IMPACT OF OPTIMIZATION OF ALS POINT CLOUD ON CLASSIFICATION

Wioleta Błaszczak-Bąk\textsuperscript{1,2}, Anna Sobieraj\textsuperscript{2}

\textsuperscript{1} Institute of Geodesy
University of Warmia and Mazury in Olsztyn
\textsuperscript{2} Chair of Geodesy
Gdansk University of Technology

Received 5 December 2012; Accepted 6 June 2013; Available on line 15 July 2013.

K e y w o r d: optimization, classification, intensity.

A b s t r a c t

Airborne laser scanning (ALS) is one of the LIDAR technologies (Light Detection and Ranging). It provides information about the terrain in form of a point cloud. During measurement is acquired: spatial data (object’s coordinates $X$, $Y$, $Z$) and collateral data such as intensity of reflected signal. The obtained point cloud is typically applied for generating a digital terrain model (DTM) and a digital surface model (DSM). For DTM and DSM generation it is necessary to apply filtration or classification algorithms. They allow to divide a point cloud into object groups (e.g.: terrain points, vegetation, etc.). In this study classification is conducted with one extra parameter–intensity. The obtained point groups were used for digital spatial model generation.

Classification is a time and work consuming process, therefore there is a need to reduce the time of ALS point cloud processing. Optimization algorithm enables to decrease the number of points in a dataset. In this study the main goal was to test the impact of optimization on the results of a classification. Studies were conducted in two variants. Variant 1 includes classification of the original point cloud where points are divided in the groups: roofs, asphalt road, tree/bushes, grass. On variant 2 before classification, an optimization algorithm was performed in the original point cloud. Obtained from these two variants object groups were used to generate a spatial model, which was then statistically analyzed.

Introduction

Airborne Laser Scanning ALS provides a data in the form of a point cloud. Obtaining such data is possible thanks to integration of a laser scanning system with GPS and INS systems (Csanyi et al. 2005, Grejner-Brzezińska...
et al. 2004, TOTH, GREJNER-BRZEZIŃSKA 2004). At this stage, data processing includes filtration or classification. Filtration is a process in which two groups of points are obtained: points representing terrain and non-terrain (BORKOWSKI, JÓZKÓW 2007, BŁASZCZAK-BĄK et al. 2010). Classification is based on point cloud division into object groups. The number of distinguished groups is related to the aim of the classification. Characteristics of each group results from the type of surface from which the laser reflected. Examples are groups like: terrain, vegetation, buildings etc. Point groups representing terrain can be used as a basis for Digital Terrain Model (DTM) generation. Point groups obtained from classification are used for Digital Surface Model (DSM) generation.

Assessment of filtration and classification algorithms can be found in many articles, for example in TÓVÁRI and PFEIFER 2005, the authors presented their proposal to divide algorithms into four group:

– Methods based on morphological filters (e.g.: VOSSelman, MAAS 2001, ZHANG et al. 2002),

– Methods which use algorithms based on active TIN (Triangulated Irregular Network) modeling, in an iterative way, the algorithm fits a modeling surface with data representing real surface (e.g.: SCHUT 1976, KASS et al. 1988, PFEIFER 1998, ELMQVIS 2002),

– Methods which include point cloud segmentation and aim at usually to three point groups: terrain, buildings and vegetation (e.g.: ROGGERO 2002, TÓVÁRI, PFEIFER 2005, SITHOLE, VOSSelman 2005).

Detailed analyses of selected filtration and classification algorithms are presented in SITHOLE, VOSSelman (2004). Existing methods have their limits, which are related to incorrect filtration of points from regions with complex topography. This is the impact for conducting studies in area where filtration and classification algorithms can be developed, modified and tested with various parameters.

Studies on intensity of laser pulse reflection utility were conducted by KATZENBEISSER 2003, SONG et al. 2002, CHARANIYA et al. 2004, WANG and LU (2009), ANTONORAKIS et el. 2008 and others. In Katzenbeisser’s studies there is a conclusion that using intensity is difficult and complex. Only for flat surfaces reliable information about a given object can be obtained. Höfle and Pfeifer focus on calibrating intensity.

In this study classification was conducted using intensity as an additional classifying factor. Due to the fact that such way of ALS point cloud processing is time and work consuming, the authors proposed to apply point cloud optimization. Optimization conducted before filtration shorten time necessary for DTM generation about 4–6 times in comparison to standard methodology (BŁASZCZAK-BĄK et al. 2010a, 2010b, 2011a). During optimization points
which are not significant for the DTM generation are removed (Błaszczyk 2006). The lack of those points and subsequently intensities related to them, was tested in the scope if it has an impact on the point set used for DTM generation.

Intensity ranges for specific types of objects were based on indicated points representing given object. Studies were conducted in two variants. In variant 1 classification was based on an original point set, while in variant 2 classification was preceded by optimization, which reduced the number of points in the dataset. Resulted object groups and models generated from them were statistically analyzed.

Proposal of ALS data processing scheme

One of the ALS data processing stages is classification. The result of a conducted classification is a group of datasets – points representing various objects. Intensity was a basic parameter for establishing groups of points. Theirs types and number resulted from analyses of cartographic data and recognition in the field. Of course, there are many methods of classification using the intensity parameter, for example object-based land cover classification (Antonarakis et al. 2008). There also are many software based on specific algorithms (e.g. VRMesh, DTMaster). Method of classification used in this publication is characterized by its simplicity. It enabled to control the process of classification and to conduct analyses on each stage of tests. Authors decided to used such solution, because they did not have access to commercial software for classification.

Distinguished groups in this study are as follows: grass, asphalt road, roofs, trees/bushes and others. The last group includes points that have not been assigned to any previous group. Obtained datasets were a basis for spatial model generation.

The basis for each spatial data model is the choice of basic (spatially defined) geometric elements used to represent objects. Depending on the size of basic geometric element used for spatial model generation, three types of models can be distinguished (Izdebski 2008):

– point models – the basic geometric element is a point,
– linear models – the basic element of the line is made up of a sequence of points,
– surface models – the basic element is the area.

Depending on the location and the shape of basic elements the surface models can be divided into:

– regular models,
– irregular patterns.
In view of the fact that airborne laser scanning provides data in discrete form (as point cloud) it is reasonable that a point model is created. A Digital Spatial Point Model generated from discrete ALS data has the value of the intensity at each point \( x, y, z \) located in the model. Digital Terrain Model (DTM) is generated on the basis of two object groups obtained in the classification representing surface of the ground: grass and asphalt road.

The processing of the ALS point cloud was conducted in two versions. In variant 1 classification was taken into account, in variant 2 classification was preceded by the optimization process aimed at reducing the size of the data set. The optimization algorithm, which was used in this study, consists of the following steps (Błaszczak 2006, Błaszczak, Kamiński 2007, Błaszczak-Bąk et al. 2010):

**Step 1:** defining test strips in the \( XY \) plane, parallel to the \( Y \) axis.

**Step 2:** choice of cartographic generalization method used for reducing the size of the measurement set.

**Step 3:** using in each profile (in the \( YZ \) plane) the chosen method of generalization.

An important step is to select the generalization method in test strips and the selection of appropriate tolerances. Threshold values has a significant impact on the degree of reducing a ALS dataset (the number of points removed). In this study, in the second stage in the optimization algorithm, the method of Visvalingam-Whyatt (V-W) (Visvalingam, Whyatt 1992) was used. The generalization procedure using comparative surface V-W has been shown in Figure 1. The implementation of the method V-W is to create the next line which is based on generalized points of triangles (e.g. points 1, 2, 3 form a triangle in the plane \( YZ \) with area \( P_1 \), Figure 1b). The calculated area of a triangle is verified with the comparative surface. Its size is defined by the user, depending on the statistical characteristics of a dataset from measurements. If the area of the triangle determined from the measured points exceeds the area of a tolerance triangle, then the second point of analyzed triangle is retained (step 2 in Figure 1b). Subsequently, the triangles are formed further (e.g. a triangle with the area \( P_2 \) – Fig. 1c or triangle with area \( P_3 \) – Fig. 1d), and theirs areas are compared with the area of the tolerance triangle. When the area of a tolerance triangle is greater than the area of determined triangle, the second point of formed triangle is removed (e.g. point 4 in Fig. 1d).

For both variants Digital Spatial Point Models (described as DSPMv1 and DSPMv2) and Digital Terrain Models (called DTMv1 and DTMv2, respectively) were generated. DTMs were statistically analyzed. A general scheme of the study is presented in Figure 2.
Fig. 1. Stages in line generalization based on comparative surface
Fig. 2. General scheme of study

ALS POINT CLOUD

VARIANT 1

CLASSIFICATION

GRASS  ASPHALT ROAD  ROOFS  TREES/BUSHES  OTHERS

DIGITAL SPATIAL POINT MODEL DSPMv1

DIGITAL TERRAIN MODEL DTMv1

VARIANT 2

OPTIMIZATION

CLASSIFICATION

GRASS  ASPHALT ROAD  ROOFS  TREES/BUSHES  OTHERS

DIGITAL SPATIAL POINT MODEL DSPMv2

DIGITAL TERRAIN MODEL DTMv2

STATISTICAL ANALYSES
Results

The study was based on a point cloud obtained from airborne scanning made by Visimind. For the measurements were used: the Riegl LMS-Q240 laser, Topcon GPS, IMU units and digital cameras. The scan angle was 60 degrees with a resolution of 10,000 Hz. Scanning took place from a helicopter moving at a speed of 50 km/h at an altitude of 700 m. The measured area is located on the outskirts of Olsztyn (Warmia and Mazury). The original data set contained 1.5 million in the format X, Y, Z, intensity. In this study raw data were used for test.

Statistical samples concerning intensity were obtained by identifying and manually selecting the points that belong to the specific object. The results of the analyses are presented in Tables 1–4 The tables use the following code: \( n \) – number of points, minimum intensity – \( I_{\text{min}} \), the maximum intensity – \( I_{\text{max}} \), the average value of the intensity – \( I_{\text{av}} \), the standard deviation – \( \delta \).

<table>
<thead>
<tr>
<th>Object</th>
<th>roof1</th>
<th>roof2</th>
<th>roof3</th>
<th>roof4</th>
<th>roof5</th>
<th>roof6</th>
<th>roof7</th>
<th>roof8</th>
<th>roof9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>31</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>22</td>
<td>24</td>
<td>11</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>( I_{\text{min}} )</td>
<td>0.234</td>
<td>0.391</td>
<td>0.425</td>
<td>0.397</td>
<td>0.438</td>
<td>0.430</td>
<td>0.416</td>
<td>0.406</td>
<td>0.385</td>
</tr>
<tr>
<td>( I_{\text{max}} )</td>
<td>0.455</td>
<td>0.458</td>
<td>0.449</td>
<td>0.447</td>
<td>0.460</td>
<td>0.452</td>
<td>0.435</td>
<td>0.448</td>
<td>0.452</td>
</tr>
<tr>
<td>( I_{\text{av}} )</td>
<td>0.415</td>
<td>0.432</td>
<td>0.438</td>
<td>0.430</td>
<td>0.447</td>
<td>0.439</td>
<td>0.426</td>
<td>0.429</td>
<td>0.435</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.040</td>
<td>0.021</td>
<td>0.007</td>
<td>0.013</td>
<td>0.006</td>
<td>0.005</td>
<td>0.006</td>
<td>0.014</td>
<td>0.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object</th>
<th>roof10</th>
<th>roof11</th>
<th>roof12</th>
<th>roof13</th>
<th>roof14</th>
<th>roof15</th>
<th>roof16</th>
<th>roof17</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>20</td>
<td>26</td>
<td>16</td>
<td>18</td>
<td>20</td>
<td>11</td>
<td>15</td>
<td>17</td>
</tr>
<tr>
<td>( I_{\text{min}} )</td>
<td>0.378</td>
<td>0.338</td>
<td>0.442</td>
<td>0.399</td>
<td>0.284</td>
<td>0.440</td>
<td>0.446</td>
<td>0.410</td>
</tr>
<tr>
<td>( I_{\text{max}} )</td>
<td>0.434</td>
<td>0.450</td>
<td>0.455</td>
<td>0.416</td>
<td>0.440</td>
<td>0.461</td>
<td>0.479</td>
<td>0.457</td>
</tr>
<tr>
<td>( I_{\text{av}} )</td>
<td>0.402</td>
<td>0.413</td>
<td>0.449</td>
<td>0.407</td>
<td>0.421</td>
<td>0.452</td>
<td>0.465</td>
<td>0.427</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.021</td>
<td>0.028</td>
<td>0.004</td>
<td>0.004</td>
<td>0.032</td>
<td>0.006</td>
<td>0.008</td>
<td>0.011</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object</th>
<th>asph1</th>
<th>asph2</th>
<th>asph3</th>
<th>asph4</th>
<th>asph5</th>
<th>asph6</th>
<th>asph7</th>
<th>Asphalt road</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>11</td>
<td>20</td>
<td>14</td>
<td>22</td>
<td>16</td>
<td>22</td>
<td>26</td>
<td>131</td>
</tr>
<tr>
<td>( I_{\text{min}} )</td>
<td>0.365</td>
<td>0.349</td>
<td>0.366</td>
<td>0.358</td>
<td>0.341</td>
<td>0.379</td>
<td>0.356</td>
<td>0.359</td>
</tr>
<tr>
<td>( I_{\text{max}} )</td>
<td>0.396</td>
<td>0.392</td>
<td>0.444</td>
<td>0.390</td>
<td>0.394</td>
<td>0.401</td>
<td>0.398</td>
<td>0.402</td>
</tr>
<tr>
<td>( I_{\text{av}} )</td>
<td>0.384</td>
<td>0.376</td>
<td>0.392</td>
<td>0.376</td>
<td>0.374</td>
<td>0.392</td>
<td>0.380</td>
<td>0.382</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.009</td>
<td>0.013</td>
<td>0.026</td>
<td>0.009</td>
<td>0.017</td>
<td>0.005</td>
<td>0.012</td>
<td>0.013</td>
</tr>
</tbody>
</table>
### Table 3

<table>
<thead>
<tr>
<th>Object</th>
<th>tree1</th>
<th>tree 2</th>
<th>tree 3</th>
<th>tree 4</th>
<th>tree 5</th>
<th>tree 6</th>
<th>tree 7</th>
<th>tree 8</th>
<th>tree 9</th>
<th>tree 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>20</td>
<td>26</td>
<td>30</td>
<td>21</td>
<td>29</td>
<td>25</td>
<td>19</td>
<td>29</td>
<td>29</td>
<td>40</td>
</tr>
<tr>
<td>$I_{\text{min}}$</td>
<td>0.173</td>
<td>0.100</td>
<td>0.022</td>
<td>0.123</td>
<td>0.245</td>
<td>0.127</td>
<td>0.086</td>
<td>0.353</td>
<td>0.145</td>
<td>0.114</td>
</tr>
<tr>
<td>$I_{\text{max}}$</td>
<td>0.420</td>
<td>0.411</td>
<td>0.419</td>
<td>0.442</td>
<td>0.434</td>
<td>0.456</td>
<td>0.413</td>
<td>0.417</td>
<td>0.426</td>
<td>0.445</td>
</tr>
<tr>
<td>$I_{\text{av}}$</td>
<td>0.362</td>
<td>0.324</td>
<td>0.305</td>
<td>0.418</td>
<td>0.384</td>
<td>0.365</td>
<td>0.278</td>
<td>0.400</td>
<td>0.359</td>
<td>0.387</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.064</td>
<td>0.113</td>
<td>0.104</td>
<td>0.067</td>
<td>0.047</td>
<td>0.088</td>
<td>0.114</td>
<td>0.014</td>
<td>0.082</td>
<td>0.063</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Object</th>
<th>Gr1</th>
<th>Gr2</th>
<th>Gr3</th>
<th>Gr4</th>
<th>Gr5</th>
<th>Gr6</th>
<th>Gr7</th>
<th>Gr8</th>
<th>Grass</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>28</td>
<td>14</td>
<td>31</td>
<td>29</td>
<td>24</td>
<td>21</td>
<td>26</td>
<td>23</td>
<td>196</td>
</tr>
<tr>
<td>$I_{\text{min}}$</td>
<td>0.425</td>
<td>0.419</td>
<td>0.417</td>
<td>0.405</td>
<td>0.420</td>
<td>0.413</td>
<td>0.418</td>
<td>0.310</td>
<td>0.403</td>
</tr>
<tr>
<td>$I_{\text{max}}$</td>
<td>0.440</td>
<td>0.444</td>
<td>0.433</td>
<td>0.460</td>
<td>0.435</td>
<td>0.434</td>
<td>0.455</td>
<td>0.454</td>
<td>0.444</td>
</tr>
<tr>
<td>$I_{\text{av}}$</td>
<td>0.430</td>
<td>0.433</td>
<td>0.426</td>
<td>0.435</td>
<td>0.428</td>
<td>0.423</td>
<td>0.434</td>
<td>0.430</td>
<td>0.430</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.004</td>
<td>0.007</td>
<td>0.004</td>
<td>0.010</td>
<td>0.005</td>
<td>0.005</td>
<td>0.009</td>
<td>0.028</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Based on the statistical data intensity, ranges for each object group were extracted. The ranges of the intensity are based on the relationship: $<I_{\text{av}} - \delta; I_{\text{av}} + \delta>$. Table 5 presents the juxtaposition of intensity ranges for each type of studied object.

### Table 5

<table>
<thead>
<tr>
<th>Objects</th>
<th>Intensity ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roofs</td>
<td>0.417–0.444</td>
</tr>
<tr>
<td>Asphalt road</td>
<td>0.369–0.395</td>
</tr>
<tr>
<td>Trees/bushes</td>
<td>0.282–0.434</td>
</tr>
<tr>
<td>Grass</td>
<td>0.421–0.439</td>
</tr>
</tbody>
</table>

Unfortunately, the result intensity ranges distinguished on the basis of statistical sampling are partially overlapping. In order to clearly define the limits of ranges for each object, ranges were modified. Correct intensity ranges were used to build DSPMs. Analysis also showed that some of the intensity values are not classified to any of the groups, and it is impossible to identify what objects they represent. Not qualified values of intensity (0.00–0.282) and (0.445–0.577) were omitted. Table 6 shows the intensity ranges for each types of objects that have been adopted to build the DSPM.
In a detailed study fragment of the point cloud 83,424 points were used. A processed subset of this ALS point cloud is shown in Figure 3. Given the ALS dataset was processed in two ways. In the first way, only classification was performed. In the secondary, point cloud was optimized and then classified. In both cases, the intensity ranges shown in Table 6 were applied.

<table>
<thead>
<tr>
<th>Objects</th>
<th>Intensity ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roofs</td>
<td>0.417–0.420 and 0.440–0.444</td>
</tr>
<tr>
<td>Asphalt road</td>
<td>0.369–0.395</td>
</tr>
<tr>
<td>Trees/bushes</td>
<td>0.282–0.368 and 0.396–0.416</td>
</tr>
<tr>
<td>Grass</td>
<td>0.421–0.493</td>
</tr>
</tbody>
</table>

Variant 1

ALS point cloud classification was conducted using the previously mentioned ranges of intensity. Obtained results of classification are shown in Figure 4.

The figures indicate that not all the points were correctly classified for the corresponding types of objects. This can be seen especially for the group “roofs” (Fig. 4b), where beside roof points, there are a lot of points representing trees and bushes. The group “grass” (Fig. 4a) includes the majority of points. For the group “asphalt road” (Fig. 4c) the road was classified correctly, but like in the group “roofs”, there also points appear representing trees. As is seen, intensity for the trees has a wide ranges and therefore appeared in all separate groups.
This resulted from the fact that the measurement was performed when there were leaves on trees. Diversity of leaf color and leaf variable angle relative to the incident laser beam led to the wide range of intensity for trees.

**Variant 2**

The size of the dataset was reduced using the optimization algorithm. During the reduction, it was assumed that the width of the strip equals 2 m and the area of a tolerance triangle is 0.02 m². This allowed a reduction to 43,715 points.
Optimized dataset were then classified.

Fig. 6. Variant 2: \textit{a} – grass, \textit{b} – asphalt road, \textit{c} – roofs, \textit{d} – trees/bushes
Figure 6 shows the spatial distribution of the optimized dataset of the point cloud. The effect of optimization can be seen for all groups. The reduction in the number of points is most visible in the groups as asphalt road and roofs. As in Variant 1 classification after the optimization does not give good final results. Manual classification is still needed.

**Result analysis**

Analyzing the presented results, it is clear that Variant 2, which corresponds to the classification performed on the optimized dataset, allows achieving similar results as in Variant 1. In both cases, asphalt road and grass were well separated. For trees/bushes and the roofs the same mistakes can be observed in both Variants. However, there is no doubt, that a smaller number of points resulting from the optimization process certainly would have an impact on the processing time. Results obtained for the two variants are presented in Table 7.

<table>
<thead>
<tr>
<th>Objects</th>
<th>Number of points used for generation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSPMv1</td>
</tr>
<tr>
<td>Roofs</td>
<td>9,891</td>
</tr>
<tr>
<td>Asphalt road</td>
<td>6,102</td>
</tr>
<tr>
<td>Trees/bushes</td>
<td>18,545</td>
</tr>
<tr>
<td>Grass</td>
<td>45,212</td>
</tr>
<tr>
<td>Σ</td>
<td><strong>83,424</strong></td>
</tr>
<tr>
<td>Others</td>
<td>3,674</td>
</tr>
</tbody>
</table>

Number of points for Digital Models generation in Variant 2 is about 50% less than in Variant 1.

**Spatial models generation**

**Digital Spatial Point Model generation**

On the basis of the sets of points from Variant 1 and Variant 2 (Table 7) digital spatial point models were generated. They represent the intensity distribution in the study area. DSPMv1 and DSPMv2 shown in Figure 7.
Analyzing the two point models, similar point distributions can be observed. The DSPMv2 is more readable due to the reduced number of points. Such models may be useful in the detailed classification, because the observer sees colors that indicates a particular object and also receives information about the height of a point.

**Digital Terrain Model generation**

Digital Terrain Models were generated from two object groups: grass and asphalt road. Due to the fact that these objects were separated only on the basis of the intensity of the reflection and as previously mentioned, some of these points were misclassified, it was decided to include an additional parameter: the height of the point. Two new subsets of points for Variant 1 and Variant 2 were obtained: a subset of data points showing the terrain generated from the original and then the optimized datasets. Separated subsets became the basis for generating DTMv1 and DTMv2 (Figure 8).
Additionally, a digital model DTMv3 was generated. It was adopted as the reference model to evaluate the accuracy of obtained two models DTMv1 and DTMv2. DTMv3 was built on the basis of the data obtained after filtration of the original dataset. The most popular method of filtration, the active method TIN model (AXELSSON 2000), was applied here.

The three generated DTMs were subjected to statistical comparison. In comparing DTMv1, DTMv2 and DTMv3, the following parameters were applied (OKSANEN, SARJAKOSKI 2005, HEJMANOWSKA et al. 2008):

a) mean error:

\[ m_0 = \sqrt{\frac{\Sigma (z_{\text{mean}} - z_i)^2}{k - 1}} \]  \hspace{1cm} (1)

where:

- \( z_{\text{mean}} \) – mean height calculated from two datasets used for DTMs generation,
- \( z_i \) (i=1,2,..,k) – height of point in data from measurement,
- \( k \) – number of points in dataset;

b) range \( R = z_{\text{max}} - z_{\text{min}} \), where \( z_{\text{max}} \) – maximum height and \( z_{\text{min}} \) – minimum height;

c) mean height difference (systematic error):

\[ \Delta h_{\text{mean}} = \frac{\Sigma_{i=1}^{k} (\tau)}{k} \]  \hspace{1cm} (2)

where:

\( \tau = \tau_1 \) or \( \tau_2 \)

\( \tau_1 = Z_{\text{NMTv2}} - Z_{\text{NMTv1}}, \quad \tau_2 = Z_{\text{NMTv2}} - Z_{\text{NMTv1}} \)
d) root mean square error (RMSE), describes absolute height accuracy of DTM:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{k} (\rho)^2}{k}}
\]

where:

\[
\rho = \rho_1 \text{ or } \rho_2
\]

\[
\rho_1 = r_1 - \Delta h_{\text{mean}}, \quad \rho_2 = r_2 - \Delta h_{\text{mean}};
\]

e) coefficient of determination (measure of model fitness):

\[
D^2 = \frac{\sum_{i=1}^{k} (Z_{NMTv2} - Z_{\text{mean}})^2}{\sum_{i=1}^{k} (Z_{NMTv1} - Z_{\text{mean}})^2}
\]

\[
D^2 = \frac{\sum_{i=1}^{k} (Z_{NMTv2} - Z_{\text{mean}})^2}{\sum_{i=1}^{k} (Z_{NMTv1} - Z_{\text{mean}})^2}
\]

A summary of the calculated parameters for assessing the accuracy of the generated DTMs are presented in Table 8.

<table>
<thead>
<tr>
<th>Parameters for assessing the accuracy</th>
<th>DTMv3 reference</th>
<th>DTMv1</th>
<th>DTMv2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m_0)</td>
<td>4,655 m</td>
<td>4,675 m</td>
<td>4,662 m</td>
</tr>
<tr>
<td>(R)</td>
<td>21,988 m</td>
<td>21,295 m</td>
<td>21,992 m</td>
</tr>
<tr>
<td>(\Delta h_{\text{mean}})</td>
<td>-</td>
<td>0,026 m</td>
<td>-0,019 m</td>
</tr>
<tr>
<td>RMSE</td>
<td>-</td>
<td>0,299 m</td>
<td>0,168 m</td>
</tr>
<tr>
<td>(D^2)</td>
<td>-</td>
<td>0,991</td>
<td>0,997</td>
</tr>
</tbody>
</table>

The mean error and range for the reference model and the model from Variant 2 have similar values, indicating that the interpolated height of GRID points are comparable. For Variant 1 the mean error is about 2 cm larger, and range is about 70 cm smaller in comparison to the reference model. However, assessment based on such measures does not enable to compare models from
two variants. For such purpose the RMSE and the coefficient of determination were used. Variant 2 achieves better values. RMSE error is about 13 cm lesser, and $D^2$ is more close to unity. By analyzing obtained values it can be stated that the model generated from the optimized dataset DTMv2 has a better fit for the reference model than the DTMv1.

Conclusions

In this study the results of the classification were presented. The classification criterion was the intensity of the reflected signal. In addition, the optimization process was analyzed as factor changing the size of original datasets. From analyses the following general conclusions can be draw:

– The intensity parameter is not sufficient for the proper distribution of the point cloud to a object group and cannot be the only parameter taken into account in the classification,

– Optimization streamlines the process of classification and generating DSPM and DTM, improves readability of DSPM, does not cause loss of important information.

Specific conclusion can be drawn as follows:

– The intensity assigns for the trees and bushes appears in all selected ranges, in both variants,

– Generated Digital Spatial Models illustrate the point distribution of the intensity parameter for each class of object. Such models may be useful in filtering and classification and in analyzing the propriety of these processes,

– Optimization resulted in a reduction of about 50% of point cloud, but this did not affect the correctness of the classification,

– Digital terrain model generated from an optimized dataset better fits to the reference model which is DTM generated form dataset filtered by means of active TIN model method.

ALS point cloud processing is time and work consuming process, mainly because of the size of dataset. Therefore conducting optimization before such processes is reasonable. Optimization has an impact on efficiency of DTM generated from ALS point cloud.

Translated by Anna Sobieraj
References


