

RESEARCH ARTICLE

Archaeological Ground Point Filtering of Airborne Laser Scan Derived Point-Clouds in a Difficult Mediterranean Environment

Michael Doneus*, Gottfried Mandlbauer^{†,‡} and Nives Doneus[§]

Digital terrain models (DTM) based on airborne laser scanning (ALS) are an important source for identifying and monitoring archaeological sites and landscapes. However, a DTM is only one of many representations of a given surface. Its accuracy and quality must conform to its purpose and are a result of several considerations and decisions along the processing chain. One of the most important factors of ALS-based DTM generation is ground point filtering, i.e., the classification of the acquired point-cloud into terrain and off-terrain points. Filtering is not straightforward. The resulting DTM is usually a compromise that might show the surface below very dense vegetation while losing detail in other areas. In this paper, we show that in very complex situations (e.g., strongly varying vegetation cover), an optimal compromise is difficult to achieve, and more than one filter with different settings adapted to the varying degree of vegetation cover is necessary. For practical reasons, the results need to be combined into a single DTM. This is demonstrated using the case study of a Mediterranean landscape in Croatia, which consists of open areas (agricultural and grassland), olive plantations, as well as extremely dense and evergreen macchia vegetation. The results are the first step toward an adaptive ground point filtering strategy that might be useful far beyond the field of archaeology.

Keywords: ALS; Classification; Adaptive Filtering; Archaeology; Mediterranean Environment; Macchia

1. Introduction

Within the last one and a half decades, airborne laser scanning (ALS) has become a widely used technology in archaeology. This is due to the possibility of deriving detailed digital terrain models (DTM – please note that we are following a Eurocentric diction) even under dense vegetation and shallow water, the growing availability of countrywide ALS-derived DTMs (at least in Europe), and its seeming ease of use. A simple hill-shaded representation of the DTM is almost self-explanatory, even allowing archaeologists untrained in remote sensing to identify previously unknown archaeological and palaeo-environmental features. As a result, ALS has become a fixed component of integrated prospection approaches and has dramatically changed our understanding of archaeological sites, monuments, and landscapes, especially in wooded areas.

The downside of this alleged ease of use is a certain ignorance of the underlying modelling and data processing

strategies combined with a naivety of expectations, followed by an incorrect use that might lead to disappointment. Knowledge about its basic principles and workflow is, therefore, a prerequisite for assessment of the suitability of ALS-based datasets for a certain purpose. Consequently, the only eligible countermeasure to this scenario is training and information about the whole process of ALS, from project planning to final visualization.

The basic steps during an ALS workflow are (Crutchley, 2010; Doneus & Briese, 2011; Opitz, 2013; Fernandez-Diaz et al., 2014; Shan & Toth, 2018):

- Project planning (including defining the purpose of the scan),
- system calibration (determination of lever arms, bore-sight angles, range and scan angle offsets, and scales),
- data acquisition (instrumentation and settings, time-frame and flight mission parameters),
- geo-referencing (direct geo-referencing of scan data by combining scanner range and deflection angle measurements and trajectory data derived from GNSS and inertial measurement devices),
- refraction correction when employing water-penetrating green lasers,
- flight strip adjustment,
- classification/ground point filtering (although today, classification (and recently ‘semantic labelling’) seems

* University of Vienna, Department of Prehistoric and Historical Archaeology, AT

† University of Stuttgart, Institute for Photogrammetry, DE

‡ TU Wien, Department of Geodesy and Geoinformation, AT

§ Austrian Academy of Sciences, Austrian Archaeological Institute, AT

Corresponding author: Michael Doneus (michael.doneus@univie.ac.at)

to be the more widely used term, we will be mainly using the term ground point filtering throughout this paper. While 'classification' is the assignment of points to different classes, 'ground point filtering' is the removal of all points from the georeferenced point-cloud that do not contribute to an archaeologically relevant DTM),

- DTM interpolation, and finally,
- visualization.

All these workflow steps require considerations and decisions that will affect the suitability of the resulting DTM for a certain purpose.

For archaeology, the most critical part of the workflow is ground point filtering – the removal of all points from the geometrically calibrated, georeferenced point-cloud that do not contribute to an archaeologically relevant DTM (see next section). It is comparatively rare that archaeologists will have the possibility to independently generate a DTM from ALS data, a process requiring access to the unfiltered point-cloud, classification software, and respective data processing expertise. Consequently, ground point filtering is a blackbox in (not only) archaeological applications. ALS data providers usually do not report on software and settings used for classification. In these cases, even when working with project-based data that were specifically acquired for archaeological purposes, reproducibility of the provided DTM is not possible and – even worse – its archaeological value is difficult to assess. This makes it even more important to understand constraints and pitfalls that come along with ground point filtering in order to evaluate the archaeological suitability of an ALS-based DTM.

Still, little has been published on this topic (Doneus & Briese, 2006; Cifani et al., 2007; Crow et al., 2007; Lasaponara & Masini, 2009; Heinzl & Sittler, 2010; Lasaponara et al., 2011; Lugmayr, 2013; Opitz & Nuninger, 2014). To the best of our knowledge, no information can be found on archaeology-oriented ALS-point cloud classification of extremely difficult situations, including varying topography with steep slopes and disparity of vegetation density. These are situations that usually cannot be filtered with a single parameter set. For geomorphological applications, terrestrial laser scans of complex scenes have been classified using a multi-scale dimensionality analysis (Brodu & Lague, 2012).

This paper investigates a suitable archaeological ground point filtering of ALS-derived point-clouds for challenging settings, as described above. After presenting some important facts about DTM generation in Section 2, the case study and datasets are introduced in Sections 3 and 4. In Section 5, the filters used are presented. Sections 6 and 7, finally, present and discuss the results, including the propagation of an adaptive filtering strategy combining multiple filtered DTMs.

2. Ground point filtering of ALS derived point-clouds

As mentioned in the introduction, a crucial step of archaeological DTM generation is filtering the point cloud into terrain and off-terrain points. Any DTM is a representa-

tion of the terrain's surface (be it bare earth or including walls and buildings). There is no single DTM that could be regarded as displaying "the" true surface. This is already expressed by the term "model," which is meant to be a simplified image of reality. This implies that one and the same ground surface can be modelled theoretically by an indefinite number of terrain models. It is, therefore, important to know that any DTM is one of many representations of a ground surface. The decision as to whether it is a good or appropriate model depends on the purpose of the terrain modelling, which influences its scale and in consequence, its detail.

It is important to assert that, while a DTM is *sensu stricto* a representation of the bare earth surface void of vegetation, buildings, cars and the like, an archaeologically relevant DTM may differ from this specification. When a DTM is used to identify archaeological features, a DTM also needs to be void of any vegetation, but buildings (e.g., ruined castles), standing stones, walls, roads, channels, earthworks and the like should survive any filtering and be represented in the final "DTM." Therefore, it might be rather called an "archaeological digital elevation model." Still, this term cannot be coined for all archaeological applications; for example, modelling palaeohydrology of a certain point in time might demand terrain models void of walls and standing stones from later periods.

For the purpose of mapping archaeological features, a suitable DTM should depict bare ground displaying archaeologically relevant micro-relief as mentioned above but at the same time excluding any non-archaeologically induced micro-relief (e.g., tree-stumps). There is usually no evidence available that would provide a proper indication of the "true" ground as well as quantity, appearance, and shape of the archaeologically relevant micro relief. Therefore, during the process of point-cloud classification, the question of whether a filtered DTM is appropriate can only be decided by the person carrying out the task of point cloud classification. In an ideal case, this would be an archaeologist or a domain expert working in close cooperation with the archaeological end-user.

Ground point filtering of point clouds for archaeology is not straightforward. There are a wide range of software packages available that have been reviewed in a small number of papers (Sithole & Vosselman, 2003; Lugmayr, 2013; Polat & Uysal, 2015; Forte & Campana, 2017). Depending on the software used, a greater or lesser number of parameters must be set in order to identify surface points for certain environmental settings. In any case, the creation of a DTM from an ALS-based point-cloud is a compromise between deleting unwanted points while at the same time keeping archaeological detail. This is, to a certain degree, achievable in areas with deciduous trees and sparse underwood in off-leave conditions (although walls could become truncated by the filtering). A dense vegetation cover, however, might prevent most laser pulses from reaching the ground, resulting in a high quantity of vegetation-points and only a few ground echoes. Filtering this kind of situation requires strong filter settings that might also remove archaeological detail in less densely vegetated areas of the dataset.

The most problematic settings are mixed situations with an extremely varying topography (e.g., a mixture of smooth and predominantly horizontal terrain, undulating terrain, and steep slopes) and disparity of vegetation density. These cases usually cannot be classified with a single parameter set. However, to our knowledge, there is no classification software available that could handle such difficult settings using adaptive parameters based on vegetation density and slope. In order to receive optimum results for such diverse areas, repeated classification with particular parameter sets for dense vegetation and more open areas are necessary, resulting in multiple DTMs. For practical reasons during visualization and interpretative mapping, the resulting DTMs should, however, be combined into a single terrain model. In the following, this approach is demonstrated using the case study of a Mediterranean landscape in Croatia.

3. Case study: The Mediterranean landscape of the Medulin Bay, Croatia

The northern Adriatic peninsula of Istria lies between the Gulf of Trieste, Italy, and the Kvarner Bay, next to Rijeka, Croatia. Along with the port of Pula and the bay of Raša, the Medulin Bay is one of the largest natural anchorages in Istria (**Figure 1**). It is situated at the southernmost tip of Istria, between Marlera and Kamenjak Cape with a total area of 22 km². The water depth has an average of less than 8 m in the inner bay. A few small, uninhabited islands are scattered around the bay. In the west, the bay of Medulin is bordered by the peninsula of Premantura, in the north and east it is connected to the villages of Pomer and Medulin. The relief is represented by hills and

limestone terraces with a maximum height of 80 m and steep slopes towards the coast in the southern part. The vegetation cover consists of open areas (agricultural and grassland), olive plantations, as well as extremely dense and evergreen macchia vegetation (**Figure 2**).

The region is mainly known for its numerous Roman sites (Koncani Uhač, 2008; Girardi-Jurkić, 2013), the most notable being the villa maritima of Vižula (Džin & Girardi Jurkić, 2008), situated in the municipality of Medulin. Remains of Roman architecture, partly submerged due to the rising sea level, have been documented here over a total length of 1.2 km. The site has been investigated through underwater and terrestrial excavations (Girardi Jurkić et al., 2012; Miholjek, 2012) since the mid-1990s, and via large-scale geophysical prospection from 2014. The EU project “Archaeological Park Vižula” (2017–2019), led by the Municipality of Medulin to preserve and present the site to the public, provided the acquisition of new data by means of ALS in 2018. The objective was to combine underwater and terrestrial research and to understand the relationship between the site and the past landscape.

4. Data

Approximately 24 km² of land and underwater terrain were scanned in the area of Medulin Bay, including Vižula and the peninsula Premantura (**Figure 3**). A topo-bathymetric laser scanner operating at a wavelength of 532 nm (i.e., visible green domain of the electromagnetic spectrum), which can penetrate clear water (see also Doneus et al., 2013), was used due to the fact that the Roman architecture is currently partly submerged (Vacchi et al., 2016). Data was acquired in March 2018 using a Riegler VQ-820-G



Figure 1: Map of the northern Adriatic coast indicating the location of the Medulin Bay (red rectangle). Data Source: SRTM. North is up.



Figure 2: On-site photographs of the scanned area from February 2019; **(a)** low to medium height macchia overgrowing large parts of the areas with interspersed fields **(b)** area with extremely dense and high macchia **(c)** area with pine trees **(d)** coastal area with partly very steep slopes.

topo-hydrographic airborne laser scanner (Steinbacher et al., 2012) operated by Airborne Technologies GmbH in calm water conditions. The entire area was covered in two blocks, with 29 flight strips overlapping at least by 50%. The scan settings are listed in **Table 1**. Additionally, a medium format digital RGB camera IGI DigiCam-H/39 (equipped with a Hasselblad lens HC 3.5/50) simultaneously documented the area with a ground sampling distance of 7 cm.

Each of the scanned strips has an average laser pulse density of 6 shots/m². As the strips have overlaps between 50 and 60%, the laser pulse density has an overall mean of 12 shots/m². This number is, however, not uniform as it varies depending on the number of strips overlapping in an area. Also, areas of deeper water did not return any echoes, and in that way, reducing the average laser pulse density (see **Figure 3b**). Within a more typical project region (on land with mainly 2-3-fold strip overlap and a mix of open area and high vegetation – as in **Figures 6** and **7**), the mean pulse density is 15 points/m² (see **Table 1**).

Due to the diverse vegetation, the classification resulted in an uneven distribution of ground points, which will be detailed in the following sections. Overall, the average of classified ground points is 11 points/m². The ratio of laser pulse density and resulting ground points typically lies between 1 (in open areas and shallow water) and 4 (dense Macchia). In the region of dense Macchia from **Figures 2b** and **10**, the average pulse density is as high as 21 (the high number is due to a cross-strip flown over this area), while the average ground point density is only 7.5 points/m².

Data pre-processing was carried out by the lead author using the software package OPALS (Mandlbürger et al.,

2009; Pfeifer et al., 2014) and comprised the following steps:

1. Echo detection: this was achieved using the scanner's online waveform processing capability (Pfennigbauer & Ullrich, 2010; Pfennigbauer et al., 2014) and the scanner manufacturer's software RiProcess.
2. Derivation of a water surface model. Due to the calm water conditions and short acquisition time, modelling in the air-water-interface with a single horizontal surface at the water level was considered sufficient. Minor influences from residual waves and tidal differences were ignored.
3. Georeferencing of each scanned strip by first combining the synchronously captured flight trajectory (GNSS/IMU) and scanner measurements (raw range and scan angle), and by subsequently performing range and refraction correction of the water echoes based on the water surface model (Snell's law).
4. Strip adjustment of georeferenced point cloud including quality control of the results (Ressl et al., 2008; Ressler et al., 2011).

The application of these four processing steps resulted in a dataset of 29 georeferenced, refraction-corrected, and adjusted point clouds in ASPRS LAS format, each representing one of the original strips.

5. Ground point filtering methodology

Strip adjustment is usually followed by the classification, which uses appropriate algorithms to classify the geometrically calibrated point cloud amongst others into

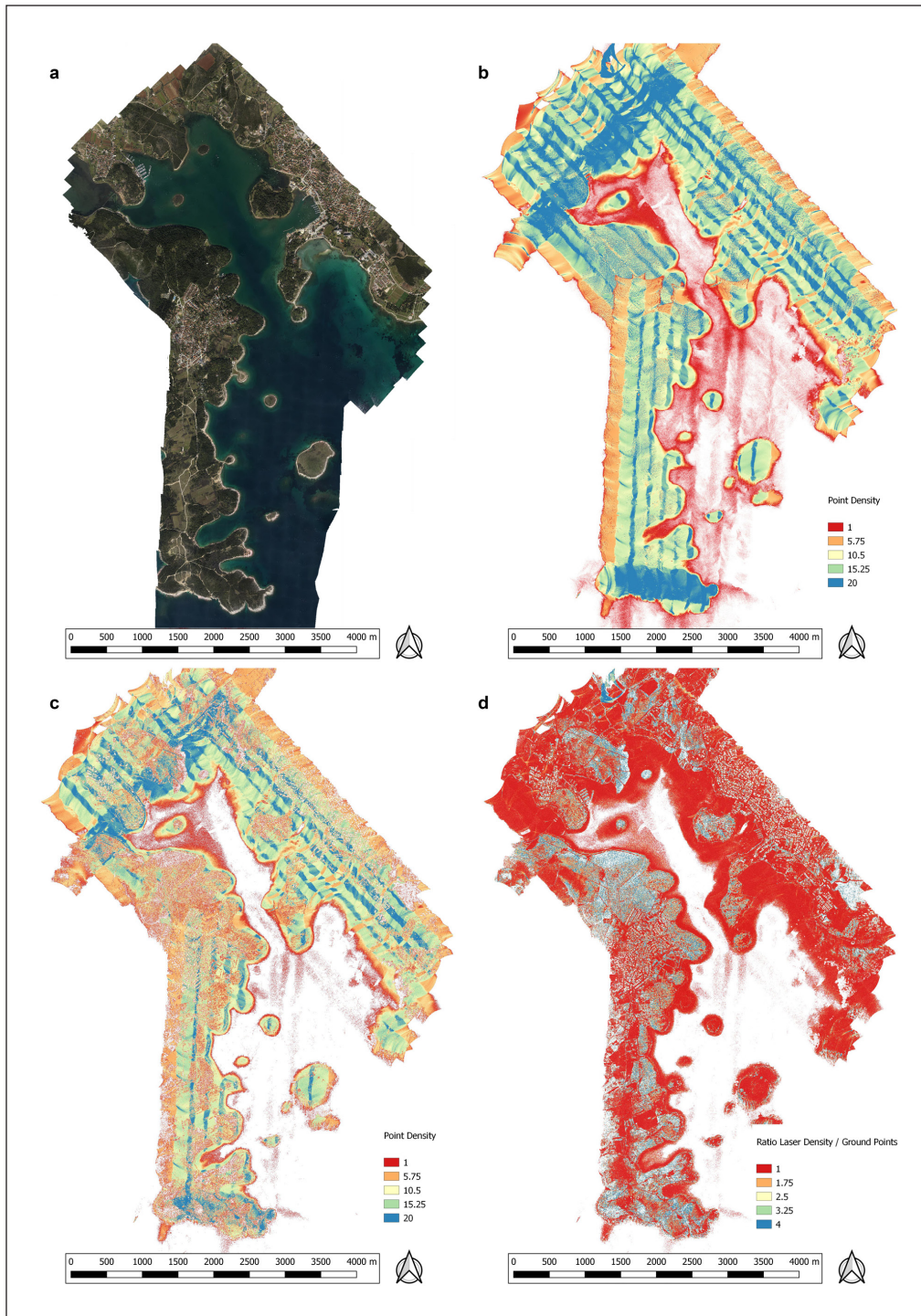


Figure 3: (a) Orthophoto mosaic of the scanned area. The aerial photographs were acquired simultaneously with the laser scan in March 2018. (b) Laser pulse density (representative pulse per square meter). (c) The density of ground points after classification using the 'Dense Vegetation' setting (see next section). (d) The ration of laser pulse density and classified ground points. Low numbers (i.e., number of ground points equals laser pulse density) are due to open areas, modern settlements, and shallow water; high numbers indicate dense macchia vegetation.

ground points, clutter points (low and high points), low, medium, and high vegetation. This process is also known as semantic labelling of the dataset. In our case, the aim was to derive an archaeological digital elevation model as specified in Section 2. The software package SCOP++ was employed, which uses robust hierarchic interpolation with an eccentric and asymmetrical weight function (Kraus & Pfeifer, 1998; Pfeifer et al., 2001; see Doneus et al., 2008:

887). Because of the extreme variance in vegetation density, two DTMs had to be derived, applying different parameter sets: (1) a filter strategy was adapted to open areas and underwater surfaces, (2) a filter for very dense vegetation.

Both filters were developed within SCOP++ in a hierarchical framework (Table 2). Each filter is composed of 13 steps arranged in five groups. Each of the first three groups contains:

Table 1: Most important metadata of the airborne laser scan covering the Medulin Bay.

Date of acquisition	March 2018
Instrument	Riegl VQ-820-G
Scanner type	Full waveform (with online waveform processing)
Pulse repetition rate (PRR) [kHz]	284
Altitude ground level (AGL) [m]	400
Footprint diameter [m]	0.4
Field of view (FOV) [deg]	42
Scan lines per second	157
Average speed [kts]	108
Average laser pulse density per m ² (in a typical area with at least 50% strip-overlap)	15
Average ground point density per m ²	11

Table 2: Filter steps and relevant parameters of both filter strategies defined within the framework of SCOP++.

	Steps	Step name	Parameters	Open area	Dense vegetation
1	Step 1	ThinOut	Cell size	10	10
			Thinning method	mean	mean
			Level k	–	–
	Step 2	Filter	Lower branch	4;4;8	4;4;8
			Upper branch	0.15;0.15;0.3	0.15;0.15;0.3
			Trend	below	below
			Prediction	below	below
			Penetration rate	80%	40%
			Interpolation	Grid width: 10m CU:20 Covariance function: bell curve	Grid width: 10m CU:20 Covariance function: bell curve
	Step 3	SortOut	Upper distance	0.7	0.7
Lower distance			7.0	7.0	
Slope dependency			2.0	2.0	
2	Step 4	ThinOut	Cell size	5	5
			Thinning method	mean	k-th lowest (4)
	Step 5	Filter	Lower branch	0.7;0.7;1.0	0.7;0.7;1.0
			Upper branch	0.3;0.3;0.3	0.3;0.3;0.5
			Trend	below	below
			Prediction	below	below
			Penetration rate	80%	40%
			Interpolation	Grid width: 4.0 CU:20 CF: bell curve	Grid width: 4.0 CU:20 CF: bell curve
	Step 6	SortOut	Upper distance	0.5	0.5
			Lower distance	1.0	2.0
			Slope dependency	2.0	2.0
	3	Step 7	ThinOut	Cell size	1.5
Thinning method				mean	k-th lowest (2)

(Contd.)

	Step 8	Filter	Lower branch	0.15;0.15;0.2	0.35;0.35;1.0
			Upper branch	0.05;0.05;0.1	0.05;0.05;0.1
			Trend	below	below
			Prediction	below	below
			Penetration rate	80%	40%
			Interpolation	Grid width: 1.5 CU:50 CF: bell curve	Grid width: 1.5 CU:20 CF: bell curve
	Step 9	SortOut	Upper distance	0.5	0.5
			Lower distance	0.2	1.0
			Slope dependency	2.0	2.0
4	Step 10	ThinOut	Cell size	0.3	0.3
			Thinning method	lowest	lowest
	Step 11	Filter	Lower branch	0.05;0.05;0.1	0.1;0.1;0.3
			Upper branch	0.1;0.1;0.2	0.05;0.05;0.1
			Trend	below	below
			Prediction	below	below
			Penetration rate	80%	40%
			Interpolation	Grid width: 0.25 CU:20 CF: bell curve	Grid width: 0.25 CU:20 CF: bell curve
	Step 12	Interpolate	Grid width	0.5	0.5
			CU	25	25
			Covariance function	adapting	adapting
	5	Step 13	Classify	Output format	las
High vegetation				5.0	5.0
Medium vegetation				1.0	1.0
Low vegetation				0.25	0.25
ground					
Below DTM				-0.25	-0.25
Slope dependency				-	-

- A *ThinOut* step, where the point cloud is thinned out using a raster with user-defined cell size. For each cell, one output point is chosen using a specified parameter (for a DTM, one would usually choose lowest, k-th lowest, or mean).
- A *Filter* step, applying a robust filter with an eccentric and asymmetric weight function to the thinned point cloud: First, an initial surface is computed using linear prediction (equivalent to ordinary Kriging) and equally weighted points. Then, each point is assigned an individual weight based on the distance from the initial surface using a user-defined weight function. Points below the computed surface receive a higher weight than points lying above the surface (= unsymmetrical weight function – see **Figure 4**). After that, a new surface is calculated based on the applied weights. The procedure is iterated until a certain threshold is reached.
- A *SortOut* step, which calculates the z-difference of each point of the original point cloud and the resulting

surface from the previous filter step. Points located outside a user-specified interval around the filtered surface are disregarded for the subsequent iteration round.

This sequence of *ThinOut* – *Filter* – *SortOut* is repeated within the first three groups with consecutive finer settings (see **Table 2**). In the fourth group, *ThinOut* and *Filter* are followed by an interpolation of the resulting DTM. Finally, a threshold-based classification is applied that allocates all original points into high vegetation, medium vegetation, low vegetation, ground, and points below DTM.

The most important parameters of both filter strategies are listed in **Table 2**. Due to many clutter points both below and above the actual terrain, the thinning method had to be set to “mean” in the first group (step 1). Such clutter points are typical for any (topo-) bathymetric scanners employing highly sensitive receivers to detect very weak echoes from the benthic layer (see **Figure 5**). This led to a surface that would be close to the actual terrain in open areas, at the same time being above the terrain

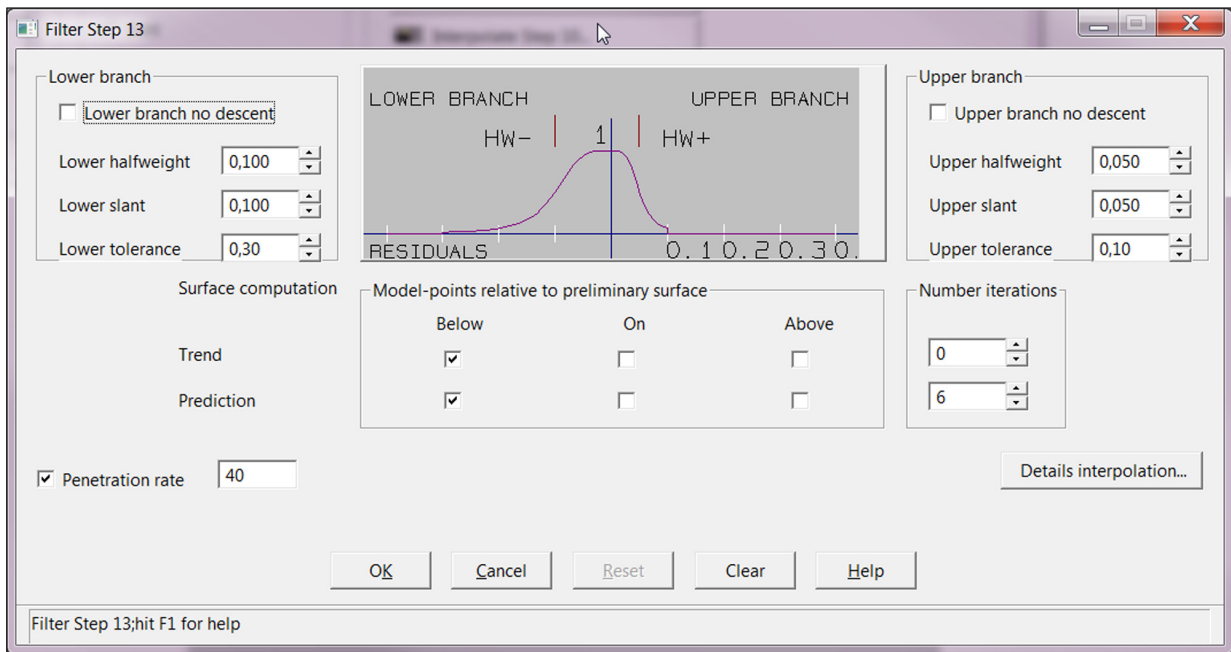


Figure 4: Screenshot from SCOP++ showing the parameters and the asymmetric weight function of a filter step.

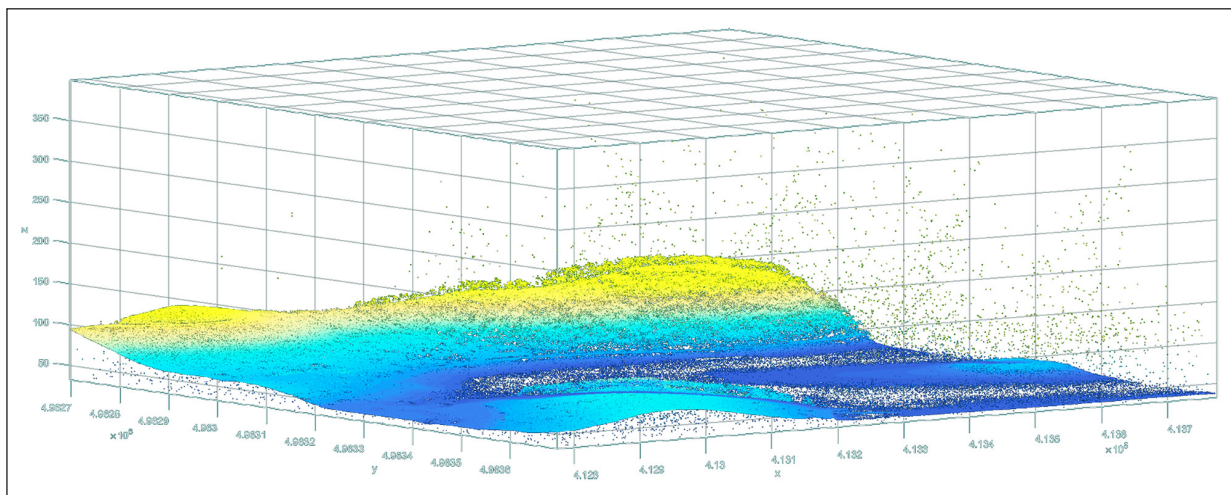


Figure 5: Perspective view on the point cloud from parts of a single laser scanning strip. Plenty of clutter points become evident above and below the actual terrain.

in vegetated areas. While the most excessive clutters had been removed by the first *SortOut* (step 3), clutter points closer to the actual terrain surface were still present. To force the surface down towards the actual terrain in vegetated areas, the thinning method is subsequently set to “k-th lowest” in steps 4 and 7 for the “dense vegetation” parameter set.

6. Result

The final DTMs differ, each displaying varying degrees of success in representing archaeologically relevant terrain, as demonstrated in **Figure 6**, which depicts a section of a coastal area. The area is characterized by small parts of open land alternating with dense macchia. The lower half of the image is dominated by the clear waters of the Mediterranean Sea, in which the green laser was able to penetrate down to 8 m depth.

The hillshade in **Figure 6a** is based on the ground point filtering that resulted from the “dense vegetation” parameter set. Here, the vegetation was removed. Consequently, the area in the upper half shows the terrain, which includes archaeologically relevant structures (a collapsed and eroded dry-stone wall – **Figure 6a/1**). However, the open area was not modelled correctly. In particular, the terrain of the seabed displays an abundance of seemingly pit-like structures. These must be regarded as artificial features resulting from clutter points below the seabed surface that could not be completely removed due to the filter settings.

The hillshaded terrain model in the lower image (**Figure 6b**) is based on the “open area” parameter-set. It exhibits a well-filtered surface in the non-vegetated areas and under shallow water. Without the disturbing pit-like structures, the hillshade now even shows traces of

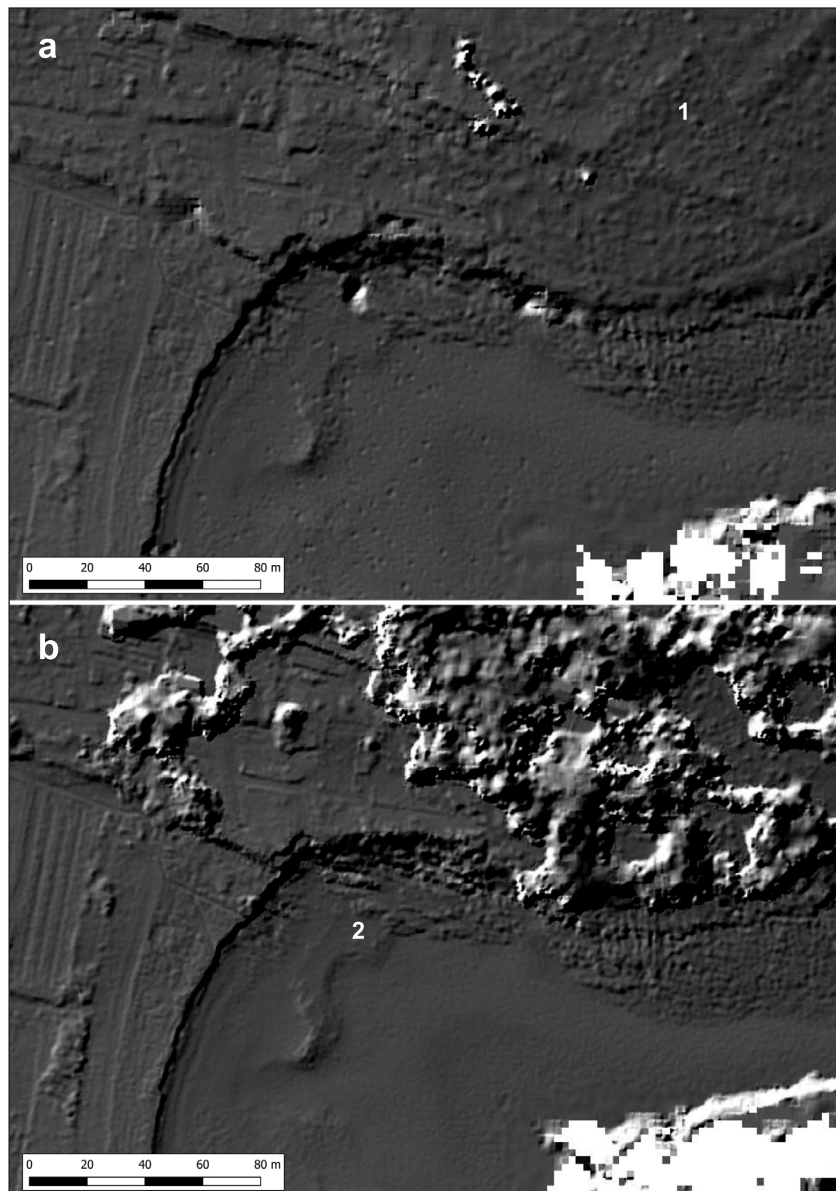


Figure 6: Hillshade of results of ground point filtering with SCOP++. **(a)** Parameter-set “open area” displaying a well-filtered surface in the non-vegetated areas and under shallow water, while dense vegetation is not filtered properly **(b)** parameter-set “dense vegetation” in which the vegetation was correctly removed, while the open area displays an abundance of seemingly pit-like structures. (1) Collapsed wall, (2) submerged wall – see text for further explanation.

submerged walls (**Figure 6b/2**). However, in this case, the dense vegetation was not filtered properly. As a result, the area in the upper half of the image is largely disturbed. The imperfectly filtered vegetation produces a seemingly raised “terrain” that obscures the actual ground surface and all the archaeologically relevant dry-stone walls.

6.1. Merging DTMs

The comparison of both datasets indicates that a single filter setting would not have been a useful approach, as it would have ended in a less than optimal compromise. When comparing both filtered DTMs, it becomes apparent that the well-filtered areas are exclusive in each DTM. Therefore, in order to obtain a single terrain model that would be entirely useful for our archaeological purpose, the most suitable solution was to merge both DTMs.

To identify a ruleset for merging both datasets into a single, archaeologically optimal DTM, both filtered models were analysed together with the orthophotographs and point density maps of the classified ground and vegetation points. During analysis, it became evident that there were two key factors behind an archaeological ground point filtering strategy:

- (1) Vegetation density was the key factor that made two different filter-sets necessary. After visual inspection of the open area DTM with the overlaid vegetation point density map, empirical testing showed that the “open area” parameter set failed to remove vegetation if the vegetation point densities (counting all echoes) exceeded 6 points/m² (**Figure 7**).
- (2) On some occasions, underwater areas were falsely classified as vegetation during classification (**Figure 8**).

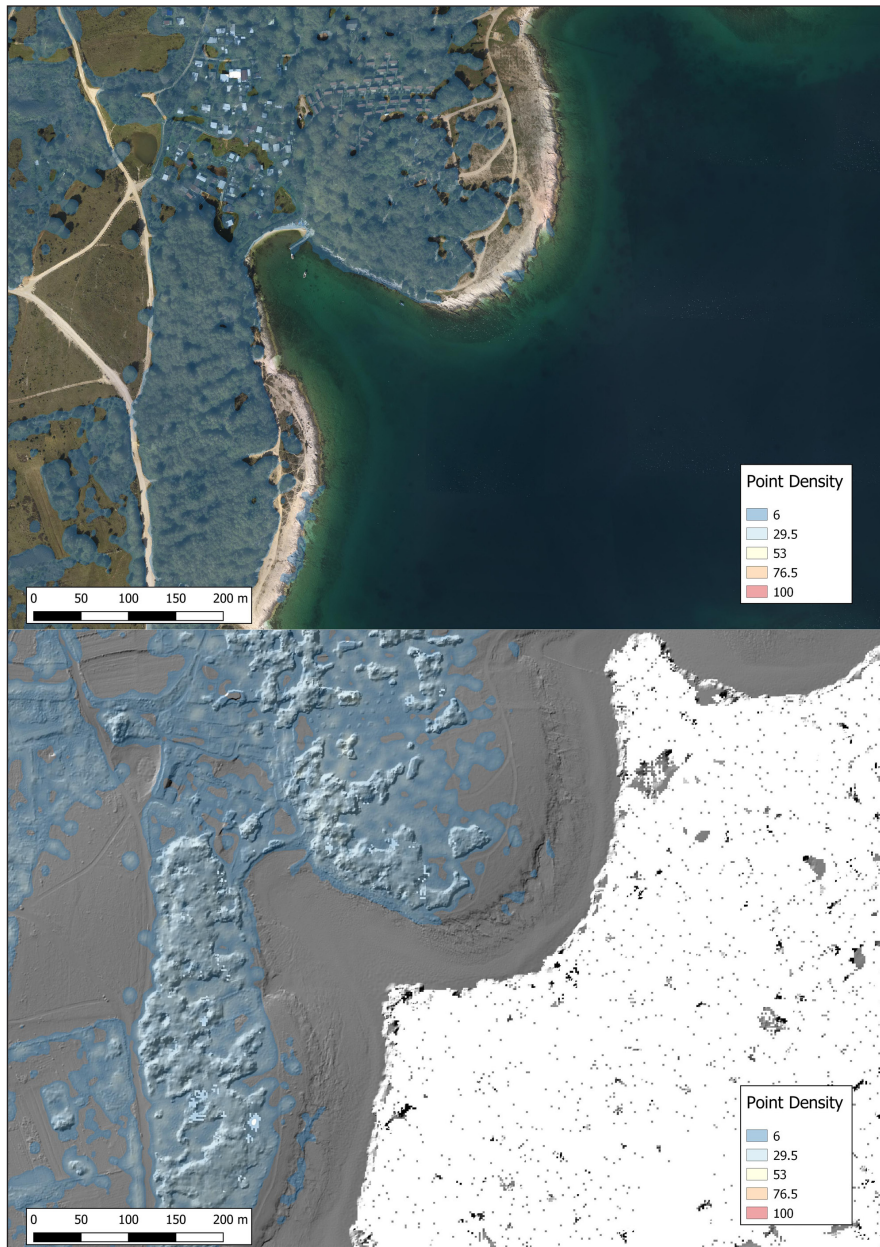


Figure 7: Orthophotograph (upper) and hillshaded DTM of parameter set “open area” (lower) are overlain by a transparent raster of the vegetation point density raster clipped at a value of 6 points per square meter. According to the merging rule explained in the text, cell values falling within a vegetation density of 6 or above (blue area) are taken from the DTM filtered with the “dense vegetation” parameter set.

Therefore, a merging rule was established in which the combined DTM would receive cell-values from the “open area” DTM, where the vegetation density is below six and cell values from the “dense vegetation” DTM where the vegetation density equals or is above six. This threshold value was determined by empirical observation, and it might differ in other parts of the scanned area. Additionally, it was necessary to apply a merging rule, where all cell values exposing a terrain height below the water surface would receive the cell values of the “open area” DTM (Figure 8). All post-processing calculations were done using the raster calculator in QGIS 3.4. The resulting DTM displays the advantages of both the “dense vegetation” and “open area” ground point filtering strategies in a single DTM (Figure 9 compared to Figure 6).

7. Discussion

The merged DTM represents a surface that is an optimal compromise of two ground point filtering processes adapted to different environmental settings: it retains micro-topographic detail and displays an archaeologically interpretable surface, even below areas of dense vegetation (Figure 9). In order to merge DTMs based on a point density layer, the underlying point density maps should be derived for cell sizes of at least 5 m edge length or smoothed with a similar kernel size. The reason for this is that densely vegetated areas often expose small areas of 1×1 or 2×2 m² containing little or no vegetation. Using point density raster maps with a smaller cell size would lead to the effect that, while the cells of densely vegetated areas receive height values from the “dense vegetation”

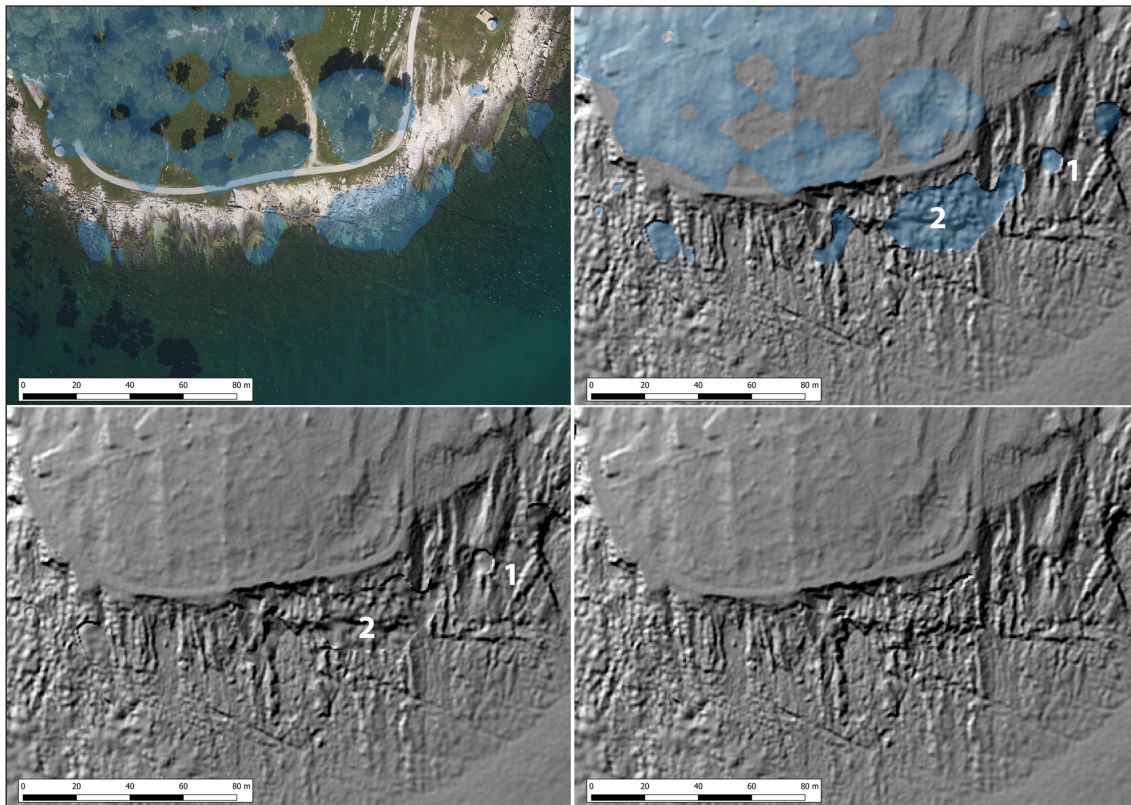


Figure 8: Orthophotograph (a – top left) and hillshaded DTM of parameter set “open area” (b – top right) are overlain by a transparent raster of the vegetation point density raster clipped at a value of 6 (compare with Figure 7). As can be seen, some underwater areas were falsely classified as dense vegetation, which would lead to errors in the merged DTM (c 1 and c 2 – bottom left). Therefore, it was necessary to apply an additional merging rule, where all cell values exposing a terrain height below the water surface would receive the cell values of the “open area” DTM (d – bottom right).

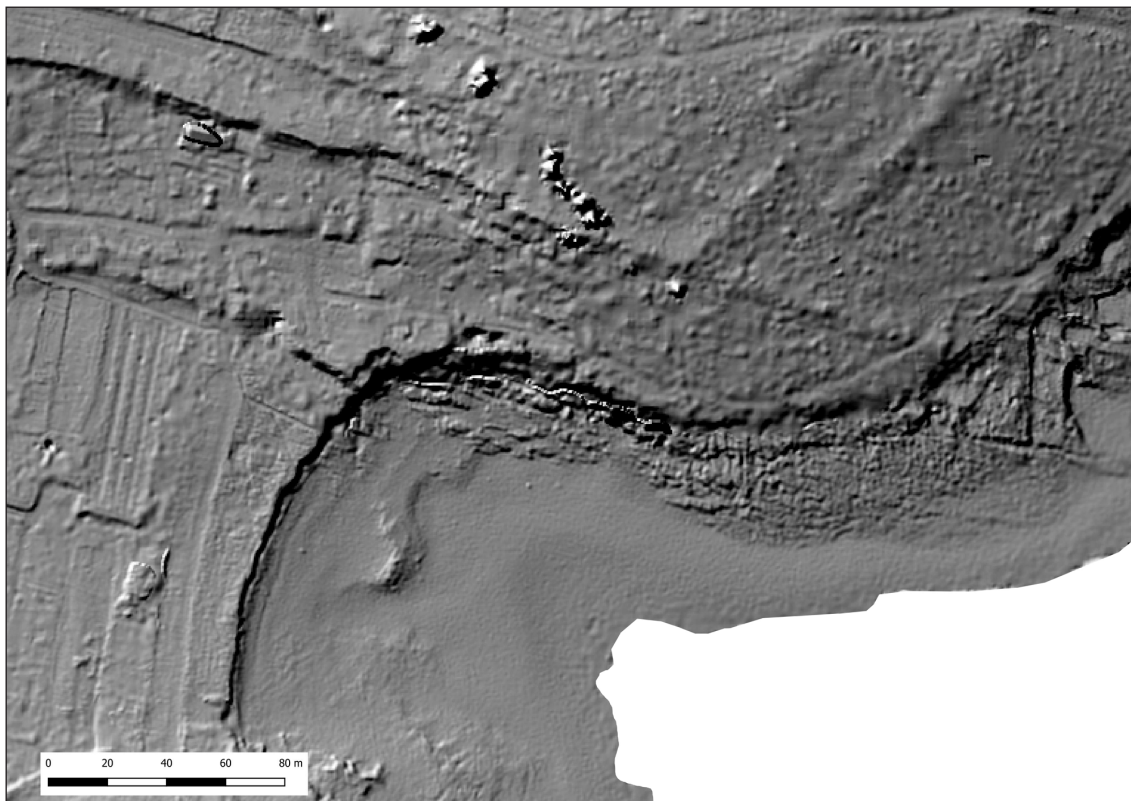


Figure 9: Final merged DTM (see text for explanation).

DTM, some smaller less vegetated patches would receive height values from the “open area” DTM, causing substantial differences (perhaps even a few meters above ground). Consequently, the small patches receive “incorrect” heights (by roughly 0.5 to 3 m) from the surrounding cells, resulting in a combined DTM that is not useable.

The quality of the ground point filtering is also demonstrated in **Figure 10**, which displays an area with extremely dense vegetation (corresponding to the on-site photograph displayed in **Figure 2b**). Still, the filter could classify enough ground points, so that past field boundaries (visible today only as shallow earthworks) become

evident in the resulting hillshade. At the same time, detailed topographic features such as the submerged wall (**Figure 10/1**) are preserved in the merged DTM. Still – as mentioned in Section 2 – we cannot ascertain that the ground was correctly classified in the entire area. Other, maybe less pronounced or non-linear archaeological features still might have been filtered out. Note that the holes in the DTM denote areas where ground points could not be classified for coherent gridding at 0.5 m resolution.

While the merged DTM combines the archaeologically relevant terrain information from (in this case) two filter settings, it still represents only a model of the

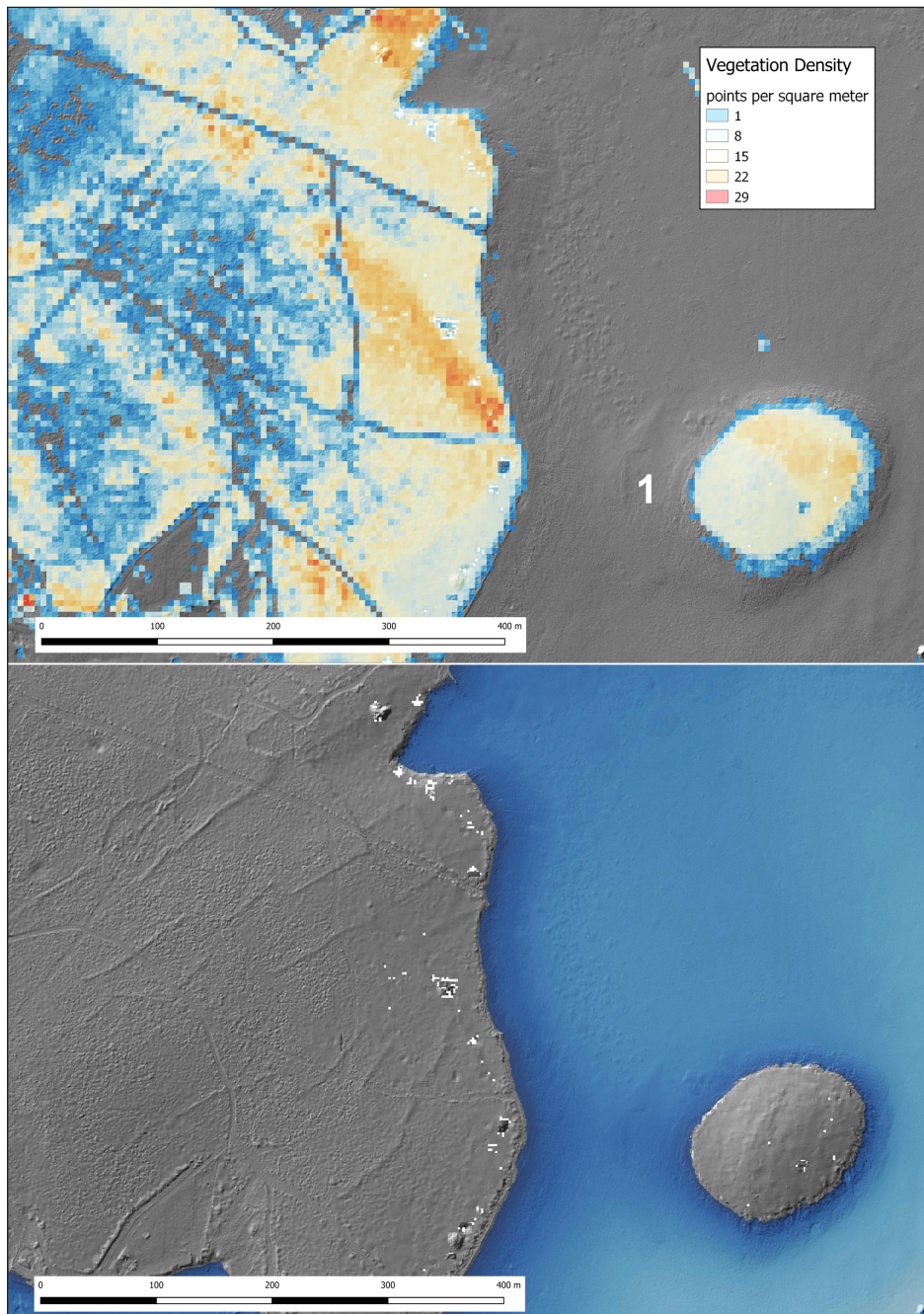


Figure 10: Upper: vegetation density map of a coastal area with extremely dense macchia in the center of the image (compare also with on-site photograph in Figure 2b). Lower: Hillshade of merged DTM. The dense vegetation could be removed by the filter (setting for “dense vegetation”), while the underwater surface (color-coded in blue) shows details of a submerged wall (No. 1 in upper image).

real-world-terrain. Although it can be regarded as an optimized version, accuracy and archaeological detail will still largely depend on the quality of the filtered input, which is affected by the original data (data acquisition, state of vegetation, quality of georeferencing, type and nature of recent and archaeological features) and on the used classification software and its settings.

This is exemplified in **Figure 11**, which displays the remains of Roman buildings submerged 2 m below the water level. The left image is noisy and shows less detail as a result of missing refraction correction and consequently, sub-optimal strip adjustment. On the right side, the same scene has become much clearer after the correction of refraction and subsequent strip-adjustment.

It is important to note that the presented method was tested in a relatively small area of 24 km² comprising a known variety of environmental settings, where the merging ruleset could be adjusted to specific and known parameters. Also, the results could be checked and verified using orthophotographs and on-site visits. Threshold-values were determined through empirical observation. This may be unsuitable for large project (Canuto et al., 2018; Evans et al., 2013) or country-wide datasets containing a wide and unknown variety of environmental settings (Cowley et al., in prep; Fernandez-Diaz and Cohen, in press).

Here, more research needs to be done towards automated adaptive ground point filtering strategies. While the presented approach is simple in the sense that it is based on established and proven standard classification techniques built around clear geometric reasoning, the current trend in semantic labelling of 3D point clouds is indisputably focussed on unsupervised machine learning techniques with a clear trend towards deep learning. Popular techniques like convolutional neural networks (CNNs) outperform most traditional classification techniques, especially in the area of pattern recognition. In the context of classification of 3D ALS point clouds, grid and voxel-based approaches based on 2D- and 3D-CNNs (Zhao et al., 2018; Schmohl & Sörgel, 2019), as well as single

point based networks (Winiwarter et al., 2019), are used. While all these approaches have already proven their competitiveness compared to standard classification techniques, they still need abundant training data. In contrast, ready-to-use DTMs for archaeological mapping purposes can be easily achieved with well-defined and easily comprehensible modifications to existing techniques using the approach presented in this paper.

Nevertheless, for smaller and known project areas, merging various DTMs derived from adapted ground point filters is a viable method to combine the advantage of each into a single terrain model. In this rather simple example, different filter settings were applied due to a difference in vegetation cover. Other scenarios might include varying filter settings induced by other factors, such as topography, tree species, or buildings. Accordingly, merging rules would differ and might include information from slope-maps or thematic maps.

The implications of this approach might be important for rethinking archaeological use of ALS-based datasets. Applying adapted ground point filters and defining rules for merging the results demand to be more explicit on filtering strategies. Consequently, the blackbox of ALS-based DTM generation will open with an improved evaluation of the archaeological usefulness of the resulting datasets. Publishing filter strategies are the only way to guarantee the reproducibility of results, at the same time reducing archaeologist's dependence on serendipity.

Normally, archaeologists will not have the skills and specialised software to process ALS data on their own. Still, to make the best use of such data for archaeological prospection, knowledge (in ideal cases), control of data acquisition, and the processing workflow are of crucial importance. When working with project-based data that were specifically acquired for archaeological purposes, georeferencing and ground point filtering of the dataset should not be left to the data providers without supervision (unless they are specialized in laser scanning for archaeological purposes). Maintaining a close contact,

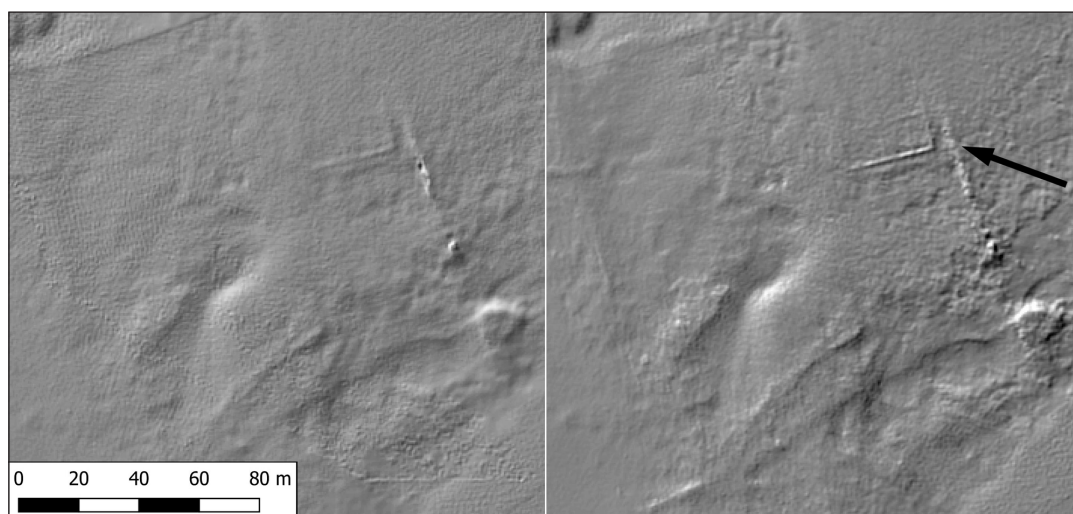


Figure 11: Hillshade of a submerged area. In the upper-right part of the images, the remains of a Roman building (see arrow) are visible in the DTM. The left image shows less detail due to a missing refraction correction and a failed strip adjustment.

including feedback loops, is necessary for a successful archaeological DTM generation. Also, it is recommended to demand for the raw data, i.e. at least the strip-wise recorded echoes from the scanner, the trajectory file(s) and information on scanner-GNSS offsets, as this data is not delivered by default (see also Payne, 2009: chapter 3.1). As methodological and software developments are constantly improved, re-processing and re-filtering might become desirable later. This will, however, require access to the abovementioned raw data (see Fernandez-Diaz and Cohen, in press).

For general purpose data, unfiltered point clouds may be available, and it might be worthwhile to invest time and expenditure into a better ground point filtering that is more archaeologically relevant. Even if this is not possible, knowledge about metadata will help to evaluate the archaeological potential of a given general-purpose dataset and therefore will assure a reasonable and successful application of ALS for archaeology (Payne, 2009; Doneus & Briese, 2011; Opitz, 2013; Grussenmeyer et al., 2016; Boardman & Bryan, 2018).

8. Conclusion

Airborne laser scanning has great potential to reveal archaeological remains in a variety of contexts. This could be demonstrated in the area of the Medulin Bay on the southern extent of Istria, a typical Mediterranean environment with varying degrees of vegetation, including open areas as well as dense and mainly evergreen macchia. Testing various filters to remove the dense vegetation while keeping archaeological detail in less overgrown areas showed that there was no single filter setting that would achieve a satisfying result. As an optimal ground point filtering compromise was unfeasible, two different filters were applied for the same dataset. These resulted in two DTMs that were subsequently merged by rule-based raster calculations and can be regarded as a first step towards an adaptive filtering strategy that might be useful far beyond the field of archaeology.

The merged DTM represents an optimal compromise of two ground point filtering processes adapted to different environmental settings: it keeps micro-topographic detail and displays an archaeologically interpretable surface even below dense vegetation. Therefore, being more explicit on ground point filtering should become best practice (at least for projects which rely on commissioned project-data), as it will enable reproducibility of processing results and make the current blackbox more transparent. This study also stresses the importance of raw data management.

Acknowledgements

The ALS data acquisition was funded by the project “Archaeological park Vižula” (2017–2019), European Regional Development Fund, Competitiveness and Cohesion OP 2014–2020. We would like to thank the Municipality of Medulin for the successful partnership.

The authors also want to express their gratitude for the excellent cooperation and support in the field:

- Igor Miholjek, head of the underwater research in Vižula and his team, Odjel za podvodnu arheologiju, Hrvatski restauratorski zavod/Department for Underwater Archaeology, Croatian Conservation Institute.
- Kristina Džin, Head of the terrestrial research in Vižula and her team, Institut društvenih znanosti Ivo Pilar/Ivo Pilar Institute of Social Sciences.

Gottfried Mandlbürger's contribution to this paper was conducted within the project “Bathymetry by Fusion of Airborne Laser Scanning and Multispectral Aerial Imagery” (SO 935/6-2) funded by the German Research Foundation (DFG).

The authors finally express their thanks to Christopher Sevara (University of Vienna, Department of Prehistoric and Historical Archaeology), for discussion and proof-reading of the text.

Publication funds were provided by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 639828).

Competing Interests

The authors have no competing interests to declare.

Author Contributions

- MD: Outline of research, setting up ground point filtering strategies, general text, images.
- GM: Setting up processing workflow in OPALS, revision of text, feedback.
- ND: Management of data acquisition, archaeological interpretation, section on case study area, revision of text, feedback.

References

- Boardman, C** and **Bryan, P.** 2018. *3D laser scanning for heritage: Advice and guidance on the use of laser scanning in archaeology and architecture*. Third edition. London: Historic England.
- Brodu, N** and **Lague, D.** 2012. 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: 121–134. DOI: <https://doi.org/10.1016/j.isprsjprs.2012.01.006>
- Canuto, MA, Estrada-Belli, F, Garrison, TG, Houston, SD, Acuna, MJ, Kovac, M, Marken, D, Nondedeo, P, Auld-Thomas, L, Castanet, C, Chatelain, D, Chiriboga, CR, Drapela, T, Lieskovsky, T, Tokovinine, A, Velasquez, A, Fernandez-Diaz, JC** and **Shrestha, R.** 2018. Ancient lowland Maya complexity as revealed by airborne laser scanning of northern Guatemala. *Science*, 361(6409): DOI: <https://doi.org/10.1126/science.aau0137>
- Cifani, G, Opitz, RS** and **Stoddard, S.** 2007. Mapping the Ager Faliscus road-system: The contribution of LiDAR (light detection and ranging) survey. *Journal of Roman Archaeology*, 20: 165–176. DOI: <https://doi.org/10.1017/S1047759400005353>

- Cowley, DC, Banaszek, L, Geddes, G and Millican, KM.** In Prep. Making LiGHT work of large area survey? Developing approaches to rapid archaeological mapping and the creation of systematic national-scaled heritage data. *Journal of Computer Applications in Archaeology*, this Special Collection.
- Crow, P, Benham, S, Devereux, BJ and Amable, GS.** 2007. Woodland vegetation and its implications for archaeological survey using LiDAR. *Forestry*, 80: 241–252. DOI: <https://doi.org/10.1093/forestry/cpm018>
- Crutchley, S.** 2010. *The Light Fantastic: Using airborne lidar in archaeological survey*. Swindon: English Heritage Publishing.
- Doneus, M and Briese, C.** 2006. Digital terrain modelling for archaeological interpretation within forested areas using full-waveform laserscanning. In *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2006)*, Ioannides, M, Arnold, D, Niccolucci, F and Mania, K (eds.), 155–162.
- Doneus, M and Briese, C.** 2011. Airborne Laser Scanning in Forested Areas – Potential and Limitations of an Archaeological Prospection Technique. In *Remote Sensing for Archaeological Heritage Management: Proceedings of the 11th EAC Heritage Management Symposium, Reykjavik, Iceland, 25–27 March 2010*, Cowley, D (ed.), 53–76. Budapest: Archaeolingua; EAC.
- Doneus, M, Briese, C, Fera, M and Janner, M.** 2008. Archaeological prospection of forested areas using full-waveform airborne laser scanning. *Journal of Archaeological Science*, 35: 882–893. DOI: <https://doi.org/10.1016/j.jas.2007.06.013>
- Doneus, M, Doneus, N, Briese, C, Pregesbauer, M, Mandlbürger, G and Verhoeven, G.** 2013. Airborne Laser Bathymetry – detecting and recording submerged archaeological sites from the air. *Journal of Archaeological Science*, 40: 2136–2151. DOI: <https://doi.org/10.1016/j.jas.2012.12.021>
- Džin, K and Girardi Jurkić, V.** (eds.) 2008. *Vižula i Burle u antici: Vižula and Burle in Roman period*. Pula: Arheološki Muzej Istre.
- Evans, D, Fletcher, FJ, Pottier, C, Chevance, J-B, Soutif, D, Tan, BS, Im, S, Ea, D, Tin, T, Kim, S, Cromarty, C, De Greef, S, Hanus, K, Baty, P, Kuszinger, R, Shimoda, I and Boornazian, G.** 2013. Uncovering archaeological landscapes at Angkor using lidar. *Proceedings of the National Academy of Sciences*, 110(31): 12595–12600. DOI: <https://doi.org/10.1073/pnas.1306539110>
- Fernandez-Diaz, J, Carter, W, Shrestha, R and Glennie, C.** 2014. Now You See It... Now You Don't: Understanding Airborne Mapping LiDAR Collection and Data Product Generation for Archaeological Research in Mesoamerica. *Remote Sensing*, 6: 9951–10001. DOI: <https://doi.org/10.3390/rs6109951>
- Fernandez-Diaz, JC and Cohen, AS.** In Press. Whose data is it anyway? Lessons in data management and sharing from resurrecting and repurposing lidar data for archaeological research in Honduras. *Journal of Computer Applications in Archaeology*, this Special Collection.
- Forte, M and Campana, S.** (eds.) 2017. *Digital Methods and Remote Sensing in Archaeology: Archaeology in the Age of Sensing*. Cham: Springer International Publishing. DOI: <https://doi.org/10.1007/978-3-319-40658-9>
- Girardi-Jurkić, V.** 2013. Povijesni i gospodarski razvitak južne Istre u antici. In *Monografija općine Medulin*, 44–77. Medulin.
- Girardi Jurkić, V, Džin, K, Paić, A and Ettinger Starčić, Z.** 2012. Vižula kod Medulina. Rezidencijska maritimna vila: Istraživačka kampanja 2011. *Histria antiqua*, 509–523.
- Grussenmeyer, P, Landes, T, Doneus, M and Lerma, JL.** 2016. Basics of Range-Based Modelling Techniques in Cultural Heritage 3D Recording. In *3D recording, documentation and management of cultural heritage*, Stylianidis, E and Remondino, F (eds.), 305–364. Caithness: Whittles Publishing.
- Heinzel, J and Sittler, B.** 2010. LiDAR surveys of ancient landscapes in SW Germany: Assessment of archaeological features under forests and attempts for automatic pattern recognition. In *Space, time, place: Third International Conference on Remote Sensing in Archaeology, 17th–21st August 2009, Tiruchirappalli, Tamil Nadu, Indien*, Forte, M, Campana, S and Liuzza, C (eds.), 113–121. Oxford: Archaeopress.
- Koncani Uhač, I.** (ed.) 2008. *Poluotok uronjen u more. Podmorska arheologija južne Istre u antici: Peninsula imersed in the sea. Underwater archaeology of southern Istria in Roman antiquity*. Pula.
- Kraus, K and Pfeifer, N.** 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53: 193–203. DOI: [https://doi.org/10.1016/S0924-2716\(98\)00009-4](https://doi.org/10.1016/S0924-2716(98)00009-4)
- Lasaponara, R, Coluzzi, R and Masini, N.** 2011. Flights into the past: full-waveform airborne laser scanning data for archaeological investigation. *Journal of Archaeological Science*, 38: 2061–2070. DOI: <https://doi.org/10.1016/j.jas.2010.10.003>
- Lasaponara, R and Masini, N.** 2009. Full-waveform Airborne Laser Scanning for the detection of medieval archaeological microtopographic relief. *Journal of Cultural Heritage*, 10: e78–e82. DOI: <https://doi.org/10.1016/j.culher.2009.10.004>
- Lugmayr, A.** 2013. ALS filtering. <http://lbi-archpro.org/als-filtering/>. Accessed 8/20/2019.
- Mandlbürger, G, Otepka, J, Karel, W, Wagner, W and Pfeifer, N.** 2009. Orientation and Processing of Airborne Laser Scanning data (OPALS) – concept and first results of a comprehensive ALS software. In *ISPRS Workshop Laserscanning '09: Paris, France, September 1–2, 2009*, Bretar, F, Pierrot-Deseilligny, M and Vosselman, G (eds.). Société Française de Photogrammétrie et de Télédétection.

- Miholjek, I.** 2012. Podmorsko istraživanje antičkih ostataka arhitekture na Vižuli – kampanja 2011. *Histria antiqua*, 525–531.
- Opitz, RS.** 2013. An overview of airborne and terrestrial laser scanning in archaeology. In *Interpreting archaeological topography: Airborne laser scanning, 3D data and ground observation*, Opitz, RS and Cowley, D (eds.), 13–31. Oxford: Oxbow Books. DOI: <https://doi.org/10.2307/j.ctvh1dqdz.7>
- Opitz, R and Nuninger, L.** 2014. Point Clouds Segmentation of Mixed Scenes with Archeological Standing Remains: A Multi-Criteria and Multi-Scale Iterative Approach. *International Journal of Heritage in the Digital Era*, 3: 287–304. DOI: <https://doi.org/10.1260/2047-4970.3.2.287>
- Payne, A.** 2009. ADS Guides to Good Practice. Laser Scanning for Archaeology: A Guide to Good Practice#. https://guides.archaeologydataservice.ac.uk/g2gp/LaserScan_Toc. Accessed 1/13/2020.
- Pfeifer, N, Mandlburger, G, Otepka, J and Karel, W.** 2014. OPALS – A framework for Airborne Laser Scanning data analysis. *Computers, Environment and Urban Systems*, 45: 125–136. DOI: <https://doi.org/10.1016/j.compenurbsys.2013.11.002>
- Pfeifer, N, Stadler, P and Briese, C.** 2001. Derivation of digital terrain models in the SCOP++ environment. In *Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Terrain Models*, OEEPE (ed.). Stockholm, Sweden.
- Pfennigbauer, M and Ullrich, A.** 2010. Improving quality of laser scanning data acquisition through calibrated amplitude and pulse deviation measurement. In *Proc.: SPIE 7684, Laser Radar Technology and Applications XV*, Turner, MD and Kamerman, GW (eds.), 76841. DOI: <https://doi.org/10.1117/12.849641>
- Pfennigbauer, M, Wolf, C, Weinkopf, J and Ullrich, A.** 2014. Online waveform processing for demanding target situations. In *Proc.: SPIE 9080, Laser Radar Technology and Applications XIX; and Atmospheric Propagation XI*, Turner, MD, Kamerman, GW, Wasiczko, LM and Spillar, EJ (eds.), 90800J. DOI: <https://doi.org/10.1117/12.2052994>
- Polat, N and Uysal, M.** 2015. Investigating performance of Airborne LiDAR data filtering algorithms for DTM generation. *Measurement*, 63: 61–68. DOI: <https://doi.org/10.1016/j.measurement.2014.12.017>
- Ressl, C, Kager, H and Mandlburger, G.** 2008. Quality checking of ALS projects using statistics of strip differences. *IAPRS*, XXXVII: 253–260.
- Ressl, C, Pfeifer, N and Mandlburger, G.** 2011. Applying 3d affine transformation and least squares matching for airborne laser scanning strips adjustment without gnss/imu trajectory data. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII-5/W12: 67–72. DOI: <https://doi.org/10.5194/isprsarchives-XXXVIII-5-W12-67-2011>
- Schmohl, S and Sörgel, U.** 2019. Submanifold sparse convolutional networks for semantic segmentation of large-scale ALS point clouds. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W5: 77–84. DOI: <https://doi.org/10.5194/isprs-annals-IV-2-W5-77-2019>
- Shan, J and Toth, CK.** 2018. *Topographic laser ranging and scanning: Principles and processing*. Second edition. Boca Raton: CRC Press/Taylor & Francis Group. DOI: <https://doi.org/10.1201/9781315154381>
- Sithole, G and Vosselman, G.** 2003. Comparison of filtering algorithms. In *Proceedings of the ISPRS working group III/3 workshop “3-D reconstruction from airborne laserscanner and InSAR data”*, Maas, H-G, Vosselman, G and Streilein, A (eds.), XXXIV, 3/W13: 71–78. Dresden, Germany.
- Steinbacher, F, Pfennigbauer, M, Aufleger, M and Ullrich, A.** 2012. High Resolution Airborne Shallow Water Mapping. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B1: 55–60. DOI: <https://doi.org/10.5194/isprsarchives-XXXIX-B1-55-2012>
- Vacchi, M, Marriner, N, Morhange, C, Spada, G, Fontana, A and Rovere, A.** 2016. Multiproxy assessment of Holocene relative sea-level changes in the western Mediterranean: Sea-level variability and improvements in the definition of the isostatic signal. *Earth-Science Reviews*, 155: 172–197. DOI: <https://doi.org/10.1016/j.earscirev.2016.02.002>
- Winiwarter, L, Mandlburger, G, Schmohl, S and Pfeifer, N.** 2019. Classification of ALS Point Clouds Using End-to-End Deep Learning. *PGF – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 87: 75–90. DOI: <https://doi.org/10.1007/s41064-019-00073-0>
- Zhao, R, Pang, M and Wang, J.** 2018. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network. *International Journal of Geographical Information Science*, 32: 960–979. DOI: <https://doi.org/10.1080/13658816.2018.1431840>

How to cite this article: Doneus, M, Mandlbürger, G and Doneus, N. 2020. Archaeological Ground Point Filtering of Airborne Laser Scan Derived Point-Clouds in a Difficult Mediterranean Environment. *Journal of Computer Applications in Archaeology*, 3(1), pp. 92–108. DOI: <https://doi.org/10.5334/jcaa.44>

Submitted: 15 October 2019

Accepted: 24 March 2020

Published: 21 April 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[

Journal of Computer Applications in Archaeology, is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 