Final report of

Limits of applications of LiDAR point clouds and photogrammetry in urban and agricultural areas

KH 130427 project

In this project we aimed to explore the possibilities of terrestrial (TLS) and aerial (ALS) laser scanning in urban areas and field circumstances, in agricultural areas along practical tasks but with scientific thoroughness. We had the following aims:

- in plain areas: how reasonable the laser scanning (TLS) in plough lands, can ALS/TLS/UAV surveys reflect the microtopography, can we plan a monitoring to follow plant growth using UAVs, can we measure the wind erosion with remote sensing techniques?
- in urban areas: building identification and reconstruction with laser scanned and photogrammetric point clouds, what is the available accuracy, does the data fusion of TLS and ALS data help to gain better results, can we extract ground points (i.e. digital terrain model, DTM) from photogrammetric point clouds, can we identify land cover from RGB imagery with acceptable accuracy?

1. Plain areas

Plains are flat surfaces and their digital reconstruction is a challenging task. Agricultural areas, especially on plot level, are the areas where small differences of the topography can mean substantial consequences in e.g. runoff or pluvial ponding, sheet erosion, etc. Producing a highly detailed DTM has different issues regarding all possible surveying methods. (i) Traditional survey needs unreasonably dense standing points, and finally, spatial interpolation is needed, which smooths the surface eliminating the sudden changes. (ii) TLS surveys need fix points for triangulation, but the area is homogenous, thus, artificial signs (targets) are needed to be placed, and scanning stations can be merged in the GIS laboratory. Another issue is that the height of the device (i.e. source of the sign, the laser beam) is about 2 m and the angle is too small; accordingly, the surveyed area is within a 30-50 m radius circle. (iii) ALS can be a good solution, but its horizontal and vertical accuracy is not fitted to the plain surfaces; however, the only method which can handle the presence of vegetation due to multiple beam returns. (iv) Photogrammetry is cheaper than laser scanning, provides a dense point cloud and is efficient, but one should consider that the outcome will be a digital surface model, showing the surface of objects and DTM can be produced only on bare soil surface. Accordingly, we tested all remote sensing techniques for accuracy in different environments (consequently, we involved the TLS as a ground-based remote sensing technique).

1.1. Monitoring possibilities of wind erosion with remote sensing techniques

Regarding wind erosion measurements, we did not expect a final methodology and quantified map of the wind-blown materials, but we intended to explore the possibilities and limitations due to the predictable low amount of eroded soil. We have investigated the accuracy of the Digital Elevation Models (DEMs), as the main photogrammetric deliverables of the UAV surveys at bare soil conditions during the first stage of the vegetation cycle. The aerial mapping had been carried out using a Zenmuse X7 high-resolution (24 MP) camera combined with a very high LW/PH 35 mm (equivalent) lens. The application of this camera mounted on a DJI Matrice 210 RTK v2 and the DJI D-RTK 2 base station made it possible to assign direct georeferencing into the image metadata. This means that with this method, we were able to omit the use of Ground Control Points (GCPs).

According to our tests, this method generated DEMs with ±4 cm of vertical accuracy. The main advantage of this type of survey made it possible to keep the bare soil surface undisturbed, which would have been impossible in case we had to place the GCPs in the traditional way. However, the horizontal and vertical accuracy of the DEMs based on the direct georeferencing are in the survey-grade, but this bias (±3-5 cm) was in the same magnitude of the observable wind erosion rates. In other words, the RTK-UAV surveys can significantly decrease the survey time and keep the original surface conditions, but for a more accurate change detection of the ~3-5 cm of soil erosion between annual blocks must be supported by additional fix control points, as a tool for precise co-registration of the dense 3D point clouds. Results are summarized in a manuscript and will be submitted to an international journal in this year.

1.2. Plant growth monitoring with UAV-based photogrammetry

We tested the accuracy of UAV-based DSMs for plant growth monitoring in an area near to Debrecen. 4 surveys were conducted with the DJI UAVs (Phantom 4, Matrice 210), and 4 plant species (cultivars: autumn wheat, sunflower, alfalfa and 2 maize plots + bare soil) were measured from 9 flight altitude (20, 30, 40, 50, 70, 85, 100, 130 and 150 m). Flight heights determined the ground sampling distances (GSD), too: 0.8, 1.2, 1.6, 2.1, 3.0, 3.5, 4.0, 4.5, 5.5 and 6.6 cm, respectively. An initial DTM was generated from a survey in the spring without vegetation cover. Plant heights were compared with 600 reference points (150/surveys) measured with a Stonex S9 RTK GNSS device. Both terrain height and plant height were measured at each point. The final aim was to provide a methodology to an automatic procedure: an accurate map for differences of a DTM and the DSMs of the plant heights (i.e. normalized DSM, nDSM) showing the plant height without (or minimized amount of) time consuming field measurements.

We determined the flight altitudes providing the least error. Flight altitude at 20 m seemed inappropriate to provide an accurate model. 0.8 cm GSD just sounded good, neighboring pictures had too large perspective angle difference, and the DSM contained too much artefacts (i.e. errors). Generally, we found that too low altitudes provided too much details on the photos and in case of plants like maize, thin stems raised bias over accuracy; i.e. GSD should be larger than 3 cm to blur the unnecessary details. Statistical analysis revealed that above 85 m altitude there were no significant differences ($p<0.05$), so 85-130 m provided similar outcomes. Accordingly, the suggested flight height should be minimum 85 m when the aim is to monitor plant height and growth.

Accuracy was different by plants. As plants have different appearance (most importantly the leaf area index, and specific density), which also changes during the growing season, the surface they form is also variegated. Accordingly, cereals and alfalfa having a constant area cover from an early stage of growing, which ensured the generation of an accurate DSM very soon. Best results were found in case of wheat (generally cereals), but for maize and sunflower precise measurements require the state when leaves form closed, continuous surface, otherwise photogrammetry-based surface models are biased by the ground points among the plants. When accuracy was justified, plant growth intensity was also possible during the survey.

1.3. Testing ALS data filtering in a floodplain

We chose a pilot area in the floodplain of Tisza beside the Rakamazi Nagy-morotva to test the ALS point filtering efficacy. The area had only 1-2 m difference in the terrain height, and swales, point bars, paleo-channels, crevasse channels, and levees were the geomorphic forms causing some diversity on the surface. Paleo-channels and swales were depressions, having better water supply, periodically (coinciding with flood events) water cover, which ensured denser vegetation than the point bars and levees at a bit higher position, running dry at first after floods. We performed a thorough analysis to reveal the relevance of the noise filtering of 3d point clouds. Two noise filtering methods (a neighborhood based, and a distance based) and the original dataset were involved as input data, then, we applied the Cloth Simulation Filter (CSF) to filter out the ground points from the aerial LiDAR point cloud with 2 setting for CSF threshold (2 and 5), and 3 parameters for cloth filter (0.5, 1.0, and 2.0). There were suggested values considering the point density, but, according to our experiences, these values often can be too general to be true for each task, and we also intended to see the direct consequences of setting smaller or larger values. After that, 5 interpolation techniques (different data preparation [filtered LAS dataset and terrain dataset] for linear interpolation, natural neighbor, Topo to Raster; altogether 5) were used to create digital terrain models. We conducted a field survey with 604 data points measured with RTK GNSS, measuring both swales and point bars in the doughtiest period, and these points were used as reference data. Finally, 180 models ([2 filtering methods + 1 original] × [6 combination of CSF filter settings] × 5 types of interpolations × 2 spatial resolutions) were tested to find the best settings to gain the most accurate digital elevation model. Welch-test with Tukey post-hoc test, robust two-way factorial ANOVA, Wilcoxon test with Monte Carlo analysis and Spearman correlation was used in the analysis.

We found that Cloth Simulation Filter (CSF) was an effective tool in removing the vegetation from the surface, but the fine tuning was an important step, wrong values (even in the suggested range) resulted in serious errors. Swales usually had larger errors than point bars due to denser vegetation. We determined the optimal range in the settings which had acceptable errors (as there was no absolute accuracy) in the resulting terrain model.

Main conclusions were the followings. We found that all types of preliminary noise filtering had significantly more accurate results related to the processing than when only applying the original database. CSF as a ground classification technique was a powerful tool, and the resulting DTMs had very low errors (from −0.03 to −0.22 m as upper and lower quartiles). CSF parameters had a significant effect on accuracy, where a coarser cloth size (5 m) and a smaller threshold (0.2) resulted in the best model performance (the suggested value was of 2 m as cloth size). In the case of interpolations, we have drawn two main conclusions: the natural neighbor method provided the most accurate model considering the medians; regarding the range of the differences, the topo to raster and terrain dataset approaches with natural neighbor interpolations provided the best DTMs. Although the density of the ALS point cloud made it possible to use a 1 m geometric resolution for the final DTM, the 2 m resolution was more accurate. We also revealed that landform elements, even when the line of sight is not limited by the topography, can decrease the models' accuracy; swales had significantly larger model errors due to denser vegetation and water absorption related to point bars. These results can serve as a guidance for ALS LiDAR point cloud preprocessing, classification, and interpolation and for choosing the right resolution in a fluvial environment.

Results were published in the 'Sensors' journal.

Szabó, Zs.; Tóth, C.A.; Holb, I.; Szabó, S. Aerial Laser Scanning Data as a Source of Terrain Modeling in a Fluvial Environment: Biasing Factors of Terrain Height Accuracy. Sensors 2020, 20, 2063. https://doi.org/10.3390/s20072063

1.4. Classification of floodplain geomorphological forms using an ALS based DTM

Using the same ALS point cloud we chose the model having the smallest error and determined 60 geomorphometric indices. We performed an analysis to reveal whether point bars, swales, crevasse channels and levees can be identified using this detailed terrain model. Pixel-based and segmentation-based techniques were compared. Having 60 variables in a study can cause high classification accuracy, but results can also suffer from the phenomenon of overfitting. Another problem can be the representativeness of the reference data, i.e. raises the question if we would have a different set of training data, do we get the same (good or poor) results. Accordingly, we investigated the overfitting and the representativeness (stability) issues on classification accuracy and we developed a method to study and quantify these phenomena. We used the Random Forest algorithm for the classification. We delineated 105 point bars, 127 swales, 20 crevasse channels and 13 levees (subdividing the existing 2) with visual interpretation and vectorization. We applied two different approach: (i) an object-oriented using the forms to extract means from each raster layer of geomorphometric indices (OO), (ii) we selected 5000 points to extract single pixel values within the forms and used the points on each raster layer of geomorphometric indices (PB). 70% of the data was used for training, 30% for testing, furthermore, models were built using the 10-fold cross-validation with 30 repetitions and hyperparameter tuning.

We found that visually segmented geomorphological forms (265 forms via the OO-approach) of the floodplain provided a very accurate classified final map with an overall accuracy of 95% accuracy. Pixel-based approach did not require so accurate visual interpretation and accurate delineation of the forms (i.e. can be regarded as more 'supervised' classification independently of field knowledge) but the classification accuracies were lower, maximum 79.5%. We were able to identify the 11 most important morphometric indexes providing 78.3% accuracy. Uncertainties were the largest for the swales and point bars. Overfitting issues were larger in case of pixel-based models.

Main findings were the following. A large number of morphometric variables can be used efficiently in the identification of levees, crevasse channels, point bars and swales. However, a larger number of variables did not ensure a relevantly better model performance. Recursive Feature Elimination, as a variable selection technique, helped to find the fewest variables making the largest contribution to obtain the grates' accuracy. Our main finding was that the selected variable set can change by model runs; the maximum overall accuracies (OAs) were almost the same. Although the variables were not the same in the repeatedly conducted models, we were able to identify the most frequent ones. Involving four variables in the case of the PB-approach and two variables in the case of the OO-approach provided sufficient accuracy, and the errors did not differ relevantly from the maximum number of geomorphometric indices. OO and PB-approaches performed differently: the object-oriented approach was more successful with 95% OA, while the 78% OA of the pixel-based approach was a weaker performance; nevertheless, all the forms were identifiable despite the misclassifications. The probability of the classifications and the pixel-based spatial uncertainty (as different classification outcomes for the same pixels) was not an appropriate tool to evaluate the classification efficiency, because the values were not in accordance with the class level accuracy metric (F1s). Overfitting was in accordance with the optimal number of variables: the lowest level of overfitting coincided with the high OAs of the optimal number of variables. We emphasize that the most important variables ensured accurate models for fluvial forms, but the selection methodology was more important. Different aims and target geomorphological forms can also be identified with the help of geomorphometry after a careful variable selection. The 78% accuracy of the PB-approach can be regarded as acceptable, as the fluvial forms studied had similar characteristics. These results, including the methodological findings, can help the water management directorates to evaluate the floodplain from a flood-management perspective and to find common points with nature conservation planners, to preserve the most valuable habitats.

Results were published in the 'Remote Sensing' journal.

Csatáriné Szabó, Z.; Mikita, T.; Négyesi, G.; Varga, O.G.; Burai, P.; Takács-Szilágyi, L.; Szabó, S. Uncertainty and Overfitting in Fluvial Landform Classification Using Laser Scanned Data and Machine Learning: A Comparison of Pixel and Object-Based Approaches. Remote Sens. 2020, 12, 3652. https://doi.org/10.3390/rs12213652

2. Urban areas

2.1. Literature review on TLS and ALS studies of geoscience

We performed a detailed literature review on the topic of LiDAR point cloud applications in the urban areas. LiDAR technology is a rather new way of surveys, which provides opportunity for object detection and feature extraction through large data clouds. We aimed to reveal the main focus areas of the applications, the relevance of geosciences in the research with laser scanning focusing on urban areas, and the ratio between the terrestrial (TLS) and aerial (ALS) techniques through a literature analysis.

Main findings are the following. We found that LiDAR has about a 50 years history, but the first applications were of meteorological measurements. The first studies were published only from 1998 in the topic of land surveying. Ratio of geosciences papers related to the total number of studies represented 30-40%. Regarding the platform of the equipment, TLS had a larger relevance, exceeding the number of studies using ALS more than three times. The difference in urban applications was only twice larger in case of TLS than ALS. Papers aiming the survey of vegetation in urban areas were 1.5 times higher using TLS techniques, but the proportionally 8% of ALS and 4% of TLS surveys dealt with this topic. Although only a few studies were published where laser scanning, urban vegetation and microclimate were the subject of the analysis, an increasing trend can be expected after the current stagnant state. First researches had promising results to help the decision makers to mitigate the hot wave issues of large cities.

Results are published in Geographica Pannonica journal.

Szabó, Zs.; Schlosser, A.; Túri, Z.; Szabó Sz. A review of climatic and vegetation surveys in urban environment with laser scanning: a literature-based analysis. Geographica Pannonica. 2019, 4, 411-421. https://doi.org/10.5937/gp23-24675

2.2. TLS as a tool in urban area surveys

We conducted a TLS survey in the Nagyerdei Békás-tó park, in Debrecen. The aim was to extract the trees from the point cloud and to gain tree parameters. We performed TLS and ALS data fusion to examine the advances. We examined several software dedicated to provide tree characteristics, but found that there were several difficulties. Thus, we developed a simple method to determine DBH (Diameter of Breast Height). The verification of the results showed ~5% error based on 39 reference field measurements. Results will be the part of the PhD dissertation of Aletta Schlosser, and we also intend to submit to an international journal.

2.3. Dense point cloud as input to extract ground points

We also aimed to explore the possibilities to use photogrammetry to create a digital terrain model (DTM). The primary product of aerial imaging is a DSM, but if we are able to extract the ground points with an acceptable accuracy, we can provide eligible data to interpolate a DTM, too. However, the parameters of ground point filtering are not obvious. We tested the different combinations of the parameters in GlobalMapper's Lidar Module. We specified the settings for Minimum Height Departure from Local Mean (MHDep: from 0.1 to 1.0), Base Bin Size (BBS: from 0.5 to 1.5), Maximum Height Delta (MHDelta: from 0.5 to 1.5), Expected Terrain Slope (ETS: from 1 to 10). Altogether, we produced 10000 models (but then reduced to 900 along multicollinearity), and tested for accuracy. An aerial LiDAR point cloud served as the reference data, and additionally, we also measured 250 ground points with RTK GNSS for reference data. We applied General Linear Modelling to reveal the effect of parameter settings on overall accuracy, sensitivity, specificity and precision. Results were evaluated by the effect sizes (ε^2p).

Our main findings were the following. The rank of the parameters showed that MHDep had the largest relevance followed by the MHDelta, and finally was the BBS. Interactions among the parameters had medium effect, but only one fourth of the main effects except the interaction of MHDelta and the MHDep, in this case the interaction had also large effect. MHDep alone explained the 86% of the variance of OA, while together with the MHDelta it was 96%. ETS and BBS had only a limited role. Suggested value for MHDep was 0.3, but we revealed that best results were achieved above 0.6 following a saturation curve (thus above 0.6 the OAs did not change relevantly). Results are summarized in a manuscript which will be submitted to an international journal this year.

2.4. Aerial images as input for land cover classification

As an urban pilot area, a district of Debrecen was chosen, the 'Úrrétje'. An orthophotograph and a photogrammetric point cloud was generated and processed. We classified five categories of the urban land cover (trees, buildings, roads/pavements, grass and other), 5 machine learning algorithms were tested with different data inputs. Beside the bands of orthophotograph spectral indices, texture indices and terrain indices (altogether 31 variables) were involved. We evaluated the input groups separately, in different combinations, using dimension reduction and feature selection algorithms. We found that the 'more predictor is better' (31 variables) was true, but using only 6 variables provided 1-2 worse overall accuracy. We also found that focusing only on building identification, textural indices had an important role and the six-variable version was among the best ones (to avoid overfitting the minimum number of predictors is desirable).

Main findings were the following. Classification performance using only one group of indices (i.e., RGB bands, texture, RGB indices or morphometric indices) varied in a wide range. Texture information was the weakest, worse when only RGB bands were used. Morphometric indices performed better on class level than on overall because DSM and its derivatives added valuable information especially in case of buildings. RGB indices had a relevant contribution in the improvement but on class level it was worse than the overall accuracy. Combination of different group of indices ensured higher accuracy both on overall and class level. Best option is to use the morphometric indices with the RGB bands; it had >90% OA, PA, and UA. Combining three types of indices provided the most efficient models, having >95% OA, PA, and UA. The RGB bands, RGB indices, morphometric indices and the 4 bit texture information had the largest (100% UA and 98% PA). In addition, 4 bit and 8 bit texture information had small differences in these combinations, and the most important to avoid their common application (both versions decrease the accuracy). Model evaluation should contain the UA and PA values, and having several model solutions, visualization of these metrics helps to find the trade-offs between omission and commission errors. In addition, F1 and IoU can express it with a single value which helps to create ranks of accuracy. RFE as a variable selection method provided an importance rank, and both the six and ten variable sets were efficient, providing the same accuracy as including all variables (100% UA and 98% PA). We suggest using the fewest number of variables to avoid overfitting. However, our most important variables (nDSM, RGBVI, GLI, blue band from RGB, slope, VARI) can be different in other study areas, so the methodology and the careful and customized variable selection is more important. Efficiency of this approach can be limited in areas where high buildings have large shadows and building roofs are flat. While shadows bias the spectral profiles, flat roofs will be identical with roads, pavements, and parking lots; thus, slope and aspect cannot discriminate between buildings.

Results confirmed that archive images can provide appropriate data for urban sprawl monitoring focusing directly on the buildings.

Results were published in the ‘Remote Sensing’ journal.

Schlosser, A.D.; Szabó, G.; Bertalan, L.; Varga, Z.; Enyedi, P.; Szabó, S. Building Extraction Using Orthophotos and Dense Point Cloud Derived from Visual Band Aerial Imagery Based on Machine Learning and Segmentation. Remote Sens. 2020, 12, 2397. https://doi.org/10.3390/rs12152397