ESTIMATION OF URBAN GREEN VOLUME BASED ON LAST PULSE LIDAR DATA AT LEAF-OFF AERIAL FLIGHT TIMES

Robert. Hecht[1], Gotthard Meinel[2] and Manfred F. Buchroithner[3]

1. Leibniz Inst. of Ecological and Regional Development (IOER), Germany; r.hecht@ioer.de
2. Leibniz Inst. of Ecological and Regional Development (IOER), Germany; g.meinel@ioer.de
3. Institute for Cartography, TU Dresden, Germany; manfred.buchroithner@tu-dresden.de

ABSTRACT

The estimation of urban green volume is getting more and more important within the frame of an ecologically orientated city planning. The difference of the first and the last pulse of LiDAR measurements provide the basis for the estimation of the green volume, but these optimal data are not always available. That's why this paper deals with the question whether LiDAR data (last pulse only) that have not been taken during the vegetation period allows a sufficient estimation of the vegetation.

The work sets up on previous results where LiDAR data have been compared to photogrammetrically determined vegetation height measurements (i). The subtraction of the laser-based Digital Terrain Model (DTM) and Digital Surface Model (DSM) in vegetated areas leads to a vast underestimation of green volume of up to 85 %, which is mainly due to the standing deciduous trees with an underestimation of 90 %. Starting from the existence of different laser response characteristics within various vegetation types the relative point density and the normalized height of classified non-ground points were analysed in-depth (ii). The results show a good separation within different vegetation types. Further a method of reconstruction of the underestimated vegetation (mainly deciduous trees) is carried out by generating a cylinder for every classified non-ground point. The point density of non-ground-points and the normalized height of the laser responses will be the input parameter for an adaptive reconstruction based on fuzzy logic techniques. Based on the reference model and external data the accuracy could be estimated. In spite of the suboptimal LiDaR data the results of the work leads to a sufficiently exact and efficient estimation of green volume compared to the costly conventional methods like field investigations.

INTRODUCTION

Since urban vegetation is an essential need for the urban population there is a growing demand on the part of the environmental protection agencies for indicators to quantitatively describe vegetation and its development. The indicators related to the vegetation like the biotope area index, soil function index or the green volume index [m³/m²] with its special ecological statement are counterparts to the indicators for the structural use of land like the site occupancy index and the floor-space index (iii, iv). Unlike these indices for calculating the green volume a three-dimensional survey is essential necessary (v).

But there is a lack of all-covering and up-to-date information about the green volume and there are hardly city-wide applications found. In only few cases where it had been recorded, its estimation was based on the mapping of biotope type areas including the capturing of a medium vegetation height on these areas (vi). This kind of determination is very complicated and costly, rather inaccurate and gives acceptable values only for larger area units.

Now, the vegetation height measurement based on LiDAR data has been tested many times and have proved to be worthwhile in forestry praxis to extract forest stand parameters like tree height, the crown surface shell or the crown radius (vii). For forestry one uses data recordings from vegetation periods of the first and the last pulse with a relatively high point density. Whereas in urban areas there are usually only laser scanner data from leaf-off aerial flight times and in some cases without the first pulse (example Dresden) since the emphasis is hereby on the creation of 3-D city models and Digital Elevation Models (e.g. for flood modelling).

Laser scanner aerial flights of urban areas during the vegetation period for estimating the green volume is desirable but considering the limited financial means of local authority districts not affordable. This leads to the question if and how existing suboptimal laser scanner data (at leaf-off aerial flight times) can be used for estimating the green volume.

Caused by a flood in Dresden (autumn 2002) the city was in serious need of an accurate DTM to simulate flooding. Therefore the whole city was captured during December 2002 with an Optech ALTM 1225 (pulse rate: 25 kHz, scan angle: 12°, flying height: 1 000 m, point density: 1.1 p/m²). The first pulse have been recorded but regrettably not processed due to the requirements of the city. An automated classification of the laser points was made by the company TopScan, which results in two data sets. Firstly the classified ground points and secondly the classified non-ground points, which describe all man-made objects and vegetation hits. By interpolation of the ground points a DTM (1 m resolution) can be derived, whereas a DSM results in using all laser points. By subtracting the two models (DSM – DTM) the normalized Digital Surface Model (nDSM) can be calculated, which is the basis for height measurements above the earth surface (buildings, trees, etc.).

METHODS

In this section the laser response characteristics in vegetation are highlighted. Further the separation of the different vegetation types based on indicators, the principle of reconstruction and its implementation will be explained.

Laser response characteristics in vegetation

A detailed analysis of the LiDAR data compared to reference measurements has been demonstrated in previous work (i, ii). Thus only a brief summary is given here. Based on photogrammetric measurements on stereo image pairs (10 cm ground resolution) a reference-vegetation-nDSM for a study area in Dresden (approximately 500 x 500 m) has been created. Further, reference areas for the three main vegetation types (shrubs, conifers and deciduous trees) were defined to provide a basis for a point based as well as grid-based comparison of the laser-model and the reference-model.

Table 1: Reference and LiDAR measurements in comparison (standard deviation in brackets).

Vegetation type	Shrub (<3 m)	Conifer (3 - 30m)	Deciduous tree (3 - 30m)
No. of reference areas	28	32	67
Total area [m²]	977	1 924	34 476
Part of non-ground points [%]	52.0 (23.6)	89.4 (7.5)	5.5 (4.3)
Reference measurement height [m]	2.4 (1.6)	12.7 (5.6)	15.6 (5.0)
Non-ground point height [m]	2.2 (1.2)	13.2 (5.3)	5.6 (6.5)
Reference green volume [m³/m²]	1.6 (0.7)	11.2 (4.3)	10.9 (5.0)
Laser green volume [m³/m²]	1.4 (0.7)	11.4 (3.5)	0.8 (0.5)
Volume difference [m³/m²]	+ 0.2	- 0.2	+ 10.2

Table 1 presents the results and shows clearly a deep lack of non-ground points (only 5 % of all points) within the deciduous tree stock caused by the high penetration rate during the leaf-off season. Whereas the differences in height and in volume are low for shrubs and conifers, the mean height of the non-ground points in deciduous trees is approximately 1/3 of the reference height. This is due to a high number of hits at shrub level. Further an intense volume underestimation of 90% could be found out for deciduous trees which is an effect of the deficiency in number of non-ground points (area loss) and the height underestimation (height loss).

In order to get a visual impression of the point clouds selected profiles are showing the reference measurements and the classified laser points for the main vegetation types (fig. 1). There are a sparse number of non-ground points for deciduous tree stock (fig. 1a) whereas the shrubs and conifers offer more densely distributed non-ground points (fig 1b, 1c).

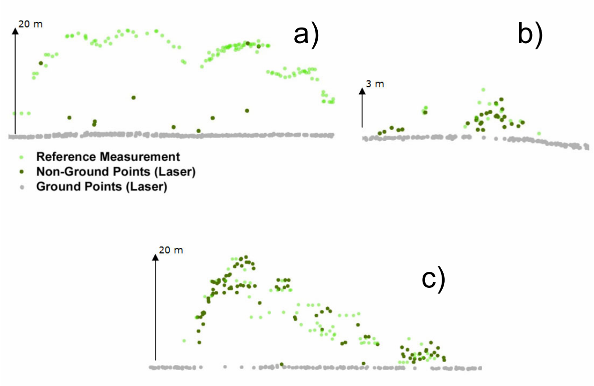

Figure 1: Side view of the reference and laser points: a) deciduous trees, b) shrub c) conifers.

Getting inspired by the good separability of the three vegetation types by using the part of the non-ground points (table 1) a more detailed point density analysis was made which follows in the next section.

Indicators for point classification

Since there is a correlation between the point density of non-ground points and the density of branches (vegetation types) the attempt follows to use this information as an indicator for an adaptive correction of the nDSM. Due to the varying total point density (swath overlaps) the calculation (F. 1)must be the ratio of the point density of non-ground points to the total point density and is defined as relative point density PD_{rel} [%].

$$PD_{rel} = \frac{PD_{non\text{ - }ground\ points}}{PD_{all\ points}} \cdot 100\ [\%] \qquad \text{(F. 1)}$$

In order to choose the right search radius the PD_{rel} for the laser points in different vegetation types was calculated with different search radii. The results in our work show the best separation using a search radius of 6 m, which accords to the mean crown radius of a deciduous tree (ii).

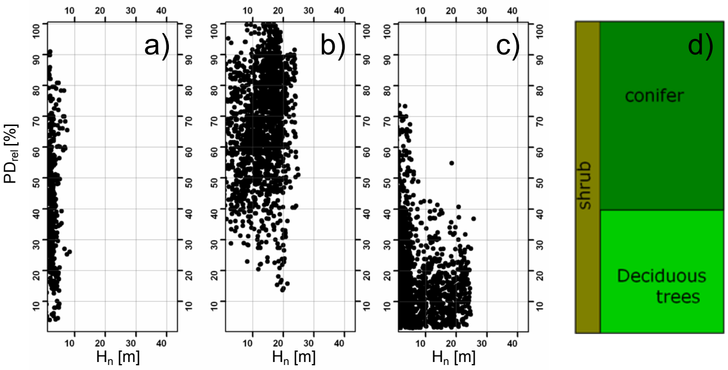

Figure 2: 2D-Feature space of non-ground points for a) shrubs, b) conifers and c) deciduous trees (vertical - relative point density, horizontal - normalized height) and d) its classification.

To enhance the separability, the normalized height of the non-ground points should be considered to separate the undergrowth with its high varying relative point density from conifers and deciduous trees. Figure 2 shows the 2D-Feature space for the different vegetation types. Shrubs beneath the canopy of deciduous trees can be seen as points with a low height in figure 2c. The depicted boxes (fig. 2d) show a classification within the feature space assuming sharp sets. Certainly, the classification has fuzzy boundaries, which need to be considered within our model.

Thus, the basis for a model controlled correction of the laser-nDSM is given by the possibility to separate the non-ground points belonging to the underestimated vegetation (deciduous trees) with the indicators normalized height and the relative point density.

Method of reconstruction

To reconstruct the underestimated part of vegetation a cylinder for every non-ground point within the deciduous tree stock was constructed and defined by a radius and a corrected height. The size of the radius *R* as well as the height correction factor *HCF* is modelled in dependency of the normalized height H_n and the local relative point density PD_{rel} by the following unknown functions.

$$R = f(H_n, PD_{rel}) \qquad \text{(F. 2)}$$

$$HCF = f(H_n, PD_{rel}) \qquad \text{(F. 3)}$$

Figure 3 shows the method of the cylinder construction in a schematic depiction. If there is a high PD_{rel} the surface is described well by the laser-nDSM and no or only a small cylinder is needed to be inserted on the nDSM. In case there exists a low PD_{rel} starting from the non-ground point a cylinder with a large radius *R* and a corrected height will be constructed before inserting it on the nDSM. To avoid the construction of too large radii for non-ground points at shrub level, the normalized height H_n will be the indicator for that. It is a simple principle which is based on the fact that both indicators are a measurement for the volume underestimation.

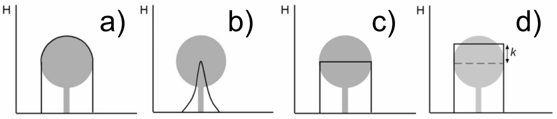

Figure 3: Schematic depiction of the signals and its reconstruction for deciduous trees: a) Reference model, b) laser model, c) Correction of the area loss, d) Correction of the height loss.

Fuzzy Model

Since there are two indicators regulating one output parameter (e.g. *R* in dependency of H_n and PD_{rel}) without having a mathematical model the fuzzy logic technique can be used to solve these sorts of problems. The fuzzy process consists of the following steps: fuzzification, definition of if-then rules, choosing of inference method and defuzzification (viii). Firstly, for every indicator the membership functions (MFs) have to be defined to get it into a linguistic form (fuzzification). The MFs are composed of piecewise linear functions, which allow a rapid calculation. Further the output (e.g. *R*) will be regulated by definition of rules in terms of "if-then" conditions. Thus, the fuzzy output can be determined based on the inference method (algebraic product operator) and subsequently transformed in a sharp output value using the singleton centre of gravity method (defuzzification).

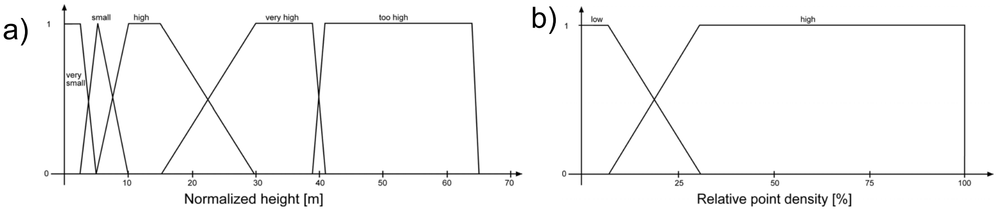

Figure 4: Membership functions of H_n (a) and PD_{rel} (b) used for fuzzy model "Radius".

Figure 4 shows the two indicators (H_n, PD_{rel}) in linguistic form and its MFs. These and the definition of 10 output rules provide the basis for the regulation of the radius *R* for the cylinder. The rules are expressed like: "IF H_n is low (shrub) AND PD_{rel} is low THEN $R = 0$" or "IF H_n is high (high tree) AND PD_{rel} is low THEN $R = 6$ m".

The same procedure of defining the MFs and setting up the output rules are arranged to model the height correction factor *HCF*.

To find out the right input, output and rule definitions statistically determined behaviours helps to understand the process and to get the knowledge to create the model. The behaviour of *HCF* could be analysed by calculating the residuals of the non-ground points compared to our reference measurements. Determining the behaviour of the radius *R* was more complex. For that purpose 760 non-ground points where chosen and associated radii were created manually in that way, that they are compensating the area loss in proper style. Both initial behaviours results in a 3D-matrix which comprises a high ratio of noise that need to be smoothed using a low pass filter. Based on these two behaviours the fuzzy models could be established initially. The fuzzy models have been created with the fuzzy tool of the Spatial Analysis and Modelling Tool (SAMT), which is an open source project and stands under the General Public Licence (www.zalf.de). The advantage to feature it in a fuzzy model allows an easy intervention of the user to make knowledge-based adjustments afterwards. Based on our reference model and the reference areas both models were calibrated in the way that the volume residuals are minimized for all main vegetation types. Fig. 5 pictures the two final behaviours of the fuzzy models for *R* and *HCF*. The output values for the radius ranges between 0.0 and 12.0 m, whereas the height correction factor can have a maximum of 2.5.

Implementation

This method of reconstruction is used operationally to calculate the green volume index for the entire city of Dresden. The workflow is briefly depicted in figure 5.

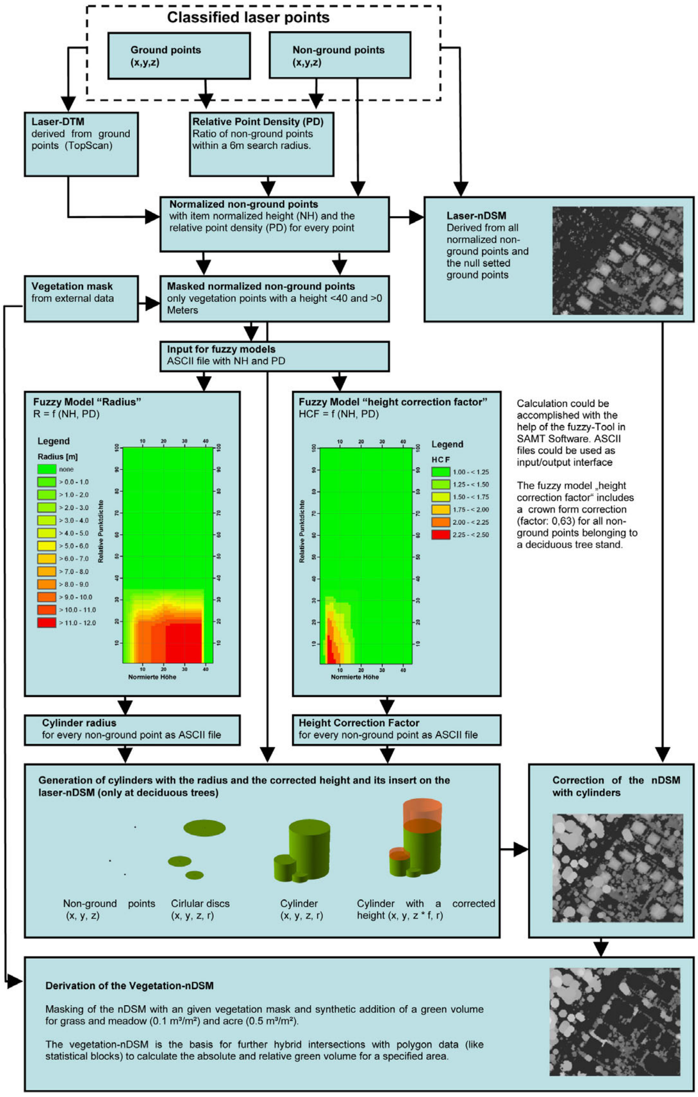

Figure 5: Workflow of Reconstruction.

Starting from the given classified laser points PD_{rel} and H_n can be calculated for every non-ground point. Further, these value pair is the input for the two fuzzy models. Since the output values *R* and *HCF* of the fuzzy models are calculated the cylinders were constructed and inserted on the laser-nDSM. Cylinders are only constructed for non-ground points belonging most likely to the underestimated vegetation (deciduous trees). This achieves relatively less data to process, which results in faster computing.

Given that the crown form for deciduous trees only is not yet considered a flat factor was carried out based on tree cadastre statistics and digitized outlines of the most frequently found deciduous tree species (23 species) which capture 95% of the whole database. Through binarizing the outlines a weighted mean factor of 0.63 could be determined, which could be applied to all deciduous trees. For the final calculations the factor was implemented in the fuzzy model *HCF*.

RESULTS

The workflow (fig 5) shows the uncorrected and corrected laser-nDSM of a part within Dresden city. The bright circular discs are representing the deciduous tree stock. Further, it proves that only the underestimated vegetation is reconstructed and all non-ground points belonging to buildings do not affect the nDSM.

The study area with the uncorrected laser-DSM, corrected laser-nDSM the reference-nDSM and the ortho photo are shown in figure 6. In that case, the non-ground points of man-made objects (e.g. buildings) are sorted out in advance using an external vegetation mask derived from a NDVI. The comparison shows the reconstruction of the standing deciduous trees in the park. The shrubs and the conifers will remain unchanged.

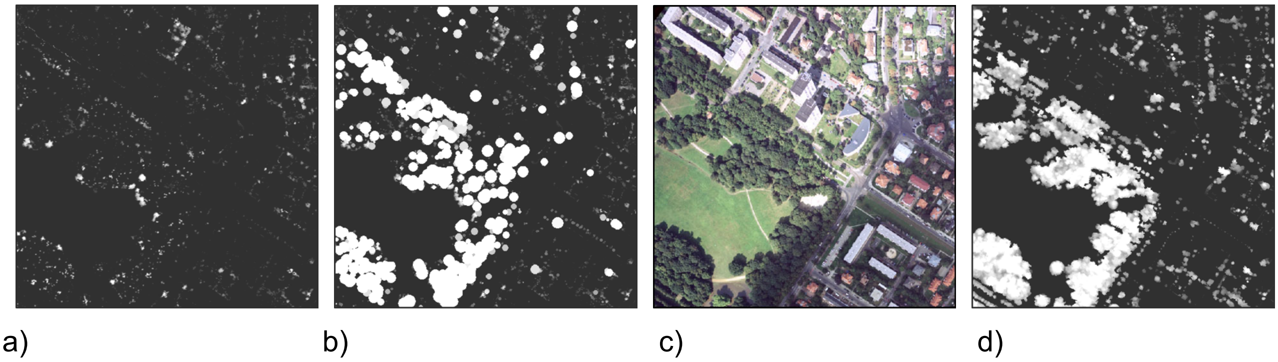

Figure 6: Study area: a) uncorrected laser-nDSM, b) corrected laser-nDSM, c) ortho photo, d) reference model.

Figure 7 presents the same study area in a 3D-view. It identifies clear visible improvements due to the application of our method of reconstruction. Nevertheless, the canopy surface at a group of deciduous trees shows still some gaps in the corrected model. For some single trees the volume will be overestimated mainly caused by to large radii.

Figure 7: 3D-view of the study area a) Laser-DSM (not corrected), b) Laser-DSM (corrected), c) Reference-DSM.

Based on the reference areas the green volume was calculated in order to evaluate the results differentiated according to the three vegetation types. The differences are shown in table 2 below.

Table 2: Comparison of the green volume for different vegetation types.

Vegetation type	Shrubs (< 3 m)	Conifer (3 - 30 m)	Deciduous trees (3 - 30 m)
Total reference green volume	1.80	12.13	14.63
Total uncorrected green volume [m^3/m^2]	1.41	11.75	0.70
Total corrected green volume [m^3/m^2]	1.84	12.09	14.60
Total volume difference compared to reference	-0.04	0.04	0.03

The vast underestimation (90 %) of the deciduous trees could be compensated (4 %) through our correction. By comparison of table 1 and table 2 the green volume for shrubs and conifer are nearly the same. In spite of it, looking at the standard deviations, high uncertainties for deciduous trees of approx. ± 60 % are found, which is partly caused by the small reference areas.

Without having comparable large-area reference data determining the accuracy is rather difficult. Although an analysis of the accuracy in dependency of the size of the basis was done. Therefore, difference models of the study area ($nDSM_{Referenz} – nDSM_{Laser}$) with various cell sizes were calculated. The standard deviations of the differences are decreasing logarithmically with the increasing size of the cell. Alike the correlation coefficient approaches to 1, with an increasing cell size. Figure 8 shows the two diagrams. A statistical block in Dresden (Ø approx. 15 000 m^2) results in uncertainties of ±25% in green volume calculation. Remarkable, the accuracy is also dependent from the locally existing vegetation type.

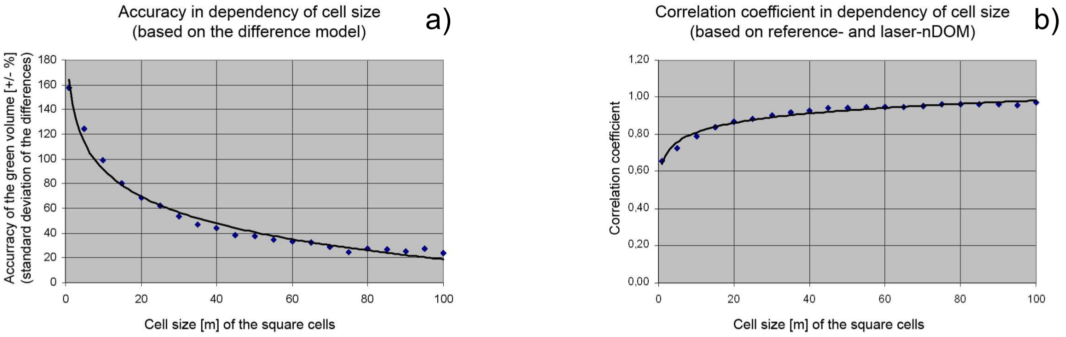

Figure 8: Accuracy of the green volume calculation: a) Standard deviation of the difference model and b) correlation coefficient in dependency of the used cell size.

CONCLUSIONS

The presented method shows a high capability to realize adaptive corrections in order to improve the nDSM based on last pulse data at leaf-off aerial flight times. The intense underestimation of 90% (for deciduous trees) could be compensated through the construction of cylinders for every laser point belonging to the deciduous tree stock. Based on fuzzy logic techniques the size of the cylinders (height, radius) could be regulated by the point density and the laser height. The cylinders are put into the nDSM, which are now compensating the underestimation. This Method is applicable to every comparable LiDAR data.

The results are indicating some problems (gaps, overestimations) on closer examinations. The reasons lie in the availability of the last pulse only and its deep lack of non-ground points with a correct height mainly for deciduous trees. Further, the point density varies for different deciduous tree species, which causes different cylinder radii and some uncertainties. But uncertainties are found in field investigations too - caused by human errors.

However, keeping in mind that the main goal is to carry out block-based green volume calculations, this method shows profits using LiDAR data which are not optimal (no first pulse, leaf-off aerial flight time, no intensity data). Even a big benefit of using LiDAR data taken during the winter season is given by the possibility to calculate the volume for every vegetation type separately. Even if it is less exact than a determination of green volume during summer season or with an available first pulse the results expected can be much more reliable than conventional area-related estimations of the green volume. Since conventional methods or field investigations are too costly it is very beneficial to use suboptimal LiDAR data which are at least available.

REFERENCES

i Meinel G & R Hecht, 2004. Determination of urban vegetation volume on the basis of laser scan data at non-leaf aerial flight times. In: Laser-Scanners for Forest and Landscape Assessment, edited by M Thies et al. (NATSCAN, Freiburg) 334-339

ii Meinel G & R Hecht, 2005. Reconstruction of Urban Vegetation based on Laser Scan Data at Leaf-off Aerial Flight Times – First Results. In: Proceedings of the 31st International Symposium on Remote Sensing of Environment 2005, (ISRSE, St. Petersburg), CD-ROM

iii Heber B & I Lehmann, 1993. Stadtstrukturelle Orientierungswerte für die Bodenversiegelung in Wohngebieten, IOER Schriften 05, ISSN 0944-114X.

iv Kenneweg H, 2002. Neue methodische Ansätze zur Fernerkundung in den bereichen Landschaft, Wald und räumliche Planung. In: Tagungsband 19. edited by S Dech et al. (DFD-Nutzerseminar), 127-137

v Großmann M, H D Schulze & W Pohl, 1984. Bodenfunktionszahl, Grünvolumenzahl, Grünzahl. Gutachten im Auftrag der Umweltbehörde Hamburg. Schriftenreihe der Umweltbehörde Hamburg. Heft 9/1984

vi Arlt G, J Hennersdorf, I, Lehmann & N X Thin, 2003. Basisindikator Vegetationsvolumen. In: Stadtforschung und Statistik. IOER Texte 2/2003. Dresden.

vii Hyyppä J, H Hyyppä, P Litkey, X Yu, H Haggrén, P Rönnholm, P Pyysalo, J Pitkänen & M Maltamo, 2004. In: Laser-Scanners for Forest and Landscape Assessment. edited by M Thies et al. (NATSCAN, Freiburg) 82-89

viii Mechler J, A Mayer, A Schlindwein & R Wolke, 1993. Fuzzy Logic – Einführung und Leitfaden zur praktischen Anwendung mit Fuzzy Shell in C++. Addison-Wesley, Bonn, ISBN 3-89319-443-6, 9-95