ACCURACY ASSESSMENT OF A LIDAR DIGITAL TERRAIN MODEL BY USING RTK GPS AND TOTAL STATION

Cornelis Stal[1], Timothy Nuttens[1], Jean Bourgeois[2], Leen Carlier[1], Philippe De Maeyer[1], and Alain De Wulf[1]

1. Ghent University, Department of Geography, Krijgslaan 281 (S8), 9000 Ghent, Belgium; {Cornelis.Stal, Timothy.Nuttens, Leen.Carlier, Philippe.DeMaeyer, Alain.DeWulf}(at)UGent .be
2. Ghent University, Department of Archaeology, Sint Pietersnieuwstraat 35 (UFO), 9000 Ghent, Belgium; Jean.Bourgeois(at)UGent.be

ABSTRACT

The aim of this research is to determine an accuracy assessment of a digital terrain model (DTM), derived from airborne laser scanning (Light Detection and Ranging or LiDAR). Samples of this LiDAR DTM with a resolution of 50 cm of the Mount Kemmel (Kemmelberg) in Belgium are compared with manually measured points using both Real Time Kinematic Global Navigation Satellite System (RTK GNSS) and total station. Airborne laserscanning is a well-known technique to acquire relatively accurate points in a very short timeframe over a large area. The Flemish Agency for Geographic Information (AGIV) provides statewide digital elevation models based on this technique. Although the resolution of the model of the Mount Kemmel (50 cm) is ten times higher than the standard models of the Agency (5 m), the same accuracy criteria are taken into account in this research, i.e. 20 cm for meadows, where the test sites are located. The proposed methodology consists of a comparison of this DTM with manually measured control points using RTK GNSS and total station. Since the last measurement techniques have a higher theoretical accuracy, it can be tested if the criterion of 20 cm is fulfilled and if LiDAR datasets are subsidiary with manually measured terrain points for these meadows, using a two sided *t*-test. The relation between these errors and the local slope of the topography are investigated as well. A full elaboration, describing the difference between and substitution of the LiDAR dataset and the total station dataset, is given in this paper.

INTRODUCTION

Before constructing a DSM (Digital Surface Model) or DTM, possible error sources should be taken into account and compensated or corrected. The most important errors are errors occurring during the acquisition of the data. The configuration of the acquisition equipment on the flying object (airplane, helicopter, etc.) contains three components, all generating its own errors. There are two components dealing with the navigation and positioning, namely a GNSS and an INS (Inertial Navigation System). There is also the LiDAR sensor itself. In addition to the sensor-specific errors, errors may occur due to electronic delays during the communication between devices, acquisition, storing data, etc. Therefore, a proper time synchronization is indispensible (1). Other errors with a more geometrical source are mainly caused by the flying height, scan angle (2) or the local topography (3). There are also important error sources during the post processing, the strip adjustment (4) and the interpolation to a discrete grid (5,6). If the point set is available without further notice about the acquisition process and strip adjustment, only the latter can be qualified. However, in this paper, the errors of a regular DTM are assessed in general, without respect to their error sources. This makes it possible to compare the LiDAR dataset with manually measured terrain data in a straightforward way. It is possible to describe the accuracy of a point cloud or interpolated DTM without any field work, by different quality parameters (7). The terrain has an important influence on the quality of the data, and the curvature is one of these parameters in an empiric-stochastic approach at motioned in (8,9), where multiple neighbouring cells play an important role in the quality assessment. Other parameters are point density or the standard deviation of the height values of different points in one cell. The distance to the nearest other point is used in a geometric approach to assess the quality of a point set at a sub cell level.

In this paper, the quality of the DTM is assessed by a control point set, acquired by RTK GPS and total station. By comparing this point set with the DTM, acquired by LiDAR, not only a quality assessment can be given, but the substitutability of LiDAR, in contrast with conventional techniques can be qualified as well, as described in the next paragraph. Thereafter, in the following paragraph, the relation between the local topography and the difference between the LiDAR and reference datasets are studied. On the south flank of the mount Kemmel (Belgium) (Figure 1), 697 points are measured, divided over an area of 2.4 hectare in order to construct a reference point set.

Figure 1: Impressions of the study area

DESCRIPTION OF THE DATASETS

The used LiDAR data set is a part of a larger airborne LiDAR campaign in the areas of the Mount Kemmel and Spiere-Helkijn in 2008, both in the Flemish region in Belgium. This campaign was ordered by the deputy of the province of West-Flanders, on behalf of archaeological research by Ghent University (10). The dataset has been acquired by a *LiteMapper 5600* at a flying height of 400 m and has a point density of approximately 4.5 points/m^2. The resulting point set is filtered into ground and non-ground points and the resulting sets are interpolated by inverse distance weighting (IDW) to respectively a DTM and DSM with a resolution of 50 cm of the test area (Figure 2). This interpolation technique is used by the AGIV as well. Since the point density and the resolution of the digital models are of comparable magnitude, this interpolation method will not result in significant errors and will not be discussed here. The used sample of the DTM is visualized in Figure 3.

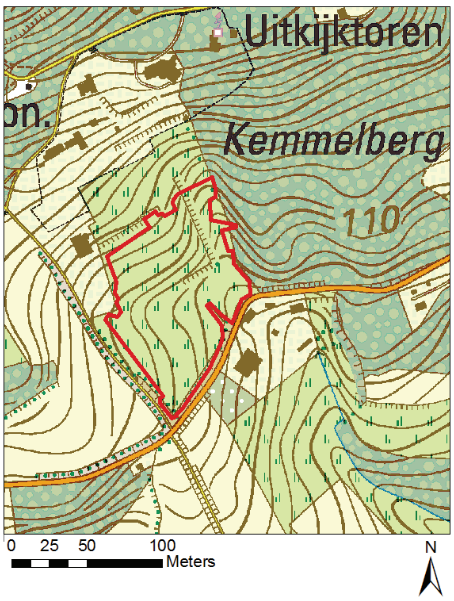

Figure 2: Overview of the study area (source: NGI).

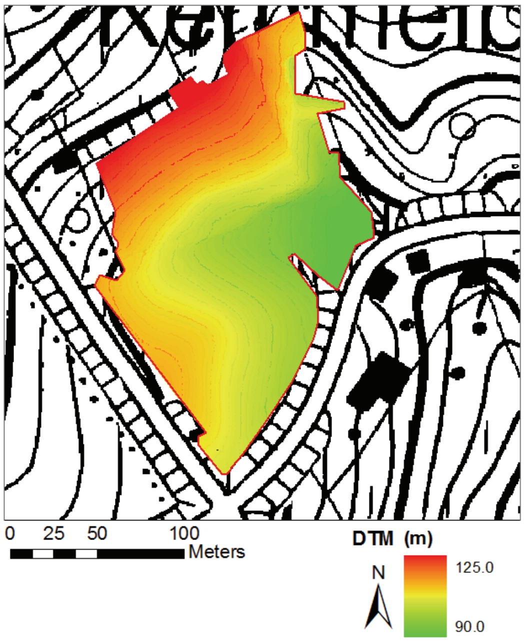

Figure 3: DTM of the study area, acquired by airborne LiDAR.

The AGIV requires an overall accuracy of 20 cm for their LiDAR measurements for meadows (11). To verify this quality criterion, 697 points are measured using RTK GPS with an accuracy of 3.2 cm planimetric and 8.6 cm altimetric. A total station has been used to support this control measurement as well, with an accuracy of 2.0 cm.

A full elaboration, describing the difference between the LiDAR dataset and the total station dataset, is given in this paper. The height difference of each point *i* and its corresponding cell *j* is calculated: $\Delta h_i = h_{ts,i} - h_{lidar,j}$. For a point *i*, the corresponding cell *j* is found by a spatial intersection. A table of all means and standard deviations is given (Table 1), based on a dataset containing 697 (*n*) measured points.

Table 1: Statistics of the datasets and its difference.

	$\bar{h}$ (m)	s^2
LiDAR	106.86	5.62
Total Station	106.88	5.61
Difference	-0.02	0.09

STATISTICAL ANALYSIS OF THE DATASETS

It is possible to measure the quality of a DTM in post processing (8). Determining the existence of a significant difference between two sets of height points can be performed using a two sided *t*-test, assuming a normal distribution of the differences. The *t*-test standardizes the values in the dataset and compares the transformed data with a theoretical student-*t* distribution. This assumption of normality is stated by the central limit theorem (12) and is visually stated by Figure 4, where a more or less equal spread is present around the peak in the middle of the histogram. The peak of the histogram does not seem to be located at $\mu_\Delta = 0$, but at $\bar{h}_\Delta = -0.02$. The significance of this difference will be tested using a *t*-test.

To perform a two sided *t*-test, the difference between two sets of heights has to be calculated. In a Geographic Information System (GIS) environment, the control sets of height points, measured by RTK GPS and total station, are draped on this raster. These datasets have respectively a mean $\bar{h}_{RTK}$ and $\bar{h}_{ts}$ and a standard deviation s_{RTK} and s_{ts}. This makes it possible to determine the height difference for every measured point, by subtracting the height of this point by the underlying cell value. Each cell, containing a point measured by RTK GPS or total station, will be a sample of the total DTM, acquired by LiDAR, for which a mean $\overline{h}_{lidar}$ and a standard deviation s_{lidar} can be calculated. A list of differences, with its own mean $\overline{h}_{\Delta}$ and standard deviation s_{Δ} is generated this way. The difference between the LiDAR dataset and the control dataset are mapped and visualized in Figure 5.

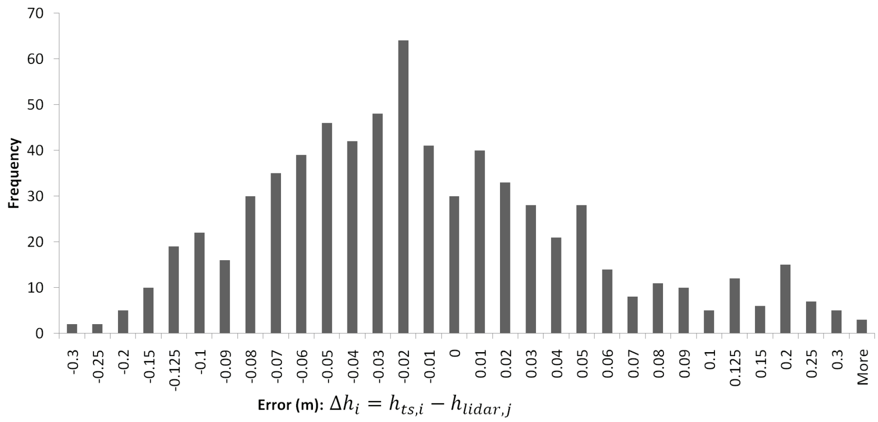

Figure 4: Distribution of the difference between total station and LiDAR measurements.

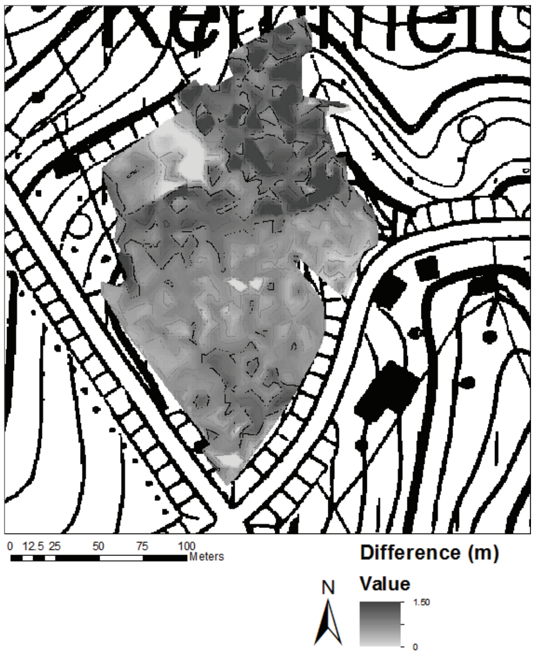

Figure 5: Spatial distribution of the differences in the study area.

Thereafter, a null hypothesis is set up to assess whether there is a significant difference in heights between the LiDAR dataset and the control sets or whether the difference between the two sets significantly differs from 0. Both datasets containing the RTK GPS and total station points are assessed separately.

$$H_0 : \mu_{\Delta} = 0$$
$$H_A : \mu_{\Delta} \neq 0$$

In words, the null hypothesis states that the height difference between the LiDAR dataset and the control dataset is equal to zero, in contrast with the alternative hypothesis, which states that the height difference between the two datasets is unequal to zero. The statistic

$$H_0 : t_{\left(v,\frac{\alpha}{2}\right)} \leq \frac{\overline{h}_{\Delta} - 0}{s_{\Delta}} \sqrt{n} \leq t_{\left(v,1-\frac{\alpha}{2}\right)}$$

is used to reformulate the null hypothesis. Using this statistic, the *t*-value of this sample is calculated:

$$t = \frac{\mu_\Delta}{s}\sqrt{n} = \frac{-0.02}{\sqrt{0.09}}\sqrt{697} = -1.50$$

The theoretical *t*-value has 696 degrees of freedom *v* and significance level α of 5% $t_{695,0.025} = 1.960$. Since $t \in [-1.960;1.960]$, the null hypothesis is accepted with confidence level of 95%. This means that the points measured by total station and LiDAR are statistically the same and no significant difference between the two sets exists. It can be concluded that at least 95% of the possible difference between the height value of the LiDAR dataset and the control dataset is caused by a random effect.

CORRELATION BETWEEN HEIGHT DIFFERENCE AND SLOPE

A possible error source of DTM's, generated using LiDAR point sets, is the horizontal displacement (3) (Figure 6). Since LiDAR is not an image-generating technique, it is very hard to determine the value of this displacement. The horizontal displacement will cause a vertical displacement as well. It is obvious that both displacements are correlated as a function of the topographic slope α (13). For each cell the slope is calculated by taking the maximal height difference of the neighbouring cells, divided by the resolution of the DTM. The correlation between this calculated value and the error of the DTM can be determined by a regression analysis. The error of the DTM is supposed to be the same as the height difference between the datasets, acquired by LiDAR and RTK GPS or total station measurements. In the following, the measurement using total station are taken into account.

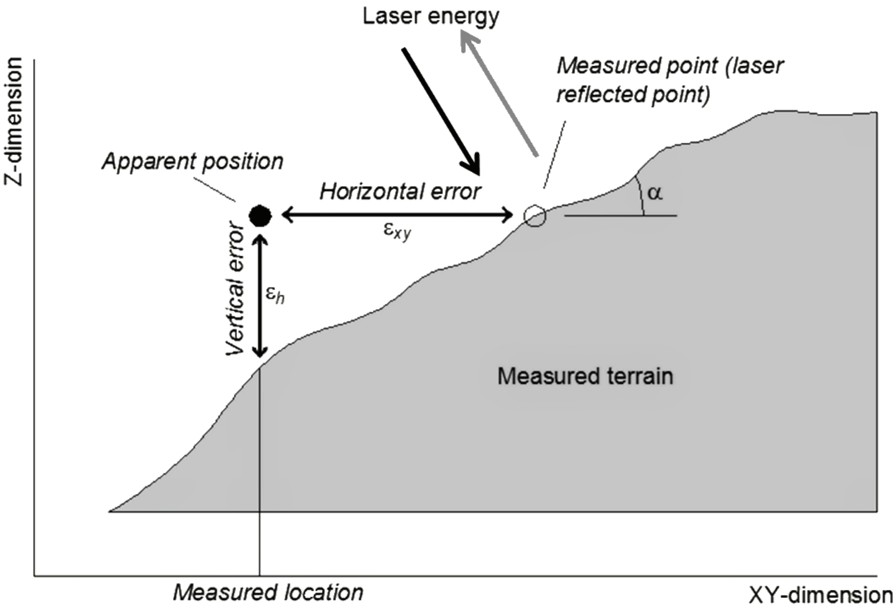

Figure 6: Horizontal and vertical error on LiDAR points, in function of the horizontal displacement and local topography (adapted from (6) with modifications).

For each measured point a slope α (in degrees) is calculated (Figure 7), within the interval of $[1.29;49.58]$, with a mean slope $\bar{\alpha} = 10.06$. Based on Figure 8, it is clear that the slope in this area does not follow a normal distribution. Although the absence of this normal distribution, a linear regression is calculated for the vertical error or height difference Δh, as a function of the slope α, so $\Delta h_i \sim \alpha_i$. Before the linear function can be calculated, an outlier detection has to be executed on the height error.

The following criterion is used for outlier detection, as mentioned in (14):

$$|\Delta h_i| \leq 2.5 \times \sqrt{2} \times s_\Delta$$

Based on the filtered data, a regression line is calculated, and both the point set and this line are plotted in a scatterplot (Figure 9). The Pearson's correlation coefficient is calculated as well, having a value of $R^2 = 0.142$. This means that only 14.2% of the total variance of the height difference is explained by the slope of the terrain. A transformation of the slopes by calculating the logarithm or calculating a second order polynominal trendline does not have a large influence on the stated conclusion (respectively $R^2 = 0.146$ and $R^2 = 0.156$). A division of the slopes into specified intervals would not be useful as well, since this will negate the randomness of the dataset. Therefore it can be stated that the relation between the two parameters has no significant influence on this dataset, in contrast with the hypothesis of Hogdson & Bresnahan (6). This could be caused by the high density of the measured control points or the high resolution of the DTM, in relation with the much lower decline of the topography. Another parameter could be the low variance in the slope data, caused by the smoothing effect of the interpolation.

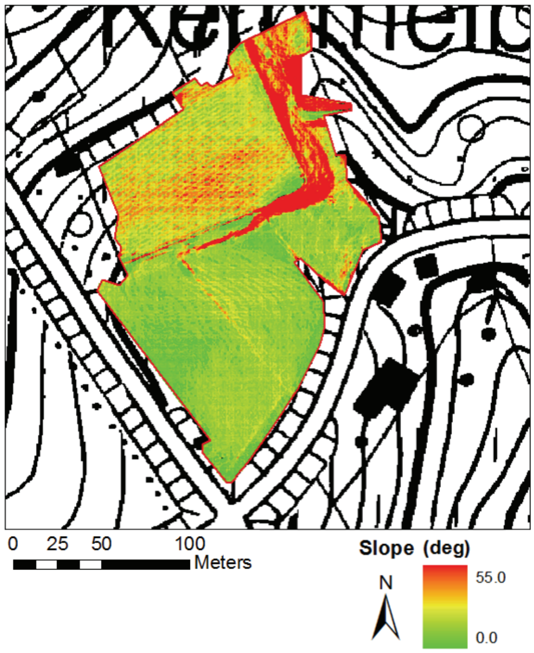

Figure 7: Slope of the study area, calculated by the LiDAR DTM.

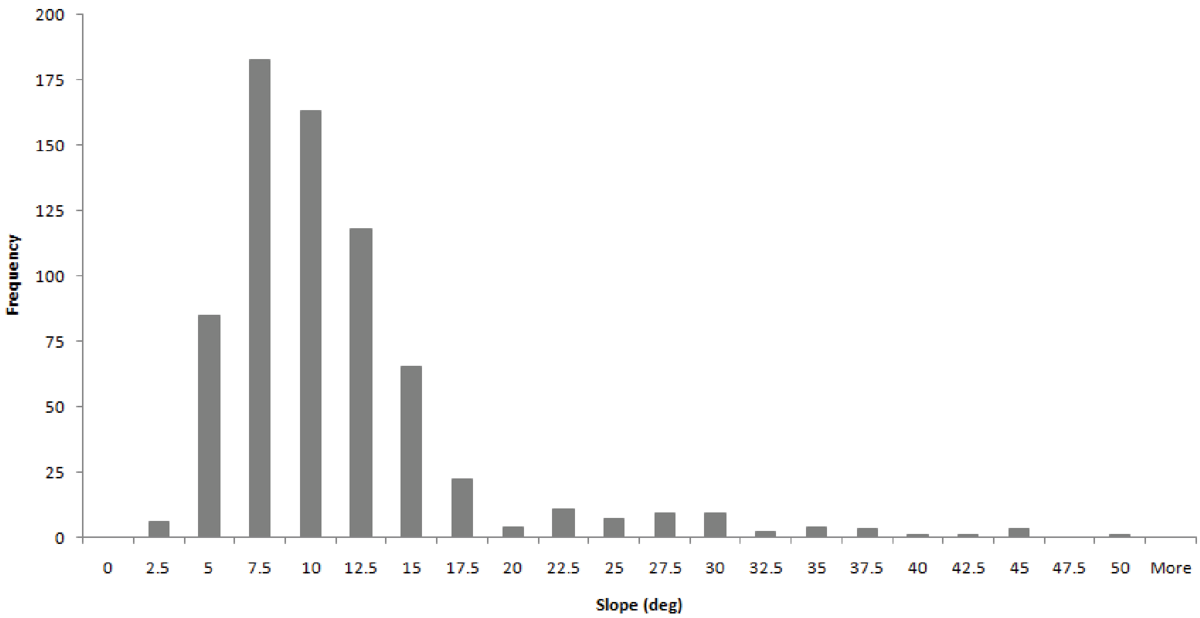

Figure 8: Distribution of the slope.

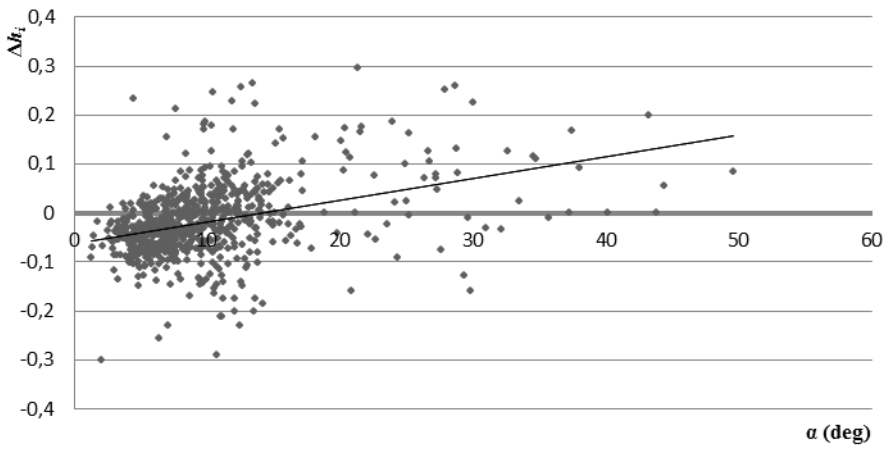

Figure 9: Squared error in relation with the slope.

CONCLUSION

Taking into account the requirements of the AGIV, it can be concluded that the interpolated DTM, acquired by LiDAR, fulfils the maximal vertical error criterion of 20 cm for meadows. The control datasets, measured by GPS RTK and total station, having a standard error much lower than this criterion, demonstrate no significant difference with the LiDAR DTM. The mount Kemmel is one of the steepest hills in the Flanders region. The test site was located on a slope with relatively high inclination. Nevertheless, no significant relation is assessed between the vertical error and the slope.

ACKNOWLEDGEMENTS

This work forms a part of a bigger archaeological research on the Mount Kemmel, executed by the department of Archaeology of Ghent University, with technical support by the Department of Geography. Our special thanks go to the province of West Flanders and provincial permanent delegate G. Perty in particular.

REFERENCES

1 Würländer R & K Wenger-Oehn (2007) Die verfeinerte Georeferenzierung von ALS-Daten - Methodik und praktische Erfahrungen. In: Strobl J, T Blaschke & G Griesebner (Eds.) Angewandte Geoinformatik 2007 (Heidelberg: Wichmann) 6 pp. http://www.wenger-oehn.at/pdf/agit2007.pdf (last date accessed: 20 Feb 2011)

2 Katzenbeisser R (2003) About the calibration of LiDAR sensors. ISPRS Workshop on 3D Reconstruction from ALS and InSAR Data (Dresden, Germany) 6 pp.

3 Hodgson M & P Bresnahan (2004) Accuracy of airborne LiDAR derived elevation: Empirical assessment and error budget. Photogrammetric Engineering and Remote Sensing 70(3): 331-339

4 Attwenger M & C Briese (2003) Vergleich digitaler Geländemodelle aus Photogrammetrie und Laserscanning. Österreichische Zeitschrift für Vermessung und Geoinformation, 91(4): 271-280

5 Sithole G & G Vosselman (2003) Comparison of filtering algorithms. ISPRS Workshop on 3D Reconstruction from ALS and InSAR Data (Dresden, Germany) 8 pp.

6 Wack R & A Wimmer (2002) Digital terrain models from airborne laser scanner data - a grid based approach. ISPRS Symposium on Photogrammetric Computer Vision (Graz, Austria) 4 pp.

7 Kraus K, W Karel, C Briese & G Mandlburger (2006) Local accuracy measures for Digital Terrain Models. The Photogrammetric Record, 21(116): 342-354

8 Kraus K, C Briese, M Attwenger & N Pfeifer (2004) Quality Measures for Digital Terrain Models. ISPRS Congress Istanbul 2004, Proceedings of Commission II: 113-118

9 Kraus K, M Attwenger, C Briese & G Mandlburger (2004) Qualitätmaße für digitale Geländemodelle am Beispiel eines Photogrammetrie- und eines Laserscannerprojektes. Tagungsband der 24. Wissenschaftlich-Technischen Jahrestagung der DGPF. Halle, Germany

10 Stal C, J Bourgeois, P De Maeyer, G De Mulder, A De Wulf, R Goossens, T Nuttens, & B Stichelbaut (2010) Kemmelberg (Belgium) case study - Comparison of DTM analysis methods for the detection of relicts from the first world war. Proceedings of 30th EARSeL Symposium (Paris, France) p. 65-72

11 AGIV (2003) Nieuwsbrief 16: Digitaal hoogtemodel Vlaanderen (in Dutch) (Brussels: GIS-Vlaanderen) 28 pp.

12 Moore D, G McCabe & B Craig (2009) Introduction to the Practice of Statistics (New York: W.H. Freeman)

13 Podobnikar T (2009) Methods for visual quality assessment of a digital terrain model. Surveys and Perspectives Integrating Environment and Society, Special Issue: Visualizing the World, Gadal S (Ed.) 2(2): 19

14 Iso/Cd 17123-8:2007 (2007) Optics and Optical Instruments - Field Procedures for Testing Geodetic and Surveying Instruments. Part 8: GNSS Field Measurements in Real Time Kinematic (RTK)