A NEW APPROACH TO DTM ERROR ESTIMATION BASING ON LAPLACIAN PROBABILITY DISTRIBUTION FUNCTION

Beata Hejmanowska [1] , Simon Kay [2]

[1] European Commission JRC, Institute for Environment and Sustainability, Via Enrico Fermi 2749, I-21027 Ispra (VA), Italy – (beata.hejmanowska simon.kay)@jrc.ec.europa.eu
[1, 2] University of Science and Technology, Faculty of Mine surveying and Environmental Engineering, Krakow, Poland, al. Mickiewicza 30, 30-059 Kraków, Poland
galia@agh.edu.pl

KEY WORDS: Aerial, LIDAR, Accuracy, Point Cloud, Laser scanning

ABSTRACT: A Digital Terrain Model (DTM) derived from Airborne Laser Scanning (ALS) was the subject of our research. In this paper, the vertical accuracy of the DTM was analyzed on the basis of the commonly used statistics, i.e. mean error and standard deviation, assuming a normal (Gauss) error distribution. The further approach, the so-called robust method (Höhle, Höhle 2009), was also tested, where the median was a substitute for the mean error and the Normalized Median Absolute Deviation (NMAD) for the standard deviation. An alternative method based on the Laplace function is proposed in the paper for describing the probability density function, where the parameters of the Laplace function are proposed for DTM error estimation. The test area was near the Joint Research Centre in Ispra, Italy; raw ALS data covering the test area were collected in 2005 and processed for DTM generation. Accuracy analysis was performed based on the comparison of DTM with the raw ALS data and with in-situ height measurements. The distribution of DTM errors calculated from ALS data was significantly non-normal, confirming other results reported in the literature. The Gauss distribution function considerably overestimated the vertical DTM errors; however, the robust method underestimated them. The Laplace function matched the error histograms the best, and accuracy parameters derived from this function could be considered as an alternative method for DTM accuracy evaluation.

1. INTRODUCTION

The accuracy of any geo-database used for decision support is an important issue, especially if the decisions have financial consequences or are connected with some risk or liability. The accuracy of Digital Terrain Models (DTM), an important data layer in most geographic information systems (GIS), is for example crucial for flood prevention and for minimizing consequences of a flood. To reduce such risks, a DTM of the appropriate accuracy specification, with determined error or, better, the spatial distribution of DTM error should be applied. Spatial distribution of DTM vertical error is described in detail by previous authors (e.g. Fisher and Tate 2006, Kraus et al. 2006, Oksansen 2006, Wechsel 2001). However, usually a representative estimator of the error is a useful indicator of this quality and no general agreement exists on this approach for Airborne Laser Scanning derived products.

Automatic measurements (matching by photogrammetry or laser scanning), applied to DTM generation, usually deliver many more measured points then the grid points in the DTM (in contrast to the methods based on the interpolation from sparse points); the DTM surface is in fact approximated (interpolated) in a cloud of measured points. The DTM error can be obtained from the statistical analysis of the differences between the height of points in the measurement cloud and the corresponding heights of the DTM. This may be considered as an internal (relative) accuracy, even though it is not suggested in relevant standards (ISO 19113) as a suitable *positional* accuracy assessment value for DTM. In addition, field height measurements can be compared with the DTM, obtaining the external (absolute) accuracy suggested in ISO 19113 for DTM positional accuracy evaluation. In the authors' opinion, both parameters internal (relative) and external (absolute) errors of DTM are useful as metadata for the DTM; internal error describes how near to the measurements points is the DTM. This approach allows a continuous internal error distribution over the DTM, permitting the identification of places where the discrepancy is significant – in other words a rich source of information on spatial accuracy distributions. By contrast, in the case of external (absolute) DTM error, we usually have a limited number of ground control points, indicating of course how precisely our DTM reflects the real surface. As always, a reference value is needed in accuracy assessment: For internal accuracy the reference height is derived from LIDAR measurements (cloud of LIDAR points), for external accuracy analysis the heights of the in-situ measurements are assumed as a reference.

Vertical DTM accuracy is usually estimated by the mean error, standard deviation or RMSE, based on the height errors calculated as differences between the reference height and the height derived from the DTM. Common approaches assume normal distribution of the errors; however, many authors report the non-normal distribution of these errors and suggest alternative accuracy estimators (ASPRS Guidelines Vertical Accuracy Reporting for LIDAR Data, Aguilar et al. 2010, Atkinson et al. 2005, Darnel et al. 2008, Zandbergen 2008). An alternative so-called "robust" method has also been proposed (ASPRS Guidelines Vertical Accuracy Reporting for Lidar Data, Höhle, Höhle 2009). This approach is based upon the median, normalized median absolute deviation (NMAD) or quantiles. Analysis of the histogram of DTM errors against the theoretical curve permits verification of how they fit to each other. The main objection for the Gauss distribution function is that it considerably overestimates vertical DTM errors, so the standard deviation, mean error and summary RMSE are not representative as DTM error. This is well demonstrated by (Höhle, Höhle 2009) where an alternative robust method is proposed, based on the Normalized Median Absolute Deviation (NMAD) or percentiles; however, analyzing the histogram and the theoretical curve derived from robust method it can be observed that this approach underestimates the error. While the Gauss distribution is much too "flat the robust is too much "sharp". This problem leads to the need to identify a more appropriate method based on the different distribution function; after over viewing the available functions we have chosen the Laplace function.

In our research theoretical curves are calculated on the basis on the errors: mean error, standard deviation and RMSE with the assumption of normal, Gauss error distribution, median and NMAD in robust method, whereas the median and b (average of the absolute

value of the discrepancy between the error and median) are the parameters of the Laplacian distribution. The three methods were tested during our research. The applied parameters are described in detail in chapter 3.

2. TEST AREA

The test area covers part of the east side of Lake Maggiore, near the European Commission Joint Research Centre (JRC) in Ispra, Italy. ALS data were acquired by Compagnia Generale Ripreseaeree S.p.A. with approximately 1 point per 1m, on the 14th April 2005. The data were delivered as ASCII files containing XYZ coordinates of the first and last pulse, together with the strength of the signal, in the UTM 32N zone (WGS-84 ellipsoid), with data coverage of ~3860ha. Ground control points (GCP) were measured by the JRC using GPS RTK in 2008 with centimetre-level accuracy.
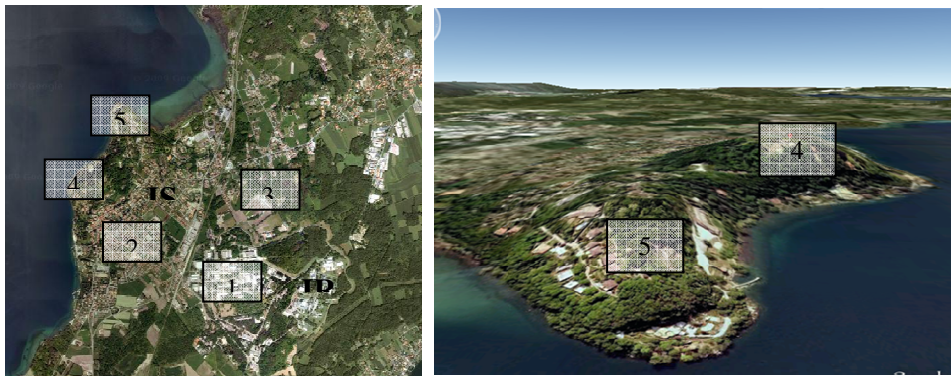

Fig. 1. Part of the area registered by ALS in 2005: ISPRA (JRC). Five sub test areas for detailed accuracy analysis (1, 3 – flat, 2 – medium, 4, 5 – hilly). Zoom of sub test areas 4, 5 from the site of Lago Maggiore

For the accuracy analysis, a test area covering the village of Ispra (JRC) was chosen. Five sub-test areas were selected on two types of surface: flat (1, 3), middle/built-up (2) and hilly near the lake (4, 5), (Figure 1). The height of the terrain changes from 240 m to 350 m, and the slopes varies, for example, in sub test area 5, slopes greater than 45^0 cover more than 50% of the test area.

3. THEORY AND CALCULATIONS

ALS data were analyzed in the project for:
- DTM, DSM (Digital surface model) generation,
- Testing of the procedure of height layer classification,
- Accuracy analysis of DTM for only the test sites in Figure 1.

In this paper, only accuracy analyses are presented. All other details can be found elsewhere (Hejmanowska et al. 2008).

ALS data were processed with the software TerraScan (TerraSolid - The Standard for Airborne and Mobile Lidar Data and Image Processing http://www.terrasolid.fi/) where we can separate the raw cloud of the LIDAR points into the following classes: ground, vegetation and buildings.

The workflow of ALS data processing was as follows: ALS raw data are classified as the ground points (automatically extraction with manual support). Then, the DTM in the tin (Axelson 2000) and grid model is calculated. Next, the heights of the raw ALS points and of the points measured in-situ are subtracted from the DTM, the differences are later analyzed for internal or external accuracy assessment, to help determine which value best describes the accuracy of the DTM.

Assume that Δh is a DTM vertical error:

$$\Delta h = Z_{REFR} - Z_{DTM} \tag{1}$$

where:
Z_{REF} – reference height,
Z_{DTM} – height of DTM corresponding to Z_{REF} (the height derived from DTM for the same horizontal position, like Z_{REF}).

The density probability of DTM error, assuming a normal distribution, can be calculated with the following equation:

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Delta h - \mu)^2}{2\sigma^2}} \tag{2}$$

where
μ - mean vertical error,
σ - standard deviation of the vertical error.
The mean error can be calculated as follows:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} \Delta h_i \tag{3}$$

where n is the number of check points, whereas the standard deviation (σ) is

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\Delta h_i - \mu)^2} \tag{4}$$

In addition, the Root Mean Squared Error (RMSE) is commonly used in accuracy assessment:

$$RMSE = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\Delta h_i)^2} \tag{5}$$

For the comparison with the percentile of 95% standard deviation was multiplied by 1.96.

The observed distribution differs from the normal; it is elongated and skewed. The kurtosis (k) and skew (s) can be calculated as follows:

$$s = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{\Delta h_i - \mu}{\sigma}\right)^3 \tag{6}$$

$$k = \left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{\Delta h_i - \mu}{\sigma}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)} \tag{7}$$

The normality of the observed distribution can be checked on the basis of Q-Q plots or using a standard statistical test (for a large amount of check points, using T-Student rather than, e.g., Shapiro-Wilk).

As an alternative to the mean error and standard deviation, a robust method based on percentile analysis was proposed (Höhle, Höhle 2009). Instead of the mean error, the median (m) (50% percentile) is assumed, and, instead of standard deviation, the Normalized Median Absolute Deviation (NMAD):

$$NMAD = 1.4826 * median\left(\left|\Delta h_i - m\right|\right) \tag{8}$$

Then, the probability density function is calculated from Gauss formula with the median and NMAD:

$$f_2(x) = \frac{1}{\sqrt{2\pi NMAD^2}}e^{-\frac{(\Delta h - m)^2}{2NMAD^2}} \tag{9}$$

Function f_2 is the basis of the robust method. Alternatively density probability can be calculated according Laplace function from the following equation:

$$f_3(\Delta h|m,b) = \frac{1}{2b}e^{-\frac{|\Delta h - m|}{b}} \tag{10}$$

$$b = \frac{1}{n}\sum_{i=1}^{n}\left|\Delta h_i - m\right| \tag{11}$$

205

where *m* is the median.

In a Laplace distribution, the variance is calculated as

$$\text{var} = 2b^2 \tag{12}$$

For a cumulative distribution Laplace function, the error for a given confidence level (*p*) can be calculated as:

$$F_3^{-1}(p) = m - b\,\text{sgn}(p - 0.5)\ln(1 - 2|p - 0.5|) \tag{13}$$

The mean error (median) and standard deviation (NMAD) are the parameters that describe the accuracy in normal distribution (Gauss/robust), whereas the median and b are the parameters of the Laplacian distribution.

4. RESULTS

4.1 Internal accuracy

Internal (relative) accuracy analysis was performed on the basis of the difference map between the raw LIDAR point heights and the DTM. Results of the two sub-test areas are presented in Figure 2. The accuracy of the DTM for hilly area is not spatially uniform. Original LIDAR points lie from 0.5m to 1.0m above the DTM on the east side (blue crosses) and the same values below on the north east side (brown crosses). For the flat area spatial accuracy is more uniform and the errors are mainly in range -0.2m to +0.20 m (green crosses). Larger errors can be found on slopes (blue and yellow crosses).

Difference maps for all sub test areas were used for the statistical analysis based on the formulas described in Chapter 3. Summary parameters are presented in Table 1.

Tab. 1. Parameters of internal accuracy for all sub test areas

Sub test area	1	2	3	4	5
No	94323	80227	69231	36369	26424
MinΔh	-1.050	-2.080	-3.630	-3.930	-7.810
MaxΔh	1.410	2.930	2.500	9.860	10.640
μ	-0.007	0.001	-0.007	-0.017	-0.049
σ	0.073	0.083	0.129	0.317	0.271
m	0.000	0.000	0.000	0.000	-0.010
NMAD	0.044	0.030	0.044	0.044	0.059
k	17.14	106.65	59.54	87.67	228.25
s	0,192	1.581	-1.875	2.321	-1.476
P 0,683	0.040	0.040	0.050	0.060	0.000
b	0.044	0.040	0.063	0.127	0.107

206

The number of points of each sub test area can be found in the first line of the table (ca. 20,000 – 80,000 points). Vertical error of DTM varies from -1.050m to as much as 10.64m for hilly areas. Mean error (μ) is rather small, from 0,001m to -0.049m showing little systematic error, whereas median is equal to zero. Standard deviation is roughly between 10cm - 30cm, whereas NMAD is less variable at 4cm - 6cm.

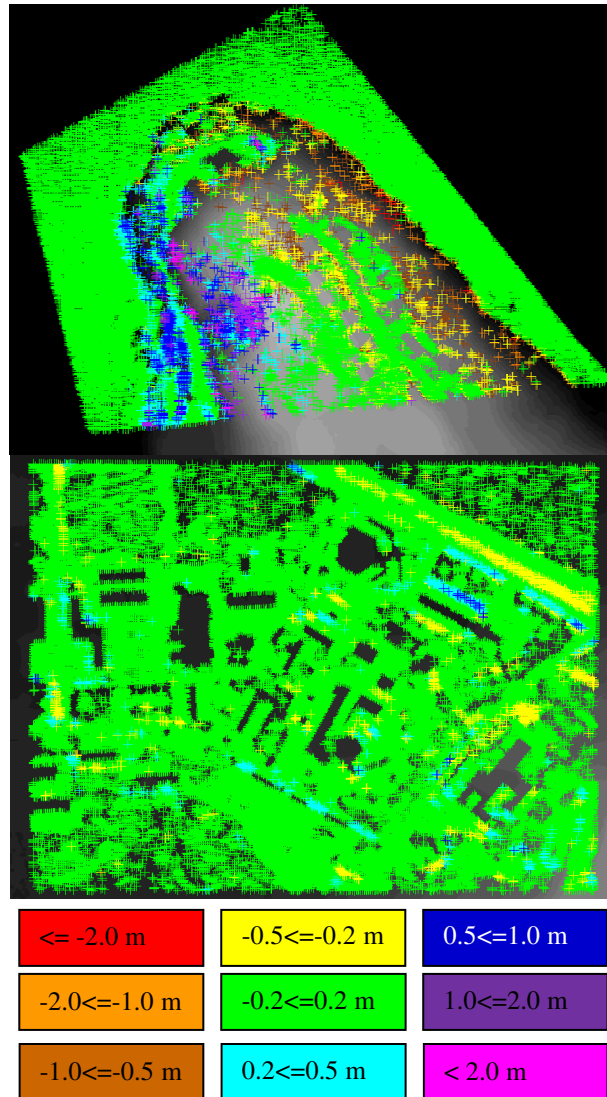| <= -2.0 m | -0.5<=-0.2 m | 0.5<=1.0 m |
| -2.0<=-1.0 m | -0.2<=0.2 m | 1.0<=2.0 m |
| -1.0<=-0.5 m | 0.2<=0.5 m | < 2.0 m |

Fig. 2 Spatial distribution of the internal (relative) accuracy of the DTM (at red, orange, brown and yellow points, the points higher than the DTM; at cyan, dark blue, violet and magenta points, the

points are below the DTM; and at green points the DTM is more or less correct). Test site 5 (hilly) above, test site 1 (flat) below.
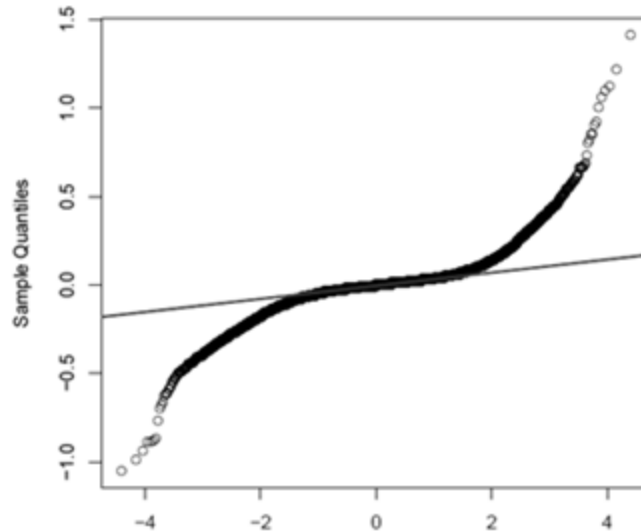

Fig. 3. Q—Q plot for the distribution of Δh for sub test 5.

A normal Q—Q plot is usually used for testing the normality of the error distribution; an example for subset 5 can be found in Figure 3. Analysis of Q—Q plots, showing the shape of histograms of the Δh errors (Figure 4), show narrow distributions with elongated tails, and where kurtosis and skew (Table 1) confirm that the observed histograms cannot be treated as a normal distribution.

The Gauss distribution function was also tested through the fitting of theoretical curves to the observed histogram. Three functions (f_1, f_2 and f_3) were calculated from the error sets for test subsites and were superimposed on the observed histograms (Figure 4). For the small Δh errors, near zero the Gauss curve (f_1) built on the mean and standard deviation values lies always the lowest, and the Gauss curve (f_2) calculated on the basis on median and NMAD (robust method) is the highest, and fits the histogram better. The second sensitive range is placed in vicinity of standard deviation, where the Gauss curve is significantly above and robust one below the observed histogram. Figure 4 also shows the Laplacian curve calculated from f_3, which is placed always between the curves f_1 and f_2.

The discrepancies between the observed errors and the modeling curves can be followed in the Figure 5 for subset 5. As was the case for all test areas, the Q-Q plots, histograms and fitting curves using functions: f_1, f_2 and f_3 were generally similar to the Figure 3 and Figure 4. The RMSE of the curves matching to the observed histogram with the indexes corresponding to the fitted function can be found for all test subsites in the Table 2.
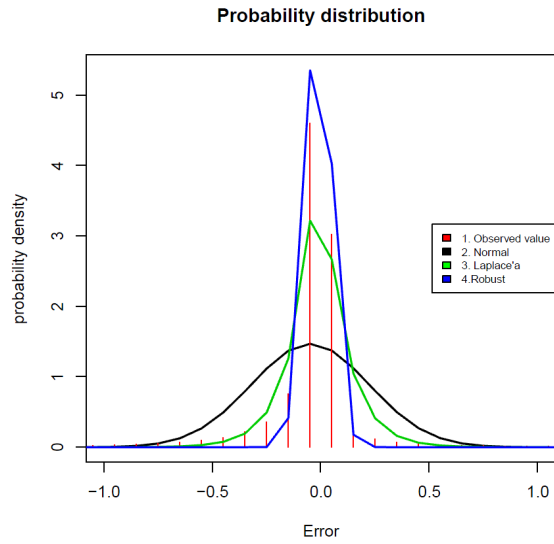
Probability distribution

Fig. 4. Histogram of the DTM errors for test subsite 5.

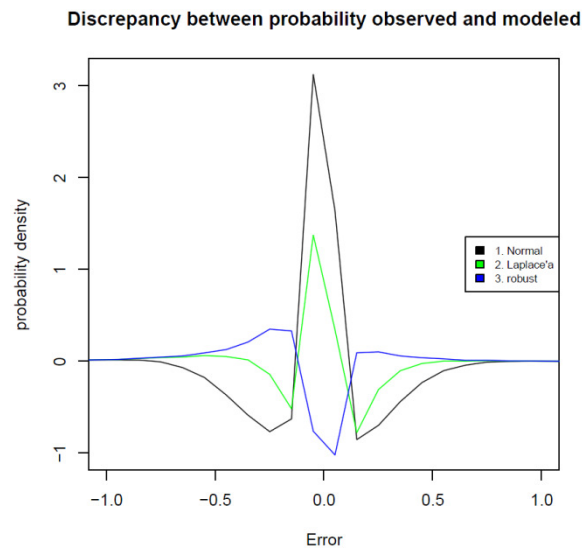**Discrepancy between probability observed and modeled**

Fig. 5. Discrepancies between the observed error Δh and modelled for sub test 5

Tab. 2. Evaluation of the fitting the theoretical distribution function to the observed histogram

	1	2	3	4	5
RMSE$_{f1}$	0.756	0.719	0.624	0.362	0.289
RMSE$_{f2}$	0.426	0.257	0.217	0.146	0.103
RMSE$_{f3}$	0.332	0.249	0.272	0.202	0.128

As shown in Figure 4 and 5 Laplace distribution curves fit the observed histogram much better than the Gauss and robust methods, especially in the two sensitive areas, i.e. near Δh equal zero and in the inflection range of standard deviation.

Besides the analysis of the quality of the histogram fitting using the three functions, some parameters of accuracy evaluation were also calculated. For the Gauss function $F_1^{-1}(0.95)$ corresponding to $1,96\,\sigma$ was computed; for the robust method, $F_2^{-1}(0.95)$ was used. Cumulative distribution of Laplace function for a 95% confidence level was also calculated ($F_3^{-1}(0.95)$ - formula 13), (Table 3).

Tab.3. Parameters of vertical error distribution

	1	2	3	4	5
$F_1^{-1}(0.95)$	0.136	0.164	0.245	0.605	0.480
$F_2^{-1}(0.95)$	0.086	0.084	0.086	0.086	0.106
$F_3^{-1}(0.95)$	0.132	0.120	0.189	0.380	0.310
P 0.95	0.150	0.130	0.200	0.630	0.180

The parameters basing on Gauss function are, as expected, the largest in all cases; the robust method percentiles 0.95 are by contrast the lowest. The error parameter derived from the Laplace function is between the two of them.

4.2 External accuracy

The test of the external accuracy of DTM was performed based on the GPS RTK measurements. These data also do not have a normal distribution; see the Q-Q plot in Figure 6(a). In Figure 6(b), the histogram of DTM vertical errors is presented with overlaid matching probability density curves. Discrepancies between the observed and modeled probability are shown in Figure 6(c). The best fit was observed for the Laplacian curve, with a worse fit for the robust and Gaussian methods. The error of Gaussian function fitting to the histogram was 0.389, for robust method 0.289; however, the Laplacian function fitting error was significantly smaller: 0.125.

A mean error of -0.165 m was observed, which indicates some bias. In other words, the DTM is systematically above the GPS points. The limit error using the 1.96 standard deviation at the 95% confidence level was 0.425 m (RMSE 0.534 m). The limit error from the robust method was 0.038 m, but from Laplacian distribution: 0.165 m. Percentile (0.95) was equal 0.625 m.
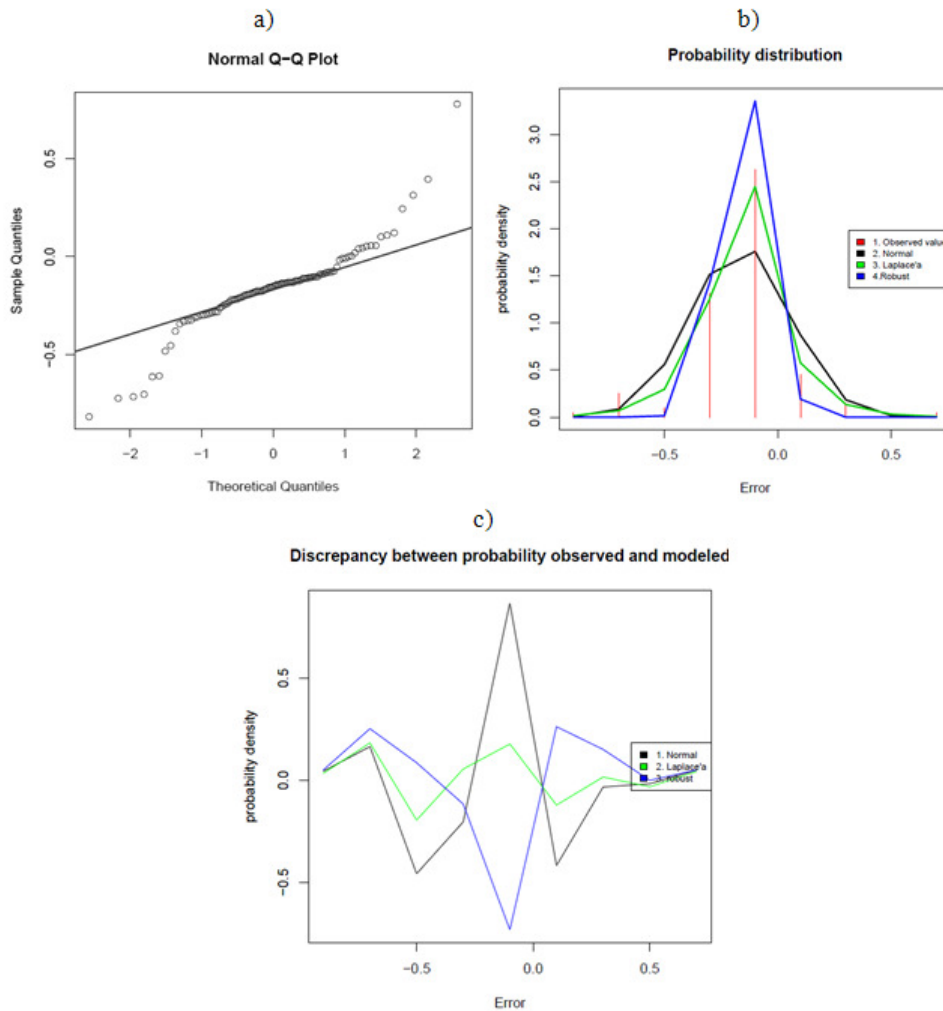
Fig. 6. Q-Q plot, histogram of Δh and discrepancy between the observed and modeled probability, external accuracy

5. CONCLUSION

According to our research:

1. Errors obtained for ALS data do not seem to have a Gaussian normal distribution.
2. The Gaussian approach overestimates the DTM errors. However, a robust method taken from the literature appears to be too "liberal" and underestimates DTM error.

3. The Laplacian function fits better that the previous two approaches to the histograms and $F_3^{-1}(0.95)$ and might be considered as an alternative for DTM accuracy estimation.
4. In our research, outliers were not removed, the histograms have a very long tails on the left and right side (Figure 3 is only zoomed to the range -1 1); a clear approach for such outlier removal will require further research.

Estimation of the DTM accuracy is crucial for many applications, for example, for flood risk assessment. However, use of the Gaussian approach to assess accuracy is too crude, and the forecasted risk area could be overestimated; conversely, the robust approach might lead to too small a flood area with a high flood risk. The results of the Laplacian approach are between the results of the other two methods and match the real error more reliably, and should be a better estimate of the area at real risk. These conclusions were, nonetheless, derived only from analyzing the fit of the curves to the histograms. A method to check the reliability of different kinds of DTM errors is therefore still an open question. Early results from ongoing research (based on LIDAR data for Krakow, Poland near Vistula River) do however confirm the results presented in this paper and will be published in a follow-up of this study.

6. ACKNOWLEDGEMENTS

7. REFERENCES

Axelson P., 2000. *DEM Generation from Laser Scanner Data Using Adaptive TIN Models.* International Archives of Photogrammetry and Remote Sensing, Amsterdam, Netherlands 32,B4/1, pp. 110 – 117

ASPRS Lidar Committee. 2004. ASPRS Guidelines Vertical Accuracy Reporting for Lidar Data,
http://www.asprs.org/society/committees/lidar/downloads/vertical_accuracy_reportriangleg _for_lidar_data.pdf (accessed 7th April 2011)

Aguilar F.,J., Mills J.P., Delgado J., Manuel A. Aguilar M.A., Negreiros J.G., Pérez J. L. – 2010, *Modelling vertical error in LiDAR-derived digital elevation models* , ISPRS Journal of Photogrammetry and Remote Sensing 65 (2010) 103_110

Atkinson Gordo A. D. J., Ariza López F.J, José L. García-Balboa J. L. 2005 *Positional accuracy control using robust estimator*, In Proceedings of the 21st International Cartographic Conference, 09-16 July, Acoruna, Spain

Darnel A. R., Nicholas J. Tate N. J., Chris Brunsdon C., 2008 - *Improving user assessment of error implications in digital elevation models*, Computers, Environment and Urban Systems 32 (2008) 268–277

Fisher P.F., Tate N.J. 2006. *Causes and consequences of error in digital elevation models*. Progress in Physical Geography 30, 4, p. 467-489

Höhle J., Höhle M., 2009 - *Accuracy assessment of digital elevation models by means of robust statistical methods*, ISPRS Journal of Photogrammetry and Remote Sensing 64 (2009) 398_406

Höhle J. The EUROSDR project "Automated checking and improving of digital terrain models" http://www.asprs.org/publications/proceedings/tampa2007/0027.pdf (accessed 7th April 2011)

Hejmanowska B., Borowiec N. Badurska M. 2008 – „Processing airborne data to Digital Surface Model and Digital Terrain Model. Final report http://home.agh.edu.pl/~galia/research/Processing%20LIDAR%20%202007%20final%20re port_5_03_2008.pdf (accessed 7th April 2011)

Kraus K., Karel W., Briese Ch., Mandlburger G., 2006. *Local Accuracy Measures for Digital Terrain Model*. Photogrammetric Record 21 (116), p. 342-354

Oksansen J. 2006. Digitale elevation model error in terrain analysis. PhD Thesis http://ethesis.helsinki.fi/julkaisut/mat/maant/vk/oksanen/digitale.pdf (accessed 7th April 2011)

Wechsler S.P., 2001. DEM Uncertainty: Evaluation and Effect on Topographic Parameters. Dissertation State University of New York, College of Environmental Science and Forestry, Syracuse, New York

Zandbergen P. A., 2008 *Positional Accuracy of Spatial Data: Non-Normal Distributions and a Critique of the National Standard for Spatial Data Accuracy*, Transactions in GIS, 2008, 12(1): 103–130

ISO 19113:2002 Geographic information - Quality principles