 trees were used for the verification of 3-D tree height model elevation accuracy of single trees.

**Figure 2: The location, diameter at breast height (1.3 m), and height of every tree were recorded. The location of every tree was measured as a reference to the centre of the sample plot. Distance and angle deviating from compass north were recorded.**

### Pre-processing of laser data

The laser scanner survey provided a cloud of points, the  $x$ ,  $y$  and  $z$  coordinates of which are known. They form a digital surface model (DSM), which includes terrain points, vegetation points, and points reflected from buildings. By processing the data and classifying the points to terrain and vegetation points, it was possible to produce a digital terrain model (DTM) and a digital vegetation model (DVM). When only the top of the vegetation is included in the model, it can be called digital crown model (DCM). The difference between the DCM and DTM models is a 3-dimensional representation of the tree heights within the target forest area.

### DTM Generation

There exist several algorithms to produce the DTM, but a new approach is briefly described here. The generation of DTM includes five phases:

1. Calculation of the original reference surface  $P$
2. Classification of vegetation and removal of vegetation from the reference surface
3. Classification of the original cloud of points
4. Calculation of the DTM based on classified ground hits
5. Interpolation of missing points

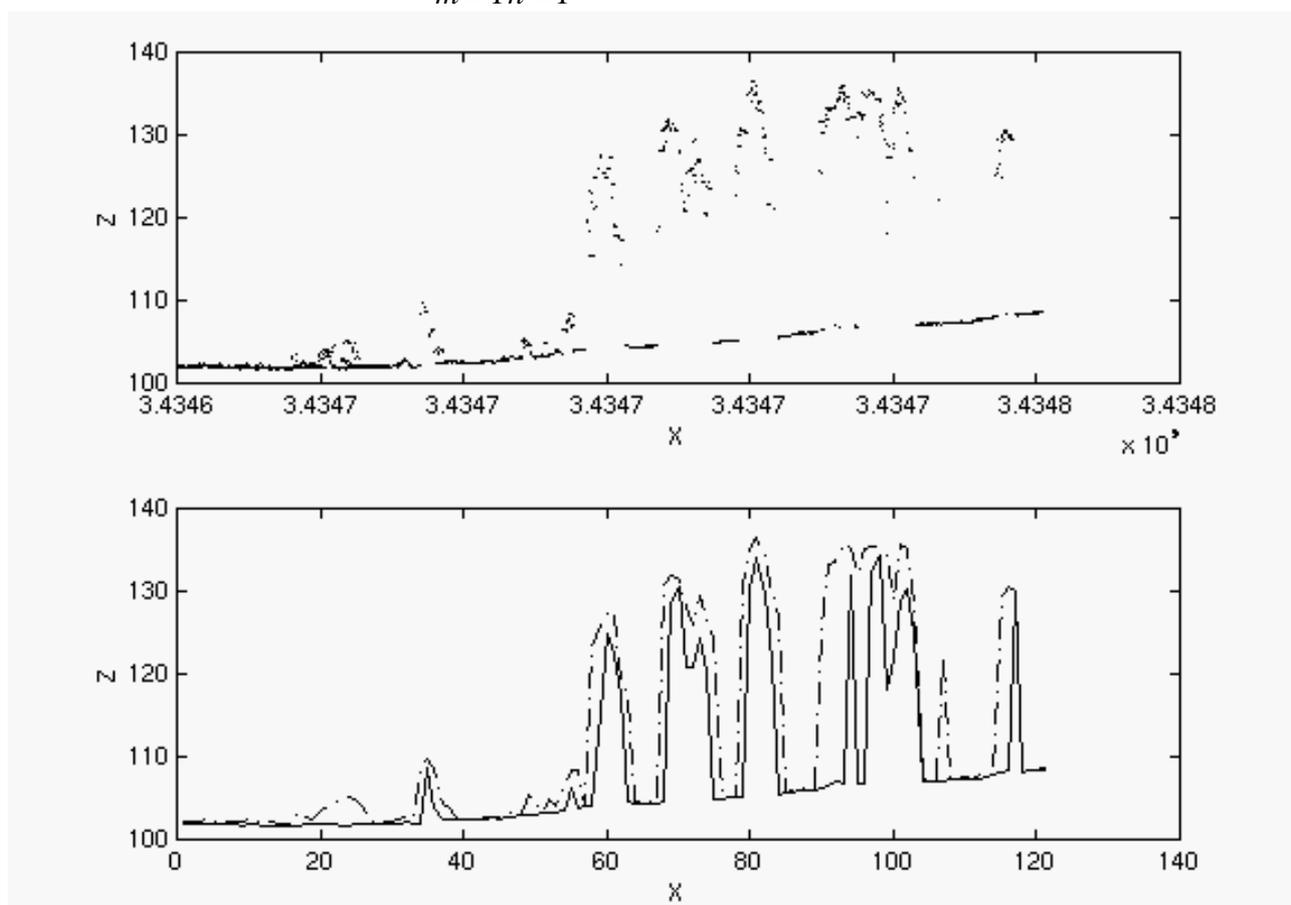
*Calculation of the original reference surface* - A matrix (pixel size 1 m) corresponding to the target area was established, and the cloud of  $x$ ,  $y$  and  $z$  co-ordinates were transformed into a grid. The minimum surface height  $z$  of all points corresponding to certain cell location in the matrix was recorded for each matrix cell and was used as original reference surface. Figure 3 depicts the process and shows the values for one line in the matrix.

*Removal of vegetation from the original reference surface* - Removal of vegetation was performed using filtering. It was assumed that the original reference surface is continuous and terrain eleva-

tions do not significantly change locally. Two gradient matrices were calculated as a sum of differences of near-by pixels and as a sum of absolute values of differences of near-by pixels according to the formulas.

$$g(j,i,1) = \sum_{m=1}^3 \sum_{n=1}^3 |P(((j-2+m), (i-2+n)) - P(j,i)| \quad (1)$$

$$g(j,i,2) = \sum_{m=1}^3 \sum_{n=1}^3 P(((j-2+m), (i-2+n)) - P(j,i) \quad (2)$$



**Figure 3: The upper image shows all elevations within 1-m wide intersection. The maximum values are represented by dashed-dotted line and minimum values by solid line in the lower image.**

The calculated gradient values were compared with a threshold value: if either of the gradient values were less than the threshold, the elevation corresponding to pixel was labelled as ground hit. Otherwise it was assumed as vegetation hit. The new elevations for the ground hits were calculated by the Delaunay interpolation algorithm and using the heights of near-by pixels. The process was iterative. In this study three iterations were made. After the second phase, we obtained a first approximation of the DTM, called rough DTM.

*Classification of original cloud of points* - Classification of the original cloud of points was performed using the rough DTM. The height of original points was compared to the rough DTM and the difference  $dz_n$  was calculated

$$dz_n = z_n - z(j,i) \quad (3)$$

where  $z_n$  is the surface height of original cloud of points, and

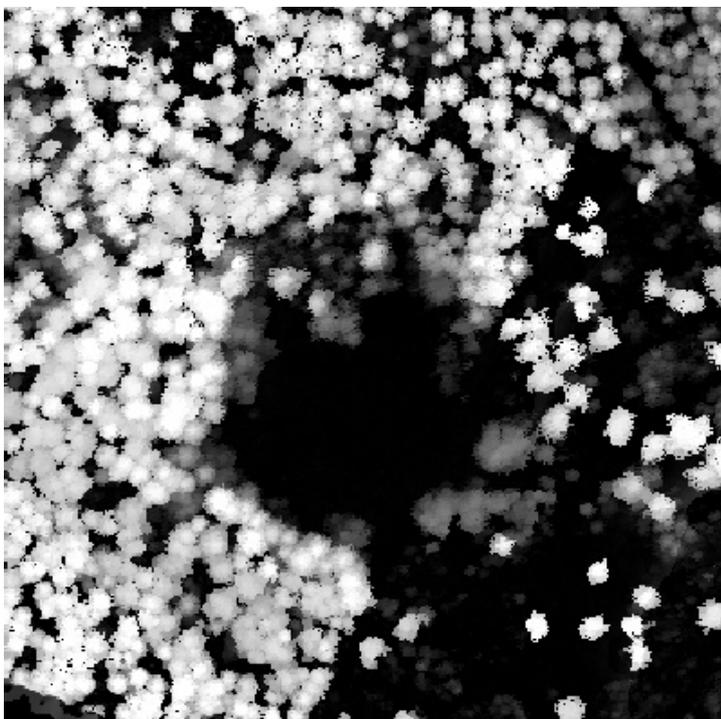
$z(j,i)$  is the surface height of corresponding pixel in the rough DTM.

According to the difference  $dz_n$ , the decision was made whether the point represents the vegetation or the ground hit. The  $dz_n$  values were classified with a 20-cm interval.

*Calculation of the final DTM* was performed using the mean and median values of classified ground hits. *The missing surface heights* were interpolated using the near-by heights and the Delaunay algorithm.

### 3-D tree height model

The maximum value of all the points within the resolution cell was calculated. It was found to represent the tree tops rather well. When there were holes (no data), the value for these points was obtained by interpolation and using the knowledge of near-by pixels. Diverging points (e.g. from birds) in the digital crown model were detected by the gradient method and thresholding. The final height model was calculated as the difference between the DCM and the final DTM. The demonstration was carried out using first pulse data, since it appeared that the first pulse mode was enough to produce rather accurate DTMs. The use of both modes would be likely to further improve the results obtained in this paper.



**Figure 4: Digital tree height model obtained for a sample area (white areas represent tree heights of 35 m, black ones representing terrain heights) . Individual tree crowns can be easily seen.**

### Evaluation

The mean squared error was calculated in order to evaluate the accuracy of the terrain and target heights. The difference between the field-measured terrain height and the laser-derived terrain height was calculated. The sum of the squares of the differences ( $g$ ) is called the sum of squares error and is denoted by  $SSE$ .

$$SSE = g_1^2 + g_2^2 + g_3^2 + \dots + g_N^2 \quad (4)$$

where  $N$  represents the number of samples. The mean squared error MSE corresponds to

$$MSE = \frac{SSE}{N-1} \quad (5)$$

The squared root of the MSE, denoted by RMSE (root mean squared error), corresponds to the accuracy of the laser-derived surface height and can be divided into the bias (systematic), denoted by  $s$ , and random error, denoted by  $r$ :

$$MSE = s^2 + r^2 \quad (6)$$

## RESULTS

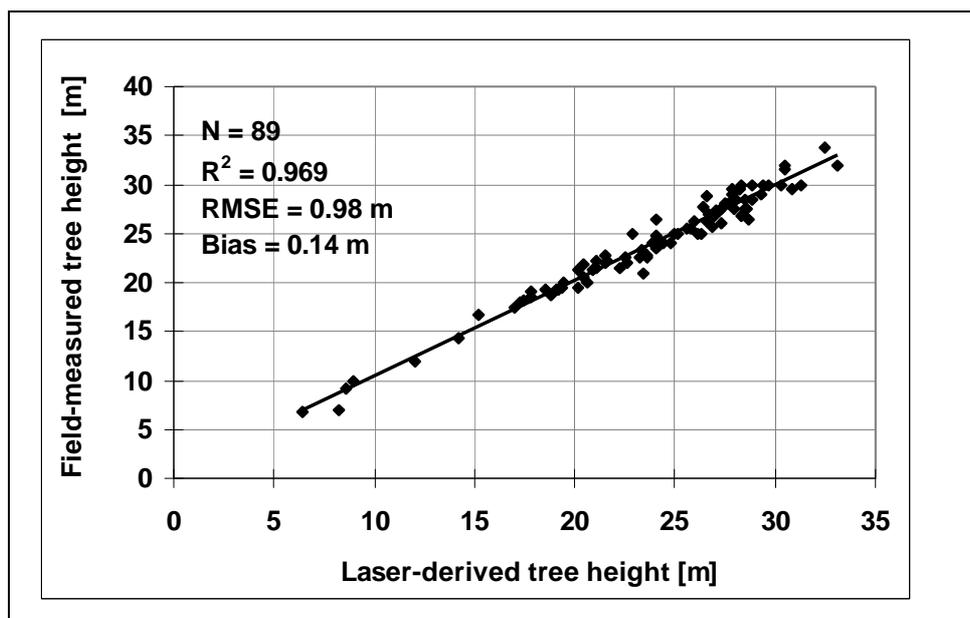
### Digital terrain model

The difference between tachymeter measurements and laser measurements was compared as a function of  $dz_n$  values. The most accurate terrain models were those taking the average or median value of all ground hits which were located at the maximum 60 cm (tolerance value) above the minimum values, resulting in a standard deviation of 22 cm. The 60 cm tolerance value can be considered typical of Finnish forest conditions. In areas of larger terrain slopes, larger values, e.g. 1 m should be used. The question of whether to use the average or the median value of the hits was marginal. Also, the rough DTM was almost as accurate (std 25 cm) as the best model (std 22 cm). This happens only for the first pulse mode data. Since the rough DTM was based on the minimum height of each pixel, it is very vulnerable to hits that may appear below the true ground level. Such bad reflections were not observed with the first pulse mode, but they exist with the last pulse mode, even though the number of them was extremely limited (0.004 %). For the last pulse mode, the use of minimum values is, therefore, a rough approximation and the described method is better applicable to last pulse mode data. The last pulse mode should be used in areas of dense vegetation (e.g. central Europe). The algorithm can be simplified for operational use, if the tolerance value is known. The present version can be used to define the optimum tolerance value.

The quality of the terrain affected the accuracy as expected. The laser-derived DTM of the forest road differed from the true height by only 8.5 cm (standard deviation). While the tachymeter measurements were conducted, two types of terrain heights were recorded: those corresponding to bare forest ground and those corresponding to the height of the forest undervegetation. It was assumed that the laser pulse reflects from the top of this layer. When these two sets were analyzed separately, it was noticed that the obtained standard error was equivalent (23 and 21 cm). The systematic shift between these two data sets was 12 cm. This confirms that the laser pulse really reflected from the top of the forest undervegetation in these cases.

The accuracy of the laser measurements was analysed as a function of terrain slope. It was found out like in (1) that the errors increased with higher slope values. By grouping the data into separate classes representing the slope, we obtained the standard error for each slope class. Over the flat areas, the average standard error was about 15 cm, and with the slope of 40 %, the standard error was more than 40 cm. The observation was found to fit with the basic photogrammetric rule (7). Huising and Gomes Pereira (8) concluded that standard errors of terrain heights are generally well within 15 cm, but for hilly and flat terrains densely covered by vegetation, the accuracy estimates do not generally fulfill the requirements.

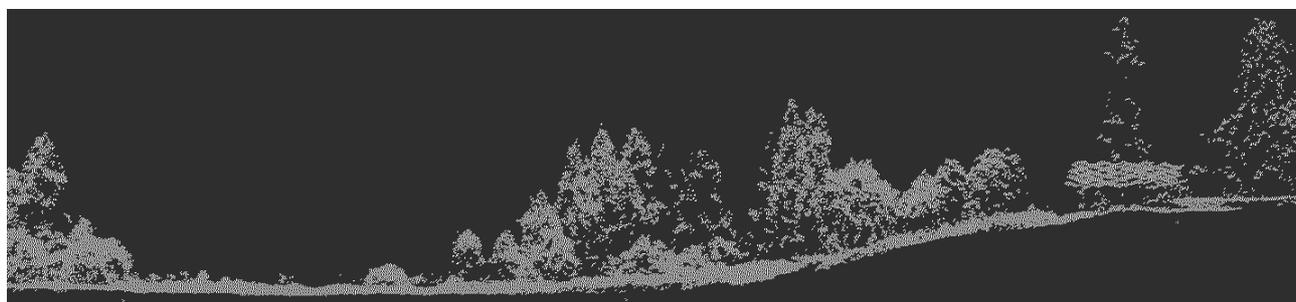
The maximum distance between two nearest laser ground hits was calculated to be 4 m for bushes, 10 m for dense coniferous forest, and 4 m for old mature coniferous forest. One should notice that these results were obtained using the first pulse mode.



**Figure 5:** Laser-derived tree heights compared to field-measured tree heights of individual trees.

### Digital tree height model

Figure 5 shows the regression plot for the derivation of individual tree heights from laser scanner data. The 89 selected trees were mainly from the dominating storey. Individual trees were measured with an accuracy better than 1.0 meter using scanning laser data. The bias with respect to hypsometer measurements was 14 cm. The coefficient of determination was 0.97. The result implies that no single or dominant tree was missed by the laser. Previously, Hyyppä et al. (9) reported a corresponding accuracy of 1.5 m. The improvement can be explained by the modifications in the DTM generation and improvements in the selection of the corresponding tree from the laser data. The accuracy of conventional clinometer measurements is typically better than 0.5 m, but can be up to 1 m for taller trees (>25 m). Therefore, the obtained accuracy seems to be incredibly promising for operational activities. Evaluating the usability of the results, one has to take into account that the time spent for manual height measurements is enormous; and with the laser, the height of all trees could be assessed within a fraction of time, but laser survey is still expensive. The costs per tree of laser surveys are just a fraction of those of manual measurements, but the measured trees can be more qualitatively selected in field surveys. The capability of laser-derived tree height models in vertical canopy probing is demonstrated in Figure 6. The conical shape of spruce trees can be recognized visually. Further, the shape of a summer cottage is clearly visible.



**Figure 6:** Forest canopy profile using a view depth of 6 m and a width of 200 m.

### CONCLUSIONS

The automatic assessment of large-scale, high-quality topographic data in forested areas has long been a difficult task to perform for traditional photogrammetry. Photogrammetric measurements have suffered from very dark shadows on the ground and from difficulties in finding the same

ground point in two aerial images. Laser scanning is a relatively new method for direct measurement of digital terrain and target models and particularly suited for forested areas. This paper evaluates and discusses the accuracy of laser scanner in DTM and target model generation in forests. Special emphasis was laid on the optimisation of the selection of ground hits used for the creation of the DTM of future high-pulse-rate laser scanners. A novel DTM algorithm is depicted in detail. Even though it has not improved the results of using minimum elevations very much (3 cm), it improves the DTM significantly if hits below the ground level exist and it also corrects the systematic errors. A standard error of 15 cm was obtained for flat forest areas and the error increased with increasing terrain slope to the value of approximately 40 cm with the slope of 40 %. The average standard error in DTM for forest area was about 22 cm. The laser-derived DTM of the forest road deviated from the true height by only 8.5 cm. An optimum performance for DTM generation was obtained by averaging the ground hits located at the maximum 60 cm above the minimum terrain values. It was also shown that tree heights of individual trees can be obtained with less than 1 m standard error using high-pulse rate laser scanners.

### ACKNOWLEDGEMENTS

The authors are grateful to the European Commission for financial support (HIGH-SCAN project). The authors want to acknowledge the HIGH-SCAN project partners: Helsinki University of Technology (Finland), Joanneum Research (Austria), University of Freiburg (Germany), TopoSys GmbH (Germany), Scherrer Ingenieurbüro AG (Switzerland) and Forestry Development Centre Tapio (Finland).

### REFERENCES

1. Kraus, K. and Pfeifer N. 1998. "Determination of terrain models in wooded areas with airborne laser scanner data". ISPRS J. of P. & Rem.& Sens. 53:193-203.
2. Hyypä, J., Inkinen, M., 1999. "Detecting and estimating attributes for single trees using laser scanner, The photogrammetric journal of Finland, 16:27-42.
3. Nelson, R. Krabill W. and Tonelli J. 1988. "Estimating forest biomass and volume using airborne laser data". Rem. Sens. Env. 24: 247-267.
4. Nässet, E. 1997. "Estimating timber volume of forest stands using airborne laser scanner data". Rem. Sens. Env. 61:246-253.
5. TopoSys. 1996 Digital elevation models. Services and Products. 10 p.
6. Ackerman, F. 1997. "Airborne laser scanning for elevation models". GIM, 2 p.
7. Kraus, K. 1987. Photogrammetrie. Band 2. 2<sup>nd</sup> edition. Dümmler. Bonn..
8. Huising, E.J. and Gomes Pereira L.M. 1998. "Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications". ISPRS J. of P. & Rem. Sens. 53:245-261.
9. Hyypä, J., Hyypä H. and Samberg A. 1999. "Assessing Forest Stand Attributes by Laser Scanner". Laser Radar Technology and Applications IV. Proceedings of SPIE. 3707:57-69.