

3D Point Cloud Classification of Natural Environments Using Airborne Laser Scanning Data

Ana-Maria Loghin¹, Valeria Ersilia Oniga¹, Catrinel-Raluca Giurma-Handley¹,

¹(Department of Terrestrial Measurements and Cadastre, Faculty of Hydrotechnical Engineering, Geodesy and Environmental Engineering / „Gheorghe Asachi” Technical University of Iasi, Romania)

Corresponding author: Ana-Maria Loghin

ABSTRACT: Airborne Laser Scanning technology has been recognized as a valuable and high precision topographic source for mapping and 3D modelling of the Earth surface. 3D point cloud classification in natural environments has become an important topic due to an increasing number of applications in geomorphology, 3D cartography, 3D modelling and environmental issues.

This article presents two different three dimensional classification approaches applied for Airborne Laser Scanning point clouds. Firstly, a supervised machine learning classification algorithm using random forests, decision trees and training datasets and secondly a semi-supervised algorithm composed by a cascade of binary classifiers, based on Support Vector Machines are used for labeling each LiDAR return of the input point cloud as one of the following categories: ground, water, vegetation and gravel.

The 3D classification results are demonstrated for a near natural reach of the Pielach River, the “Neubacher Au” area, located in Lower Austria, with Airborne Laser Scanning data acquired from 26 February 2015, consisting in 2.8 million LiDAR returns over 45150 m², with a point density of 21 points/m².

KEYWORDS: classifier, decision tree, machine learning, random forests, supervised classification

Date of Submission: 09-02-2018

Date of acceptance: 26-02-2018

I. INTRODUCTION

Light Detection and Ranging also known as LiDAR is an optical remote sensing technology that uses light pulses from lasers, illuminating the target objects, that are not in a direct contact to the sensor. As a main result of this technology is the geometrical distance measurement, but also additional information like: Amplitude, EchoNumber, Number of Echos, Red, Green, Blue, InfraRed channels, Scan direction.

Airborne laser scanning is an important surveying technology used in a multitude of applications, in fields such as topographic, environmental, industrial and cultural heritage 3D data acquisition. One of the main advantages of ALS is the acquisition of three-dimensional (3D) data with a high precision, based on polar measurements (angles and distances) [1].

Airborne laser scanners are non-contact measurement instruments, that can capture and record the geometry and the textural information of visible surfaces of objects and sites. The resulting point clouds from laser scanning surveys contain points both on the terrain and also above it, represented by vegetation, trees, buildings.

The measurement principle of a laser scanner is based on the emission of short pulses of infrared light which travel through the atmosphere, are scattered and partially absorbed by any objects in the instantaneous field of view of the laser beam. In case of many object surfaces like: vegetation, bare ground, building surfaces, the phenomenon of diffuse reflection takes place. In this case, a part of the incident light is scattered back to the sensor and the signal is recorded. Therefore, having the time phase difference between transmission of the pulse until the detection of its echo and the well known speed of light, the distance from sensor to the object can be computed [2]. The resulted collection of scattered points with a known geometrical 3D position in space $p_i = (x, y, z) \in P$ is a so-called point cloud.

II. REVIEW OF EXISTING CLASSIFICATION METHODS

Classification of point cloud data is a major application met in Computer Vision, being used in the automated processing of point clouds. This process involves the labeling of each point from the original input data, in order to define different class types of points. For example points located on buildings, streets, vegetation, trees, water, earth surface, gravels.

In literature there are mentioned two main types of classification methods: the first one is the cluster analysis and the second one is the object-oriented classification. Cluster-based classification is used in image processing, being based on the geometrical features of objects or their spectral characteristics [3]. Li et al [4] used an approach that combines high-resolution aerial imagery with LiDAR data for an object-oriented urban land-cover classification, based on Support Vector Machines. Object-based classification is more popular and uses the geometrical and spectral object characteristics that show spatial and spectral differences for object recognition.

In Carlberg et al [5] is presented a multi-category classification system that works in 3D domain, for identifying water, ground, roof and trees in airborne LiDAR. It is organized as a cascade of binary classifiers, each of them performing an unsupervised region growing, followed by supervised, segment-wise classification, using a 3D shape analysis and region growing technique.

The classification of aerial LiDAR point clouds plays an important role in point cloud processing and analysis, in 3D modelling and object recognition applications. In the point classification step, each LiDAR point is classified into the most meaningful categories such as ground, vegetation, building and water, based on the data properties. The accurate classification is crucial for achieving accurate 3D city models, building extraction, DEM generation, because the result of the classification is directly used in these applications [6]. Regarding the natural environments, in these areas, 3D point cloud classification has become an important topic, due to an increasing number of applications in: geomorphology, 3D cartography, 3D modelling and environmental issues.

III. PRESENTATION OF THE STUDY AREA, MATERIALS AND EQUIPMENT

The study area Neubacher Au is located at the lower course of the pre-alpine Pielach River near the Lower Austrian town Loosdorf ($48^{\circ}12'53.2''N$ $15^{\circ}22'25.1''E$, WGS 84), at about 100 km west of Vienna.

Neubacher Au is a Natura2000 conservation area featuring a meandering river course. It is a riparian area, covered by grassland and alluvial forests (mostly deciduous trees with heights up to 35 m). On the northern side of the river, open grassland with high and low shrubs is present [7].

The Lidar measurement campaign took place on 26th of February 2015 with a Riegl LMS-Q1560 sensor, providing a full waveform with a laser wavelength of 1064nm and a point cloud density of approximately 21 echoes / km². This full-waveform lidar has the capability to record the complete waveform of the backscattered signal echo, which provides in addition to range measurements other attributes namely echo amplitude, echo width and information on multiple echoes from one shot.

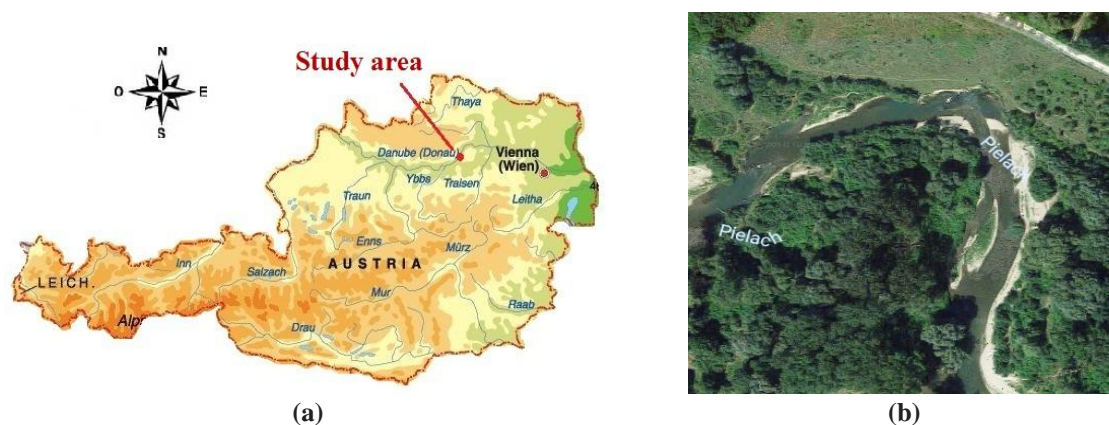


Figure 1: The study area, Neubacher Au, Lower Austria; (a) Overview map of Austria; (b) location of the study area



Figure 2: Terrestrial photos of Neubacher Au study area

IV. APPLIED ALGORITHMS FOR 3D POINT CLOUD CLASSIFICATION

Classification of point clouds in natural environments is a difficult task, due to the heterogeneity and complexity of natural surfaces. This complexity arises from the non-uniformity of individual objects, different grain sizes, different types and age of the vegetation (grass, bushes, medium and high trees), the variable lithology, riverbed dynamics morphology with depositions, debris and erosions. Therefore, in contrast to point cloud classification in urban environments, characterized by artificial structures with regular geometries, like: streets, buildings with simple geometry – plane surfaces and sharp angles, the classification of point clouds in natural environments is more complex, involving more processing steps and a deep analysis.

In this study two different 3D classification approaches are used for labelling each LiDAR return of an input point cloud corresponding to a complex natural environment into four different classes: ground, gravel, water and vegetation.

Therefore there is presented a comparison of two different classification methods applied to a point cloud corresponding to Neubacher Au study area. The first one is a supervised machine learning algorithm, which is based on a Supervised Classification combined with Decision Trees and Random Forests and the second one is a semi-supervised algorithm, containing a cascade of binary classifiers generated with the Support Vector Machines (SVM) technique.

V. PROCESSING STEPS

Preprocessing of the ALS point clouds (direct georeferencing, strip adjustment) were performed using the Riegl software suite RiProcess, software designed for managing, processing, analyzing, batch registration, filtering tools and visualizing data acquired with airborne laser scanning systems.

As a pre-processing step, a point cloud filtering of the outliers was applied. In this step, single lidar points having heights over or under the medium value were considered outliers and were eliminated. The remaining point cloud is then treated separately with the two classification methods implemented in OPALS and in Canupo respectively. In the final step, the results were compared and validated.

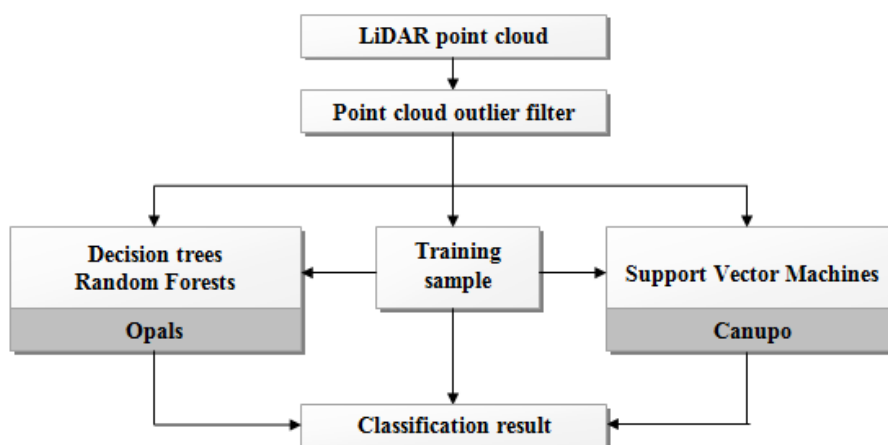


Figure 3: Overview of the processing steps

VI. SUPERVISED CLASSIFICATION ALGORITHM

The supervised classification algorithm applied to the LiDAR point cloud is based on Decision Trees and Random Forests. Random Forests is known as an ensemble machine learning algorithm, which contains a group of classifier learners with improved performance, based on multiple decision trees. Random Forests are classifying objects, by using their attributes, having an excellent performance in classification tasks [8]. They run efficiently on large data sets, are able to handle thousands of input variables and have multiple applications in domains such as: medicine, market industry, Computer Vision (CV) for image classification.

A Decision Tree is a graph that uses a branching method to illustrate a model of decisions and their possible consequences. It is a decision support, based on a branching series of boolean tests, containing conditional control statements. The main components of a decision tree are: nodes – these are the decision points where the attribute is being tested, branches – the two different outcomes of the test and leaf nodes – these are the class labels (the final decision taken after computing all the attributes).



Figure 4: Workflow for the supervised classification algorithm

In this work, Random Forests are applied to classify airborne lidar data on a natural scene and to select the most important attributes that are going to be used in the classification process. The entire processing chain, written in Python programming language is implemented in the OPALS laser scanning software (Orientation and Processing of Airborne Laser Scanning data) developed by the Institute of Photogrammetry and Remote Sensing (I.P.F.) from the Technical University of Vienna [9]. As a first step, training datasets were selected and manually classified in Trimble DTMaster software.

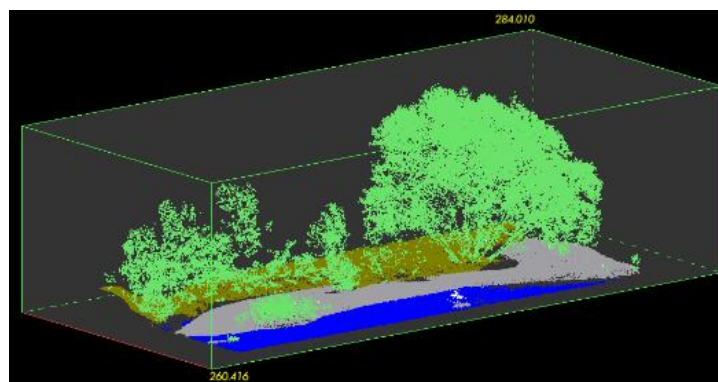


Figure 5: Manually classified point cloud;
green colour – vegetation, brown – ground, grey – gravel, blue - water

The second step of the workflow is the attribute computation in each lidar point. The features used in this study are classified into four categories, as follows: geometrical features (NormalX, NormalY, NormalZ, NormalSigma0, NormalizedZ, dZRange), echo features (Number of echoes, Normalized number of echoes), eigenvalue features (Planarity, Sphericity, Linearity, Anisotropy) and full-waveform features (EchoRatio, EchoWidth). Based on the computed attributes for each point, the classification model is created, by means of machine learning algorithms on the training datasets. This model is further applied to the full dataset, obtaining as a result the classified point cloud.

VII. SEMI-SUPERVISED CLASSIFICATION ALGORITHM

The semi-supervised classification algorithm applied in this study is based on a cascade of binary classifiers and on Support Vector Machines, also known as SVM. This classification method, described in [10] and especially developed for 3D point clouds in complex natural environments is implemented in Canupo suite, but it can be also found as a plugin in CloudCompare software. This tool is coded to handle large point cloud datasets and can be used by non-specialists of machine learning both in an automatic way, but also having a control of the entire classification process. The main idea that represents the basis of his classification procedure is the definition of the best combination of scales at which the dimensionality is measured, allowing the maximum separability of two or more categories. The used linear classifier is based on Support Vector Machines (SVMs), which are a set of related supervised learning methods that analyze data and recognize patterns for classification and regression analysis. SVM is developed from the neural network algorithm [3]. The algorithm takes a set of input data and predicts, for each given input, in which of the two possible classes belongs to, making the SVM a non-probabilistic binary linear classifier.

As it can be seen in Fig. 6, the method uses a cascade of binary classifiers, in three different processing phases. In order to assess the accuracy of each classifier result, two parameters were used: balanced accuracy (ba) for quantifying the performance of the classifier and Fisher Discriminant Ratio (fdr) for assessing the class separability. For example, a large value of ba, over 95 % indicates a very good classification and a large value of the fdr parameter shows that the classes are well separated.

In contrast to the supervised classification algorithms, this method minimizes user input of manually extraction and labelling of the points in the scene and maximizes the generalization and automation ability of the classifier. At a first beginning step there were defined samples for each object class: vegetation, water, ground and gravel.

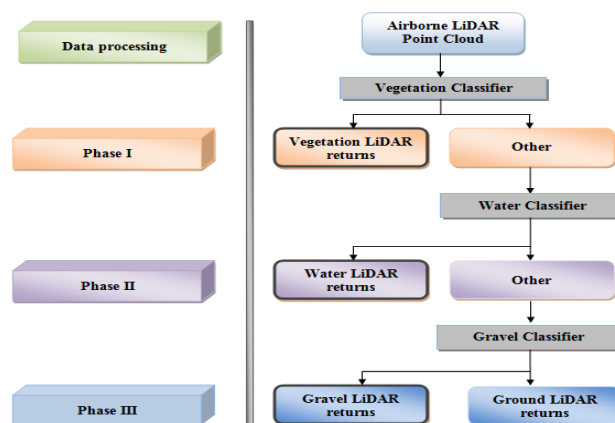


Figure 6: Workflow for the semi-supervised classification algorithm

In Fig. 7 there are presented the three classifiers applied to the point cloud: Vegetation classifier, Water classifier and Gravel classifier, with their accuracy assessment. The first classifier separates vegetation points from the other three classes, having a ba of 98.57% and a level of separability fdr of 9.32. Classifier 2 separates water points from ground and gravel surfaces. The automated training procedure resulted in a ba of 98.87%, value that indicates a very good classification and the large value 12.46 of the fdr parameter reflects a high level of class separability. Classifier 3 separates gravel surfaces from ground points, with an accuracy of 92.25% and a fdr value of 10.57.

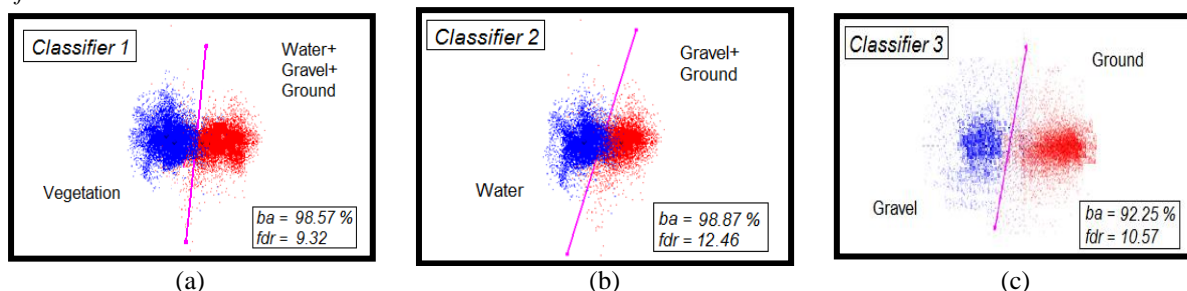


Figure 7: Semi-supervised classifiers applied in the classification process

VIII. CLASSIFICATION RESULTS

The final result of the applied 3D classification algorithms is the labelling of each LiDAR return of the input point cloud as one of the four following categories: vegetation, water, ground and gravel. As it can be seen in Figure 8, the vegetation points are represented in a light green colour, ground in brown colour, gravels with grey and water in blue colour.

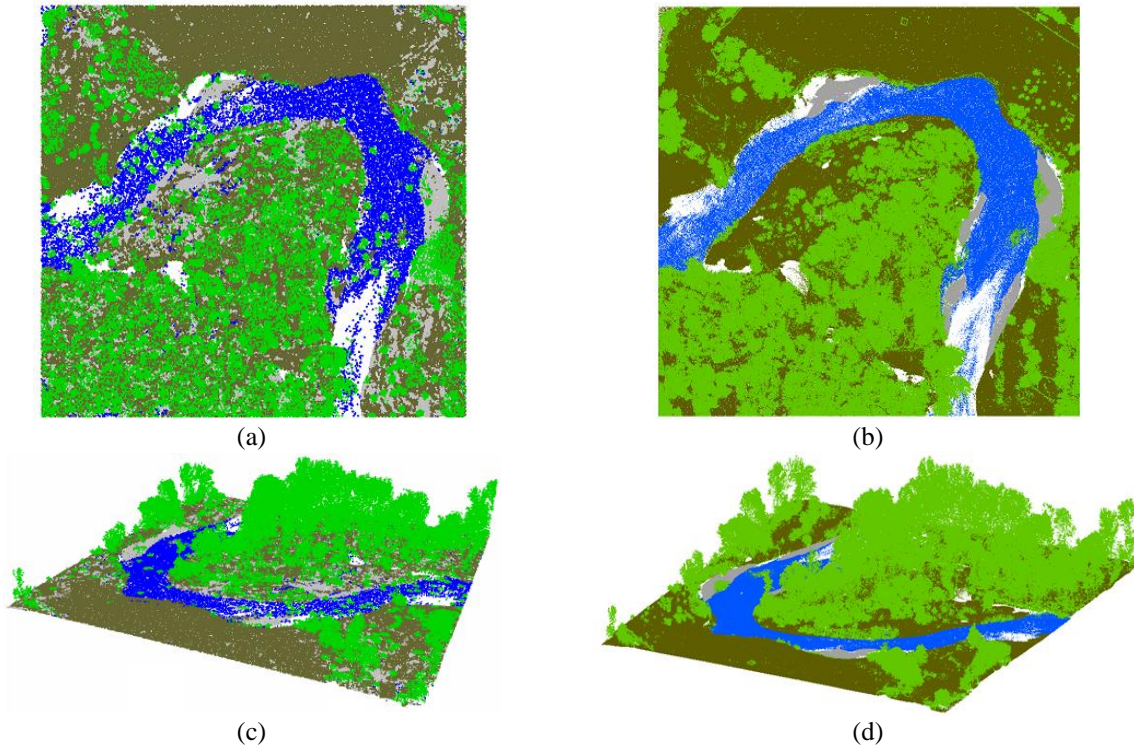
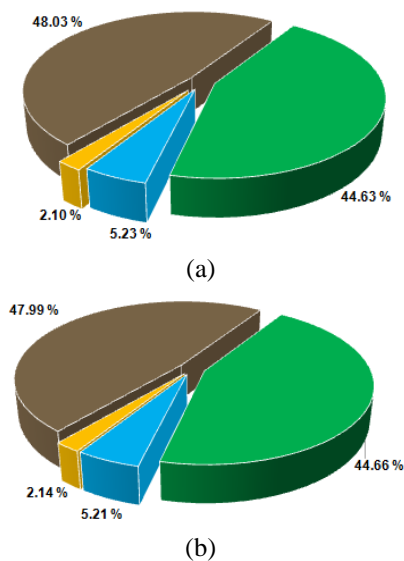


Figure 8: Result of the classification process for the Neubacher Au study area -top and 3D view (a), (c) using OPALS software, and(b), (d) using CANUPO suite

The graphics below show that from a total of approximately 2.8 millions of echos, the resulted percentages of points per class are as follows: 48.03% ground points, 44.63% vegetation points, 5.23% water points and 2.10% gravel points using the Canupo suite software and 47.99% ground points, 44.66% vegetation points, 5.21% water points and 2.14% gravel points using the Opals software respectively.



Legend

- Ground
- Vegetation
- Water
- Gravel

Figure 9: Percentage of points per class (a) Using semi-supervised algorithm (b) Using supervised algorithm

IX. CONCLUSIONS

In this work two different methods: a supervised and a semi-supervised classification algorithms were successfully applied to a LiDAR point cloud in a natural environment with a complex morphology. The ALS data is collected by a Riegl LMS-Q1560 instrument, providing a full waveform point cloud. The applied classification algorithms provide many advantages, including the following: they are simple to understand and to interpret, are able to handle both numerical and categorical data, require little data preparation, use a white box model containing boolean logic operations and they are able for handling large datasets with a higher dimensionality. Both classification methods provide high accuracy, having similar results, the greatest differences of 0,05 % being registered for laser points from with class “ground”.

The Canupo suite is efficiently used in the process of automatically point cloud classification, allowing the generation of own classifiers that works as binary classifiers, separating the laser points into two different categories. The classifiers have a general character and can be successfully transferred and independently applied to any other point cloud, with the main condition that the measurement unit to be the same as that one, for which the classifier was defined. The classification accuracy is computed in each point, allowing the possibility of creating a new class of uncertainty, corresponding to the points where the classification is uncertain. The semi-supervised classification method uses the local geometry of objects and it can be applied for large point clouds containing from ten to hundreds of millions of points, on a normal computer and with a reasonable processing time.

Compared to 2D and 2.5D classification algorithms, the 3D classification algorithms described in this paper assign a label to every 3D point in the LiDAR point cloud without projecting it, thus allowing a correct identification of structures, components and natural elements.

The classified point clouds, with labelled points, offer a better visualization and can be further used in a variety of applications like: geomorphology, 3D cartography, 3D modelling and environmental issues.

REFERENCES

- [1]. A. Loghin, I. Giurma, E. Oniga, Digital Surface Models Derivation from Airborne Laser Scanning Data, “1 Decembrie 1918” University of Alba Iulia RevCAD, 22, 2017, 141-150.
- [2]. C. Waldhauser, R. Hochreiter, J. Otepka, N. Pfeifer, S. Ghuffar, K. Korzeniowska, G. Wagner, Automated classification of Airborne Laser Scanning Point Clouds, *Solving Computationally Expensive Engineering Problems. Springer Proceedings in Mathematics & Statistics, Volume 97*, 2014, 269-292.
- [3]. Q. Zhan L. Yu, Objects classification from Laser Scanning Data based on Multi-Class Support Vector Machine, *Remote Sensing, Environment and Transportation Engineering (RSETE)*, 2011, 520-523
- [4]. H. Li, H. Gu, Y. Han, J. Yang, Fusion of High-Resolution Aerial Imagery and LiDAR Data for Object-oriented Urban Land-cover Classification Based on SVM, *ISPRS Workshop on updating Geospatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs*, August, 2007, 179-184.
- [5]. . Carlberg, P. Gao, G. Chen, A. Zakhor, Classifying Urban Landscape in Aerial LiDAR Using 3D Shape Analysis, *Video and Image Processing (ICIP)*, 16th IEEE International Conference, 2009.
- [6]. N. Yastikli, Z. Cetin, Classification of LiDAR data with Point Based Classification Methods, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B3*, 2016, XXIII Congress, 12-19 July 2016, Prague, Czech Republic, 441-445.
- [7]. M. Wieser, M. Hollaus, G. Mandlbürger, P. Glira, N. Pfeifer, ULS LiDAR supported analyses of laser beam penetration from different ALS systems into vegetation, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-3*, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic, 233-239.
- [8]. N. Chehata, L. Guo, C. Mallet, *Airborne Lidar feature selection for urban classification using Random Forests, Laser scanning 2009, IAPRS- International Archives of Photogrammetry and Remote Sensing, Vol. XXXVIII, Part3/W8*, September 1-2, 2009, 207-2012.
- [9]. G. Mandlbürger, J. Otepka, W. Karel, W. Wagner, N. Pfeifer, Orientation and processing of airborne laser scanning data (OPALS)—Concept and first results of a comprehensive ALS software, *IAPRS- International Archives of Photogrammetry and Remote Sensing, Vol. XXXVIII, Part3/W8, September 1-2*, 2009, 55-60.
- [10]. N. Brodu, D. Lague, 3D Terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology, *ISPRS Journal of Photogrammetry and Remote Sensing* 68, 2012, 121-134.

Ana-Maria Loghin “3D Point Cloud Classification of Natural Environments Using Airborne Laser Scanning Data” American Journal of Engineering Research (AJER), vol. 7, no. 2, 2018, pp. 191-197.