REIN ALGORITHM AND THE INFLUENCE OF POINT CLOUD DENSITY ON NDSM AND DEM PRECISION IN A SUBMEDITERRANEAN FOREST

A. Kobler [a, *], P. Ogrinc [a]

[a] Slovenian Forestry Institute, Vecna pot 2, SI - 1000 Ljubljana, Slovenia - (andrej.kobler, peter.ogrinc) @gozdis.si

KEY WORDS: lidar, forest, DEM, nDSM, point cloud density, REIN algorithm

ABSTRACT:

The REIN algorithm makes use of the redundancy in lidar point cloud to generate bare ground DEM and vegetation canopy nDSM. The influence of the input lidar point density on the DEM precision and consequently also on the nDSM precision in the context of REIN have been analyzed in a rough-relief submediterranean karstic forested site. Different lidar point densities were simulated by thinning the density of the basic lidar dataset by factors of 2, 4, and 8. The DEMs and nDSMs were calculated separately from entire dataset and from the thinned lidar data. Strong smoothing effect of lidar data thinning was found in the study area both for DEM and nDSM. Based on the preset minimum precision criteria, the three highest point densities (i.e., 2.71, 5.43 and 10.85 last and only returns per m^2) may be used for DEM generation in the study area, and the two highest point densities (i.e., 8.29 and 16.56 of all returns per m^2) may be used to generate the nDSM. A coarser DEM raster resolution than 1 m is advised for all the lidar point densities except the highest one.

1. INTRODUCTION

Since nineteen-nineties the 3D representations of the bare ground relief under the forest canopy and of the forest canopy itself have been often captured and modeled using aerial laser scanning. Digital elevation models (DEMs) have been extracted from lidar data using a number of different approaches (e.g., Axelsson 2000, Kraus and Pfeifer 1998, Pfeifer et al. 2001, Sithole 2005, Vosselman 2000). Some of them were compared in Sithole and Vosselman (2004). Lidar DEMs are utilized among others in forest road construction and in archaeological studies. By subtracting the DEM from the corresponding forest DSM (digital surface model of the forest canopy) the nDSM (normalized DSM) can be computed, reflecting the relative forest vegetation heights. The nDSM are utilized, e.g., to detect tree tops and to analyze forest canopy closure and forest stand structure.

The REIN (REpetitive INterpolation) algorithm used in this study to extract the DEM and consequently the nDSM, was presented in detail in Kobler et al. (2007). Briefly, REIN is especially applicable in steep, forested areas where other filtering algorithms typically have problems distinguishing between ground returns and off-ground points reflected in the vegetation. REIN is applied after an initial filtering (Figure 1a) of the point cloud, which involves removal of all negative outliers and removal of many, but not necessarily all, off-ground points by some existing filtering algorithm (e.g., using the morphological filtering, Vosselman 2000). REIN makes use of the redundancy in the initially filtered point cloud (FPC) in order to mitigate the effect of the residual off-ground points. Multiple independent random samples are taken from the initial FPC. From each sample, ground elevation estimates are interpolated at individual DTM locations (Figure 1b). Because the lower bounds of the distributions of the elevation estimates at each DTM location are almost insensitive to positive outliers, the true ground elevations can be approximated by adding the global mean offset to the lower bounds, which is estimated from the data (Figure 1c). While other filters behave deterministically, always generating a filter error in special situations, in REIN, because of its random aspects, these errors do not occur in each sample, and typically cancel out in the final computation of DTM elevations.

As the REIN algorithm makes use of the redundancy in the initially filtered point cloud, the input lidar point density has an influence on the DEM precision and consequently also on the nDSM precision. These influences have not yet been analyzed specifically in the context of REIN algorithm, so it is the aim of this study to estimate the influences of lidar point density on (1) the precision (i.e., spatial detail) of DEM extracted in a submediterranean karstic forested relief using the REIN algorithm, and on (2) the precision of the corresponding nDSM.

2. LIDAR DATA AND STUDY AREA DESCRIPTION

The study area (Figure 2) measures 400 m (E-W) by 250 m (N-S), spanning elevations between 71 m and 233 m. It is located in the submediterranean region of Kras in western Slovenia, 5 km from the Gulf of Trieste. The local Gauss – Krueger coordinates are: UL = 5394730, 5075590, LR = 5395130, 5075340. The relief of the study area is rough with slopes ranging up to 60°, the average slope being 22°. The micro-relief features include rock outcrops up to 1 m in size, due to karstic limestone geology, a narrow gorge, and some remains of frontline trenches (the Doberdob section of the World War 1 Isonzo front), which have been heavily vegetated since (Figure 2c). The study area is covered by submediterranean coppice forest. The main tree species are Ostrya carpinifolia, Pinus nigra, Corylus avellana, and Ulmus minor, the latter being found especially within the gorge. The average tree height in the study area is 9 m with the highest trees exceeding 20 m, estimated from the nDSM, using maximum available lidar point

* Corresponding author

density. The vegetation canopy coverage CC is between 5% and 91%, average value being 64%. CC was estimated as the ratio of the first and the only returns for each 10 by 10 sq. m area: CC = Nfirst / (Nfirst + Nonly) * 100. The discrete lidar data of the study area were acquired on April 27, 2005, after beginning of vegetation, as part of a larger campaign, covering 2 by 20 sq. km. The aerial laser scanning was performed by an Optech ALTM-3100 lidar mounted onto a helicopter. The ground speed was 120 km/h and flying height was 1000 m above ground. The lidar pulse rate was 100 kHz, scan frequency 30 Hz, scan angle ?20°, beam divergence 0.3 mrad, and up to 4 measurements including the last one were collected for each pulse. The following lidar point densities were obtained within the study area: 5.15 first returns / m^2, 0.56 intermediate returns / m^2, 7.64 last returns / m^2, and 3.21 only returns / m^2.

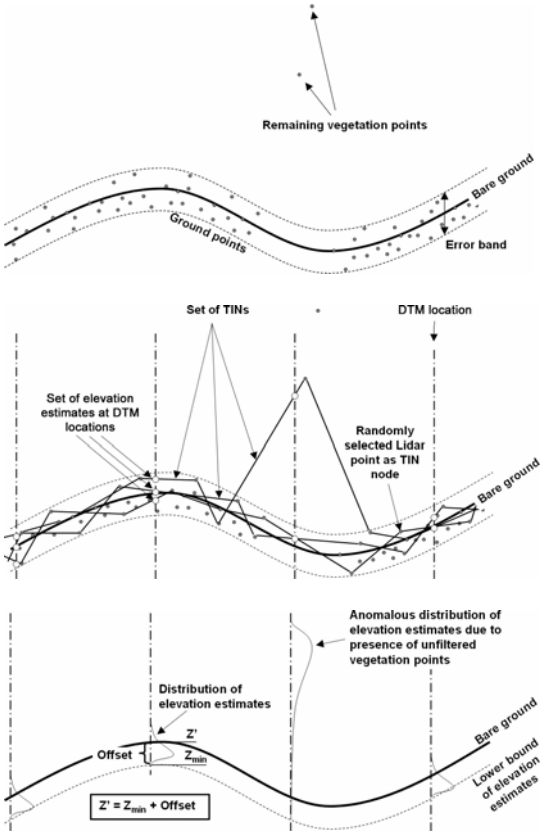

Figure 1. REIN algorithm is used after the initial filtering by any suitable deterministic filter. (a) The result of the initial filtering stage are ground points with few remaining unfiltered vegetation points and no negative outliers. Note the redundancy of ground points within the error band. The lidar point scattering within the error band is caused by measurement errors, grass and low herbal vegetation. (b) Repeated random selections of lidar points are used to build a set of TINs, out of which sets of elevation estimates are interpolated at the locations of DTM grid points. Note that also the remaining unfiltered vegetation points may become TIN nodes. (c) DTM elevations are approximated by adding global mean offset to the lower bounds of elevation distributions, which are unaffected by the unfiltered vegetation points.

3. METHODS

Different lidar point densities were simulated by repeatedly thinning the density of the basic lidar dataset by a factor of 2, yielding the thinning factors of 2, 4, and 8, respectively. The points to be retained in the thinned dataset were selected by first ordering the points according to their respective GPS time-stamps and then selecting every second point. This procedure was performed separately for each point type. The resulting point densities are presented in Table 1. Total point densities used to calculate DEMs (last returns + only returns) and DSMs (all four return types) are given in the bottom two lines of Table 1.

Point density [m^{-2}]		Data thinning factor			
		1	2	4	8
Point type	First	5.15	2.58	1.29	0.64
	Intermediate	0.56	0.28	0.14	0.07
	Last	7.64	3.82	1.91	0.96
	Only	3.21	1.61	0.80	0.40
	L + O	10.85	5.43	2.71	1.36
	F + I + L + O	16.56	8.29	4.14	2.07

Table 1. Lidar point densities obtained by thinning the basic dataset. The starting point density is given in the first column (thinning factor = 1). Total point densities used to calculate DEMs (last returns + only returns) and DSMs (all returns) are given in the bottom 2 lines.

The DEMs, DSMs, and nDSMs were calculated separately from thinned lidar data corresponding to each thinning factor. The raster resolution of all the DEMs, DSMs, and nDSMs were 1 by 1 sq. m, the grids thus measuring 400 columns by 250 rows. The lidar DEMs were calculated from the last and the only returns, using the REIN algorithm, as presented in Section 1. The following REIN parameter values (see Kobler et al. 2007 for their detailed treatment) were used:

- threshold slope = 60° for the initial slope filtering,
- numsamples (i.e., number of repetitive TINs used to interpolate DEM elevations) = 20,
- samplesize (i.e., percentage of lidar points used to build a TIN at each repetition) = 10% of the last and the only returns, i.e., 1.09, 0.54, 0.27, and 0.14 points / m2 respectively, corresponding to thinning factors of 1, 2, 4, and 8 respectively.

Each nDSM was calculated by subtracting the bare ground DEM from the corresponding forest canopy DSM. The elevation of each DSM pixel was estimated from the highest point of any type (first, intermediate, last, only) within each 1 by 1 sq. m. The DEM and nDSM calculated from the non-thinned dataset (DEM1, nDSM1) were used as the reference against which the "thinned" DEMs and nDSMs (DEMx, nDSMx; x = 2, 4, 8) were compared to estimate the decrease of precision due to lower point density.

The DEM precision was estimated by statistics of the image differencing DEM1 – DEMx, and by visual comparison and evaluation of the difference images, and of the wireframe DEM renderings. The nDSM precision was estimated using percentage of the empty pixels, i.e., pixels containing no lidar points, and using statistics of the image differencing nDSM1 –

nDSMx, and correlations between pixel values of nDSM1 and nDSMx within the nonempty pixels.

The minimum criteria for an acceptable DEM and nDSM precision were as follows:

- vertical DEM standard error ≤ 15 cm
- vertical DEM bias ≤ 5 cm
- percentage of nDSM empty pixels ≤ 2 %
- vertical nDSM standard error ≤ 150 cm
- vertical nDSM bias ≤ 50 cm
- vertical nDSM correlation to the reference nDSM ≥ 0.9

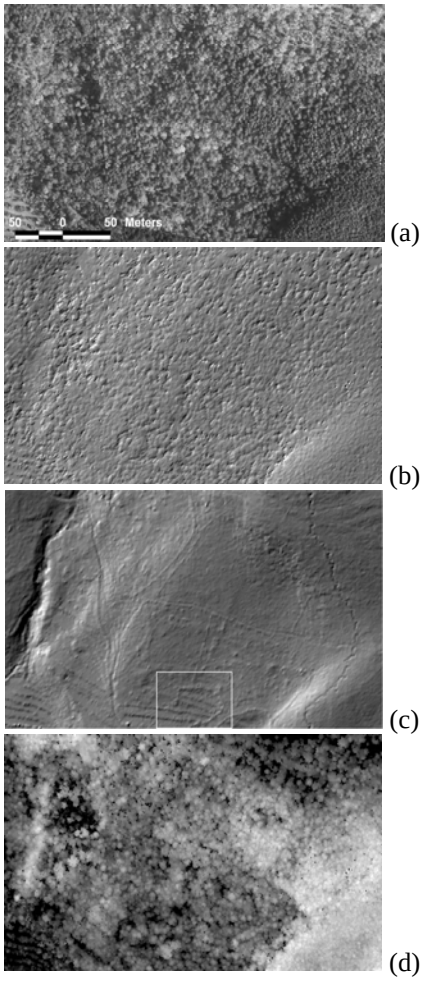

Figure 2. The study area measuring 400 m (E-W) by 250 m (N-S) is shown (a) on an aerial orthophoto, (b) on a shaded forest canopy DSM, (c) on a shaded DEM of the bare ground (the latter being computed using the REIN algorithm), and (d) nDSM. The maps shown in (b), (c), and (d) were computed using the highest available point densities. The white rectangle in (c) denotes the detail rendered as wireframe model in Figure 4. The DEM and the DSM have 1 m raster resolution. The average tree height in the study area is 9 m and the average forest canopy coverage is 64%. The average relief slope is 22°. There are several relief features to note in (c): the gorge on the left, the low manmade walls and footpaths appearing as crisscrossing lines, rock outcrops appearing as the rough surface, the abandoned agricultural terraces in the bottom part, and the jagged line on the right denoting the remains of the WW1 frontline trenches.

4. RESULTS

The DEM difference images DEM1 – DEMx are given in Figure 3 and the corresponding statistics are given in Table 2. The DEM subsets are compared as wireframe models in Figure 4. The percentage of empty nDSM pixels due to data thinning is illustrated in Figure 6. The vegetation height images nDSM1 – nDSMx are given in Figure 5. The statistics of the differences and the correlations between nonempty nDSM1 and nDSMx pixels are given in Table 3 and Figure 7, respectively.

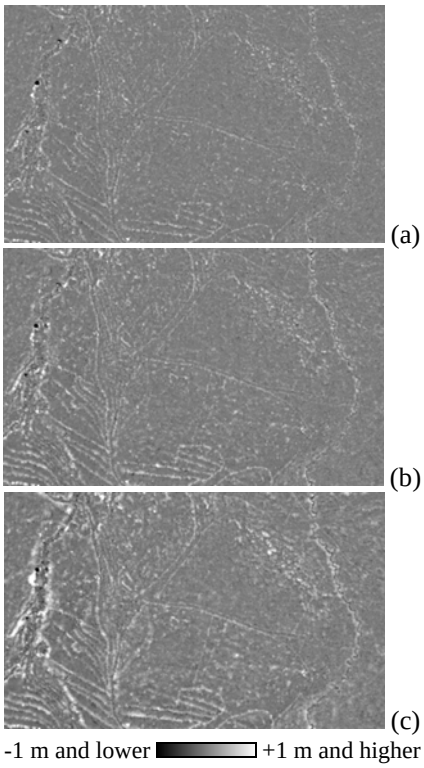

Figure 3. The elevation difference images due to data thinning: (a) DEM1 – DEM2, (b) DEM1 – DEM4, (c) DEM1 – DEM8. The differences are the greatest at sharp break-lines and at locations of pronounced micro-relief, e.g., low manmade walls, rock outcrops, or terraces (cp. Figure 2c).

[m]	x = 2	x = 4	x = 8
Minimum error	-2.34	-2.07	-1.98
Maximum error	1.92	1.85	2.87
Bias	0.02	0.03	0.05
Standard error	0.11	0.13	0.17

Table 2. The statistics for the difference images DEM1 – DEMx (x in table header) shown in Figure 3.

[m]	x = 2	x = 4	x = 8
Minimum error	-18.04	-18.62	-20.54
Maximum error	1.63	1.55	2.13
Bias	-0.41	-1.39	-2.62
Standard error	1.18	2.72	3.95

Table 3. The statistics for the vegetation height difference images nDSM1 – nDSMx (x in table header) shown in Figure 5. Only the non-empty nDSM pixels were taken into account.

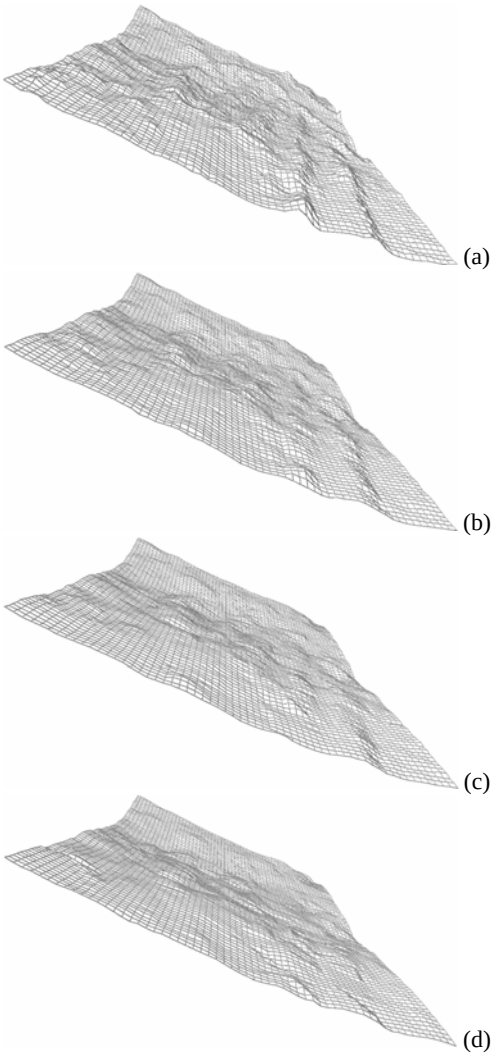

Figure 4. The wireframe DEM rendering of the DEM subset indicated in Figure 2c, as seen from the west. (a) DEM1 – based on lidar point density (i.e., last + only returns) 10.85 m-2, (b) DEM2 – density 5.43 m-2, (c) DEM4 – density 2.71 m-2, (d) DEM8 – density 1.36 m-2. The subset area is 80 m by 60 m. All wireframes are shown using 1 m raster. The main feature of the relief in the subset are the abandoned and overgrown agricultural terraces, which are increasingly smoothed out by REIN at greater data thinning factors.

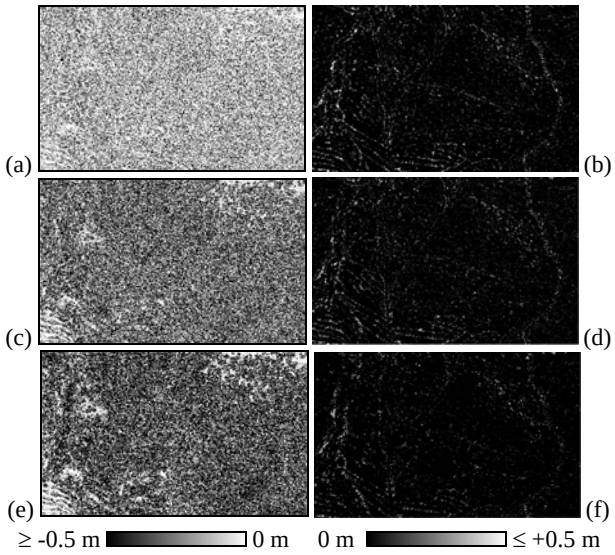

Figure 5. The vegetation height difference images: (a) and (b) nDSM1 – nDSM2, (c) and (d) nDSM1 – nDSM4, (e) and (f) nDSM1 – nDSM8. Due to different grayscale legends, the figures in (a), (c), and (e) in the left column highlight the negative difference values, and the figures in (b), (d), and (f) in the right column highlight the positive ones. Note the quasi-random pattern of differences in the left column, and spatial coincidence of differences with the microrelief features in the right column. The differences at the empty pixels were set to 0. Compare also with Figure 2d.

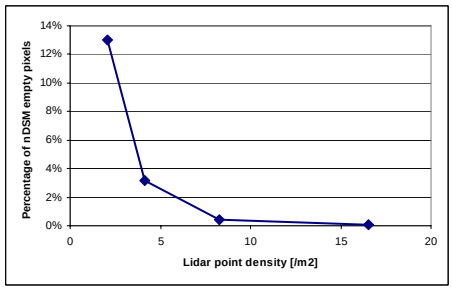

Figure 6. Percentage of the nDSM empty pixels (pixel size 1 m2) due to data thinning. Empty pixels are pixels containing no lidar points. Lidar point densities include all point types (first, intermediate, last, only).

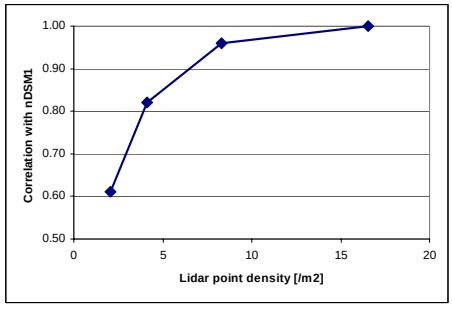

Figure 7. Correlations between the vegetation heights in the reference nDSM and the “thinned” nDSMs. Only the non-empty nDSM pixels have been taken into account.

5. DISCUSSION AND CONCLUSIONS

The REIN algorithm was designed to generate DEM in steep forested relief, where other filtering algorithms typically have problems distinguishing between ground returns and off-ground points reflected in the vegetation. REIN takes multiple independent random subsets of the initially filtered point cloud, making use of the redundancy in dense point clouds. Lowering the density of the input point cloud also reduces the the size of the randomly selected point subsets that are used as nodes for TINs, generated in each repetition of REIN. This in turn affects the precision (i.e., spatial detail) of the generated DEMs, as illustrated in Figure 4. If one uses the DEM1 (i.e., the DEM based on the non-thinned point cloud) as the reference, this effect is mapped in Figure 3 and correspondingly quantified in Table 2. The largest elevation differences between DEM1 and the "thinned" DEMs (i.e., DEMS based on variously thinned point clouds) are found at sharp break-lines and at locations of pronounced micro-relief, e.g., low manmade walls, rock outcrops, terraces, and sharp depressions (Figure 3, Figure 2c). The standard deviation of the difference thus increases from 11 cm for DEM2 to 17 cm for DEM8 (Table 2). There is also a slight increase of the bias (i.e., average difference) from 2 cm to 5 cm for DEM2 and DEM8 respectively, which is due to a more biased estimate of REIN's global mean offset calculation at lower point densities (Figure 1c). If a vertical DEM standard error of less than 15 cm, and a vertical DEM bias of less than 5 cm, respectively (Table 2), are used as criteria for acceptable DEM precision, then all but the most thinned lidar point cloud densities are suitable for REIN-based generation of DEM in the study area of rough relief covered with dense forest. However, the visual evaluation of the wireframe models (Figure 4) suggests coarser DEM raster resolutions than 1 m would be advised for all the lidar point densities except the highest one. This is partly due to the decision of the analyst to use aggressive REIN filtering in this study in order to exclude all DEM errors related to positive outliers (i.e., vegetation points). Less aggressive REIN operating parameters would yield more detailed micro-relief even given less dense lidar point clouds, however at the cost of some remaining vegetation errors in the DEM.

The smoothing effect of lidar point cloud thinning can also be observed in the resulting nDSMs. The comparison of nDSMs at different lidar point cloud densities reveals a strong influence of point density on the proportion of empty pixels, i.e., pixels containing no lidar points, where nDSM height has to be interpolated from the surrounding pixels. At a 1 m raster resolution, the number of empty nDSM pixels when low point density is used in the study area is proportionally much higher, compared to high point density (Figure 6). If 2 % are taken as the maximum allowable percentage of empty pixels, than only the two highest point densities should be used for the study area (considering all point types). Similarly as in DEM, the point cloud density influences the precision of forest canopy rendered in a nDSM. Because the vegetation height in a nDSM is calculated as the difference DSM – DEM, an additional factor in nDSM precision is also the underlying DEM precision. If one uses the nDSM1 (i.e., the nDSM based on the non-thinned point cloud) as the reference, these effects can be illustrated in Figure 5. Note the quasi-random pattern of differences due to varying fidelity of forest canopy in the left column of Figure 5, and the spatial coincidence of differences with the the break-lines and with the micro-relief features, reflecting imprecision of the underlying DEMs, in the right column of Figure 5. These effects are quantified in Table 3, where the underlying DEM imprecision is reflected in the maximum differences (1.63 m for nDSM2 and 2.13 m for nDSM8), and the forest canopy precision is reflected in the minimum differences (-18.04 m for nDSM2 and -20.54 m for nDSM8). The strong effect of lidar point cloud thinning on nDSM precision can also be observed in decreasing correlation of the reference nDSM1 with the "thinned" nDSMs (Figure 7). If a vegetation height bias of ±50 cm, a standard error of 150 cm, and a vegetation height correlation of 0.9, respectively, are taken as the minimum criteria for the nDSM precision, then only the two least thinned point clouds (i.e., thinning factors 1 and 2) are suitable for nDSM generation.

In conclusion, the following can be summarized for the study area on the basis of the mentioned DEM and nDSM minimum precision criteria. The three highest point densities (i.e., 2.71, 5.43 and 10.85 last and only returns per m^2) may be used for the REIN-based DEM generation in the study area. Coarser DEM raster resolutions than 1 m is advised for all the lidar point densities except the highest one. However, note that a less aggressive REIN operating parameters would yield more detailed micro-relief given even a less dense lidar point cloud, at the cost of some remaining vegetation errors in the DEM. To generate a precise enough REIN-based nDSM, the two highest point densities (i.e., 8.29 and 16.56 of all returns per m^2) may be used.

6. ACKNOWLEDGEMENTS

The study was funded by the Slovenian Research Agency and by the Slovenian Ministry of Agriculture, Forestry, and Food within the research project "Processing of Lidar Data" (contract 3311-04-8226575), and by the Slovenian Ministry of Defense and by the Slovenian Research Agency within the research project "Forecasting GIS Model of Fire Danger in Natural Environment" (contract 3311-04-828032). The authors are thankful for fruitful comments by the anonymous reviewers.

7. REFERENCES

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. In: International Archives of Photogrammetry and Remote Sensing, 33 (B4/1) 110-117.

Kobler A., Pfeifer N., Ogrinc P., Todorovski L., Oštir K., Džeroski S., 2007. Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. In: Remote Sensing of Environment 108 (2007) 9–23.

Kraus, K., Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne laser scanner data. In: ISPRS Journal of Photogrammetry and Remote Sensing, 53, 193-203.

Pfeifer, N., Stadler, P., Briese, C., 2001. Derivation of digital terrain models in the SCOP++ environment. In: OEEPE workshop on airborne laserscanning and interferometric SAR for detailed digital elevation models, Stockholm, Sweden.

Sithole, G., 2005. Segmentation and Classification of Airborne Laser Scanner Data, Dissertation, TU Delft, ISBN 90 6132 292 8. In: Publications on Geodesy of the Netherlands Commission of Geodesy, Vol. 59.

Vosselman, G., 2000. Slope based filtering of laser altimetry data. In: International Archives of Photogrammetry and Remote Sensing, Vol. 33, B3/2, 935–942.