

Progress in LiDAR Sensor Technology – Chance and Challenge for DTM Generation and Data Administration

GOTTFRIED MANDLBURGER, CHRISTIAN BRIESE, NORBERT PFEIFER, Vienna

ABSTRACT

Airborne laser scanning (ALS, also referred to as LIDAR (light detection and ranging)) became a standard method for the acquisition of topographic data. However, since the first commercial flights in 1994 a significant development in the sensor technology can be observed. These developments lead to a high increase in data volume due to a significant higher number of determined echoes as well as due to an increased number of information stored per echo.

Next to a short presentation of the current sensor developments this paper focuses on the current challenges in management of ALS data (including the storage of additional information provided by full-waveform ALS sensors). Furthermore, advanced procedures for DTM generation integrating geometric criteria with these additional attributes per echo are discussed.

1. INTRODUCTION

Airborne laser scanning (ALS, also referred to as LIDAR (light detection and ranging)) became the standard method for acquisition of topographic data, especially for the reconstruction of digital terrain models (DTM) (Kraus and Otepka, 2005). For terrain determination so-called filtering algorithms were developed that perform an automatic classification of the acquired point cloud into terrain and off-terrain points (Sithole and Vosselman, 2004). However, for high quality requirements a certain amount of human control is still necessary, i.e. manually correcting classification errors of the automatic procedures. In any case, the DTM from ALS is a standard product, offering higher quality than other methods. This holds especially, but not only, in forested terrain due to the vegetation penetration capability of laser ranging. This property also builds the base for the successful application of ALS in forestry (Hyypä et al., 2004). Next to terrain reconstruction, topographic applications as building or city model reconstruction from airborne laser scanning are a major topic of research but did not mature to the state of DTM extraction yet.

Since the first commercial flights worldwide in 1994 (personal communication with J. Lindenberger and P. Friess) the development of the sensor technology was tremendous (e.g. increase in point density, etc., cf. section 2). These developments, but also the acceptance of the measurement method by surveying authorities and users, e.g. hydraulic engineers, lead to an increase in data volume. In the beginning point clouds were provided in ASCII (list of *xyz*-triples) on CDs, now the binary exchange via hard disks (capacity e.g. 200GB) is common. Additionally to the increase in the number of points, also the attributes of data delivered has changed. The new data are: intensity resp. amplitude measures, GPS-time stamps per emitted pulse, trajectory information (after Kalman filtering for combined GPS and IMU processing), and the full-waveform information of the backscatter. At least some of the mentioned parameters are contained in the binary LAS file format (LAS, 2007) released by the ASPRS, which is gradually becoming a standard for LIDAR data exchange. This is now available to photogrammetric researchers. However, with the development of processing algorithms this information can also be used beneficially for applications. Filtering algorithms for terrain reconstruction may use, e.g. the full-waveform information in order to increase accuracy and reliability in the filtering process (cf. section 5). Also, trajectory information supports strip adjustment of laser scanning data with extended ground control for increasing the accuracy of the point cloud product (Kager, 2004, Csanyi and Toth, 2007). It is expected that in this

way precision well below the decimetre will be reached for the single point, if the backscattering object surface is defined appropriately.

While some of the chances offered by the sensor development are becoming apparent, the challenges are much clearer. The entire data, including point attributes, flight path information, ground control, and meta data has to be managed. As project data with 1 Tera-Byte, e.g., 10^{10} points, are not uncommon, quick data access requires high-performance spatial indexing. There is no standard for full-waveform data, and it should be discussed, if a) the entire digitized waveform shall be stored, or b) echo width and amplitude, which are a result of a more advanced echo analysis approach, or c) the cross section (m^2) of each echo which can be obtained with a very limited set of assumptions (Wagner et al., 2006), or d) ground reflectance (Höfle and Pfeifer, 2007), which require assumptions on target, i.e., ground surface, shall be used. Maintaining the information (on echoes, the flight path, etc.) throughout the entire workflow of airborne laser scanning data processing can be seen as good practice, but is subject to discussion as well. The most important challenge is verifying that the full-waveform and the other attributes listed above carry indeed information (and not only data), and developing methods for exploiting it automatically in a reliable way.

This paper is structured as follows. Section 2 gives more details on the sensor development and the current state of the technology. This builds the basis for Section 3, where the first steps in exploiting the attribute information, namely echo decomposition and strip adjustment, are presented shortly. In Section 4 a proposal for management of airborne laser scanning data, under consideration of the previous sections, is made. Section 5 presents developments on DTM extraction exploiting the new possibilities. Finally, the paper is summarised and an outlook into future research for advanced DTM generation from ALS data is provided.

2. SENSOR DEVELOPMENT

As already mentioned in the introduction, a significant progress in the ALS sensor development took place in the last years. ALS systems still use almost solely the round trip time measurement principle for ranging (one exception is the research system ScaLARS, which applies the phase difference measurement principle (Hug and Wehr, 1997)), but in the last years, the frequency of range measurements developed from a few kHz to about 200kHz. This progress led to a significant increase in the data volume and allows nowadays a much more efficient and detailed acquisition of the topography. Likewise the range spectrum increased. The maximum flying height increased from 1km to 5km.

Furthermore, while the first systems either allowed to store the range to the first or last echo, current systems allow the simultaneous storage of both echoes. More advanced systems even allow to determine intermediate echoes. Next to the range, intensity recording, characterising the reflectance behaviour, was introduced. This development expanded to the full-waveform digitisation (Wagner et al., 2006).

We can nowadays distinguish between two different types of commercial ALS sensor systems: discrete echo and full-waveform scanners. While discrete echo scanners detect a representative trigger signal for multiple echoes in real time using analogue detectors, full-waveform ALS systems digitise the entire analogue echo waveform, i.e. the time-dependent variation of received signal power, for each emitted laser beam (typically with an interval of 1 ns). In contrast to the "classical" discrete echo systems, the full-waveform ALS systems allow an echo determination in post-processing adapted to the individual needs based on the stored waveforms.

Additionally, in order to increase the data coverage rate by an increase in the flying height, pulse repetition rate and maximum range, the so-called multipulse technology was introduced. While standard ALS systems just allow a consecutive pulse emission and detection, this technique allows

to detect the echoes of multi laser pulses travelling through the atmosphere simultaneously. Moreover, some data providers provide multiple scanners on one airborne platform, e.g. for increased side-looking or for forward and backward looking.

Finally, also the hardware developments of the other systems components, such as GPS and IMU (inertial measurement unit) as well as advanced software solutions for direct georeferencing and subsequent fine-georeferencing (cf. section 3.2) have to be mentioned. While one could recognise an increase in accuracy and measurement frequency (e.g. kHz IMU frequency) of the position and orientation sensors, advanced software solutions help to determine accurate trajectory information. Moreover, calibration procedures (e.g. for the determination of the relative position of the sensors on the moving platform) were improved in a way that current ALS point clouds have an absolute accuracy in plane of about $\pm 20\text{-}40\text{cm}$ and in height of about $\pm 10\text{-}15\text{cm}$.

3. PREPROCESSING AND GEOREFERENCING

This section focuses on the determination of the echoes and their additional parameters as well as on the georeferencing of the data. While the echo and attribute (typically just an intensity value) determination of discrete echo ALS systems is performed in real time during data acquisition this procedure, which is discussed in the following subsection, has to be done in a separate processing phase in the case of full-waveform ALS. Independent of the final application of the ALS data, the trajectory of the laser scanner has to be determined based on the position and orientation information (typically GPS and IMU observations) in order to georeference the ALS observations. Subsequently, with the resulting trajectory the position of the synchronised echoes can be determined in one coordinate frame. Then, the georeferencing of the data may be checked by an analysis of the height differences between overlapping strips and/or to external reference data. In the case of higher accuracy requirements a fine-georeferencing process has to follow (cf. section 3.2). Finally, the georeferenced ALS echoes can be stored together with their attributes (echo ID, GPS-time, intensity, echo width, etc.) in a topographic database, which can be accessed for further processing (cf. section 4).

3.1. Determination of echo parameters

This subsection describes the determination of the echoes and their additional parameters in the preprocessing step of full-waveform ALS data. The digitized full-waveform signal per emitted pulse is the result of a convolution of the emitted laser pulse with the illuminated object surface (vgl. Wagner et al., 2006). Usually it is assumed that the emitted pulse can be well approximated by a Gaussian bell curve and that the resulting waveform can be well modelled by the summation of Gauss curves that represent the backscattered energy (Hofton et al., 2000, Wagner et al., 2006, c.f. fig. 1).

In this modelling phase of the waveform, the first step is typically the determination of the number of echoes and their approximate position and amplitude by searching all local maxima in the recorded signal. Based on these approximate values a model of the waveform can be fitted by minimising the residuals to the recorded waveform information. In the case of a Gaussian decomposition the individual Gaussian curves describe the backscattered signal per echo (dotted curves in fig. 1).

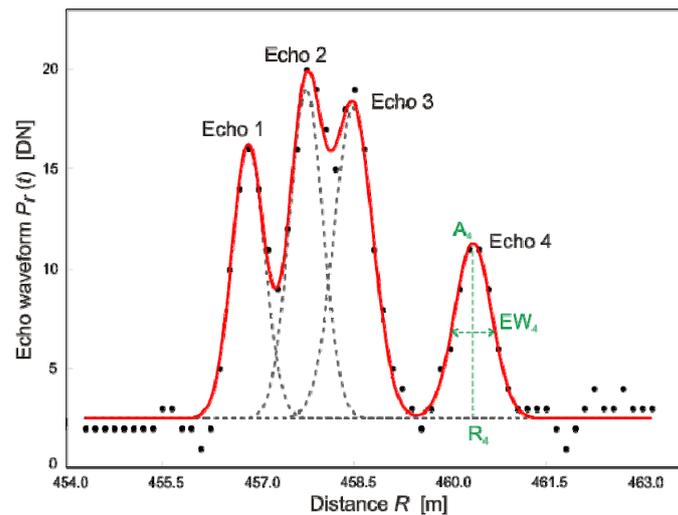


Fig. 1: Modelling of the recorded waveform (black dots) by Gaussian decomposition; Red: Sum of all Gaussian functions, Black dotted lines: individual Gaussian functions per echo; Next to the Distance R_i further echo parameters, like the amplitude A_i and the echo width EW_i , which characterise the individual echoes, can be determined.

Based on the individual curves further characteristic attributes, like the maximum amplitude value A_i and the echo width EW_i , can be determined for each echo. While the amplitude describes the backscattering properties of the illuminated surface, the echo width describes the range variations of all small surface elements that contributed to that echo. Therefore, a big echo width (in relation the width of the emitted pulse) indicates that surface elements at a different range contributed to that echo, while a small echo width indicates a rather flat (in respect to the beam direction) surface. Based on these attributes the so-called backscatter cross section, which is a measure of the electromagnetic energy intercepted and reradiated by objects backwards towards the sensor, can be determined (cf. Wagner et al., 2006).

3.2. Georeferencing

For the transformation of the ALS data (range and angle observations) into one common coordinate system the time dependent position and orientation of the laser scanner has to be determined. As mentioned in the introduction this requires a Position and Orientation system (POS). Together with the laser scanner the POS forms a multiple sensor system. In order to determine the time dependent position and orientation of the scanner the relative orientation, i.e. shifts and rotations, between the components (GPS antenna, IMU, and laser scanner) and time lags in the synchronization have to be determined. To some extent these parameters can be determined by total station measurements on the ground, but a number of parameters, e.g. the IMU orientation or time lags, are better determined during flight within a calibration procedure. Approaches for the calibration have been presented in (Burman, 2002), (Filin, 2003), and (Kager, 2004).

Based on the mounting calibration the coordinates of the echoes can be determined within one coordinate frame by direct georeferencing. However, in practice often deficits of the direct georeferencing of the point cloud can be found (e.g. when analysing strip differences between overlapping ALS strips). Therefore, for advanced accuracy requirements strip adjustment techniques are necessary in order to fine-georeference the ALS data. This can be done either by calibration (Burman, 2002, Kager, 2004), or by applying corrections to the points directly (Crombaghs et al., 2000, Kraus and Pfeifer, 2001). The resulting fine-georeferenced data (coordinates of all echoes and their attributes) is the basis for the determination of subsequent models. Due to the typical high amount of data an adequate data management is essential (see the next section).

4. ALS DATA MANAGEMENT IN TOPDM

It has been pointed out in the previous sections that modern ALS sensors provide data for a more dense and detailed description of the earth's surface at the price of an increased data volume. Even small projects of several dozens or hundreds of square kilometres require nowadays an efficient and flexible system for data handling. For administration of huge amounts of elevation data, the program system TopDM (Topographic Data Management) was developed at the Institute of Photogrammetry and Remote Sensing (I.P.F.) in Vienna. Initially designed for storage of countrywide elevation data from photogrammetry and other traditional surveying techniques, TopDM has proved to be well suited for management of ALS data over the last years.

4.1. General concept and current state

TopDM (Hochstöger, 1996, Warriner and Mandlbürger, 2005) is basically a database management system (DBMS) founded on a geo-relational database like TOPDB (Loitsch, Molnar, 1991) or other, commercially available systems (e.g. Oracle Spatial). The communication between TopDM and the DB is realized via an abstract database interface utilizing the database language TOPSQL - a query language based on standard ANSI-SQL with extensions regarding geometric/topological types and operators. Using TOPSQL geometric and attribute related conditions can be combined in a single SQL statement.

The main goal of TopDM is an efficient long-term storage and archiving of primary topographic data (original data) and derived products (e.g. DTM). Apart from geometry (coordinates) and coding (geometric type and meaning) this requires also storage of information about other data properties (accuracy, compilation method, creator of data, etc.) to allow detailed data selection. Additional important aspects are flexible support of data formats for data import and export as well as rigorous geo-coding of the administrated data. The latter includes also transformations between different coordinate systems and geodetic dates, which gain special importance for ALS data as they are originally captured in a geocentric reference system (e.g. WGS84) and normally to be processed in the national coordinate system. In this sense TopDM mainly focuses on regional or national surveying authorities and is e.g. used by the Federal Office of Metrology and Surveying (A), swisstopo (CH), Ordnance Survey (GB) and several surveying authorities in Germany. Above that, with the increasing amount of topographic data as provided by modern ALS sensors TopDM becomes a basic component of today's large DTM projects (e.g. ALS DTM of Tyrol, Lower Austria, Vienna, Neusiedlersee (Attwenger and Chlapek, 2006)). Within a sophisticated DTM production system, TopDM first serves as data source of the original point cloud. From here data are passed to applications for filtering and DTM modelling. The end products (e.g. DSM, DTM, classified points, hill shadings) are finally returned to the TopDM database for storage and delivery to customers.

4.2. New development

The general concept as described above is based on the administration of the xyz-coordinates only. As pointed out in previous sections (and as it will be shown in the next section), additional point-related attributes like scan angle, echo width, amplitude and the more can be used to improve the derived DTM. Therefore, an extension of the existing TopDM table system was designed and finally implemented allowing storage of additional attributes per point (=per echo). According to the available sensor data different table types may be applied to preserve e.g. information provided by LAS data files (LAS, 2007) or data from full-waveform laser scanners. Examples of table definitions are shown in Tab. 1.

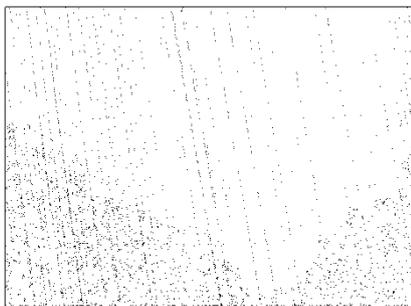


Fig. 3a: Point cloud (all echoes)

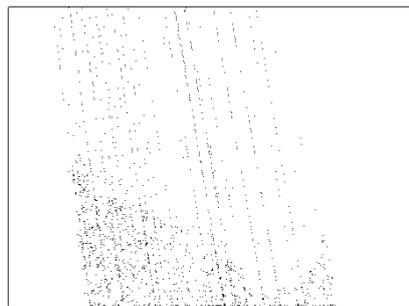


Fig. 3b: Selection by GPS time

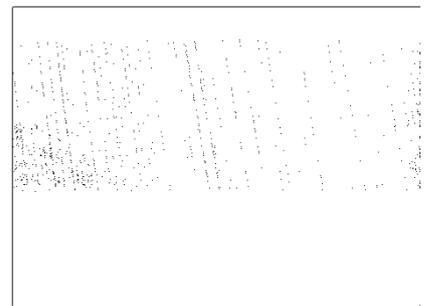


Fig. 3c: Selection by scan angle

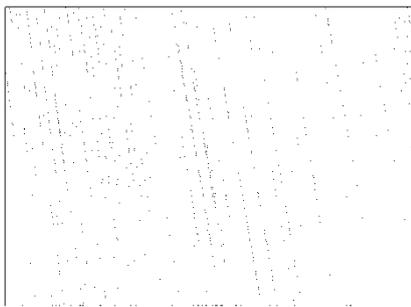


Fig. 3d: Selection by amplitude

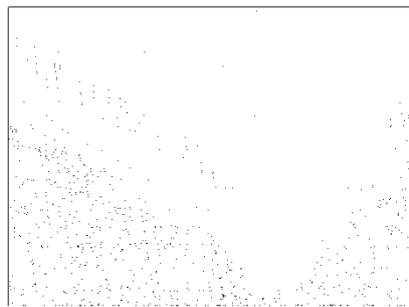


Fig. 3e: Selection by echo width

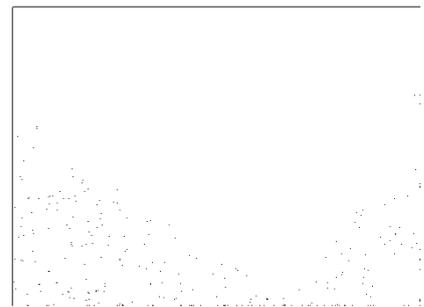


Fig. 3f: Selection by echo class

The respective SELECT conditions for fig. 3a-f are as follows:

- 3a WHERE (COORDINATES .X. WINDOW (20525. 313395. 20660. 313495.)) AND (AGGREGATE='STRIP4')
- 3b WHERE (COORDINATES ... AND (GPSTIME > 555675.5 AND GPSTIME < 555678.0))
- 3c WHERE (COORDINATES ... AND (SCANANGLE >= -1.5 AND SCANANGLE <=6.5))
- 3d WHERE (COORDINATES ... AND (AMPLITUDE >= 75))
- 3e WHERE (COORDINATES ... AND (ECHOWIDTH >= 1.7))
- 3f WHERE (COORDINATES ... AND (ECHOCLASS = 2))

5. CHANCE AND CHALLENGE FOR DTM GENERATION

During the process of data acquisition and echo determination no interpretation resp. classification of the determined echoes, which were reflected from different object surfaces, is performed. However, for the generation of a DTM the classification of the ALS data into terrain and off-terrain points is essential. In the past, many different solutions for the filtering of the ALS data were published (cf. Sithole and Vosselman, 2004). All these approaches have in common that they rely only upon geometric criteria (typically the height relation of neighbored points) for the elimination of off-terrain points.

Within the following subsection visualisations of the additional echo parameters from a small test area in the Leithagebirge 2006 (cf. Doneus and Briese, 2006) are presented. In the subsequent subsection the potential of the additional echo parameters for DTM generation are presented.

5.1. Surface characteristics from echo parameters

In figure 4, next to a shading of the digital surface model (DSM) of the first echo point cloud, visualisations of the amplitude and the echo width of the last echo point cloud are presented. One can recognise that the echo width in areas with low vegetation is – compared to open or wooded terrain with high trees – significantly increased. Furthermore, in the amplitude visualisation vegetation areas can be recognised quite accurately.

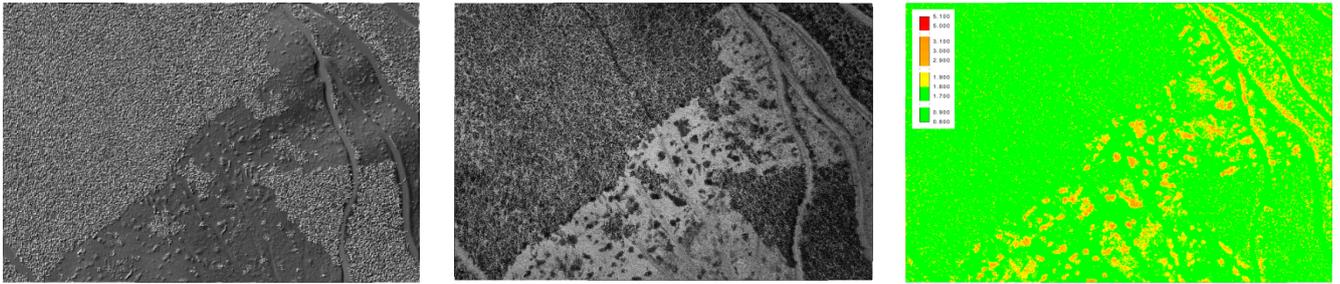


Fig. 4: Left: Digital surface model determined from the first echo point cloud; middle: amplitude (so-called intensity) of echoes extracted from the waveform, right: echo width model of the last echo point cloud

Using this test suite Doneus and Briese (2006) demonstrated that the classification of the ALS points into terrain and off-terrain points with standard algorithms considering the geometric relation between neighbored points did not work properly for the areas with low vegetation. Therefore they proposed an additional pre-filter step that eliminates all echoes with a significantly higher echo width. This process lead to a significant improvement of the DTM in the areas with low vegetation. However, within this simple procedure the definition of the threshold for the echo width might be critical. Therefore, in the following subsection an advanced concept for the consideration of additional full-waveform information into the filtering and classification process for DTM generation is provided.

5.2. Echo parameter exploration in DTM generation by Robust Interpolation

Robust Interpolation (Kraus and Pfeifer, 1998) is a well known approach for the classification of ALS data into terrain and off-terrain points. Within this section we expand this approach to the usage of individual a-priori weights of the ALS point cloud. These weights can be determined based on the additional echo parameters (e.g. based on the echo width). Using these a-priori weights for robust interpolation allows to combine the additional knowledge of the echo attributes with the geometric criteria within the robust surface estimation.

This procedure is demonstrated in figure 5 for a profile extracted from the full-waveform ALS campaign Leithagebirge 2006. For this example the a-priori weights are defined using the echo width parameter. Points with a small echo width get a high a-priori weight, whereas the weight of points with higher echo width is significantly decreased. As can be seen in figure 5a-b the widening of the echo width parameter can be observed in regions with low vegetation. In these areas the DTM determined without individual weights is too high due to the fact that the last echo is a mixture of reflections of small objects at different ranges above the DTM surface. Using robust interpolation with individual weights leads to a more reliable DTM (especially near the road in the middle of the profile). It is less affected by reflections in low vegetation.

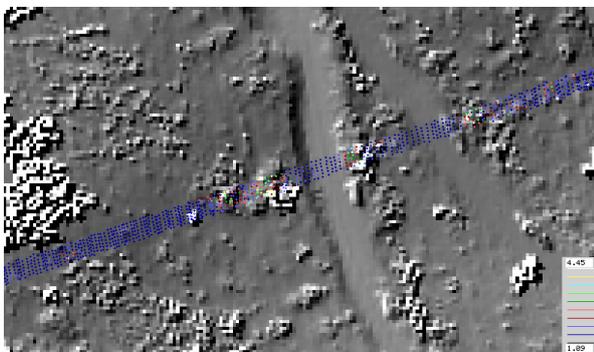


Fig. 5a: Shading of the DSM and selected last echo points within a profile coloured by the echo width

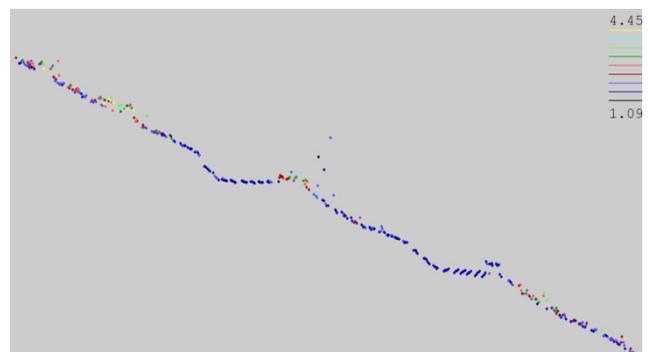


Fig. 5b: Profile view: Last echo points coloured by the echo width

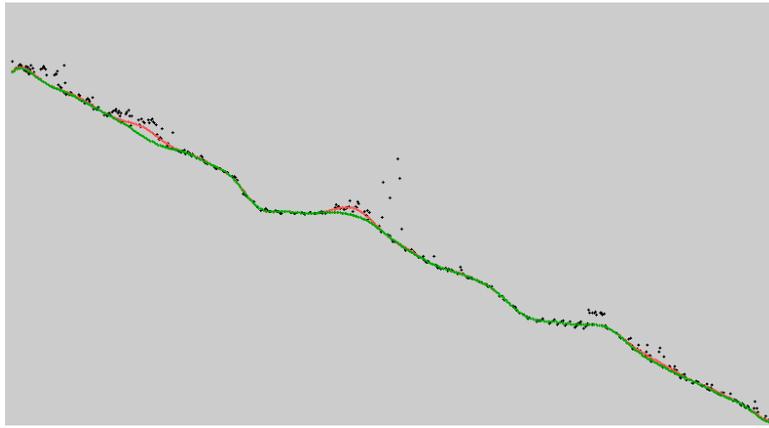


Fig. 5c: DTM generation by robust estimation with a-priori weights determined from the echo attributes (in this case just using the echo width) demonstrated on a profile; Black dots: Last echo point cloud, Red surface: DTM without considering individual weights. Green surface: Improved DTM by robust interpolation with individual weights.

6. CONCLUSION

The paper presents some of the current chances and challenges for DTM generation and data management. It especially focuses on the additional echo attributes that can be determined from full-waveform ALS data. Within the area of data administration the paper presents a concept that allows to store and access further information of the ALS flight mission as well as for each individual ALS echo. In this way it allows to link each echo with all stored attribute data, which is the basis for tracing the echo information within all the subsequent processing steps.

Next to advanced consideration in data management some thoughts for the improvement of DTM generation based on these additional attributes are presented. However, we have just presented some preliminary ideas and concepts which have to be extended and prove in the future. Still there is a need to study the additional parameters provided by the full-waveform sensors in much more detail. However, the presented results show that the provided information can be useful for DTM determination and other modelling tasks. In the future, additional attributes per echo might be additionally provided by other sensors (e.g. passive image data). Therefore, algorithms which allow to integrate geometric information provided by the ALS point cloud with additional information from active and passive sensors will be important for the future.

7. REFERENCES

- Attwenger, M., Chlaupek, A. (2006): Erstellung eines grenzüberschreitenden digitalen Geländemodells im Nationalpark Neusiedler See-Seewinkel/Fertö Hanság aus Laserscanner-Daten. *Angewandte Geoinformatik 2006, Beiträge zum 18. AGIT-Symposium*", J. Strobl, T. Blaschke, G. Griesebner (ed.), ISBN3-87907-437-2, pp. 16-23.
- Burman, H. (2002): *Laser Strip Adjustment for Data Calibration and Verification*. International Archives of Photogrammetry and Remote Sensing, Vol. 34/3, Graz, Austria.
- Csanyi, N., Toth, C. (2007): Improvement of Lidar data accuracy using lidar-specific ground targets. *Photogrammetric Engineering & Remote Sensing* 73(4), pp. 385-296.

- Crombaghs, M., Brügelmann, R., de Min, E. (2000): On the adjustment of overlapping strips of laseraltimeter height data. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, 3A, Amsterdam, Netherlands, pp. 230-237.
- Doneus, M., Briese, C. (2006): Digital terrain modelling for archaeological interpretation within forested areas using full-waveform laserscanning. In: M. Ioannides, D. Arnold, F. Niccolucci and K. Mania (Editors), The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST.
- Filin, S. (2003): Recovery of systematic biases in laser altimetry data using natural surfaces. *Photogrammetric Engineering & Remote Sensing*, 69.
- Hochstöger, F. (1996): Software for Managing Country-Wide Digital Elevation Data. In: International Archives for Photogrammetry and Remote Sensing, Vol. XXXI, Part B2, Vienna, Austria, 1996, pp. 160-163
- Hofton, M.A., Minster, J.B., Blair, J.B. (2000): Decomposition of Laser Altimeter Waveforms. *IEEE Transactions on Geoscience and Remote Sensing*, 38: pp. 1989-1996.
- Höfle, B., Pfeifer, N. (2007): Correction of laser scanning intensity data: Data and model-driven approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*. Accepted for publication.
- Hug, C., Wehr, A. (1997): Detecting and identifying topographic objects in imaging laser altimetry data. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXII/3-4W2, pp. 16-29.
- Hyyppä, J., Hyyppä, H., Litkey, P., Yu X., Haggrén, H., Rönholm, P., Pyysalo, U., Pitkänen, J., Maltamo, M. (2004): Algorithms and Methods of Airborne Laser-Scanning for Forest Measurements. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, Part 8/W2 (ISSN 1682-1750). Freiburg, Germany.
- Kager, H. (2004): Discrepancies Between Overlapping Laser Scanning Strips - Simultaneous Fitting of Aerial Laser Scanner Strips. International Archives of Photogrammetry and Remote Sensing, Vol. XXXV, B1. Istanbul, pp. 555-560.
- Kraus, K., Otepka, J. (2005): DTM Modeling and Visualization - The SCOP Approach. *Photogrammetric Week '05*, D. Fritsch (ed.). Herbert Wichmann Verlag, Heidelberg, pp. 241-252.
- Kraus, K., Pfeifer, N. (1998): Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 53, pp. 193-203.
- Kraus, K., Pfeifer, N. (2001): Advanced DTM generation from LIDAR data. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV, 3/W4, Annapolis, MD, USA, pp. 23-30.
- Loitsch J., Molnar L., (1991): A Relational Database Management System with Topological Elements and Topological Operators. In: *Proceedings of Spatial Data 2000*, Department of Photogrammetry and Surveying, University College London, 1991, pp. 260-269

Sithole, G., Vosselman, G. (2004): Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59. (1-2), pp. 85-101.

Wagner, W., Ulrich, A., Ducic, V., Melzer, T., Studnicka, N. (2006): Gaussian Decomposition and Calibration of a Novel Small-Footprint Full-Waveform Digitising Airborne Laser Scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60 (2), pp. 100-112.

Warriner, T., Mandlbürger, G. (2005): Generating a New High Resolution DTM Product from various Data Sources. *Proceedings of the 50th Photogrammetric Week*, D. Fritsch (ed.). Herbert Wichmann Verlag, Heidelberg. 3-87907-421-6; pp. 197-206.

www:

LAS (2007): LAS file format specification, <http://www.lasformat.org>