Comparison of Height Models from High Resolution Aerial Images and from LiDAR

Karsten Jacobsen
Leibniz University Hannover, Institute of Photogrammetry and Geoinformation, Germany;
jacobsen@ipi.uni-hannover.de

Abstract. A digital Surface Model (DSM) has been generated with an Optech Gemini laser scanner with approximately 4 points/m² in an area also covered by an UltraCam Eagle with 5cm GSD. Based on the digital images DSM and digital terrain models (DTM) were automatically/interactively produced with BAE SYSTEMS "NGATE" (New Generation of Automatic Terrain Extraction) and also SimActive's "Correlator3D" with 50cm point spacing. A direct comparison of the overlapping LiDAR strips as well as the comparison of the LiDAR data with the height models based on the aerial data showed differences especially in areas covered by vegetation as well as at building limits, requiring a filtering for the correct comparison. As for all height models it is not so simple to specify the accuracy figures correctly. The standard deviation as well as the normalized median absolute deviation (NMAD), LE90 and LE95 by theory is based on normal distributed discrepancies. Normal distributed height discrepancies require corresponding point definition in compared height models and within the height models similar conditions as e.g. flat terrain. Operationally it is very difficult to filter height models in a manner leading to normal distributed values, so the frequency distribution of height discrepancies has to be investigated. The standard deviation is very sensitive for higher number of larger discrepancies which usually cannot be avoided. NMAD is not as sensitive as the standard deviation and describes the frequency distribution of height discrepancies in a better manner. The threshold values LE90 and LE95 are even depending upon larger discrepancies as the standard deviation and should be avoided. LiDAR and image matching both have advantages and disadvantages, with both methods qualified height models can be generated.

Keywords. DEM, digital aerial images, LiDAR, analysis

1. Introduction

The geometric analysis of height models includes the accuracy analysis by comparing one height model with a reference height model having higher or at least same accuracy. With LiDAR in areas with not too dense vegetation the height of the bare ground can be determined, but not any last pulse height value is located on the bare ground, requiring filtering for points not belonging to the bare ground. In any case buildings are included in the last pulse height model, requiring filtering for buildings. If a DSM is requested, first pulse height values are used, but the height definition of the vegetation is difficult – in overlapping LiDAR strips larger height discrepancies can appear in vegetation areas. Also at building limits problems with height definition may exist – with same horizontal position from one flight line a point may be located on roof overhang, from another flight line it may be located on the ground or on the façade (figure 3).

If two height models are compared, usually points have not exactly the same position, requiring an interpolation at least in one height model. The loss of accuracy by interpolation depends upon the terrain roughness. The Hannover program DEMANAL for of digital height model (DHM) analysis is generating a raster of height values based on the reference data if the reference height model includes randomly distributed points. The height corresponding to a random point in the data set to be analyzed will be interpolated by bilinear interpolation in the reference height model. Because of the point spacing between 0.5m in the analyzed files, the accuracy loss by interpolation is limited.

2. Analysis of Lidar Data

Figure 1 shows a typical part of a LiDAR-DSM taken by an Optech Gemini in New Jersey in a single house settlement. Beside individual houses typical vegetation as bushes, hedges and trees can be seen together with terraces and cars on the streets.

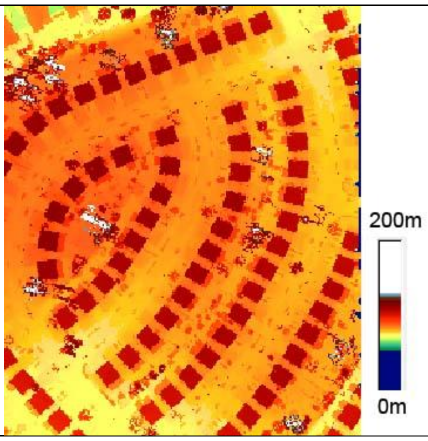
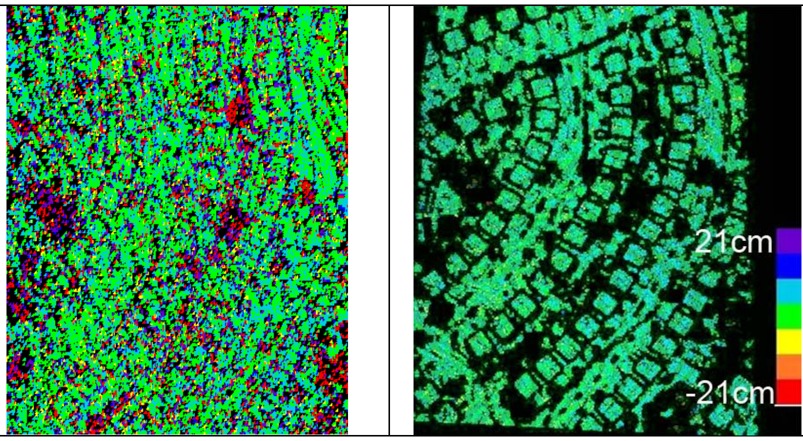


Figure 1: color coded sub-area of original LiDAR-DSM



Original LiDAR data   LiDAR data after filtering

Figure 2: color coded height differences LiDAR strip 1 against strip 2, left color scale +/-42cm

The height differences for a sub-area of LiDAR strip 1 against strip 2 of the New Jersey data set are illustrated in figure 2. Especially in the vegetation and around buildings the original data are showing a higher percentage of larger discrepancies. 0.3% of the discrepancies are exceeding 20m and 2.3% 10m. The root mean square difference of 2.7m has nothing to do with the accuracy of the DSM, it just express the different determination of the vegetation and building limits (figure 3) in neighbored LiDAR strips. The NMAD (Höhle et al. 2009) with 24cm is far away from the standard deviation and LE90 (threshold for 90% of absolute differences) with 81cm is quite below the standard deviation, demonstrating that we are far away from normal distribution of discrepancies (figure 5, left).

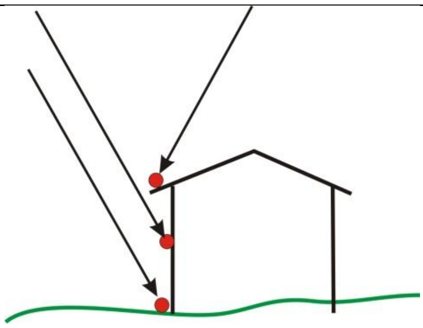


Figure 3: object points at building limits

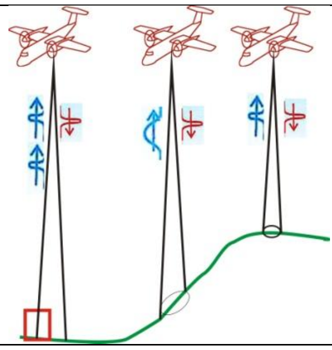
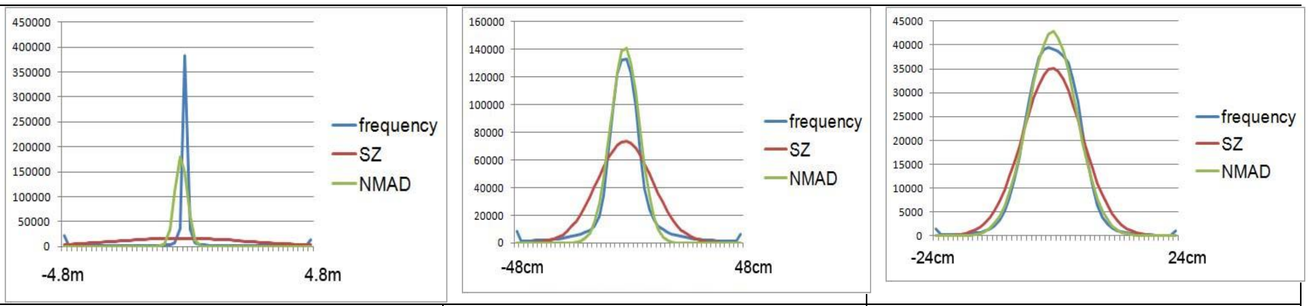


Figure 4: LiDAR pulse and echo from footprint

The threshold limits LE90 and LE95 do not present the whole population of height discrepancies. According to the Joint Committee for Guides in Metrology (JCGM) – a working group of some international metric organizations, including the International Standard Organization (ISO) – by this reason LE90 and LE95 should not be used for the accuracy definition (JCGM 100:2008). In addition the variation of these threshold limits from data set to data set varies stronger as SZ and NMAD and as mentioned above, it differs quite a lot from the expected relation to SZ and NMAD.



Original data, 20m threshold	Original data, 70cm threshold	Filtered data

Figure 5: frequency distribution of height discrepancies LiDAR strip 1 against strip 2 together with normal distribution based on SZ and NMAD

The frequency distribution of the original height discrepancies (figure 5, left), shown as blue line, demonstrates that it is far away from normal distribution plotted based on the standard deviation as red line and based on NMAD as brown line. If only height discrepancies up to 70cm are accepted (figure 5, center) the normal distribution based on NMAD is close to the real frequency distribution. In this case the standard deviation is reduced to 12.2cm and NMAD to 6.3cm. As function of the terrain inclination α the standard deviation can be expressed as SZ = 4.8cm + 22.4cm $*$ tan α and NMAD = 3.9cm + 17.3cm $*$ tan α for this case. The dependency upon terrain inclination is also caused by the not so good definition of the return pulse in inclined area (figure 4). For flat terrain – not influenced by the rough vegetation and building limits - SZ is limited to 4.7cm and NMAD to 3.7cm. Nevertheless it is better to filter the DSM for vegetation and building limits. The Hannover program RASCOR (Passini et al. 2002, Day et al. 2013) has been extended by a special vegetation filter based on the root mean square differences of neighbored heights in a specified template. This filter eliminates the height values of vegetation areas and at building limits as it can be seen by the black pixels in figure 2, right.

The vegetation filter reduces the discrepancies between LiDAR strip 1 and strip 2 to a standard deviation of 6.1cm or 5.0cm NMAD. LE90 is reduced to 8.0cm and LE95 to 9.8cm. As function of the terrain inclination: SZ = 5.1cm + 28cm $*$ tan α and NMAD = 2.5cm + 30cm $*$ tan α. For the filtered DSM the differences between strip 1 and strip 2 are not far away from the normal distribution (figure 5, right). The normal distribution based on NMAD is very close to the frequency distribution, only the higher number of larger discrepancies avoids that the normal distribution based on the standard deviation (red line) does not fit to the frequency distribution.

Table 1 gives an overview about the comparison of 7 overlapping LiDAR DSM after vegetation filter. The results are similar for all strip combinations. The average of the absolute values of the bias is negligible, caused by a pre-adjustment of the LiDAR-strips. This pre-adjustment is based on flat areas with a minimal value for the inclination; that means roofs are used. Not in any case enough roofs are available, requiring a fit of the LiDAR strips based on the vegetation filter. Nevertheless a significant error of the inclination in flight direction of up to 10cm over 2.5km was remaining after fitting based on roof height values while the roll error was negligible.

The rotation of strip 2 in relation to strip 1, after vegetation filter, improves the standard deviation of the differences just from 6.1cm to 6.0cm, but the NMAD from 5.0cm to 2.5cm. That means the systematic orientation errors have a strong influence to NMAD, while the standard deviation is dominated by the larger discrepancies. The improvement of the individual LiDAR strips by the rotation depends upon the size of the rotation. In the average (table 2) the standard deviation is just reduced by 8% while NMAD is reduced by 29% to 3.0cm. For the flat terrain the standard deviation is improved from 4.8cm to 4.5cm as well as NMAD for the flat terrain from 2.6cm to 1.9cm. This size can be explained by interpolation over the 50cm point spacing and the terrain roughness.

combination	SZ	NMAD	SZ as F(slope)	NMAD as F(slope)
Strip 1 – 2	6.1cm	5.0cm	5.1cm + 28cm ∗ tan α	2.5cm + 30cm ∗ tan α
Strip 2 – 3	6.2cm	4.9cm	5.1cm + 15cm ∗ tan α	2.5cm + 17cm ∗ tan α
Strip 3 – 4	4.6cm	3.3cm	3.9cm + 14cm ∗ tan α	2.5cm + 15cm ∗ tan α
Strip 4 – 5	5.5cm	3.5cm	4.1cm + 21cm ∗ tan α	2.5cm + 21cm ∗ tan α
Strip 5 – 6	6.9cm	3.6cm	4.3cm + 22cm ∗ tan α	2.5cm + 24cm ∗ tan α
Strip 6 – 7	9.6cm	5.1cm	6.3cm + 8cm ∗ tan α	3.0cm + 9cm ∗ tan α
average	**6.5cm**	**4.2cm**	**4.8cm + 18cm ∗ tan α**	**2.6cm + 19cm ∗ tan α**

Table 1: comparison of overlapping LiDAR DSM after vegetation filter – New Jersey

	SZ	NMAD	SZ as F(slope)	NMAD as F(slope)
average	**6.0cm**	**3.0cm**	**4.5cm + 18cm ∗ tan α**	**1.9cm + 23cm ∗ tan α**

Table 2: comparison of overlapping LiDAR DSM after vegetation filter and rotation of strips

For normal distributed discrepancies SZ and NMAD have the same value. This is not the case for the listed results, indicating the higher number of larger discrepancies as shown in figures 5. Because of the square sum formula large discrepancies influence the standard deviation strongly, while the median nature of NMAD is dominated by majority of the frequency distribution. The frequency distributions of the LiDAR strip differences are very similar for all strip combinations. In the color coded images of height differences of strip 1 against strip 2 (figure 2) as main reason for larger height differences the remaining influence of the vegetation is obvious.

The relation between the threshold value for 90% of the observations (LE90) to the standard deviation after rotation and filtering (table 2) is just 1.07. In case of normal distributed values this relation is 1.65. For LE95 the computed relation is 1.94, close to expected relation of 1.96. LE95 is 3.5 times larger as NMAD and LE95 6.4 times larger. This confirms the not normal distributed character of the discrepancies and demonstrates that the threshold values LE90 and LE95 are not reliable.

The question, what is the accuracy of the LiDAR heights, has to be raised. According to the fit of the normal distribution based on NMAD to the major part of the height discrepancy frequency distribution the answer is: the relative system accuracy of the LiDAR heights corresponds approximately to NMAD. The higher number of larger discrepancies is caused by different object point definition in neighbored LiDAR flight lines depending upon the view direction. In case of vegetation and building limits it is a random result, which part of the footprint of approximately 30cm diameter is determined as first pulse reflection.

3. DSM from high resolution aerial images

In the same area as the above mentioned LiDAR flight, images with 5cm GSD have been taken with an UltraCam Eagle with 60% endlap and 60% sidelap. DSM are automatically/interactively produced by BAE SYSTEMS "NGATE" (New Generation of Automatic Terrain Extraction) and SimActive's "Correlator3D" with 50cm point spacing. Of course with 5cm GSD a denser grid can be generated, but it was not requested for the operational project.

The image orientation by bundle block adjustment based on a limited number of ground control points resulted at independent check points to 5.6cm standard deviation in Z. This vertical accuracy of approximately 1.0 GSD is operationally and can be handled as system accuracy for points with satisfying contrast in flat areas. The automatic image matching and filtering was made by two companies with different image combinations. As it can be seen in figure 6, the area is a combination of single house settlement, a large shopping mall and small forest areas.

The height differences of the DSM generated by NGATE against DSM from Correlator3D (figure 7) show clear differences especially in the vegetation areas. The used digital images are from one photo flight, so the differences only can be explained by different methods used by the matching programs and different image combinations. In addition small problems of image orientation are indicated, which may be caused by different handling of systematic image errors. On top of large roofs with poor contrast, especially the shopping mall located in the centre of figures 6 and 7, larger discrepancies can be seen. With a threshold for accepting height discrepancies of 1m, the standard deviation reaches 24cm and NMAD 13cm. The large difference between SZ and NMAD is a clear indication of not normal distributed discrepancies. There are quite more large discrepancies as according to normal distribution. In addition 20.8% of the discrepancies are exceeding the threshold of 1m.

Also a comparison of the DTM from NGATE and Correlator3D did not lead to optimal results. Even with a threshold for accepting discrepancies of 0.7m the standard deviation is just 22cm and NMAD 17cm. 17.7% of the discrepancies are exceeding 0.7m. By this reason both DSM have been filtered by Hannover program RASCOR from DSM to DTM. Even without manual interaction a better result by comparing both DTM was achieved.

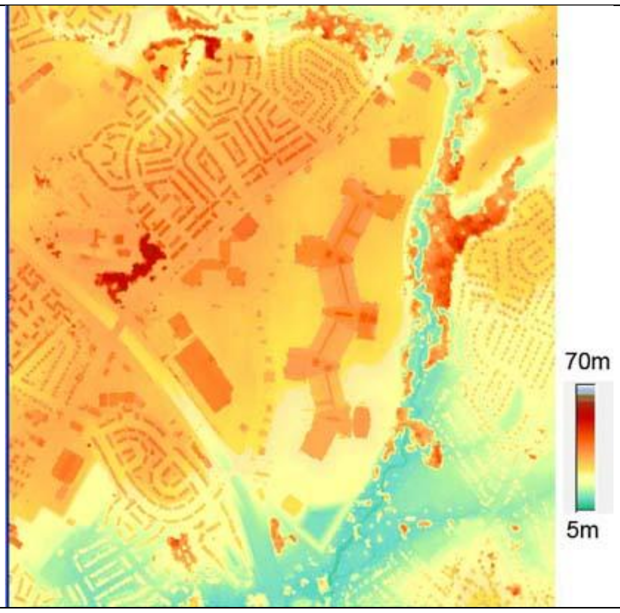


Figure 6: DSM generated by image matching with NGATE

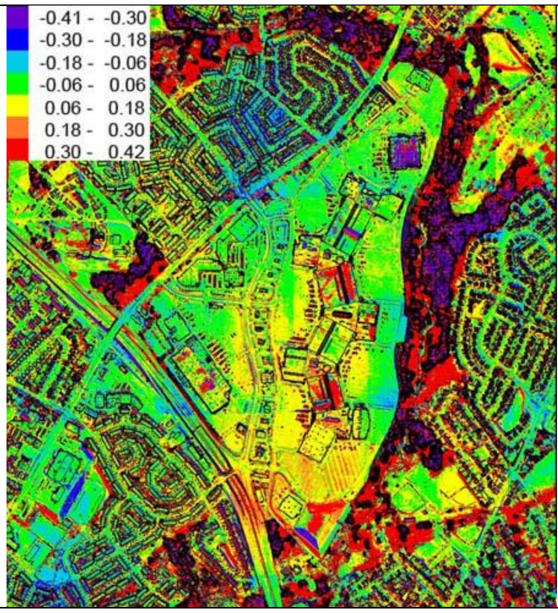


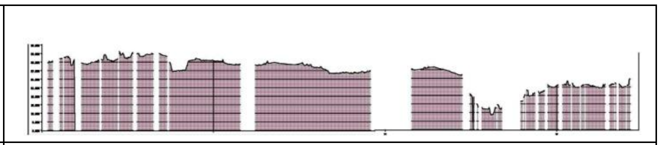
Figure 7: color coded height differences of DSM generated by NGATE and by Correlator3D

Not all elements not belonging to the bare ground have been eliminated by the filtering with RASCOR, but the result is still better as achieved by the filtering of the companies handling NGATE and Correlator3D. Figure 8 demonstrates the effect of filtering by RASCOR from DSM to DTM – all obstacles in the profile have been eliminated.

Height model	SZ	NMAD	>0.7m
DSM	18.5cm	12.5cm	23,4%
DTM	21.8cm	17.4cm	17.7%
DTM filtered	19.6cm	13.1cm	6.5%

Table 3: comparison of NGATE and Correlator3D DHM

Figure 8: profile through NGATE DSM



profile after filtering with RASCOR

The frequency distribution of the discrepancies between NGATE and Correlator3D filtered by RASCOR to DTM (figure 9) shows a quite better fit to the normal distribution based on NMAD, nevertheless a higher number of larger discrepancies exist. This is also obvious by the 6.5% of discrepancies exceeding 0.7m (table 3). Of course this percentage is quite below the 17.7% of the original DTM.

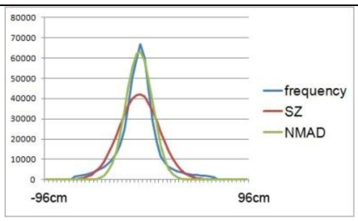
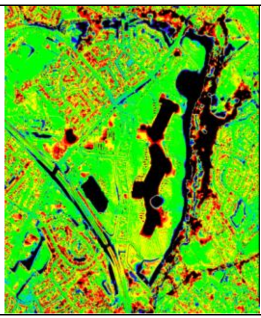
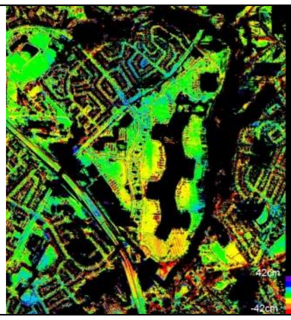


Figure 9: frequency distribution of the discrepancies between NGATE and Correlator3D filtered by RASCOR to DTM



originally from companies,



DSM filtered by RASCOR

Figure 10: color coded height differences between NGATE and Correlator3D DTM

height range of color scale: -42cm up to 42cm

Height model	SZ	NMAD	>0.7m
NGATE DSM	17.3cm	12.5cm	7.1%
NGATE DTM	17.5cm	12.9cm	12.1%
NGATE DTM filtered by RASCOR	14.5cm	10.5cm	1.1%
Correlator3D DSM	16.9cm	12.0cm	4.5%
Correlator3D DTM	17.9cm	12.8cm	10.0%
Correlator3D DTM filtered by RASCOR	12.5cm	13.9cm	1.1%

Table 4: comparison of DHM from aerial images with LiDAR

Height model	SZ	NMAD	>0.7m
NGATE DSM	16.6cm	11.3cm	4.0%
NGATE DTM	16.6cm	11.1cm	10.2%
NGATE DTM filtered by RASCOR	13.7cm	8.9cm	1.7%
Correlator3D DSM	13.4cm	9.3cm	1.9%
Correlator3D DTM	15.7cm	10.0cm	18.9%
Correlator3D DTM filtered by RASCOR	18.6cm	12.7cm	10.5%

Table 5: comparison of DHM from aerial images with object points from bundle block adjustment

A comparison of the different height models based on aerial images with the LiDAR-data (table 4) shows accuracy numbers slightly better as the comparison of the height models from NGATE and Correlator 3D. The DSM filtered by RASCOR – shown in table 4 as “DTM filtered” - is better as the original data sets from the companies. In both cases the percentage of discrepancies exceeding the used threshold of 0.7m is quite smaller for the RASCOR results as for the originally DTM generated by the companies.

In the investigated area object coordinates from tie points of the bundle block adjustment are available - they also have been used for the accuracy analysis (table 6). The comparison of the height models based on aerial images against tie point object coordinates is leading to slightly better results as the comparison with the LiDAR points. This is not a surprise because the tie points are also based on matching of aerial images; nevertheless it confirms the achieved results.

Again the question about the accuracy number characterizing the data sets has to be raised. In general the normal distribution based on NMAD (see also figure 9) fits to the frequency distribution of the height discrepancies quite better as the normal distribution based on the standard deviation. Only the number of larger discrepancies exceeds the normal distribution. This can be explained by remaining effects of the vegetation which cannot be described by an accuracy number.

4. Conclusion

Digital surface models, even based on the same method, cannot be compared directly because of different definition of the visible surface, especially in areas covered by bushes and trees. Also at building limits the height definition in LiDAR data from neighbored flight lines is not the same, requiring for a comparison a filtering of vegetation and building limits. The automatic filtering from DSM to DTM is complex. Better results have been achieved by Hannover program RASCOR as with the operator supported methods used by NGATE and Correlator3D.

For reaching the highest relative accuracy of the LiDAR data the data strips required an improvement of (remaining) rotations across and in flight direction.

The accuracy description of the generated height models is dominated by remaining influence of the vegetation causing more large discrepancies as according to normal distribution. This strongly influences the standard deviation, which by the used square sum is dominated by larger discrepancies. This is not so much the case for the normalized median absolute deviation (NMAD), based on the 68% median of the discrepancies corresponding to 68% probability as the standard deviation for normal distributed values. The normal distribution based on NMAD describes the frequency distribution of the analyzed discrepancies in a satisfying manner, while the standard deviation leads to too pessimistic results. The threshold values LE90 and LE95 do not present the whole population of discrepancies; they are sensitive for the number of larger differences and are strongly varying from data set to data set. They cannot be recommended. The most realistic description of the accuracy of height models can be made with NMAD.

Based on the comparison of overlapping LiDAR strips the system accuracy of the LiDAR height models is in the range of NMAD =3cm or for flat areas even 2cm ($NMAD/\sqrt{2}$) while the system accuracy of the height models based on digital aerial images with 5cm GSD is in the range of 9cm to 12cm ($NMAD/\sqrt{2}$); for well defined points with good contrast the accuracy is even in the range of 5cm. For inclined terrain in any case the accuracy is lower and independent upon the data source the accuracy in vegetation areas is not as good as in open areas. The accuracy requirement of natural ground is in the range of 10cm, only for artificial areas as streets, flat grassland and buildings the accuracy can be higher because of low roughness.

The used LiDAR data set have approximately 4 points/m² corresponding to 50cm point spacing and a foot print diameter of 30cm. The same point spacing was also used for the matching of the aerial images because of the practical application of this data set. Nevertheless by area based matching of the 5cm GSD images a point spacing of 15cm would be possible or with Semi Global Matching even 10cm point spacing; that means, the location of well defined objects can be defined more precise with aerial images as with LiDAR.

Height models determined by LiDAR or matched aerial images both have advantages and disadvantages. Both methods can be used; the decision for the method finally is only justified by secondary reasons.

References

[1] Day, D., Jacobsen, K., Passini, R., Quillen, S., 2013: A Study on Accuracy and Fidelity of Terrain Reconstruction after Filtering DSMs produced by Aerial Images and Airborne LiDAR Surveys, ASPRS Annual Convention Baltimore 2013

[2] Höhle, J. and Höhle, M., 2009. Accuracy assessment of digital elevation models by means of robust statistical methods, ISPRS Journal of Photogrammetry and Remote Sensing, 64, pp. 398-406

[3] JCGM 100:2008 : Evaluation of measurement data – Guide to the expression of uncertainty in measurement, http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf

[4] Passini, R., Betzner, D., Jacobsen, K., 2002: Filtering of Digital Elevation Models, ASPRS annual convention, Washington 2002