Article

3D Modeling of Urban Area Based on Oblique UAS Images—An End-to-End Pipeline

Valeria-Ersilia Oniga [1,*], Ana-Ioana Breaban [2], Norbert Pfeifer [3] and Maximilian Diac [1]

1   Department of Terrestrial Measurements and Cadastre, Faculty of Hydrotechnical Engineering, Geodesy and Environmental Engineering, "Gheorghe Asachi" Technical University of Iasi, Professor Dimitrie Mangeron Boulevard 67, 700050 Iași, Romania; maximilian.diac@academic.tuiasi.ro

2   Department of Hydroamelioration and Environmental Protection, Faculty of Hydrotechnical Engineering, Geodesy and Environmental Engineering, "Gheorghe Asachi" Technical University of Iasi, Professor Dimitrie Mangeron Boulevard 67, 700050 Iași, Romania; ioana.breaban@tuiasi.ro

3   Department of Geodesy and Geoinformation, Technische Universität Wien, Karlsplatz 13, 1040 Vienna, Austria; norbert.pfeifer@geo.tuwien.ac.at

*   Correspondence: ersilia.oniga@tuiasi.ro; Tel.: +40-745-634-472

Abstract: 3D modelling of urban areas is an attractive and active research topic, as 3D digital models of cities are becoming increasingly common for urban management as a consequence of the constantly growing number of people living in cities. Viewed as a digital representation of the Earth's surface, an urban area modeled in 3D includes objects such as buildings, trees, vegetation and other anthropogenic structures, highlighting the buildings as the most prominent category. A city's 3D model can be created based on different data sources, especially LiDAR or photogrammetric point clouds. This paper's aim is to provide an end-to-end pipeline for 3D building modeling based on oblique UAS images only, the result being a parametrized 3D model with the Open Geospatial Consortium (OGC) CityGML standard, Level of Detail 2 (LOD2). For this purpose, a flight over an urban area of about 20.6 ha has been taken with a low-cost UAS, i.e., a DJI Phantom 4 Pro Professional (P4P), at 100 m height. The resulting UAS point cloud with the best scenario, i.e., 45 Ground Control Points (GCP), has been processed as follows: filtering to extract the ground points using two algorithms, CSF and terrain-mark; classification, using two methods, based on attributes only and a random forest machine learning algorithm; segmentation using local homogeneity implemented into Opals software; plane creation based on a region-growing algorithm; and plane editing and 3D model reconstruction based on piece-wise intersection of planar faces. The classification performed with ~35% training data and 31 attributes showed that the Visible-band difference vegetation index (VDVI) is a key attribute and 77% of the data was classified using only five attributes. The global accuracy for each modeled building through the workflow proposed in this study was around 0.15 m, so it can be concluded that the proposed pipeline is reliable.

Keywords: UAS point cloud; segmentation; classification; machine learning; 3D model; urban area

1. Introduction

Urban infrastructure and the appearance of cities have become topics of great interest in recent years, as the number of people living in urban areas is constantly growing. According to United Nations report from 2018, the planet's population in urban space has a proportion of 55% and will increase by 2.5 billion people in 2050, about 68% of the planet's population at that time [1]. Facing this demographic explosion, precise methods of visualizing and analyzing spatial information are needed to help manage and plan the expansion of urban areas. According to a preliminary study [2], more than 90% of the interviewed users considered that there is a real need for the reconstruction and visualization of urban 3D spatial data.

On an international level, the scientific community and the commercial sector are concerned about creating 3D digital models of urban areas that meet some basic requirements: be available for the studied area, have sufficient accuracy and a reasonable cost. Through well-known applications (Google Earth, NASA World Wind and Microsoft Virtual Earth) available via the Internet, anyone can access satellite and aerial images and virtual models of many important urban areas around the world. However, it is important to note that although users can have a realistic view of the object space, the accuracy provided by these applications is low, not sufficient for applications such as urban planning, urban management, construction of sidewalks, placement of public lighting poles, visibility analysis for military operations and the telecommunications sector. Usually, digital models of urban areas that have a high accuracy are very expensive.

3D reconstruction of urban areas is a complex and challenging research topic and there are many methods for this purpose developed in literature that can be classified according to five main criteria, according to [3]:

– Type of data used: optical (monoscopy, stereoscopy, multiscopy), LiDAR data: terrestrial (TLS) and airborne (ALS), UAS images, topographic measurements;
– System automation level: automatic, automatic with cadastral data, semi-automatic;
– Complexity of reconstructed urban areas: dense urban areas, peripheral urban areas, areas of activity and collective housing;
– The degree of generality of 3D modeling: prismatic, parametric, structural or polyhedral;
– 3D modeling methodologies: top-down or model-driven methods, bottom-up or data-driven methods and hypothesize-and-verify.

The bottom-up approaches reconstruct building models from the primitives extracted from the point cloud, 3D segments and/or 3D planes, without taking into account the a priori knowledge about the buildings to be rebuilt. In the top-down approaches, a-priori knowledge about buildings is introduced. It reconstructs building models from a set of previously known models to choose the one that has the best consistency with the data. These approaches are generally more robust than bottom-up approaches but are limited by the a priori model library. The hypothesize-and-verify approaches are hybrid approaches between the bottom-up and the top-down approaches. They consist, in a first stage, of the formulation of hypotheses starting from primitives through a bottom-up approach, then in the second stage, of the verification of these hypotheses through a top-down approach [3,4].

In addition, two modelling approaches can be distinguished: approaches based on finding the best fitting geometric primitives and approaches based on meshing methods [5]. However, a mesh 3D model can be used only for visualization or segmentation purposes since it does not contain semantic information [4]. Therefore, it is very difficult to go through and compare all methods for 3D reconstruction of urban areas, since they use different data, different levels of automation, etc. [6].

1.1. Level of Detail and CityGML

In order to assure the interoperability and the data heterogeneity of the real-world entities of the cities, their three-dimensional representation should be defined by the level of detail (LOD) concept included in CityGML, the main international OGC standard used for 3D urban models. Defined as a quality benchmark for accuracy and complexity [7,8], the characteristics of LOD are considered for each spatial feature from their acquisition, modeling, generalization up to data exchange and being applicable at different scale representations from city object to large areas.

CityGML version 2.0 includes five types of LOD starting from LOD 0 to LOD 4, where LOD 0 includes the digital terrain model with the ground footprint of the buildings. LOD 1 and 2 bring improvements in geometric and semantic complexity, and LOD 4 keeps the exteriors defined at LOD 3 but includes also details of the interiors of the structural elements. As mentioned by [9], a new improved version of the standard, CityGML version 3.0, refines the concept of LOD by removing the LOD 4 that is used for indoor modeling and integrating it in LOD 0 to 3 [10]. Regardless, of special interest is LOD 2 that relates with

the degree of detail of the structural elements specific to buildings that require a geometric resemblance close to reality and manages to maintain a suitable abstraction level.

1.2. Related Work

Some pipelines for the process of building modeling have been developed in the past, but in some cases the reconstruction process may fail because of the scene complexity, even if the generalization rate is high [11]. However, for the majority of these pipelines the input data is ALS point clouds [12–14] not UAS point clouds, which are very noisy compare to ALS or TLS point clouds and subject to occlusions due to vegetation and angle of incidence of the digital camera optical axis. Even so, the advantages of using UAS images for 3D building modeling are multiple: low cost compared to traditional aerial photogrammetry, low flight altitude resulting in high-resolution images and a small Ground Sample Distance (GSD) of less than 3 cm, and fast and flexible image acquisition and processing.

A workflow for 3D building modeling based on UAS images was proposed by [4]. However, only nadiral UAS images were used, being acquired for a rural area at a flight altitude of 50 m. Moreover, the segmentation and planar patches detection processes were done only for the buildings' roofs. An automated workflow starting from UAV-derived point clouds to LOD 2-compatible 3D model was designed in [15]. The research is limited to roof planes' extraction based on nadiral images. The final building 3D model is obtained by extruding the 3D roof segments to the Digital Elevation Model (DEM) to create a 3D mesh of the building. The conversion to a CityGML file is done by a Python script that creates the building facades using an extruding process of the roofs. In [16] a methodology for 3D buildings reconstruction is proposed starting from a dense UAS point cloud, using the point cloud segmentation into planar surfaces and 2D buildings boundaries from UAS derived orthomosaic by manual digitization. The compiled methodology of [16] is based on an automatic 3D building reconstruction algorithm developed by [17,18]. A pipeline for 3D buildings reconstruction was proposed by [19], the result being a LOD3 parametrized top-down model. Althought, the pipeline is based on a fusion of terrestrial and UAS images.

1.3. Proposed Pipeline

This paper presents a semi-automatic data-driven approach for the 3D modeling of a dense urban scene based on low-cost oblique UAS images, without additional information such as cadastral data, the result being a parametrized 3D model with the OGC CityGML standard, Level of Detail 2 (LOD2). Open data such as construction foot prints can be handy in the process of 3D building modeling, but in some regions freely available data is not yet implemented. In Romania, high quality cadastral data layers are not publicly available since can only be extracted by an authorized person from the Integrated System of Cadastre and Real Estate official national database. The data can be found only for the buildings that have been registered in the land book system.

For this purpose, five main pieces of software have been used: Opals (for approximately 80% of the process) [20], 3DF Zephyr [21], CloudCompare [22], LAStools [23] and AutoCAD Map 3D [24]. According to the authors' knowledge, the proposed pipeline is the first end-to-end pipeline for buildings' 3D creation based on UAS images starting from flight planning to the creation of the final parametrized building 3D model.
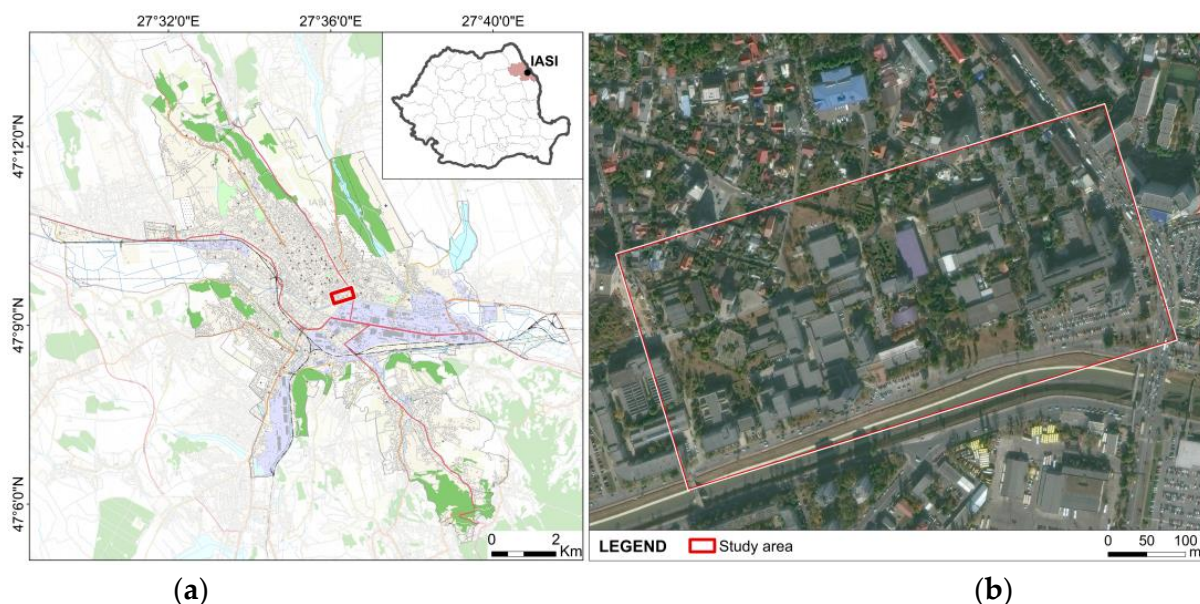
Our contributions can be summarized as follows:

1. Using a very high number of Ground Control Points (GCPs) and Check Points (ChPs), i.e., 320, to find the best scenario for UAS images orientation and georeferencing (artificial, well pre-marked by paint and natural points);
2. Taking measurements for buildings' roof corners using Global Navigation Satellite System-Real Time Kinematic positioning (GNSS-RTK) technology, with the help of a handcrafted dispositive, to test their influence in the BBA process and to assess the 3D buildings models' accuracy;

3. Presenting an end-to-end pipeline for 3D building models generation using only UAS point cloud, from raw data to final 3D models, not taking use of additional information such as buildings' footprints;
4. Presenting a new workflow for automatic extraction of vegetation UAS points, using a combination of point clouds and rasters;
5. Finding a key attribute in the process of UAS point cloud classification by random forest algorithm, i.e., the visible-band difference vegetation index (VDVI).

2. Study Area

The study area is a part of the university campus of the "Gheorghe Asachi" Technical University of Iasi, Romania, located on the left bank of Bahlui river, covering approximately 20.6 ha. It has a complex character, with the following characteristics: buildings of different sizes (more than 30), built architectures in a period of approx. 30 years with different functionalities (academic spaces, laboratories, water castles, weather stations, collective and individual dwellings, terraces, etc.), plot street, parking lots, street furniture, green area with low, medium and high vegetation and recreation areas (Figure 1).

Figure 1. Location of the study area: (**a**) Regional context; (**b**) "Gheorghe Asachi" Technical University of Iasi, Romania.
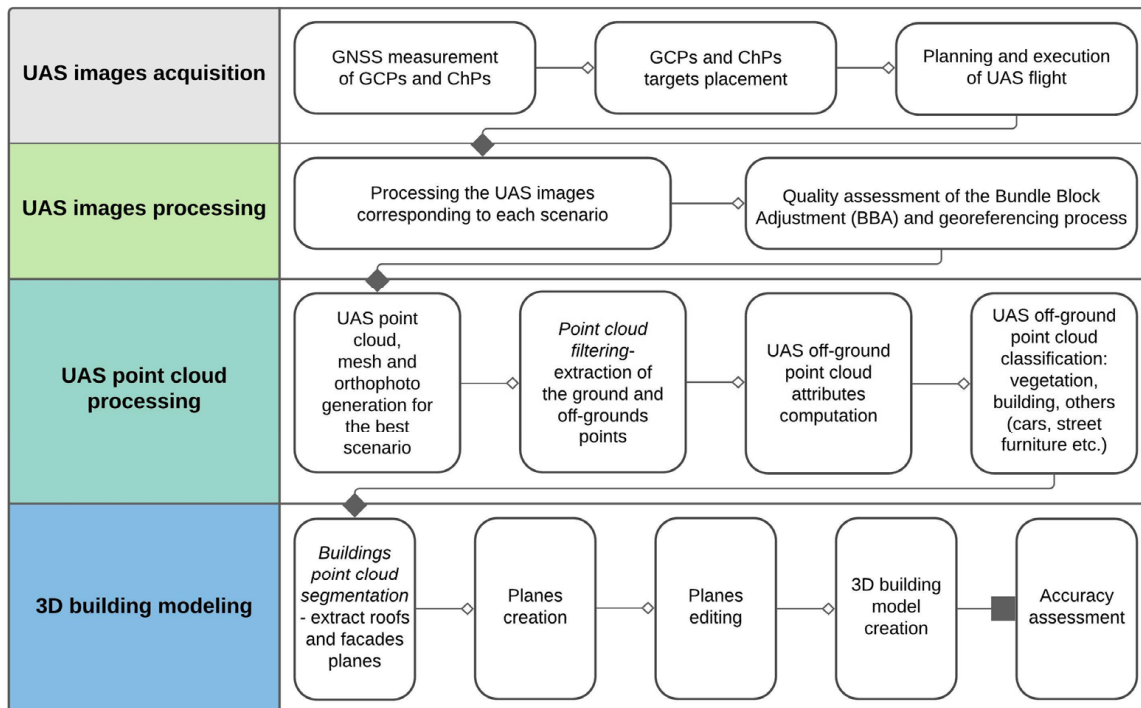
3. Materials and Methods

As stated in [25], the suitable number and distribution of the GCPs for the indirect georeferencing process is 20 GCPs with a stratified random placement. One of the conclusions reported in [25] was that using up to 20 control points improves the accuracy strongly, but for higher number of GCPs the improvements are only marginal. Consequently, this research was conducted based on previous findings, having the same number of GCPs pre-marked for the UAS flight.

The workflow applied for this study area can be structured in the following parts: field work with measurements for determining the coordinates of GCPs and ChPs, placement of targets for materialisation of GCPs and ChPs, planning and execution of UAS flight, choosing and creating different files with the desired GCPs and ChPs for each scenario, processing the UAS images corresponding to each scenario, quality assessment of the Bundle Block Adjustment (BBA) and georeferencing process, UAS point cloud, mesh and orthophoto generation for the best scenario, point cloud filtering to extract the ground and off-grounds points, attributes computation for the UAS off-ground point cloud and its classification into: vegetation, building and others (cars, street furniture, etc.), buildings

point cloud point segmentation to extract roofs and facades planes respectively, plane creation using region growing algorithm, plane editing using the AutoCAD 3D software, 3D building model creation and accuracy assessment (Figure 2).

Figure 2. Flow cart of research methodology.

3.1. GNSS Measurements

For this study, a total of 320 points have been measured in approximately one and a half weeks by GNSS real-time kinematic method (RTK) technology which proved to be a reliable and efficient method for determining the position in open areas [26], with centimeter-level accuracy when operating in RTK-fixed mode [27].

The measurements were performed with a South S82-T GNSS receiver with double frequency GPS + GLONASS + GALILEO + COMPASS + SBAS, using the Romanian ROMPOS positioning service. The planimetric transformation was performed using TransDatRo source code version 1.02 and the altitude transformation using the Romanian_ETRS89v1.gsf geoid and the Network Transport of RTCM via Internet Protocol (NTRIP) the Nearest_3.1 option which is less than 2 km from the Iasi permanent station in the National Geodetic Network.

The observations for each individual point took at least one minute, so the precision of the measured GNSS ground points fell within the range:

- Position Dilution of Precision (PDOP) $1.15 \div 3.99$ (<5), mean PDOP = 1.80,
- Horizontal Dilution of Precision (HDOP) $0.63 \div 3.16$ (<4), mean HDOP = 0.99,
- Vertical Dilution of Precision (VDOP) $0.95 \div 3.06$ (<4), mean VDOP = 1.49,

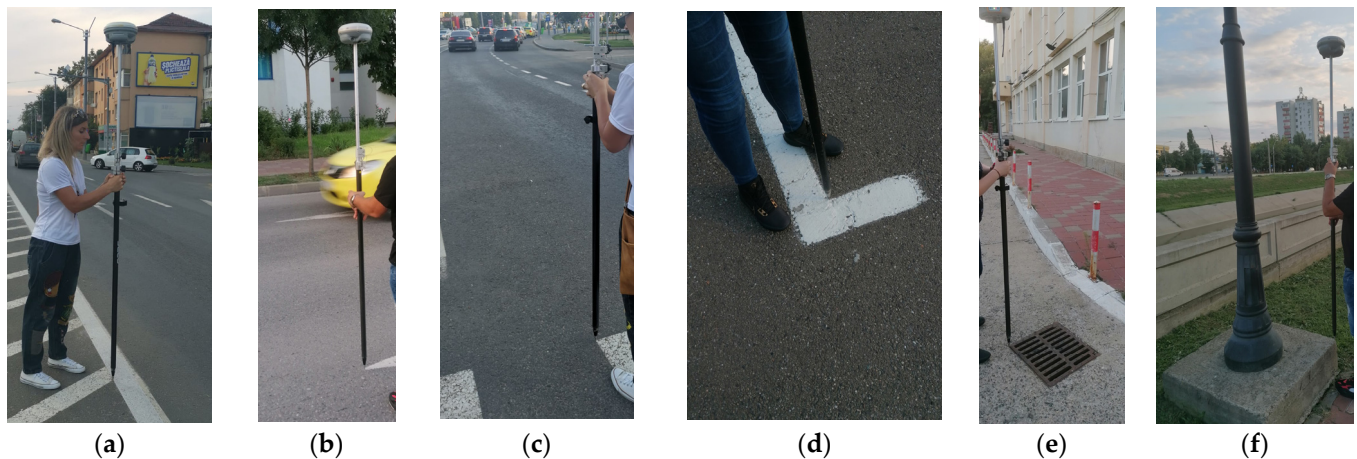and the precision of the measured GNSS roofs corners fell within the range:

- Position Dilution of Precision (PDOP) $1.16 \div 3.89$ (<5), mean PDOP = 1.73,
- Horizontal Dilution of Precision (HDOP) $0.62 \div 2.70$ (<4), mean HDOP = 0.93,
- Vertical Dilution of Precision (VDOP) $0.96 \div 3.05$ (<4), mean VDOP = 1.45,

Which proves a good configuration of satellite geometry with a good effect on positioning accuracy for both group of points [28].
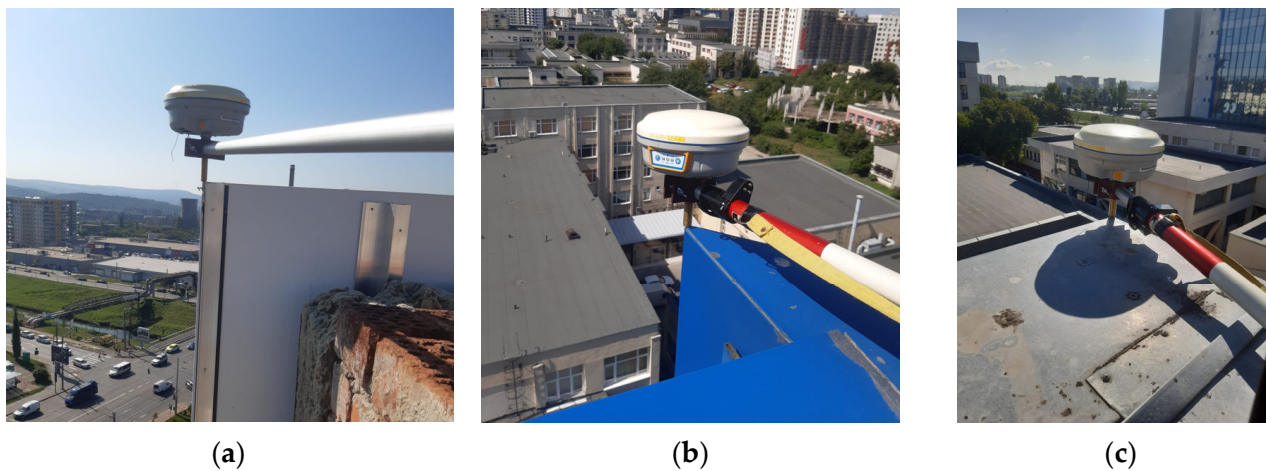
The most provocative part of RTK measurements in urban areas is that the GNSS signals are affected by high buildings, trees or other objects which introduce multipath errors [29], which is the main source of error in a constrained environment [30]. For this

reason, the ground measured points have been chosen as much as possible in open areas, obtaining the following positioning accuracy: mean _H = 0.012 m and mean _V = 0.018 m, and for the roof corners: mean _H = 0.011 m and mean _V = 0.016 m.

The nature of the 320 points is as follows: a number of 20 points have been ground marked with metallic bolts and paint, 89 points representing well pre-marked points by paint, i.e., street markings (Figure 3a,b), crosswalks (Figure 3c) and parking markings (Figure 3d), 77 points are natural points, representing sewer corners (Figure 3e) or lighting pole corners (Figure 3e) and a total of 134 points represents corners of buildings' roofs (Figure 4).

| (**a**) | (**b**) | (**c**) | (**d**) | (**e**) | (**f**) |

Figure 3. Ground control points (GCPs) and check points (ChPs) location: (**a**,**b**) street marking, (**c**) crosswalk, (**d**) parking marking, (**e**) sewer corner, (**f**) corner of lighting pole.
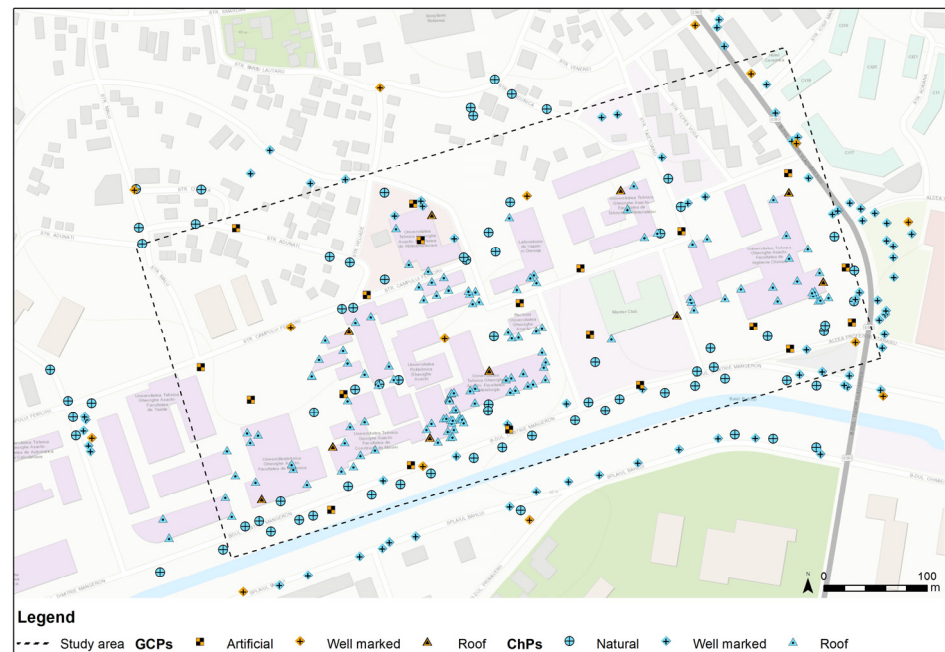

| (**a**) | (**b**) | (**c**) |

Figure 4. GNSS measurements of roofs' corners with a specifically crafted dispositive: (**a**–**c**) example of roofs' corners for different buildings.

For the buildings roofs' corners measurements, a dispositive was specifically crafted (Figure 4). The dispositive consists of an aluminium topographic pole and a part that allows for the lateral mounting of the GNSS receiver and the safe placement in the measuring point. To eliminate a lot of possible errors caused by deviations from the vertical, the dispositive ensures a very small height of 10 cm for the receiver.

The artificial, natural and roof points' distribution over the study area can be seen in Figure 5.
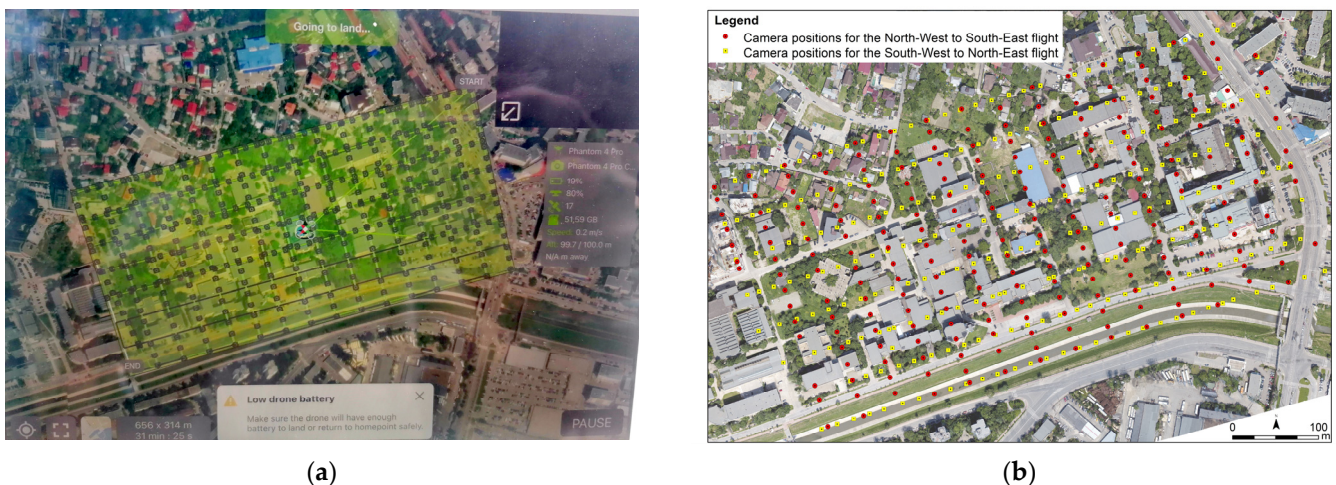
Figure 5. The spatial distribution of the 320 ground control points (GCPs) and check points (ChPs).

3.2. UAS Images Acquisition

The flight over the study area was made with a low-cost UAS, i.e., a DJI Phantom 4 Pro Professional (P4P), whose characteristics can be found in [31], at 100 m height at the beginning of June, with two flight itineraries (from North-West to South-East and from South-West to North-East) (Figure 6), the camera being oriented at 45°. A number of 492 images with 5472 × 3648 pixels resolution were acquired with a frontlap of 80% and sidelap of 70%, respectively, using the "Pix4D Capture" software for automatic flight planning, resulting a mean GSD of about 2.7 cm. The characteristics of the FC6310 camera mounted on the DJI Phantom 4 Pro UAS are summarized in [32].

(**a**)

(**b**)

Figure 6. (**a**) UAS image acquisition in automatic manner with "Pix4D Capture" software, (**b**) Camera positions for the North-West to South-East flight (red) and for the South-West to North-East flight (yellow) on the Unmanned Aerial Systems (UAS)-derived orthophoto.

During the flight, the 20 artificial GCPs have been marked with plates made by orange plexiglass and two black triangles, 3 mm thickness, 40 cm × 40 cm dimension, with a hole in the center of 5 mm diameter.

4. Results and Discussions

4.1. UAS Images Processing

For UAS image processing, the 3DF Zephyr Pro software was used, testing the accuracy of image georeferencing process in different scenarios, considering a constant number of 275 ChPs, representing natural points (74 ChPs), well-marked points (77 ChPs) and roof points (124 ChPs). The maximum number of GCPs was 45, as follows:

- 20 artificial points uniformly distributed over the study area,
- 10 roof corners uniformly distributed over the study area,
- 8 well marked GCPs situated in the exterior part of the study area,
- 7 natural and well-marked GCPs uniformly distributed over the interior part of the study area.

For UAS image processing, an initial project was created specifying the type of camera. Therefore, the calibration parameters have been automatically downloaded due to the high-quality pre-calibration of the Phantom 4 Pro camera. After importing the images, the software makes an approximation for the exterior orientation parameters for each camera position based on the EXIF information. All 320 GCPs and ChPs were manually measured on each oriented image that they appeared on (a minimum of 3).

Based on this one, new projects were created considering each time the same ChPs, while the number of GCPs varied: (1) a minimum of 3 GCPs, (2) 5 GCPs, (3) 10 artificial GCPs, (4) 10 artificial GCPs and 8 exterior GCPs, (5) 10 artificial GCPs and 10 roof GCPs, (6) 20 artificial points, (7) 20 artificial points and 8 exterior GCPs, (8) 20 artificial points and 10 roof GCPs, (9) 20 artificial points, 8 exterior GCPs and 10 roof GCPs, (10) 20 artificial points, 8 exterior GCPs, 10 roof GCPs and 7 interior GCPs (45 GCPs) and (11) only 8 exterior GCPs, each time importing the same image measurements for the GCPs.

In order to bring the results into the Romanian national coordinate system, a bundle block adjustment (BBA) process was performed, including only the GCPs as constraints (the "Constraint" and "Control" boxes were checked). At the end of the process the ChPs have been imported, so that they don't affect in any way the BBA process, the residual being automatically displayed. The file containning the computed X, Y and Z coordinates, for both GCPs and ChPs have been exported and the errors along each axis: $RMSE_X$, $RMSE_Y$ and $RMSE_Z$, together with the planimetric errors $RMSE_{X,Y}$ and the total error $RMSE_T$, have been calculated (Table 1). Appendix A individually highlights the spatial representation of the total $RMSE_T$ of the Ch.Ps for each scenario and the location of the GCPs.

Table 1. The residuals of 275 ChPs for 11 scenarios.

No. of GCPs	Residuals				
	$RMSE_X$ [cm]	$RMSE_Y$ [cm]	$RMSE_Z$ [cm]	$RMSE_{X,Y}$ [cm]	$RMSE_T$ [cm]
3	11.9	12.4	25.9	17.2	31.1
5	9.3	13.2	12.7	16.1	20.6
10	7.2	6.7	8.1	9.8	12.8
10 + 8	7.4	6.2	8.0	9.6	12.5
10 + 10	6.7	5.8	6.5	8.8	11.0
20 artificial	6.7	5.6	6.7	8.7	11.0
20 + 8	6.3	4.9	6.9	8.0	10.5
20 + 10	6.4	5.5	6.2	8.4	10.4
20 + 8 + 10	6.2	5.0	6.6	7.9	10.3
45	6.1	5.0	6.4	7.9	10.2
8 exterior	22.6	34.2	38.1	41.0	55.9

The results of the present study are in concordance with the findings of our previous study [25], even if the flight height (100 m vs. 28 m), the area (20.6 ha vs. 0.4 ha) and the UAS system (DJI Phantom 3 Standard vs. DJI Phantom 4 Pro Professional) are different, obtaining a planimetric accuracy of around 9 cm and a total accuracy of 11 cm when using
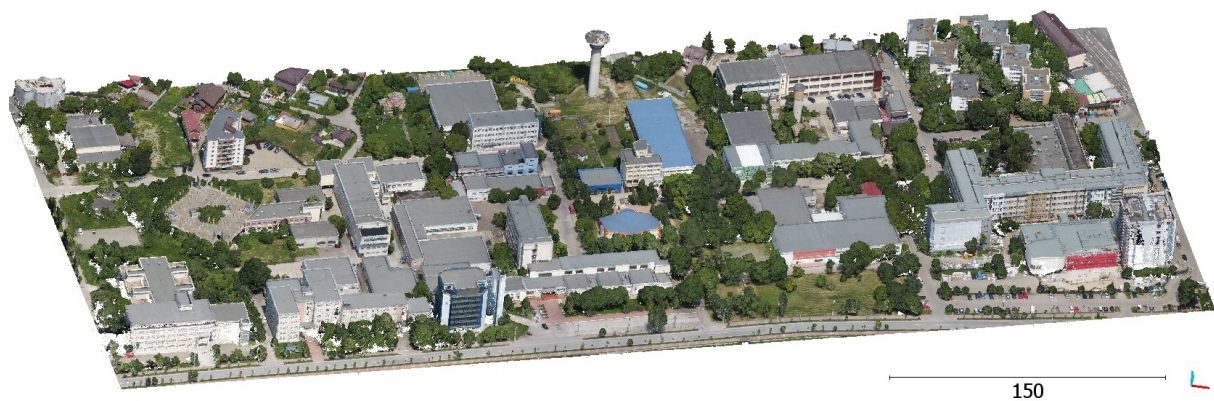
20 GCPs, uniformly distributed over the study area, comparable with 8.7 cm total accuracy in the case of 20 GCPs placed in a stratified random distribution. We have mentioned that 45 Ch.Ps out of 275 Ch.Ps are located in the exterior part of the study area, their residuals influencing the total error of each scenario.

The planimetric values ($RMSE_{X,Y}$) ranged between 7.9 cm to 41 cm and the total errors between 10.2 cm to 55.9 cm with the minimum value for the 45 GCPs scenario and the maximum error for the eight exterior GCPs scenario. The high values of errors corresponding to this scenario of around half a meter are reasonable, since no GCP was included in the interior part of the study area. It is worth mentioning that the differences between the 10 GCPs scenario and the 45 GCPs scenario are small, of 2.6 cm, and the differences between (10 + 10) GCPs scenario and (20 + 8 + 10) GCPs scenario in the range of a few millimeters. Further, when using 20 artificial well-distributed GCPs at ground level versus 20 GCPs (10 artificial and 10 roof corners), equal values for the total errors were obtained. As such, when the urban environment does not allow ground measurements due to GNSS errors, roof corners can be used instead.

4.2. UAS Point Cloud Generation

To derive a 3D dense point cloud, the "3D model generation" function was used, which allows different settings for the point clouds densification, i.e., presets, advanced and custom. For this research, the following parameters settings were applied from the advanced options: 50% image resolution, 5 number of near cameras and 50% noise filtering, obtaining a dense point cloud with 99,982,693 points (Figure 7), with a point density of 657 points/m^2.

(**a**)

(**b**)

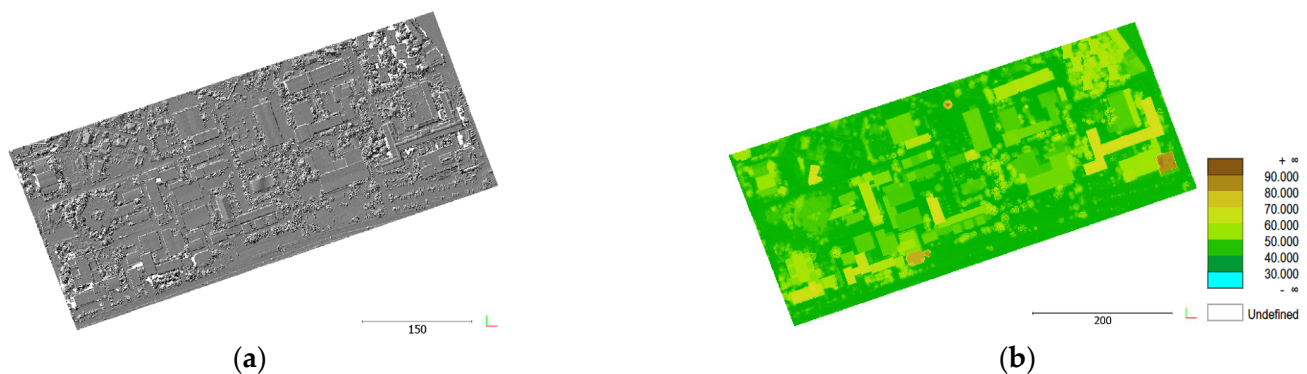**Figure 7.** (**a**) UAV point cloud of the study area, (**b**) detail.

4.3. UAS Point Cloud Processing

4.3.1. UAS Point Cloud Classification in Ground and Off-Ground Points

To classify the UAS point cloud into ground and off-ground points, two methods have been applied: the terrain mask algorithm [33] implemented into Opals software and Cloth Simulation Filter (CSF) [34] implemented in CloudCompare software.

When using the terrain mask algorithm, the basic input is the Digital Surface Model (DSM) of the study area in raster format.

For the DSM derivation, all UAS points have been interpolated using the "robust moving plane" interpolation method, implemented into "Grid" module from Opals software. The DSM resulted as a regular grid structure, with the grid cell dimension of 20 cm, for each grid cell the height being estimated by finding the best fitting tilted plane of the 20 nearest points of the grid point (x,y) found in the 60 cm search radius (3*cell size) with a quadrant oriented data selection, without taking into consideration the points detected as outliers. For each neighbor point, an individual weight was calculated by applying a weight function defined as the inverse of the squared distance (Figure 8a).

Figure 8. DSM of the study area in raster format, (**a**) using the robust moving planes interpolation method (0.2 m cell size, 0.6 m search radius, 20 neighbors), (**b**) obtained through a combination of interpolation methods, represented with the standard palette color coding.

As expected, information gaps remained (void pixels) and in order to fill them, another interpolation method was applied, namely Delaunay triangulation (OPALS- "delaunay-Triangulation"), with the following parameters: search radius 2 m and 20 neighboring points. To generate the final DSM without information gaps, the two rasters have been combined, by keeping the cells from "robust moving planes DSM" and taking the cells from the "Delaunay triangulation DSM" only when there are gaps of information in the first DSM, using the "Algebra" module from Opals software with the following formula:

opalsalgebra -inf DSM_rmp.tif -inf DSM_delaunay.tif -outf DSM_combined.tif -formula "r [0] ? r[0] : r[1]".
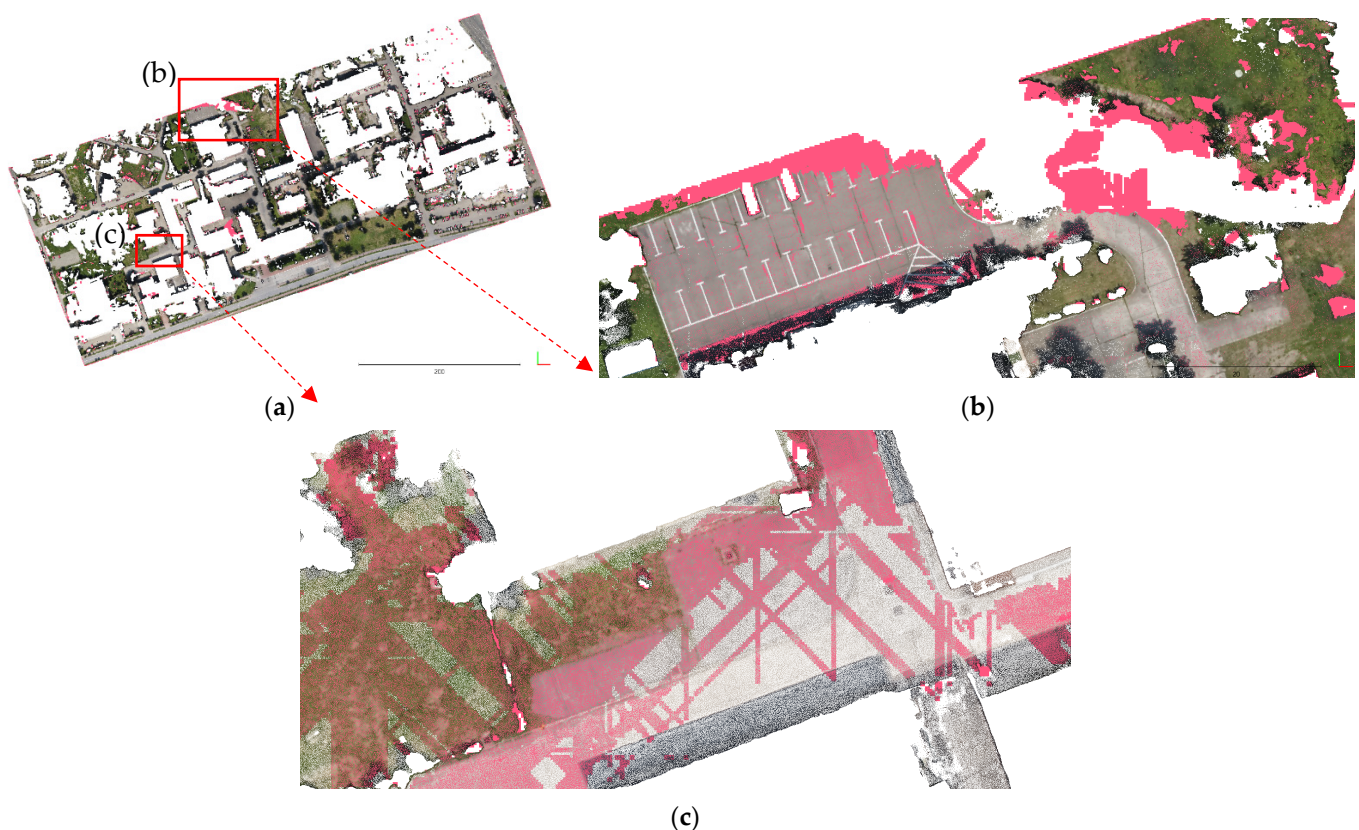
The final DSM can be seen in Figure 8b.

For this step, the "FillGaps" module can also be used, which detects gaps in raster models, and with the use of several interpolation methods the values for the void pixels are calculated.

The "TerrainMask" module is used next to identify the open terrain parts of the final DSM raster, the result being a mask in binary format where the 0 pixel values represent open terrain areas and 1 off-terrain elements. The minimum height parameter was set to 0.5 m to classify the elements above the bare earth surface and the maximum width parameter was set to 120 m, representing the maximum building length/width in the study area, after testing different values for these parameters.

The CSF algorithm is very easy to apply requiring three parameters, i.e., the type of terrain surface (steep slope, relief or flat), the cloth resolution (grid size) and the threshold for the off-ground points' classification. Nevertheless, for obtaining optimal results an important aspect is to carefully choose the values for the parameters as also reported in [35].

For the present study, the "flat" option was chosen, since the terrain level difference is about 6 m, the cloth resolution was set to 1 m and the threshold to 0.3 m. Visually analyzing the result, it can be seen that same points belonging to the bare earth surface (for example, a part of the parking lot and of the green area from the north side of the study area), have been falsely classified as off-ground points. If the cloth resolution is set to 0.2 m, the points belonging to the building roofs are classified as ground points. By overlying the terrain mask obtained in Opals over the ground points classified with the CSF algorithm, the differences are shown with magenta (Figure 9). Visually, it can be seen that, using the terrain mask algorithm, the open terrain parts have been correctly identified in comparison with the CSF algorithm in the north side of the study area, but there are terrain parts that were incorrectly not identified as open terrain, as can be seen in Figure 9c.

Figure 9. (**a**) Terrain mask overlaid over ground points classified with the CSF algorithm, (**b**) detail with open terrain parts correctly identified by terrain mask algorithm and unidentified by the CSF algorithm and (**c**) detail with open terrain parts correctly identified by the CSF algorithm and unidentified by terrain mask algorithm.

To obtain the final file that contains the ground points, the points correctly identified as ground using terrain mask and misclassified when using the CSF algorithm have been added to the ground points obtained through CSF. For the areas highlighted with magenta, close polylines have been manually created that were overlaid on the unclassified UAS point cloud. The corresponding points have been extracted automatically using the CloudCompare software. As such, a total number of 154,370 points have been added to the 28,114,937 points classified as ground points using the CSF algorithm, representing only 0.5%. For this particular case study, the difference is not notable, but for other areas the improvement could be significant.

4.3.2. DTM Creation

In the process of the 3D modeling of an urban area, the Digital Surface Model (DSM) and the Digital Terrain Model (DTM) are high importance products, and therefore must be generated with high accuracy.
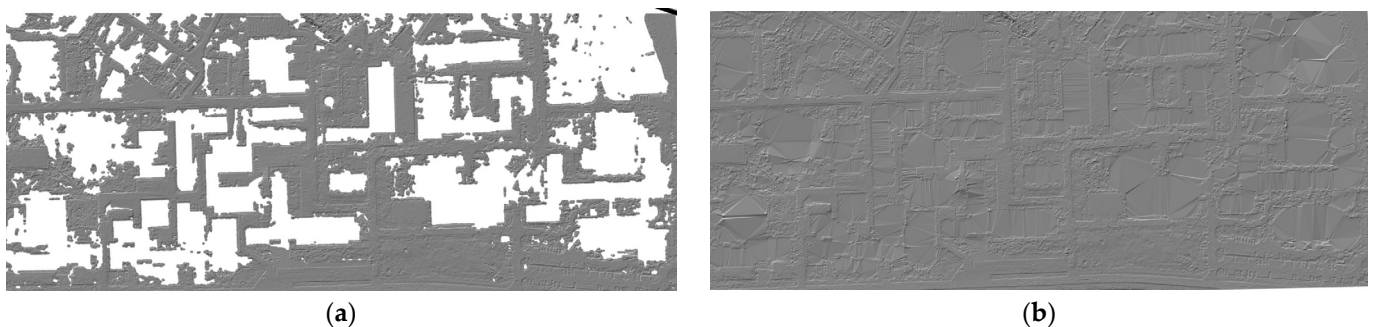
For the Digital Terrain Model derivation, all points classified as terrain points have been interpolated using the interpolation methods implemented in Opals in three different scenarios, changing the search radius and the number of neighboring points. The standard deviation of each resulted DTM has been calculated based on 59 points placed at ground level measured by GNSS technology for UAV images orientation, the values being centralized in Table 2.

Table 2. Standard deviation for different interpolation methods in three scenarios.

Interpolation Method	Standard Deviation [m]		
	0.6 m Search Radius, 20 Neighbours	1 m Search Radius, 20 Neighbours	1 m Search Radius, 50 Neighbours
Snap grid	0.0473	0.0473	0.0473
Nearest neighbour	0.0429	0.0429	0.0429
Delaunay triangulation	0.0402	0.0402	0.0402
Moving average	0.0338	0.0333	0.0306
Moving planes	0.0358 *	0.0355 *	0.0314 *
Robust moving planes	0.0356	-	-
Moving paraboloid	0.0419 *	0.0419 *	0.0377
Kriging	0.0339	0.0335	0.0311

* NoData values included—rasters extent does not cover the GCP location.

Analyzing the results reported in Table 2, we can see that the differences are in the range of millimeters, and some GCPs' locations have not been covered by pixels. As such, the Kriging interpolation method was chosen with the following parameters: 1 m search radius and 20 neighbours. As expected, information gaps remained (void pixels) (Figure 10a) and, in order to fill them, another interpolation method was applied, namely Delaunay triangulation (OPALS- "delaunayTriangulation"), with the following parameters: search radius 2 m and 20 neighboring points. To generate the final DTM without information gaps, the two rasters have been combined (Figure 10b).

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 10. (**a**) DTM in raster format with void pixels, (**b**) final DTM obtained with two interpolation methods: Kriging and Delaunay triangulation.
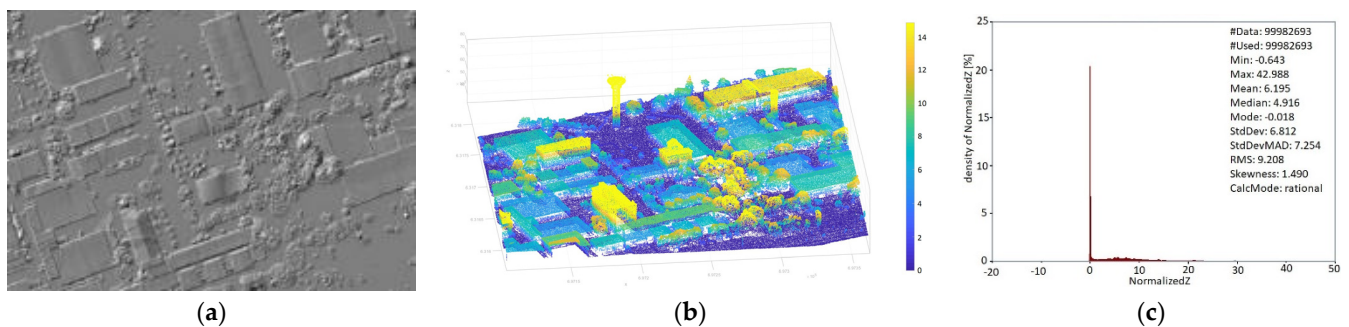
4.3.3. Normalized Digital Surface Model (nDSM) Creation

The Normalized Digital Surface Model (nDSM) is a digital model that comprises the heights of objects (e.g., buildings or trees) from the bare earth surface, obtained by subtracting the DTM from the DSM. The role of the nDSM is to remove all points below a certain height [36].

The nDSM for the study area have been obtained by interpolating all off-ground points using the "moving average" method (horizontal plane) and the "NormalizedZ"

attribute, calculated as a difference between the altitude of the UAS point and the cell altitude corresponding to the DTM (Figure 11).

Figure 11. (**a**) Shaded nDSM in raster format-a detail, (**b**) the point cloud containing the heights above the earth surface created based on nDSM raster, (**c**) the histogram of the off-ground points heights.

4.3.4. Attributes Computation (Feature Extraction) for Each UAS Point

To describe the local 3D structure at a certain 3D point, the distribution of 3D points in Euclidian 3D space within a local neighborhood is usually analyzed. In general, different strategies can be applied to recover local neighborhoods for the 3D points. These types of neighborhoods can be defined as a spherical neighborhood parameterized by a radius, a cylindrical neighborhood parameterized by a radius, a spherical neighborhood parameterized by the number of $k \in N$ nearest neighbors that respects the Euclidean distance in 3D space or a cylindrical neighborhood parameterized by the number of nearest neighbors in relation to the Euclidean distance in 2D space [37]. These neighborhood types are parameterized with a single-scale parameter, which is represented by either a radius or the number of nearest neighbors and allow the description of the local 3D structure at a specific scale, or with a multi-scale parameter [38,39] which implies building a classifier to establish the best combination of scales leading to the best separation of classes [38]. When a single scale parameter is used, in order to select an appropriate value, prior knowledge of the data is usually involved. Furthermore, identical values for the scale parameter are usually selected for all points in the 3D point cloud.

In order to reduce the extra computational burden in terms of both time and memory consumption, the point cloud has been downsampled with 0.1 m minimum distance between points using the Cloud Compare software, resulting a point density of 120 points/m². The test done by [39] on different data sets to study the effect of the downsampling process on the classification results using the Random Forest algorithm showed that a resolution of 0.7–0.8 m is suitable for all data sets. Nevertheless, for our data set the 0.2 m resolution led to a point density of 34 points/m², which does not assure a qualitative result for 3D building modeling. The neighborhood selection it is difficult and challenging [39] and different approaches are used to calculate the optimal neighborhood [40,41]. After various tests, a cylindrical neighborhood parameterized by a radius of 1 m was chosen for attributes calculation and a script was written using the Opals software. The attributes are grouped in four main categories: height-based attributes, local-plane based attributes, echo-based attributes and eigenvalues-based attributes (Table 3).

The height-based attributes are: NormalizedZ, calculated as a difference between the UAS point altitude and the corresponding DTM cell altitude, Zmin, the lowest point in the neighbourhood, DeltaZ, the difference between the point altitude and the lowest point altitude, Range, the difference between the highest and the lowest altitude in the neighbourhood, Rank, the relative height in percent in the neighbourhood and height variance (dZVariance), a statistical feature representing the variance of all Z values within the neighbourhood, being also a local roughness measure.

Table 3. Attributes computed for each UAS off-ground point in a cylindrical neighborhood.

Height-Based Attributes	Formula	Echo-Based Attributes	Formula
NormalizedZ Zmin	$Z_P - Z_{DTM}$ Z_{min}	EchoRatio	$n_{3D} / n_{2D} * 100$
Delta Z	$Z_P - Z_{min}$	**Eigenvalues-based attributes**	
Range	$Z_{max} - Z_{min}$	Linearity	$(\lambda_1 - \lambda_2)/\lambda_3$
Rank	relative height in the neighborhood	Planarity	$(\lambda_2 - \lambda_3)/\lambda_1$
Height Variance	variance of all Z values in the neighborhood	Anisotropy	$(\lambda_1 - \lambda_3)/\lambda_1$
Local plane-based attributes		Omnivariance	$\sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$
Normal vector	NormalX, NormalY, NormalZ	Eigenentropy	$-\sum\limits_{i=1}^{3} \lambda_i \ln \lambda_i$
Standard dev. of normal vector estimation	NormalSigma0	Scatter	λ_3/λ_1
Variance of normal vector components	Variance of NormalX, NormalY, NormalZ	Sum of eigenvalues	$\lambda_1 + \lambda_2 + \lambda_3$
Offset of normal plane	$Dist_{P-plane}$		
Sigma X	σ_X	Change of Curvature	$\lambda_3/(\lambda_1 + \lambda_2 + \lambda_3)$
Sigma Y	σ_Y	EV2Dratio	λ_2/λ_1
Sigma Z	σ_Z	Sum of eigenvalues in 2D space	$\lambda_1 + \lambda_2$
Verticality	1-normalZ		
Points density	$Npoints/Volume$		

The local plane-based attributes are: the normal vector at a computed local surface (for this case robust plane fit which detects blunders) with its components along the X, Y and Z axes (normal, normalY and normalZ), the standard deviation of normal vector estimation (NormalSigma0), which is high in rugged areas and low in smooth areas, the standard deviation of the normal vector components (dNormalXVariance, dNormalYVariance and dNormalZVariance), the offset from the current point to the estimated local plane (NormalPlaneOffset), the standard deviation of local plane offset estimation (NormalPlaneOffsetVariance), and the verticality, derived from the vertical component of the normal vector and the point density, which is calculated as the ratio between the total number of points in a neighborhood and the volume of the neighborhood.

The echo-based attributes category contains only one attribute, i.e., the EchoRatio attribute, which is an indicator for the degree of penetrability and roughness of a surface [42].

The eigenvalues-based attributes are computed based on the eigenvalues of the normal vector estimation covariance matrix ($\lambda1,i$, $\lambda2,i$, $\lambda3,i \in R$, where $\lambda1,i \geq \lambda2,i \geq \lambda3,i \geq 0$). The eigenvalues also provide additional features and help estimate plans, edges, corners, lines and volumes and describe the local spatial distribution of 3D points and can be exploited to characterize the local 3D shape: linearity, planarity, anisotropy, omnivariance, eigenentropy, scatter, sum of eigen values, change of curvature, the ratio between the second and the first eigenvalues and the sum of eigenvalues in 2D space.
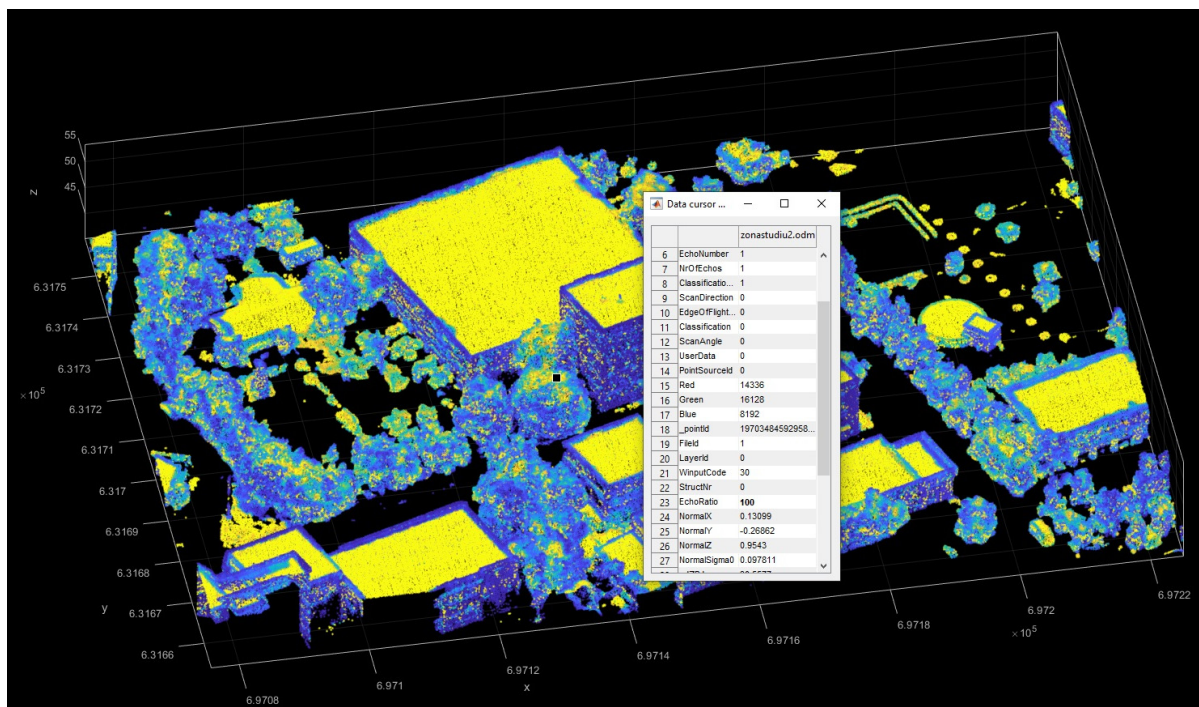
4.3.5. UAS Point Cloud Filtering Using Point Attributes

Many studies concentrate on ALS point cloud filtering and classification, taking advantage of characteristics such as "echowidth" for "ground" and "off-ground" points' separation, for example. The UAS point cloud is very different from an ALS point cloud, making the classification process more challenging.

Vegetation is a very important layer for a digital 3D city model and its monitoring process plays an important role for urban spatial data management [43]. A lot of studies

concentrate on vegetation detection, especially for forest areas, using LiDAR data [44]. Therefore, for this case study, a workflow for UAS vegetation points filtering by evaluating the existing attributes is presented and tested.

The usefulness of the "EchoRatio" attribute was demonstrated in the creation of a vegetation mask in an urban area [43], or for detection of buildings' regions [42] using high density full-waveform LiDAR data. However, in the case of UAS point clouds, the "EchoRatio" attribute cannot be simply applied for automatically detecting the vegetation, as most of the points situated on roofs' edges and on buildings' facades have values in the same range as vegetation points. Moreover, the points situated on top of the trees have values of 100% for ER as the roof points (Figure 12).

Figure 12. UAV point cloud colored based on EchoRatio attribute, yellow color indicates values of 100% for the ER attribute.
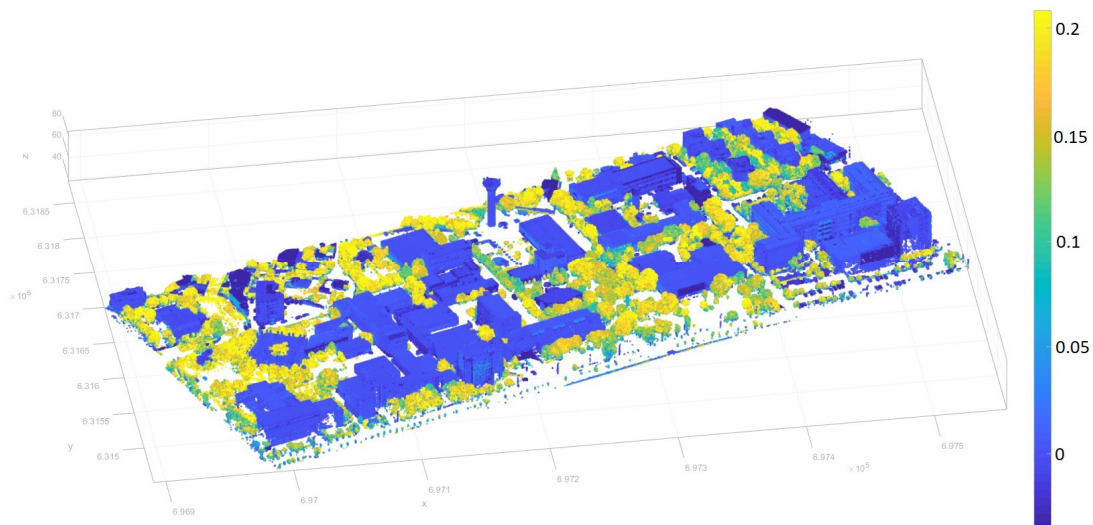
Thus, along with the attributes calculated above, a new one is introduced, making use of the RGB code of each UAS point, i.e., the visible-band difference vegetation index (VDVI), which is designed to extract green vegetation [45]. The VDVI was calculated for the off-ground points with the following formula:

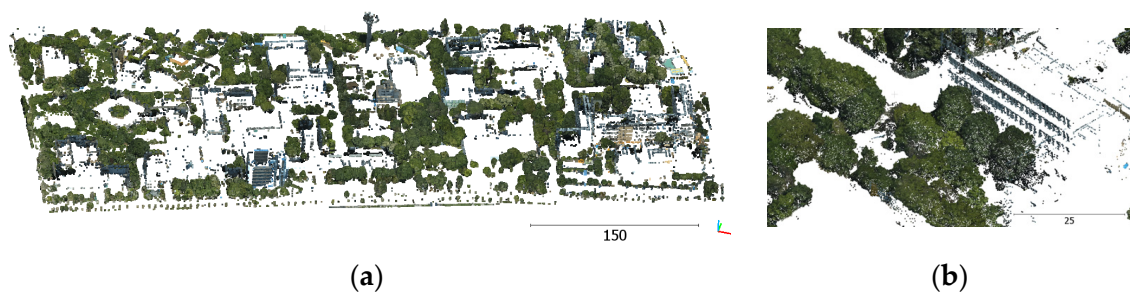$$VDVI = (2 \times G - R - B)/(2 \times G + R + B), \tag{1}$$

and the UAV point cloud colored by VDVI can be seen in Figure 13.

Analyzing the values of VDVI, it can be seen that most of the buildings' points have values of 0 or close to 0, but some points representing windows or elements with yellow color have the VDVI in the same range as vegetation points. A threshold of 0.02 for the VDVI was set and a point cloud representing vegetation was exported (Figure 14).

Figure 13. UAV point cloud colored based on the visible-band difference vegetation index (VDVI), blue color indicates values of 0 for the VDVI attribute.

(**a**)

(**b**)

Figure 14. (**a**) Vegetation point cloud extracted using the VDVI attribute, (**b**) detail.
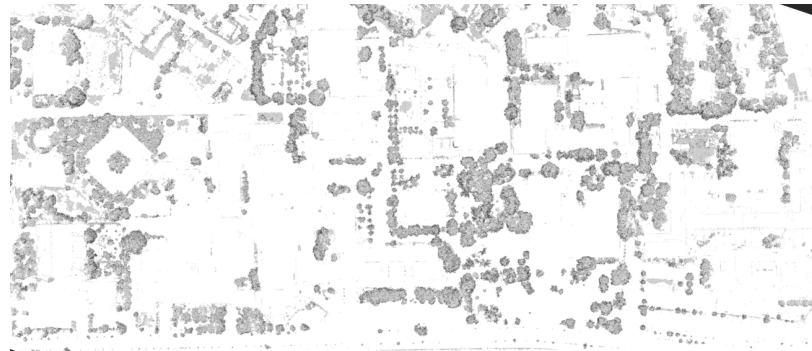
The result is not ideal, since there are points belonging to buildings that passed the threshold, although the vegetation points have been 100% extracted. A lower value for VDVI will lead to more building points, while a higher value will reduce the number of vegetation points. Further analyses are needed to automatically filter the vegetation points, so for each point the local surface normal vector was estimated, using a simple plane fit for 12 neighbors in a search radius of 2 m. The point cloud colored by the normal vector component on Z axis (NormalZ), can be seen in Figure 15.

Figure 15. The filtered vegetation points colored by the normal vector component on Z axis.
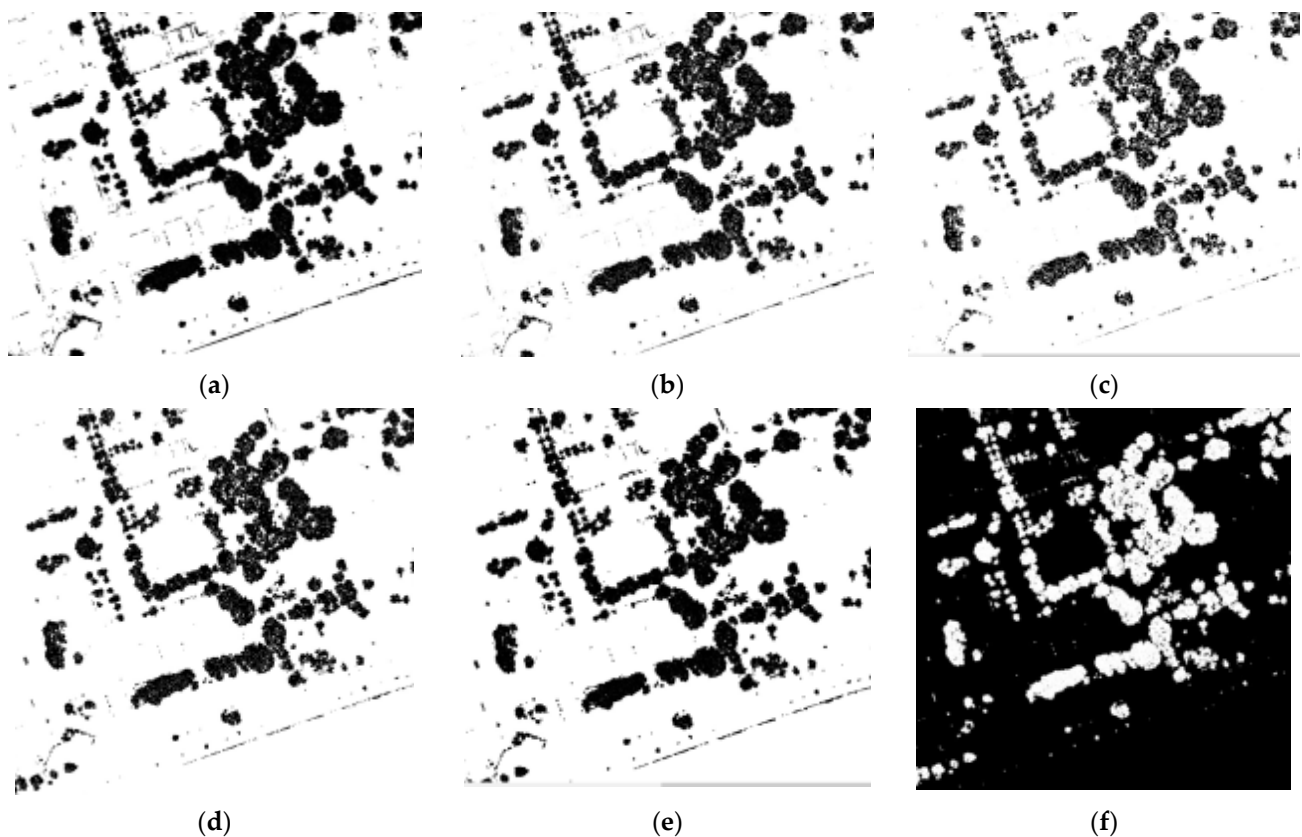
Using the minimum value of NormalZ attribute per 0.1 m cell, a raster is derived, which is further transformed in a binary raster imposing the condition that cells with values greater than 0.2 for NormalZ are set to 0, so as to eliminate some pixels representing buildings' facades (Figure 16).
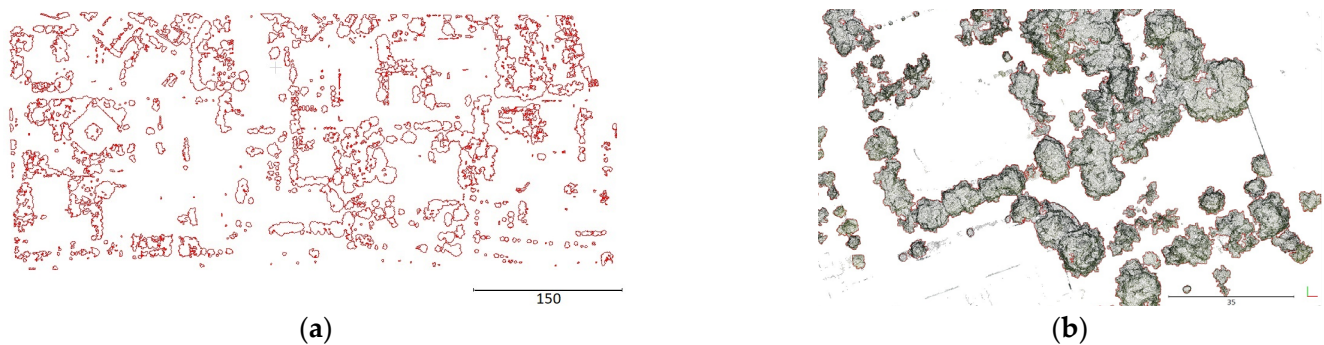

Figure 16. The binary raster for NormalZ attribute values greater than 0.2.

Subsequent steps include a series of morphological operations of erosion and dilation, applied in this order: erosion, dilation, dilation, erosion and erosion, using the "opalsmorph" module. In order to exclude the low vegetation points, falsely classified as off-ground points when using CSF algorithm, a constraint on normalized height (nDSM > 1.0 m) was used, the vegetation mask being obtained. The results are presented in Figure 17.

Figure 17. The binary raster obtained by applying the morphological operations erosion and dilation: (**a**) erosion, (**b**) dilation, (**c**) dilation, (**d**) erosion, (**e**) erosion, (**f**) vegetation mask.

Additionally, in order to exclude the remaining pixels on the building facades, a criterion for the surface area of 1 m^2 is introduced using "opalsalgebra" module, which removes the isolated polygons and closes the small holes. The resulting 587 polygons with a total area of 31,218 m^2 can be seen in Figure 18, along with a detail of polygons overlaid on the vegetation points firstly extracted with the VDVI attribute.

(**a**)   (**b**)

Figure 18. (**a**) The polygons representing vegetation areas with the high greater than 1 m and area greater than 1 m^2, (**b**) detail with polygons overlaid on the vegetation points.

To obtain the vegetation points, the polygons have been superimposed on the unsampled off-ground points, the points lying inside the polygons being automatically extracted using the LAStools software's Lasclip function. Taking into account the fact that there are still a few points on the facades and on the roof of a yellow terrace, which can be easily removed manually, the accuracy of the proposed method is about 99%.

4.3.6. UAS Point Cloud Classification Using Supervised Machine Learning Algorithms

Currently, in the online environment, both point clouds as well as 3D models which are not interpreted are available to users. 3D point clouds are the simplest, but at the same time the most important, collection of geometric elementary primitives, able to represent the shape, size, position and orientation of objects in space. A single point from the point cloud does not contain information about the connection to other points or the structure of the objects, so in order to benefit from the informative value of point clouds and for a better understanding of them, it is necessary to perform the segmentation and the classification processes. Segmentation refers to the division of a group of points into subgroups (called a segment) characterized by having one or more features in common (geometric, radiometric), while classification means defining and dividing points into specific classes, respecting several criteria.

There are three main approaches for point clouds classification [46] from the methodological perspective:

- supervised: based on a training data set, a trained model is obtained, which is applied on the entire data set to label each class;
- unsupervised: the data set is classified without user interaction;
- interactive: the user is actively involved in the classification process through feedback signals, which can improve the results of the classification.

Based on the data type used for classification, we can distinguish: point-based, segment-based or voxel-based classification methods. Supervised classification methods can be differentiated into two main classes: standard classifiers, which are based only on extracted individual features, and contextual classifiers, which exploit both the extracted individual features and the relationships between 3D points within the neighborhood. The most commonly used supervised classifiers are Nearest Neighbor, Decision Tree, Random Forest, Support Vector Machines, Maximum Likelihood and AdaBoost.
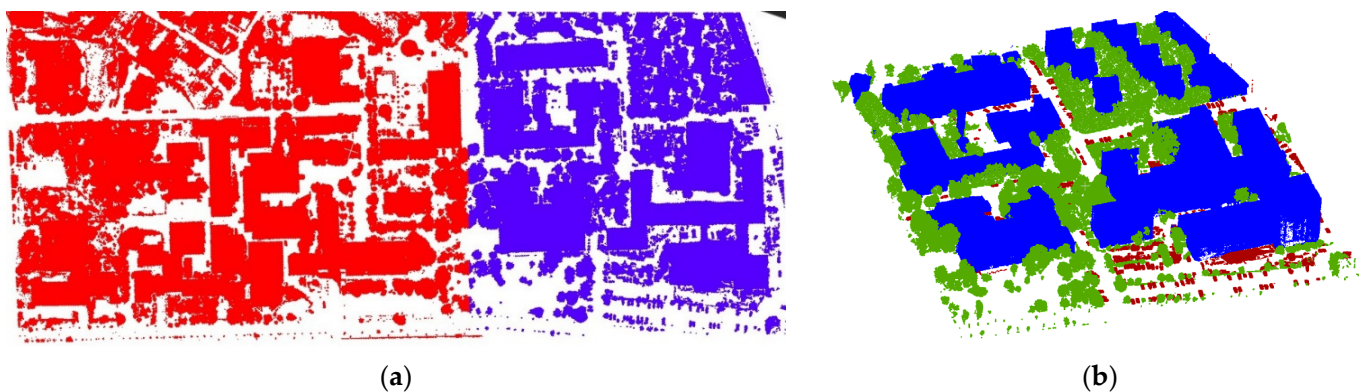
For the UAS off-ground point cloud, the supervised classification method was applied, which is based on Decision Trees and Random Forests algorithms. The Random Forests

algorithm is known as a set of techniques which contains a group of high-performance classifiers, based on a multitude of decision trees and has multiple applications in fields such as medicine, computer vision, the marketing industry and image classification. The reason for choosing these algorithms for the present case study is that they perform efficiently on large data sets, are able to process thousands of variables and the parameters such as the number of prediction variables chosen at the node level, the decision trees number and the decision trees dimension, do not significantly influence the result of the classification. By comparing the SVM, Random Forest and Logistic models, [47] found that the Random Forest-based model has the best prediction ability, which is better than the other two models [48].

The first step in the classification process is the manual selection and classification of training data sets, which will be used to train and learn the algorithm to recognize different types of classes. From the original data set, approximately 39% of the study area has been selected and manually classified using CloudCompare software in three classes: buildings, vegetation and others (cars, street furniture, etc.). The summary of the training data set classes can be found in Table 4 and the graphical representations in Figure 19.

Table 4. Summary of the training data set classes.

Training Data Set	Number of Points	Percent
Building	835,631	42.8
Vegetation	675,687	52.9
Others	68,261	4.3
Total	1,579,579	100

(**a**)   (**b**)

Figure 19. (**a**) The data set separated into training (~39%) (blue) and test (~61%) (red), (**b**) the training data set and its three classes (blue-buildings, green-vegetation, red-others).

The second step in the classification process is the attributes computation. A number of 30 attributes described above have been extracted for the training data set and for the test data set, respectively. Based on these attributes, the mathematical classification model is created with the help of machine learning algorithms, using as input data the 95% of the training data set and the Python script implemented into Opals (Figure 20).

The overall accuracy of the classifier is 92%, with an overall miss-classification rate of 8%. The confusion matrix is presented in Table 5.
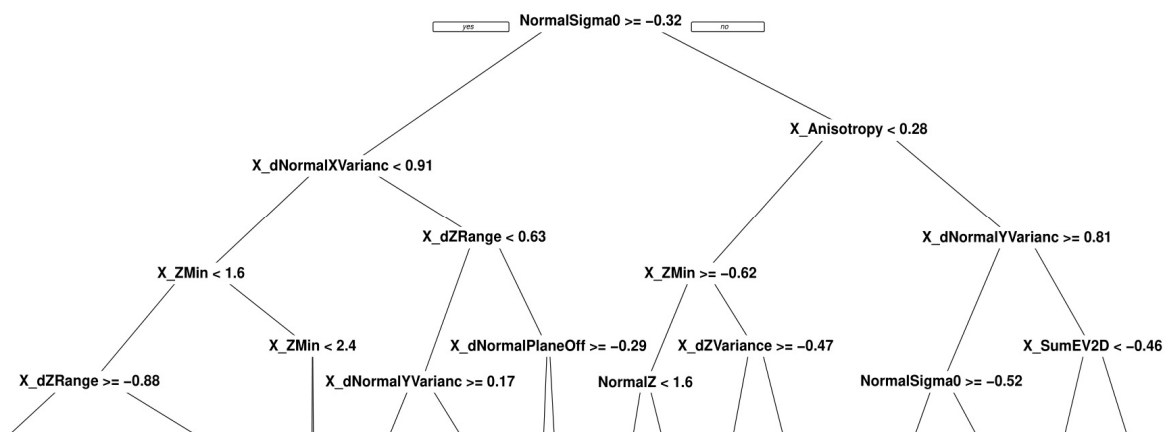
Figure 20. Classification model-detail of decision trees.

Table 5. Confusion matrix for the training data set after applying the 30 attributes classification model.

	Vegetation	Building	Others	Sum_Ref.	EoC	Completeness
Vegetation	39.7	2.5	0.5	42.8	7.1	92.9
Building	3.7	49.0	0.2	52.9	7.4	92.6
Others	0.8	0.3	3.3	4.3	24.7	75.3
Sum_estim.	44.2	51.8	4.0	100		
EoC	10.2	5.4	18.9			
Correctness	89.8	94.6	81.1			

The values are expressed in percentages.

The last step in the classification process is to apply the classification mathematical model on the entire data set. The result is presented in Figure 21a,c. Analyzing the results, it can be seen that points on the facades are classified as vegetation points. In order to overcome this aspect, the VDVI attribute was added to both data sets (training and test) and an overall accuracy of the classifier was increased, reaching 97.1%. The result presented in Figure 21b,d clearly shows improvements regarding the previously wrong classified façade points that were correctly classified by the model based on 31 attributes.
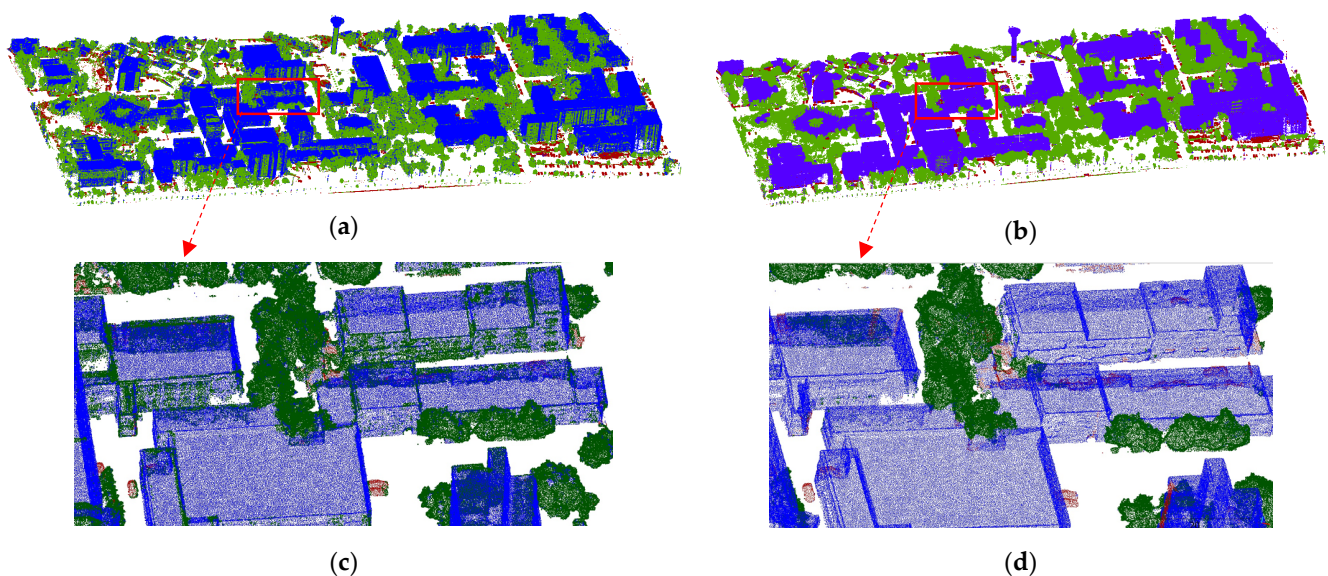

Figure 21. The classification result (**a**) with 30 attributes and (**b**) with 30 attributes and VDVI, (**c**,**d**) details.

A detail with the mathematical classification model obtained after including the VDVI attribute is presented in Figure 22.
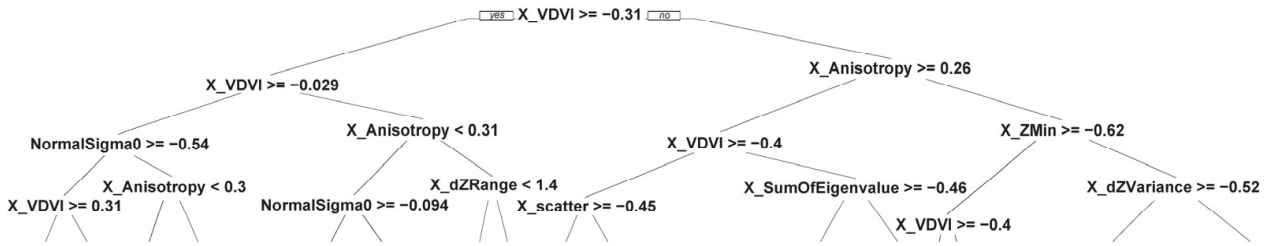

Figure 22. The mathematical classification model with 30 attributes and VDVI.

The confusion matrix for the training data set with the VDVI attribute is presented in Table 6.

Table 6. Confusion matrix for the training data set after applying the 31 attributes with the VDVI classification model.

	Vegetation	Building	Others	Sum_Ref	EoC	Completeness
Vegetation	42.1	0.6	0.1	42.8	1.5	98.5
Building	1.0	51.3	0.6	52.9	3.1	96.9
Others	0.3	0.3	3.7	4.3	14.6	85.4
Sum_estim	43.4	52.2	4.4	100		
EoC	3.0	1.8	15.6			
Correctness	97.0	98.2	84.4			

The values are expressed in percentages.

The importance of the RF model variables for both tested scenarios (with the VDVI and without the VDVI attribute) is presented in Figure 23. The one based on VDVI potentiated the other previously identified attributes (the model without the VDVI) and the aspect that can be seen is that a percentage of 77% is covered by the five most relevant attributes, compared to the other model without VDVI that identifies eight notable attributes but which answers only for 69% of the model.
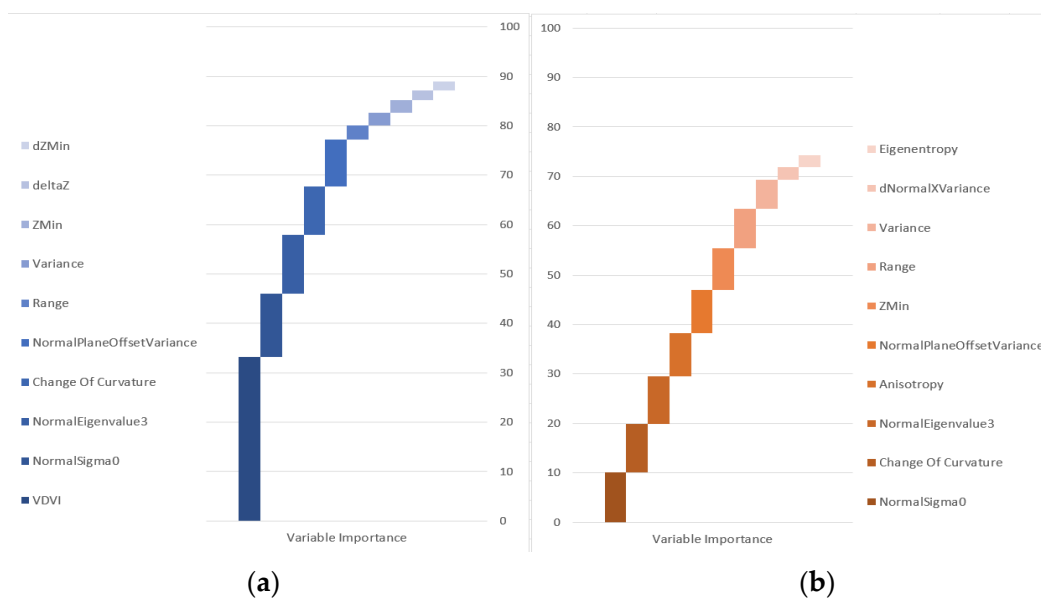

Figure 23. The variable importance for the RF models (**a**) with the VDVI attribute and (**b**) without the VDVI attribute.

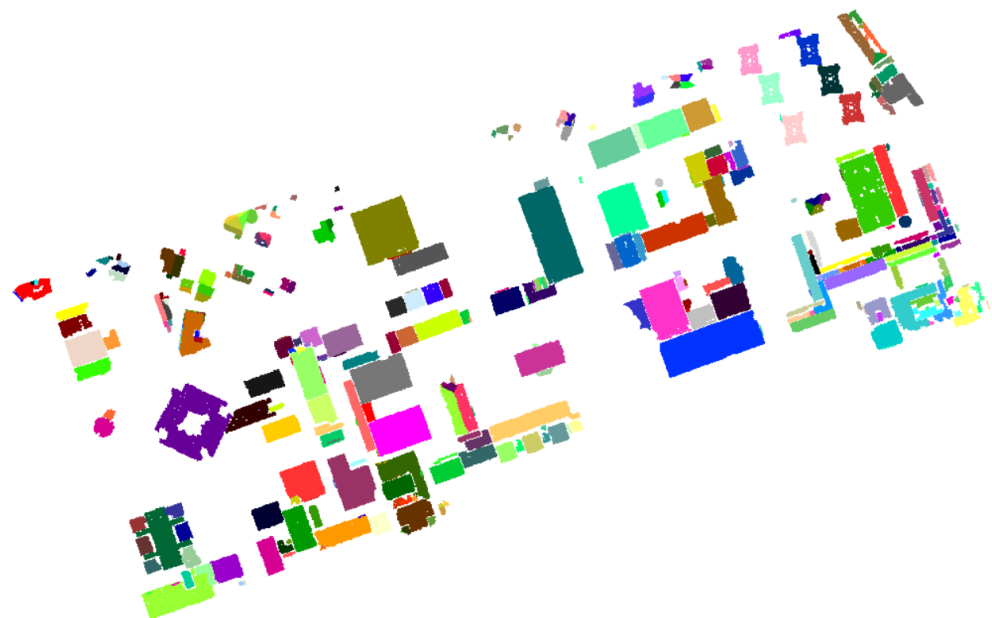### 4.3.7. UAS Buildings Point Cloud Segmentation

Segmentation is the process of grouping points with similar properties into multiple homogeneous regions that are spatially connected [49]. Ideally, each surface should generate a segment. Among the errors of the segmentation process are over-segmentation and sub-segmentation. Over-segmentation occurs when a single surface is erroneously divided into more segments than necessary, and sub-segmentation occurs when several different surfaces are grouped and labeled as belonging to a single surface.

The methods for point clouds segmentation can be classified as follows [50]: methods based on line segments, on attributes, on defining a surface (bottom-up or top-down approach), on finding the best fitting geometric shape, hybrid methods and machine-learning approaches (k-means algorithm and hierarchical grouping). The most frequently used algorithm for plane detection are random sample consensus (RANSAC) [51,52], the 3D Hough transform [53] and the region growing algorithm [54].

The process of point cloud segmentation can have as input either a point cloud or a mesh that can be used to build a graph and cluster the graph to produce segmentation by using information such as normal direction, smoothness or concavity along boundaries [49].
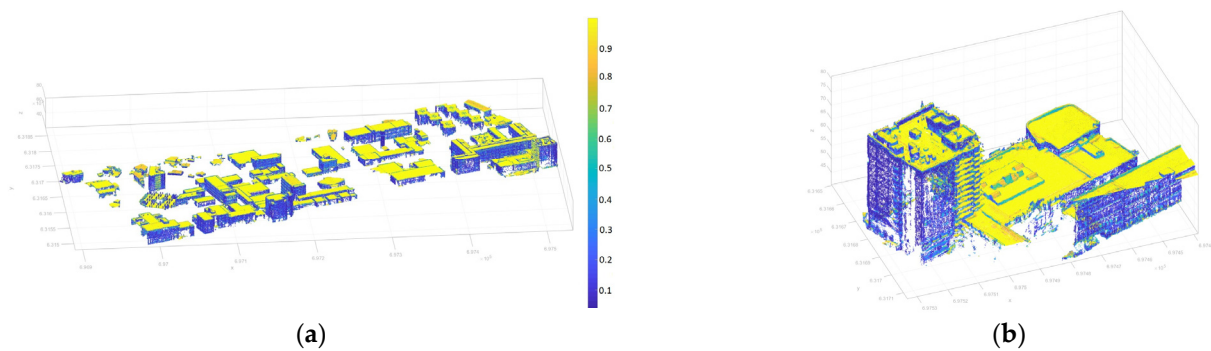
For this step, the UAS point cloud classified as "building" and the "segmentation" module provided by Opals have been used, which provides methods for point cloud segmentation based on the seeded region growing approach and a local homogeneity criterion [55].

For the roofs' segmentation, the homogeneity criterion was defined as a sum of two logical conditions that the neighboring points must meet: the angle between the normal vectors calculated in the seed point, respectively in the neighboring point to be less than 5° and the height of the point above ground (normalizedZ attribute) to be greater than 2 m. It resulted in a number of 407 segments that can be seen differently colored in Figure 24.
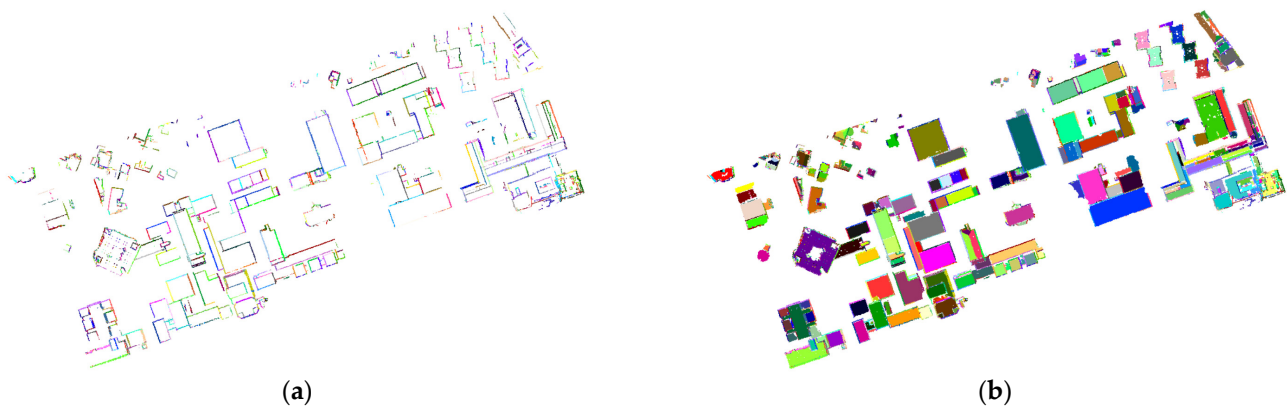

Figure 24. The segments representing the roofs planes with the high greater than 2 m.

Next, points belonging to buildings' facades have been automatically extracted by thresholding the "normalZ" attribute. The points of "building" class are colored based on "normalZ" attribute values (Figure 25), the yellow color showing values in range of 0.9 to 1, the blue color values close to 0 and orange color values in range of 0.8 to 0.9. All points having a value of "normalZ" attribute less than 0.9 have been extracted and are the subject of segmentation.

Figure 25. (**a**) The buildings' point cloud colored based on normal vector component along the Z axis (normalZ attribute), (**b**) detail.

Further, the second method implemented into the "segmentation" module was used, i.e., "planeExtraction". In this segmentation mode, the planar surfaces are identified in the point cloud, based on the normal vectors calculated at each point. A neighboring point will be added to the segment if it belongs to the best fitting plane, namely, the threshold value is not exceeded. Two types of thresholds can be defined: the point to the plane distance and the angle between the normal vector calculated at that point and the normal to the plane. Considering the fact that attached to the facades usually there are architectural elements, the standard deviation when fitting a plane to this point cloud, has a high value, greater than 0.3 m. As also mentioned in [5], the threshold should be chosen depending on the point cloud thickness, which is given by the point cloud noise and, as for this case, should be considered also the architectural elements of the facades, or the fact that the windows and the beams are placed in different planes. As such, the value for the maximum distance between the point and the plane and for the standard deviation was set to 0.5 and for the "seedCalculator" parameter, which is used for the calculation of the seed point order in the segmentation process, the value of 0.2 was used. The search radius was also increased to 4 m and a total number of 2044 segments were generated (Figure 26).

Figure 26. (**a**) The segments representing the façades planes, (**b**) all segments corresponding to the buildings from the study area.
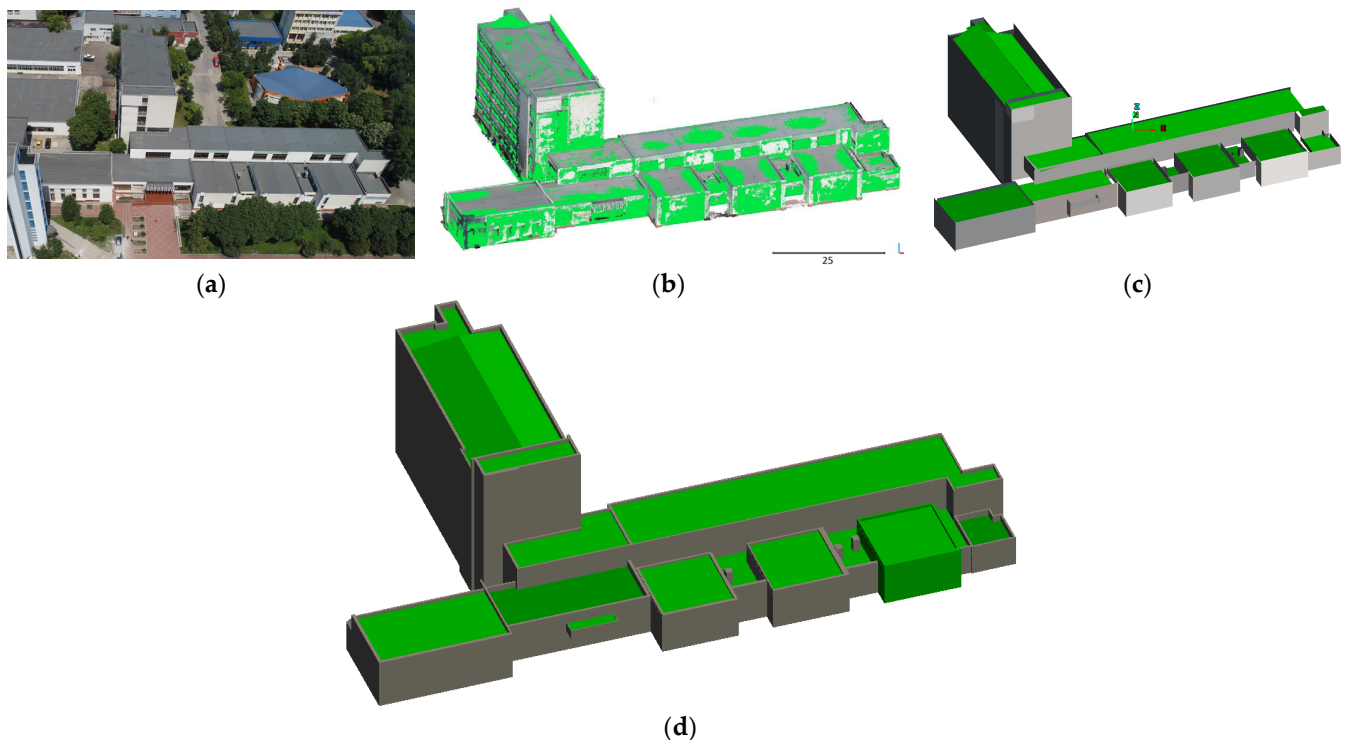
4.3.8. 3D Buildings Model Creation

The process of buildings extraction consists of three main steps: DTM creation, building detection and building reconstruction [55,56]. After obtaining the segments representing the buildings roofs and facades, information defined as attribute "SegmentID", an individual point cloud was exported for each corresponding segment using the "LAStools" software. Then, with the "region growing" algorithm, the planes have been generated for further processing. For planar approximation, the "fit plane" function can also be used, which can be found in Cloud Compare software.
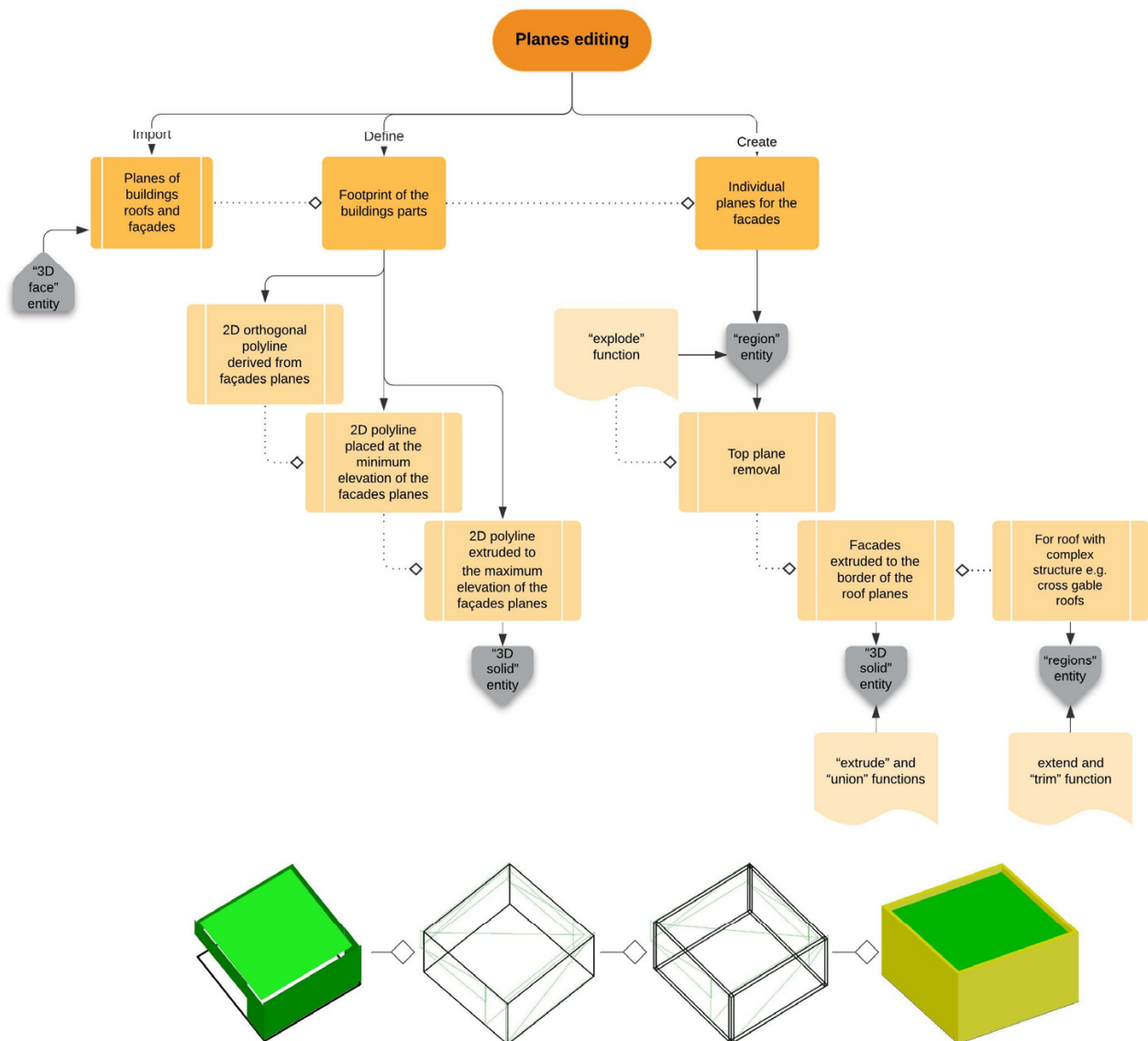
To exemplify the process of editing the roofs and facades extracted planes, four buildings have been chosen: our faculty building, a building with laboratories, the rectory building and a hotel. The first two buildings have terraced roofs, the third one has a combination of terraced roofs and cross gable and the last one has a flat roof with different structures, such as staircases. The workflow for manually editing the planes was performed in AutoCAD platform, namely AutoCAD Map 3D, but there are a lot of dedicated software which can be used for this step, such as dedicated software for point cloud modelling, including: Leica Cyclone 3DR, Terrasolid, Faro Scene, or architecture dedicated software: AutoCAD Civil 3D, Autodesk 3ds Max, Revit, etc.

The planes created for buildings' roofs and façades were imported in AutoCAD Map 3D as a "3D face" entity. The first step in the planes editing process was to define a 2D orthogonal polyline for the façades planes from a top-view point obtaining the foot print of the buildings parts. The polyline was placed at the minimum Z altitude of the facades planes then was extruded to the maximum Z altitude of the façades' planes obtaining a "3D solid" entity. The "explode" function was used to create separate planes for the facades as "region" entities and the top plane was removed. The regions corresponding to the facades were extruded to the border of the roof planes which have been converted to "regions" entities using the "region" and "union" functions. The same operation was performed also for the cross-gable roofs. In this case, an extend operation and "trim"function were used to refine the shape of the roof. In Figure 27 the final parametrized 3D model of the rectory building is presented and in Appendix B, Figure A1 the 3D models for the others three models.

(**a**)

(**b**)

(**c**)

(**d**)

Figure 27. (**a**) UAS image of the rectory building, (**b**) the corresponding planes of the building integrated into the UAS point cloud, (**c**) the building planes for further editing, (**d**) detailed model of a building determined by piece-wise intersection of planar faces.
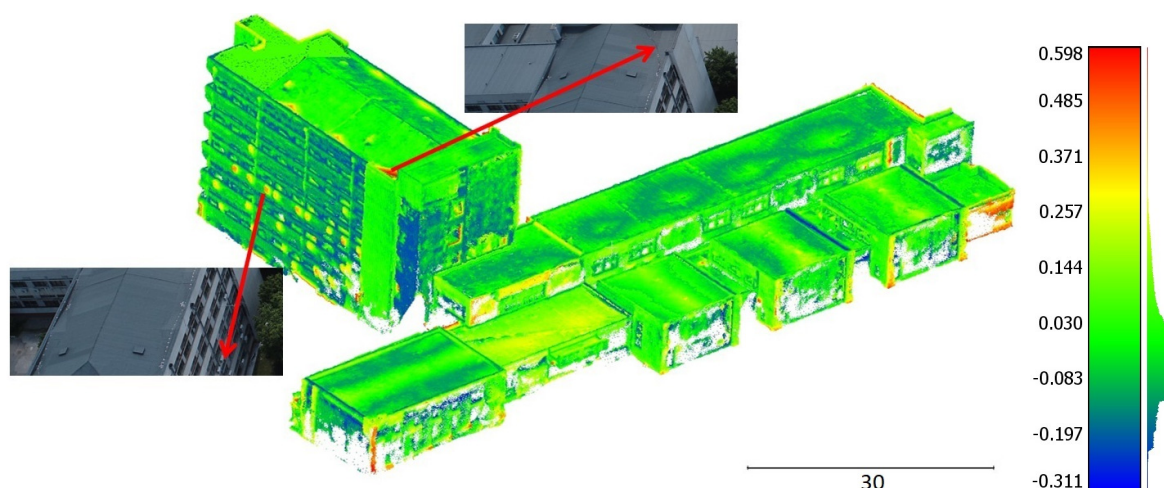
The workflow can be summarized as follows (Figure 28):

Figure 28. The workflow for 3D model reconstruction as a parametrized model in AutoCAD Map 3D.

4.4. Accuracy Assessment of 3D Buildings Models

The accuracy assessment was performed by comparing the original building corresponding point cloud with the final parametrized 3D model and by comparing the roofs corners 3D coordinates measured using GNSS-RTK technology with the coordinates extracted from the final 3D model, obtaining the Euclidean distances between the two positions of the roof corner

The original point cloud for the rectory building, i.e., the point cloud before downsampling, was compared with the 3D parametrized model using the "cloud/mesh" distance from CloudCompare software, resulting the Hausdorff distances between each UAS point and the corresponding triangle surface of the 3D model, that was previously exported in OBJ format (Figure 29).

Figure 29. The distances in meters between the UAS point cloud and the rectory building parametrized 3D model.

Analyzing the color scale we can easily obtain the value for the computed distances, for example, some points on the roof are colored in red tones, corresponding to a value of around 60 cm on the color scale, because of the irregular shape of the roof. On the other hand, some points on the facades are colored in blue, located in the windows plane, the distances being in the range of -31.1 cm $\div$ -10 cm, and the rest of the points are colored in green, the distances being in the range of -10 cm $\div$ $+22$ cm.

The comparison for the rest of the buildings from the study area can be found in Appendix B, Figure A2 and in Table 7 the values of the standard deviation obtained for the buildings' UAS point clouds and for the roofs' corners measured by GNSS technology.

Table 7. Summary of the standard deviation obtained for the buildings UAS point cloud after computing the distances between each point and the final building 3D model and the standard deviation computed based on GNSS points coordinates.

Building	Standard Deviation Computed Based on Original Point Clouds σ [cm]	Standard Deviation Computed Based on GNSS Points σ [cm]	No. of GNSS Points
Our faculty building	13.3	25.6	5
Laboratories	10.7	29.2	6
Rectory	11.9	24.4	10
Hotel	24.6 (with balconies) 8.4 (without balconies)	-	-

The values of the standard deviations computed based on GNSS points coordinates, are two to three times larger than the ones computed based on the distances between each point of the UAS point cloud and the final 3D model, because these errors combine the georeferencing errors of UAS images with those of modeling. The errors obtained based on the UAS point clouds are in accordance with the errors reported in [4] even if we have compared the entire 3D model and not only the roofs.

5. Conclusions

3D modelling of urban areas is an attractive and challenging research topic, buildings being the most prominent category. There are multiple geospatial data sources for buildings' 3D reconstruction, such as LiDAR point clouds and aerial or satellite images, and over the years algorithms and methods have been developed for this purpose. However, with

the developments of UAS systems they have become an alternative data source for the 3D building modeling process.

This paper presented in detail an end-to-end pipeline to reconstruct 3D buildings using a semi-automatic data-driven approach based on the classification and segmentation of UAS building point clouds, without additional information such as cadastral data, starting from flight planning for UAS image acquisition, to the creation of the final parametrized building 3D model, in accordance with the OGC CityGML standard, LOD2.

To find the best scenario for UAS images acquisition, 275 ChPs have been considered with different locations in the study area: on top of the buildings, in the interior or exterior part of the flight zone, having also a different nature, such as: artificial and natural points pre-marked by paint or natural features. The total RMSE calculated for the 275 ChPs showed that the best scenario was obtained when using 45 GCPs, but only 20 GCPs can ensure an accuracy of 3 GSD in planimetry, 2.5 GSD in elevation and 4 GSD in total.

By resampling the UAS point cloud with a minimum distance of 0.1 m, leading to a density of 120 points/m^2, the time complexity and the hardware requirements for the UAS point cloud classification and segmentation processes were significantly reduced. Computing only three attributes, VDVI, normal and normalized Z, a new workflow for the automatic extraction of vegetation UAS points using a combination of point clouds and rasters was proposed, providing an accuracy of approximately 99%. After testing two scenarios for the supervised classification of the UAS point cloud, including the VDVI attribute or not, it can be concluded that the VDVI is a key attribute in this process as a percentage of 77% classified data is covered by the five most relevant attributes, compared to the other model without VDVI that identifies eight notable attributes but which answers only for 69% of the model.

The workflow used for the final step of the building reconstruction uses only one architectural constraint, i.e., the perpendicularity of the building facades. Considering the global accuracy of less than 15 cm for each modeled building through the workflow proposed in this study, it can be concluded that the method is reliable. In addition, the manual operator intervention is at the end of the process when the final 3D model of the building is created with minimum effort, time and a-priori information about the building.

Appendix A

Figure A1. *Cont.*

Figure A1. The spatial distribution of GCPs and ChPs and the residuals of 275 ChPs for each scenario.
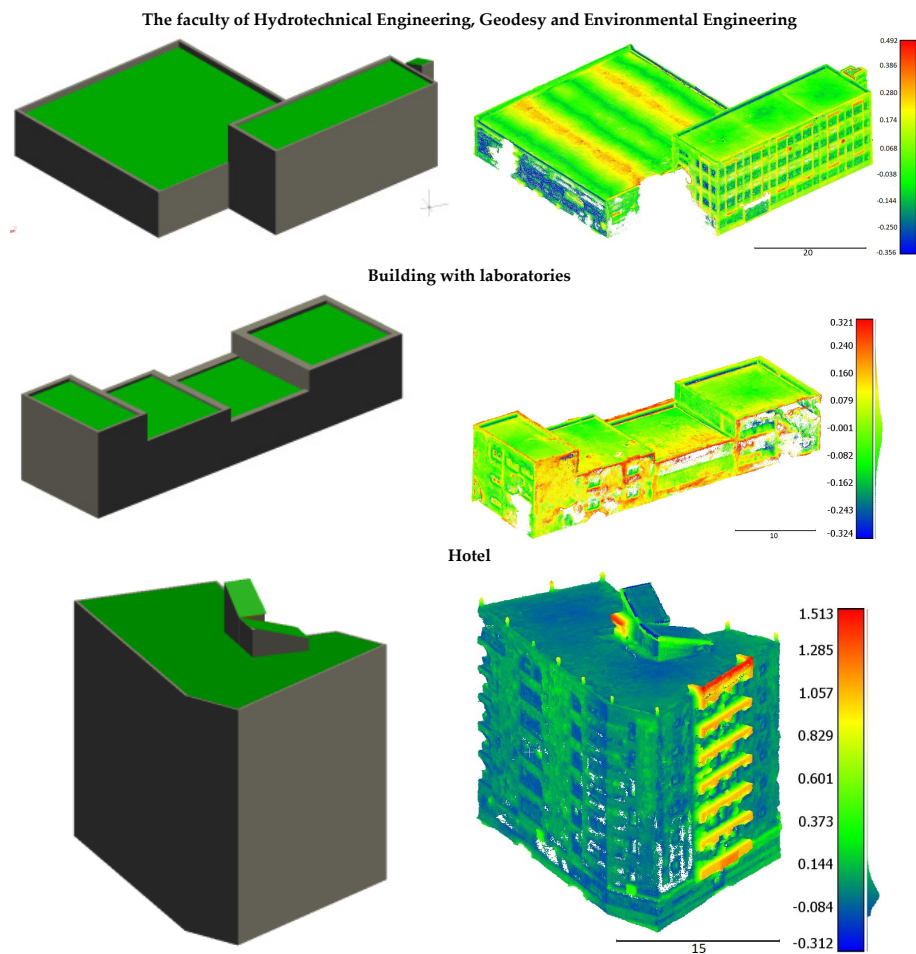
Appendix B

Figure A2. The errors of UAS point clouds corresponding to three analyzed buildings.

References

1. United Nations. *Revision of the World Urbanization Prospects*; Population Division of the United Nations Department of Economic and Social Affairs (UN DESA): New York, NY, USA, 2018.
2. Carneiro, C. Communication and Visualization of 3D Urban Spatial Data Acoording to User Requirements: Case Study of Geneva. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* **2008**, *37*, 631–636.
3. Lafarge, F. Modèles Stochastiques Pour la Reconstruction Tridimensionnelle D'environnements Urbains. Ph.D Thesis, Ecole Nationale Supérieure des Mines de Paris, Paris, France, 2007.
4. Drešček, U.; Kosmatin Fras, M.; Tekavec, J.; Lisec, A. Spatial ETL for 3D Building Modelling Based on Unmanned Aerial Vehicle Data in Semi-Urban Areas. *Remote Sens.* **2020**, *12*, 1972. [CrossRef]
5. Boulaassal, H.; Landes, T.; Grussenmeyer, P.; Tarsha-Kurdi, F. Automatic segmentation of building facades using terrestrial laser data. In Proceedings of the ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo, Finland, 12–14 September 2007; pp. 65–70.
6. Lafarge, F.; Descombes, X.; Zerubia, J.; Deseilligny, M.P. Structural approach for building reconstruction from a single DSM. *IEEE Trans. Pattern Anal.* **2010**, *32*, 135–147. [CrossRef]
7. Kolbe, T.; Groger, G.; Plumer, L. *Interoperable Access to 3D City Models, First International Symposium on Geo-Information for Disaster Management*; Springer: Delft, The Netherlands, 2005.
8. Kolbe, T.H. Representing and exchanging 3D city models with CityGML. In *3D Geo-Information Sciences*; Zlatanova, S., Lee, J., Eds.; Springer: Berlin, Germany, 2009; pp. 15–31. [CrossRef]
9. Kutzner, T.; Chaturvedi, K.; Kolbe, T.H. CityGML 3.0: New Functions Open Up New Applications. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* **2020**, *88*, 43–61. [CrossRef]
10. Löwner, M.O.; Gröger, G.; Benner, J.; Biljecki, F.; Nagel, C. Proposal for a new LOD and multi-representation concept for CityGML. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *4*, 3–12. [CrossRef]
11. Hammoudi, K. Contributions to the 3D City Modeling: 3D Polyhedral Building Model Reconstruction from Aerial Images and 3D Facade Modeling from Terrestrial 3D Point Cloud and Images. Ph.D. Thesis, Institut Géographique National (IGN), Saint-Mandé, France, 2011.
12. Dorninger, P.; Pfeifer, N. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors* **2008**, *8*, 7323–7343. [CrossRef] [PubMed]
13. Rottensteiner, F. Automatic generation of high-quality building models from LiDAR data. *IEEE Comput. Graph. Appl.* **2003**, *23*, 42–50. [CrossRef]
14. Nys, G.-A.; Poux, F.; Billen, R. CityJSON Building Generation from Airborne LiDAR 3D Point Clouds. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 521. [CrossRef]
15. Murtiyoso, A.; Veriandi, M.; Suwardhi, D.; Soeksmantono, B.; Harto, A.B. Automatic Workflow for Roof Extraction and Generation of 3D CityGML Models from Low-Cost UAV Image-Derived Point Clouds. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 743. [CrossRef]
16. Mwangangi, K. 3D Building Modelling Using Dense Point Clouds from UAV. Master Thesis, University of Twente, Enschede, The Netherlands, March 2019.
17. Xiong, B.; Oude Elberink, S.; Vosselman, G. Building modeling from noisy photogrammetric point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 197–204. [CrossRef]
18. Xiong, B.; Oude Elberink, S.J.; Vosselman, G. Footprint map partitioning using airborne laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 241–247. [CrossRef]
19. Huang, H.; Michelini, M.; Schmitz, M.; Roth, L.; Mayer, H. LOD3 building reconstruction from multi-source images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2020**, *XLIII-B2-2020*, 427–434. [CrossRef]
20. Pfeifer, N.; Mandlburger, G.; Otepka, J.; Karel, W. OPALS—A framework for Airborne Laser Scanning data analysis. *Comput. Environ. Urban Syst.* **2014**, *45*, 125–136. [CrossRef]
21. 3DFlow 3DF Zephyr Official Web Site. Available online: https://www.3dflow.net/ (accessed on 21 June 2021).
22. CloudCompare Official Web Site. Available online: http://www.danielgm.net/cc/ (accessed on 10 June 2021).
23. LAStools, Efficient LiDAR Processing Software (unlicensed). Available online: http://rapidlasso.com/LAStools (accessed on 10 June 2021).
24. *AutoCAD Map 3D v*; academic license; Autodesk, Inc.: San Rafael, CA, USA, 2016; Available online: https://www.autodesk.pl (accessed on 19 November 2021).
25. Oniga, V.E.; Breaban, A.I.; Pfeifer, N.; Chirila, C. Determining the suitable number of ground control points for UAS images georeferencing by varying number and spatial distribution. *Remote Sens.* **2020**, *12*, 876. [CrossRef]
26. Fan, P.; Li, W.; Cui, X.; Lu, M. Precise and Robust RTK-GNSS Positioning in Urban Environments with Dual-Antenna Configuration. *Sensors* **2019**, *19*, 3586. [CrossRef] [PubMed]
27. Morales, J.; Martínez, J.L.; García-Cerezo, A.J. A Redundant Configuration of Four Low-Cost GNSS-RTK Receivers for Reliable Estimation of Vehicular Position and Posture. *Sensors* **2021**, *21*, 5853. [CrossRef]
28. Gabara, G.; Sawicki, P. Multi-Variant Accuracy Evaluation of UAV Imaging Surveys: A Case Study on Investment Area. *Sensors* **2019**, *19*, 5229. [CrossRef]
29. Xie, P.; Petovello, M.G. Measuring GNSS Multipath Distributions in Urban Canyon Environments. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 366–377. [CrossRef]

30. Lesouple, J.; Robert, T.; Sahmoudi, M.; Tourneret, J.; Vigneau, W. Multipath Mitigation for GNSS Positioning in an Urban Environment Using Sparse Estimation. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 1316–1328. [CrossRef]

31. Rogers, S.R.; Manning, I.; Livingstone, W. Comparing the Spatial Accuracy of Digital Surface Models from Four Unoccupied Aerial Systems: Photogrammetry Versus LiDAR. *Remote Sens.* **2020**, *12*, 2806. [CrossRef]

32. Gašparović, M.; Seletković, A.; Berta, A.; Balenović, I. The evaluation of photogrammetry-based DSM from low-cost UAV by LiDAR-based DSM. *South-East Eur. For. SEEFOR* **2017**, *8*, 117–125. [CrossRef]

33. Piltz, B.; Bayer, S.; Poznanska, A.M. Volume based DTM generation from very high resolution photogrammetric DSMs. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2016**, *41*, 83–90. [CrossRef]

34. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [CrossRef]

35. Štroner, M.; Urban, R.; Lidmila, M.; Kolář, V.; Křemen, T. Vegetation Filtering of a Steep Rugged Terrain: The Performance of Standard Algorithms and a Newly Proposed Workflow on an Example of a Railway Ledge. *Remote Sens.* **2021**, *13*, 3050. [CrossRef]

36. Zhu, L.; Shortridge, A.; Lusch, D. Conflating LiDAR data and multispectral imagery for efficient building detection. *J. Appl. Remote Sens.* **2012**, *6*, 063602. [CrossRef]

37. Weinmann, M.; Schmidt, A.; Mallet, C.; Hinz, S.; Rottensteiner, F.; Jutzi, B. Contextual classification of point cloud data by exploiting individual 3D neigbourhoods. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2*, 271–278. [CrossRef]

38. Brodu, N.; Lague, D. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS J. Photogram. Remote Sens.* **2012**, *68*, 121–134. [CrossRef]

39. Özdemir, E.; Remondino, F.; Golkar, A. An Efficient and General Framework for Aerial Point Cloud Classification in Urban Scenarios. *Remote Sens.* **2021**, *13*, 1985. [CrossRef]

40. Demantké, J.; Mallet, C.; David, N.; Vallet, B. Dimensionality based scale selection in 3D LiDAR point clouds. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* **2011**, *3812*, 97–102. [CrossRef]

41. Lalonde, J.; Unnikrishnan, R.; Vandapel, N.; Herbert, M. *Scale Selection for Classification of Point-Sampled 3-D Surfaces*; Technical Report CMU-RI-TR-05-01; Robotics Institute: Pittsburgh, PA, USA, 2015.

42. Höfle, B.; Mücke, W.; Dutter, M.; Rutzinger, M.; Dorninger, P. Detection of building regions using airborne LiDAR—A new combination of raster and point cloud based GIS methods. In Proceedings of the GI-Forum 2009-International Conference on Applied Geoinformatics, Salzburg, Austria, 7–10 July 2009; pp. 66–75.

43. Höfle, B.; Hollaus, M.; Hagenauer, J. Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 134–147. [CrossRef]

44. Reitberger, J.; Schnörr, C.; Krzystek, P.; Stilla, U. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 561–574. [CrossRef]

45. Wang, X.; Wang, M.; Wang, S.; Wu, Y. Extraction of vegetation information from visible unmanned aerial vehicle images. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 152–159.

46. Grilli, E.; Menna, F.; Remondino, F. A review of point clouds segmentation and classification algorithms. *Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci.* **2017**, *42*, 339. [CrossRef]

47. Coussement, K.; Van den Poel, D. Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Syst. Appl.* **2008**, *34*, 313–327. [CrossRef]

48. Xue, D.; Cheng, Y.; Shi, X.; Fei, Y.; Wen, P. An Improved Random Forest Model Applied to Point Cloud Classification. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *768*, 072037. [CrossRef]

49. Nguyen, A.; Le, B. 3D point cloud segmentation: A survey, 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM). In Proceedings of the 2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM), Manila, Philippines, 12–15 November 2013; pp. 225–230. [CrossRef]

50. Xie, Y.; Tian, J.; Zhu, X.X. Linking Points with Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 38–59. [CrossRef]

51. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model fitting with application to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]

52. Schnabel, R.; Wahl, R.; Klein, R. Efficient RANSAC for Point-Cloud Shape Detection. *Comput. Graph. Forum* **2007**, *26*, 214–226. [CrossRef]

53. Satari, M.; Samadzadegan, F.; Azizi, A.; Maas, H.G. A Multi-Resolution Hybrid Approach for Building Model Reconstruction from Lidar Data. *Photogramm. Rec.* **2012**, *27*, 330–359. [CrossRef]

54. Rabbani, T.; Van Den Heuvel, F.; Vosselman, G. Segmentation of point clouds using smoothness constraint. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2006**, *36*, 248–253.

55. Pöchtrager, M. *Segmentierung Großer Punktwolken Mittels Region Growing*; Technische Universität Wien: Vienna, Austria, 2016; Available online: http://katalog.ub.tuwien.ac.at/AC13112627 (accessed on 12 November 2021).

56. Weidner, U.; Forstner, W. Towards automatic building extraction from high resolution digital elevation models. *ISPRS J. Photogram. Remote Sens.* **1995**, *50*, 38–49. [CrossRef]