

Article

A Workflow to Extract the Geometry and Type of Vegetated Landscape Elements from Airborne LiDAR Point Clouds

Ine Rosier , Jan Diels, Ben Somers  and Jos Van Orshoven 

Department of Earth and Environmental Sciences, KU Leuven, Celestijnenlaan 200E-bus 2411, B-3001 Leuven, Belgium; jan.diels@kuleuven.be (J.D.); ben.somers@kuleuven.be (B.S.); jos.vanorshoven@kuleuven.be (J.V.O.)

* Correspondence: ine.rosier@kuleuven.be

Abstract: Rural European landscapes are characterized by a variety of vegetated landscape elements. Although it is often not their main function, they have the potential to affect river discharge and the frequency, extent, depth and duration of floods downstream by creating both hydrological discontinuities and connections across the landscape. Information about the extent to which individual landscape elements and their spatial location affect peak river discharge and flood frequency and severity in agricultural catchments under specific meteorological conditions is limited. This knowledge gap can partly be explained by the lack of exhaustive inventories of the presence, geometry, and hydrological traits of vegetated landscape elements (vLEs), which in turn is due to the lack of appropriate techniques and source data to produce such inventories and keep them up to date. In this paper, a multi-step methodology is proposed to delineate and classify vLEs based on LiDAR point cloud data in three study areas in Flanders, Belgium. We classified the LiDAR point cloud data into the classes ‘vegetated landscape element point’ and ‘other’ using a Random Forest model with an accuracy classification score ranging between 0.92 and 0.97. The landscape element objects were further classified into the classes ‘tree object’ and ‘shrub object’ using a Logistic Regression model with an area-based accuracy ranging between 0.34 and 0.95.

Keywords: landscape elements; agricultural landscapes; LiDAR point clouds; random forest; logistic regression; classification



Citation: Rosier, I.; Diels, J.; Somers, B.; Van Orshoven, J. A Workflow to Extract the Geometry and Type of Vegetated Landscape Elements from Airborne LiDAR Point Clouds. *Remote Sens.* **2021**, *13*, 4031. <https://doi.org/10.3390/rs13204031>

Academic Editor: Dimitrios D. Alexakis

Received: 2 September 2021

Accepted: 4 October 2021

Published: 9 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The European agricultural landscape is characterized by a range of landscape elements, both vegetated (vLE) like hedges and lines of trees and non-vegetated like ditches and sunken roads. vLEs include natural elements growing spontaneously across the landscape, but encompass predominantly vLEs from anthropogenic origin such as hawthorn hedges planted on plot boundaries. vLEs are often linearly shaped and are frequently situated on the border between agricultural fields where they provide a multitude of agroecosystem services [1]. They act as a habitat and movement corridor for animals and plant species and have a key role in the maintenance of biodiversity in intensive agricultural landscapes [2–4]. Furthermore, their typical geometrical arrangement following the edges of agricultural parcels and their association with non-vegetated elements like ditches and fences creates networks of landscape elements [5]. Such networks affect the parcel's and catchment's hydrology since they create both hydrological discontinuities and connections with variable intensity across the landscape. vLEs decrease or enhance runoff concentration and hence affect river peak discharges and the frequency, extent, depth and duration of floods downstream [6–9]. Despite this is an important topic as extensive areas throughout Europe are regularly affected by river flooding, and the prevalence of extreme rainfall events and episodes is expected to increase due to climate change [10], the magnitude and conditionalities of the effects have not been studied extensively. vLEs exert multiple effects on the water balance and transfer in an agricultural catchment. They promote

infiltration because their rooting zone is more porous than the one of the surrounding, often compacted agricultural soils [11,12] and decrease runoff by interception of rainwater, especially in the foliation period [13,14]. Further, by increasing the flow resistance, vLEs can temporarily retain more water, promote its infiltration over longer periods of time and hence reduce the surface runoff. The level of interception and flow resistance is dependent on the foliage density of vLEs [15–17].

In this paper, we make a distinction between two types of vLEs based on their foliage density near the ground surface: tree elements characterized by a leafless trunk and shrub elements having dense foliage down to the ground level (Figure 1). We define tree elements as vegetation with a crown structure and an unbranched trunk for at least 1 m above the ground (Figure 1A). These elements can appear as individual trees in the landscape, as a grove or as a tree line. Shrub elements are here defined as woody vegetation with no defined central trunk that has branches close to the ground (Figure 1B). Shrub elements include both small spherically shaped vegetation types like bushes and elongated types like hedges.

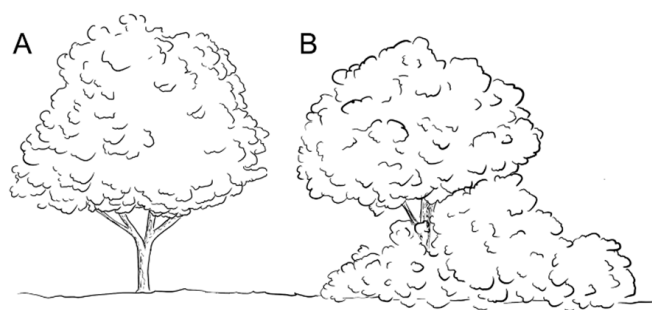


Figure 1. Illustration of a tree element (A) and a shrub element (B).

Few scientific reports are available about the extent to which individual vLEs and their spatial location affect peak river discharge and flood frequency and severity in agricultural catchments under specific meteorological conditions. Apart from the lack of appropriate Rainfall-Runoff models that are able to take into account the presence and characteristics of vLEs in a spatially explicit way, this knowledge gap can be explained by the lack of exhaustive inventories of the presence, geometry, and hydrological traits of vLEs, which in turn is due to the lack of appropriate techniques and source data to produce such inventories and keep them up to date. High-resolution multi-spectral imagery from airborne and spaceborne platforms which became available some two decades ago has been used to map vLEs by visual interpretation of the images backed up by time-consuming field campaigns [18,19]. Later methods were developed to automatically delineate vegetation elements from such spectral imagery [20–23]. The use of high-resolution spectral imagery for the detection of vLEs is however impeded by variations in space and time of reflectance characteristics of the elements and the occurrence of surrounding pixels with similar reflectance values [24]. Further, applications based on only spectral imagery have the disadvantage that they do not provide clear information about the three-dimensional structure of the vegetation elements which is important for hydrological studies as outlined above.

Thanks to programs producing freely available high-resolution airborne light detection and ranging (LiDAR) remote sensing data, a complement or even alternative for multi-spectral imagery is now available. LiDAR data is not only useful to map and characterize the built environment, it also showed a large potential for vegetation-related applications. The X, Y and Z coordinates and the intensity of the returned laser pulse, and the multiple return characteristics can be used to characterize vegetation vertical structure and derive various physical attributes such as leaf area [25,26]. A number of applications of airborne LiDAR data for mapping vLEs already exist. These applications however do not make a distinction between different types of vLEs [27], or generate classifications of vegetation based on their overall height and not on their vegetation density profile [28,29]. Therefore, the usability of such classifications in hydrological studies remains suboptimal.

In this paper, we describe, illustrate and evaluate an automated workflow to extract the geometric characteristics of shrub and tree-based types of vLE from airborne high-resolution LiDAR point clouds with a view to study in follow-on research not reported here, the impact of the vLE on the runoff concentration and the river discharge in agricultural catchments. It was hypothesized that the spatial and structural information derived from the LiDAR data would result in an effective identification and classification of vLEs.

2. Study Area and Data

2.1. Study Area

Three study areas were selected in Flanders, Belgium, all dominated by agricultural land use and with a total surface area of 385.92 ha (Figure 2). All three study areas show a variation in land use types, vLE density and vLE types. The first study area (SA1) is the Heulegracht catchment situated in the municipalities Gingelom and Sint-Truiden and has a surface area of 244.15 ha. In SA1, 69% is under cropland, 21% under fruit cultivation and 5% under agricultural grassland [30]. The land use of the remaining 5% consists out of forest, roads, private gardens and buildings, and flood retention installations. The area is characterized by a network of vegetative and non-vegetative landscape elements that were installed starting from 2004 to reduce soil erosion and the area's contribution to muddy floods affecting the downstream settlements [31]. Therefore, this study area is mainly characterized by rather young vLEs with a less dense vegetation cover. The second study area (SA2) is located in the municipality Huldenberg and has a surface area of 88.78 ha. In SA2, 56% of the area is under cropland and 31% is under agricultural grassland. The land use of the remaining 13% consists out of forest, roads, and private gardens and buildings. The third study area (SA3) is located in the municipality Tervuren and has a surface area of 52.99 ha. In SA3, 71% of the area is under cropland and 7% is under agricultural grassland. The land use of the remaining 22% of the area consists out of forest, roads, and private gardens and buildings. SA2 and SA3 are both part of the Langegracht catchment situated on the territory of four municipalities: Overijse, Tervuren, Bertem, and Huldenberg.

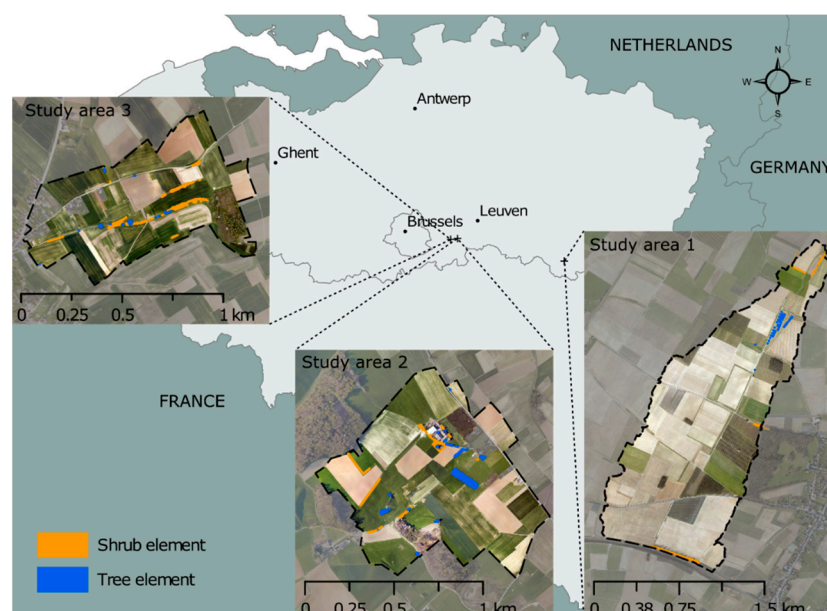


Figure 2. Location of the three study areas within the Flanders region, Belgium. The orange and blue polygons on the inset orthophotographs represent the shrub and tree elements measured in the field.

2.2. Reference Data

Reference data were collected in July and August 2019. The position and geometric characteristics of all tree and shrub elements in SA1, SA2 and SA3 were recorded to assemble a full vLE inventory. The recording was done by using a real-time kinematic (RTK)

GNSS device (Septentrio Altus APS3G) with a horizontal accuracy of 0.006 m (+0.5 ppm) and a vertical accuracy of 0.01 m (+1 ppm) [32] (Table 1). SA1 and SA3 are dominated by shrub elements while SA2 is dominated by tree elements. vLEs are more sparsely distributed over SA1 compared to SA2 and SA3. The positional data were differentially corrected by making use of the reference station network of the Flemish Positioning Service (FLEPOS) [33]. The set of recorded point locations bounding the vLEs were converted into polygons representing the vLEs (Figure 2). Finally, the geometry of the polygons was adapted by visually inspecting RGB imagery acquired simultaneously with the LiDAR data. In this way, the time lag between the acquisition of the LiDAR data and the reference data was accounted for.

Table 1. Aggregated surface area of the reference elements in m² and as percentage of the study area (between brackets) inventoried in SA1, SA2 and SA3.

Study Area	Shrub Elements (m ²)	Tree Elements (m ²)
SA1	2196 (0.09)	229 (0.01)
SA2	3774 (0.43)	8943 (1.01)
SA3	7180 (1.36)	601 (0.11)

2.3. Airborne LiDAR and Ancillary Data

Airborne LiDAR data was collected between 4 February 2015 and 13 March 2015 in SA1, between 30 March 2014 and 5 April 2014 in SA2 and between 6 April 2014 and 12 April 2014 in SA3. The data collection was part of the DHMVII LiDAR campaign over Flanders that was carried out by Flanders Information Agency (AIV) in the winter seasons, i.e., between 15 November and 15 April of 2013, 2014 and 2015. The flight strips have a minimal forward and sideward overlap of 50% and the average point density of the final point cloud is 16 points per square meter.

The LiDAR data is freely available through the EODaS Open LiDAR portal in a LAS 1.4 format and contains for each point the X, Y, and Z coordinates, the returned intensity, the RGB values, a classification label, information on the laser pulse return number and additional scanning parameters [34]. The LAS files are provided in the standard Belgian Lambert 72 projected coordinate reference system. The Z-values are referenced to the Second General Levelling (Tweede Algemene Waterpassing, TAW), i.e., the reference for vertical height measurements in Belgium expressed in meters. The point cloud has an XY accuracy of 0.10 m and a Z accuracy of 0.05 m [35]. The laser scanner that was used can record up to five returns from each emitted laser pulse. Simultaneous acquisition of digital aerial RGB imagery with a spatial resolution of 10 cm provided RGB values for the LiDAR points.

Three ancillary datasets were used for classification of the point data: a vector dataset representing the roads within the study area, part of the High Resolution Reference Database for Flanders (Grootschalig referentiebestand, GRB) [36], a 1 m resolution Digital Elevation Model (DEM) derived from the LiDAR data [37], and a vector dataset representing the agriculture parcels and their boundaries in 2015 for SA1, and in 2014 for SA2 and SA3 [38,39].

3. Methodology

The workflow that was developed and applied to extract the geometric characteristics of shrub and tree-based types of vLE from airborne high-resolution LiDAR point clouds consists of seven steps (Figure 3): (1) preprocessing of the LiDAR point data, (2) characterisation of each of the LiDAR points by a set of features extracted from the neighbouring LiDAR points, (3) classification of the LiDAR points into one of two classes: vLE and other non-ground points, (4) clustering of the vLE points, (5) segmentation of the vLE points into 2D objects, (6) characterisation of the 2D objects with features extracted from the corresponding set of LiDAR points, and (7) classification of the 2D objects into the classes ‘tree element’ and ‘shrub element’. All data were processed and analyzed using free and

open-source software. Scripting was done in Python (3.7.9), using the libraries NumPy (1.19.1), Pandas (1.1.3), and SciPy (1.5.1).

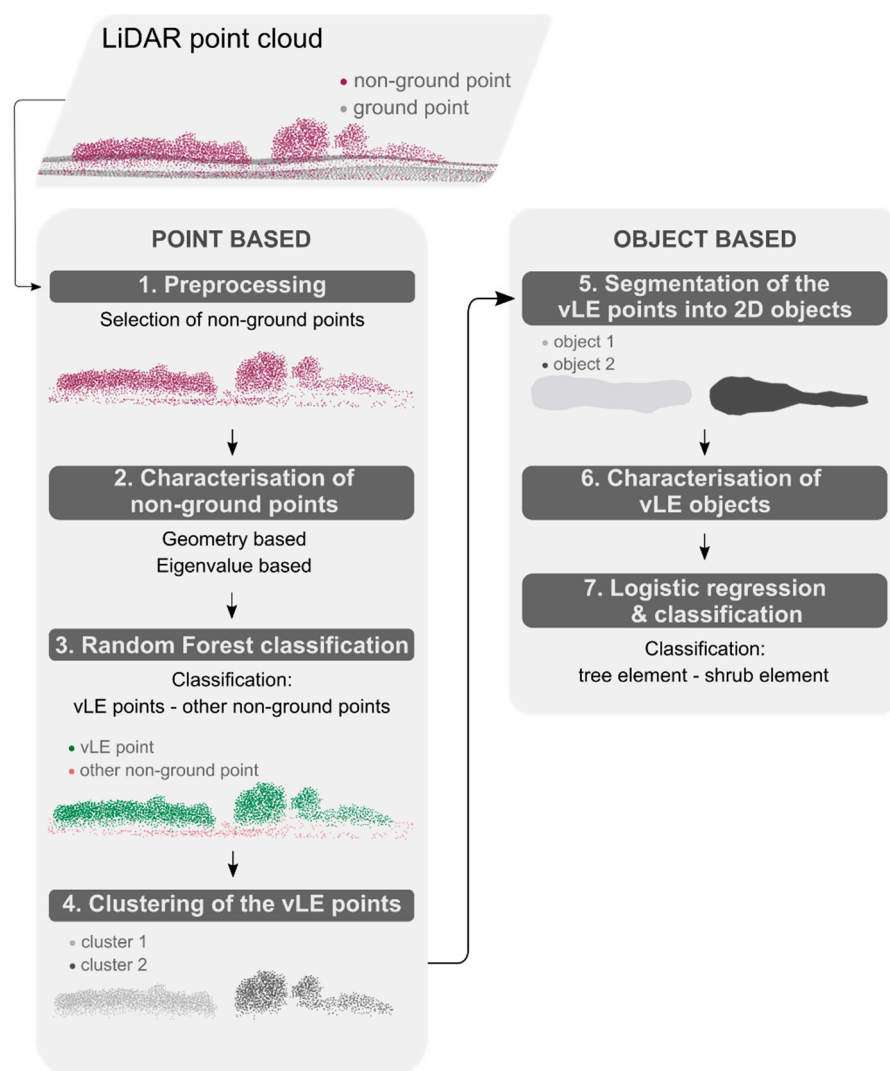


Figure 3. Schematic representation of the workflow applied to classify LiDAR points into vLE objects.

3.1. Preprocessing of the LiDAR Point Data

The LAS points received from the data provider were already classified as ‘ground level’, ‘non-ground level’ or ‘water’ using an iterative algorithm of the Terrascan software [35]. From this classified point cloud, all LiDAR non-ground level points were selected. Points within private gardens were omitted because of the lack of reference data. Private gardens were identified through field visits. Points belonging to wooded areas larger than 500 m² in which the trees are not planted in a linear arrangement were omitted as detection of these areas is out of the scope of this study. These wooded areas were identified from aerial RGB imagery acquired simultaneously with the LiDAR data. Lastly, LiDAR points located within fruit tree orchards were omitted because their typical multi-row planting system could cause confusion with vLEs such as tree rows. Orchards were identified from the vector dataset representing the agriculture parcels and their boundaries in 2015 for SA1, and in 2014 for SA2 and SA3 [38,39].

3.2. LiDAR Point Data Feature Extraction

For each point that is part of the point cloud, the number of returns, the normalized return number (i.e., sequential return number divided by the number of returns), the return intensity, the RGB values, the normalized height value (i.e., the elevation above

ground level), the shortest euclidean distance towards roads, and the euclidean distance towards agricultural parcel boundaries were considered a feature. The normalized height value (h) was determined as the height difference between each raw LiDAR point and the corresponding DEM value extracted from the 1 m resolution DHMVI product [37]. The euclidean distance towards roads and the euclidean distance towards agricultural parcel boundaries were selected because vLEs often appear alongside parcel boundaries and roads [18]. Further, seven geometry-based and eleven eigenvalue based features were calculated (Table 2). This was done by defining for each point p_i its local neighborhood in the 3D space with a certain radius r_l , where $l \in \{1 \text{ m}, 2 \text{ m}, 5 \text{ m}, 10 \text{ m}\}$, and the k points within this neighborhood $\{n_1, n_2, \dots, n_k\}$. The radii were in line with the dimensions of the vLEs effectively present in the study areas.

Based on these neighboring points, seven geometric features were calculated: normalized height difference (Δh), standard deviation of normalized height values (σ_h), the variance of the normalized height values (σ_h^2), the mean of the normalized height values (\bar{h}), the minimum of the normalized height values, the maximum of the normalized height values, and local point density (D). Further, eleven eigenvalue based features were computed. The eigenvalue-based features contain information about the relative spatial arrangement of a LiDAR point and its k neighbors. To do this, the 3D covariance matrix was used to derive the eigenvalues λ_j with $j \in \{1, 2, 3\}$, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ and $\lambda_1 + \lambda_2 + \lambda_3 = 1$ from the 3D coordinates of the points in the k -neighborhood of p_i . These eigenvalues itself were used as features, and they were also used to calculate: anisotropy (A_λ), curvature (C_λ), eigentropy (E_λ), linearity (L_λ), omnivariance (O_λ), planarity (P_λ), sphericity (S_λ) and sum of eigenvalues (Σ_λ). This was done for each of the four radii [40].

Table 2. Geometry and eigenvalue based features calculated for classification of LiDAR point cloud data into vLE classes.

Geometry Based Features			
Feature	Formula		
Normalized height difference	$\Delta h = \max_{i:1 \rightarrow k} h_i - \min_{i:1 \rightarrow k} h_i$		[41]
Normalized height values standard deviation	$\sigma_h = \sqrt{\frac{1}{k} \sum_{i=1}^k (h_i - \bar{h})^2}$		[41]
Normalized height values variance	$\sigma_h^2 = \frac{1}{k} \sum_{i=1}^k (h_i - \bar{h})^2$		[42]
Normalized height values mean	$\bar{h} = \frac{1}{k} \sum_{i=1}^k h_i$		[43]
Normalized height values minimum	$\min_{i:1 \rightarrow k} h_i$		[43]
Normalized height values maximum	$\max_{i:1 \rightarrow k} h_i$		[43]
Local point density	$D = \frac{k}{\frac{4}{3}\pi r_l^3}$		[41]
EigenvalueBased Features			
Feature	Formula		
eigenvalue 1	λ_1		[42]
eigenvalue 2	λ_2		[42]
eigenvalue 3	λ_3		[42]
Anisotropy	$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}$		[41]
Curvature	$C_\lambda = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$		[41]
Eigentropy	$E_\lambda = - \sum_{i=j}^3 \lambda_j \ln(\lambda_j)$		[41]
Linearity	$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$		[41]
Omnivariance	$O_\lambda = (\lambda_1 + \lambda_2 + \lambda_3)^{1/3}$		[41]
Planarity	$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}$		[41]
Sphericity	$S_\lambda = \frac{\lambda_3}{\lambda_1}$		[41]
Sum of eigenvalues	$\Sigma_\lambda = \sum_{j=1}^3 \lambda_j$		[41]

Finally, the mean and the variance of the RGB values and the mean and the variance of the return intensity within each radius were calculated.

3.3. Classification of Point Data

To classify the LiDAR points into the classes ‘vLE point’ and ‘other non-ground point’, a Random Forest classifier was trained and subjected to 10-fold cross-validation. To train and validate the classifier, a ground truth dataset was created by manually labeling the LiDAR non-ground points as ‘vLE point’ or ‘other non-ground point’ by using CloudCompare (version 2.10.2) [44] (Table 3). This was done by comparing the reference object polygons to the point cloud data to ensure no low vegetation points underneath tree canopies were included in the vLE classes. Training and validation were done using the features of the SA1 and SA2, after which the trained Random Forest model was applied to the features of the points of SA3 in the testing phase. To check for the reproducibility of the method, this was repeated twice, using the features of either SA1 or SA2 for testing and the features of the other two study areas for training and validation.

Table 3. Number of non-ground LiDAR points by class.

Study Area	vLE Points	Other Non-Ground Points
SA1	71161	264800
SA2	336605	308998
SA3	258155	280288

Because the Random Forest algorithm is constructed to minimize the overall error rate of the classification results, the prediction accuracy of the majority class affects the final model more than that of the minority class [45]. This often results in a reduced classification accuracy of undersampled classes. Therefore, sample imbalances between different classes often result in an overall decline in the performance of the Random Forest classification [46]. To account for class imbalances, we used a balanced Random Forest classifier implemented in the open-source Scikit-learn Python module [47].

The efficiency and accuracy of the Random Forest classification results were evaluated by calculating the recall (completeness), precision (correctness) and the overall accuracy of the classified points from the confusion matrix. These evaluation metrics were defined as:

$$\text{recall} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{overall accuracy} = \frac{TP}{TP + TN + FP + FN} \quad (3)$$

where TP, TN, FP and FN represent respectively the number of true positives, true negatives, false positives and false negatives in the confusion matrix.

3.4. vLE Object Segmentation

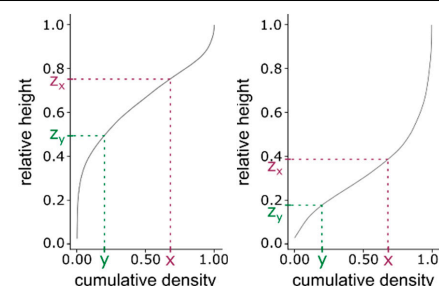
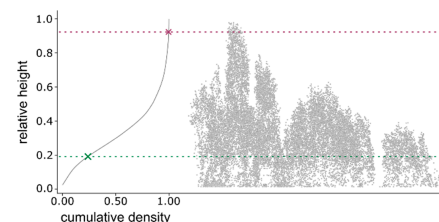
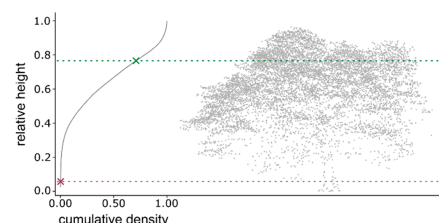
The LiDAR points classified as vLE points by the Random Forest classifier in the testing phase were used to create spatial clusters by means of the density-based clustering algorithm DBSCAN [48]. The DBSCAN algorithm requires two input parameters, i.e., the minimum number of samples in a point cluster and the epsilon parameter. The minimum number of samples (i.e., points) in a point cluster parameter was set at 15 for SA1 and at 62 for SA2 and SA3. These values were based on the number of points per vLE in our study areas. The epsilon value represents the maximum distance between two points above which they cannot belong to the same cluster. The epsilon value was automatically estimated based on the ‘knee’ (point of maximum curvature) in the k-distance plot, in which k equals the minimum samples in a point cluster [49,50]. The clustered points were then converted into 2D objects (i.e., the modelled 2D objects) by applying the alpha-shape algorithm [51] per cluster of points, with the index α set to 0.10. This algorithm describes the shape of a finite point set, with the α controlling the desired level of details.

3.5. Extraction of Features for the vLE Objects

For all 2D vLE objects, a number of features were calculated (Table 4): the minimum, maximum and mean normalized height values (h) of the LiDAR points within the segment, the mean number of returns of the LiDAR points within the segment, the surface area of the object and the 3D point density. The 3D point density was calculated as the number of LiDAR points within the segment, divided by the area of the segment multiplied by the range of normalized height values. Further, a number of features based on the LiDAR point density distribution within a segment were calculated for each segment. To do this, a cumulative point density plot was calculated for each segment in which the cumulative point density was expressed as a function of the normalized height within the segment. The relative height at which the slope of the cumulative density plot was minimal and maximal were both used as features. Further, the normalized height at which the cumulative density equaled x , divided by the normalized height at which the cumulative density equaled y , was calculated for the combinations $x = 0.5$ and $y = 0.1$, $x = 0.7$ and $y = 0.1$, $x = 0.7$ and $y = 0.2$, $x = 1.0$ and $y = 0.2$ and $x = 1.0$ and $y = 0.1$.

Table 4. Candidate features for the classification of the 2D vLE objects and justification for their consideration.

Predictor Variable	Justification
Minimum normalized height (h_{\min})	Shrub elements likely contain LiDAR points closer to the ground surface.
Maximum normalized height (h_{\max})	Shrub elements are likely smaller.
Mean normalized height (h_{mean})	Tree elements likely have a zone of low LiDAR point density near the ground surface where the trunk is located and thus a higher mean normalized elevation.
Mean number of returns (n_{mean})	Shrub elements likely have a higher mean number of returns because of their larger foliage density.
Area	Shrub elements likely have a higher surface area.
3D point density (P_{Dens})	Shrub elements have dense foliage down to the ground level while tree elements are characterized by a leafless trunk. Shrub elements thus likely have a higher 3D point density.
Relative height at highest point density ($z_{P_{\max}}$)	The slope of the cumulative density curve is minimal (dark green) at the relative height where the point density is maximal. This relative height is more likely to be higher for tree elements.
Relative height at lowest point density ($z_{P_{\min}}$)	The slope of the cumulative density curve is maximal (dark red) at the relative height where the point density is minimal. This relative height is more likely to be lower for tree elements.
Elevation ratio (z_x/z_y)	The LiDAR point density is likely distributed differently over the elevation range of shrub elements compared to tree elements. This is indicated by the ratio of z_x/z_y where z_x is the elevation where the cumulative density equals x (dark red) and z_y is the elevation where the cumulative density equals y (dark green).



3.6. Classification of the vLE Objects

A Logistic Regression model was trained to calculate the probability of the 2D objects being either a shrub element or not, in which case the vLE object was classified as a tree element. A probability threshold value of 0.5 was applied to distinguish the two classes. The model was trained and validated by means of 10-fold cross-validation applied on reference 2D vLE objects of SA1 and SA2 that were created by applying the alpha-shape algorithm on the ground-truth dataset. Data normalization was performed using the sklearn StandardScaler on the training data. The StandardScaler standardizes the data by removing from each sample the mean of all training samples scaling to unit variance [52]. The correlation matrix was calculated. For pairs of features for which a correlation coefficient above 95% was calculated, one of both features was dropped. To reduce the set of candidate predictor variables, the best features were selected according to the Select Percentile feature selection model [53] with the percentile equal to 70 (SP70) and the recursive feature elimination (RFE) feature selection model [54]. The Logistic Regression model was then applied to the modelled 2D objects of SA3 to classify them as a tree element or a shrub element. The predictors were scaled by applying the sklearn StandardScaler. The scaler was trained on the reference objects. This was repeated twice, using the modelled 2D objects of either SA1 or SA2 as testing set and the reference objects of the other two study areas for training and validation.

The efficiency and accuracy of the performance of the Logistic Regression classifier were evaluated by considering the overlapping parts of the reference 2D vLE objects and the modelled 2D objects. The recall (completeness), precision (correctness) and the overall accuracy of the classified objects were calculated from the confusion matrix. These evaluation metrics were defined as:

$$\text{recall} = \frac{\text{area}_{\text{TP}}}{\text{area}_{\text{TP}} + \text{area}_{\text{FN}}} \quad (4)$$

$$\text{precision} = \frac{\text{area}_{\text{TP}}}{\text{area}_{\text{TP}} + \text{area}_{\text{FP}}} \quad (5)$$

$$\text{overall accuracy} = \frac{\text{area}_{\text{TP}}}{\text{area}_{\text{TP}} + \text{area}_{\text{TN}} + \text{area}_{\text{FP}} + \text{area}_{\text{FN}}} \quad (6)$$

in which area_{TP} , area_{FP} , area_{FN} represent respectively the true positive, false positive and false negative surface areas of the classified objects in the confusion matrix.

The overall performance of the procedure was evaluated by also taking into account the non-overlapping parts of the reference 2D vLE objects and the modelled 2D objects when calculating the recall, precision, and overall accuracy. This potentially increases area_{FP} and area_{FN} .

4. Results

4.1. Point Classification

The efficiency and accuracy of the point classification of non-ground LiDAR points as 'vLE point' or 'other non-ground point' by using a Random Forest classifier were evaluated in the testing phase by calculating the recall, precision and overall accuracy from the confusion matrix (Table 5). An overall accuracy between 0.92 and 0.97 was calculated in the testing phase over the three study areas. A recall of the vLEs between 0.86 and 0.92 and a precision of the vLE class between 0.94 and 1.00 was calculated. The classification results of the 'vLE point' class are shown in Figure 4.

Table 5. Confusion matrix of the Random Forest classification of non-ground points.

Study Area	Precision		Recall		Overall Accuracy
	vLE	Other Non-Ground	vLE	Other Non-Ground	
SA1	0.94	0.98	0.91	0.98	0.97
SA2	0.97	0.87	0.86	0.98	0.92
SA3	1.00	0.94	0.92	1.00	0.96

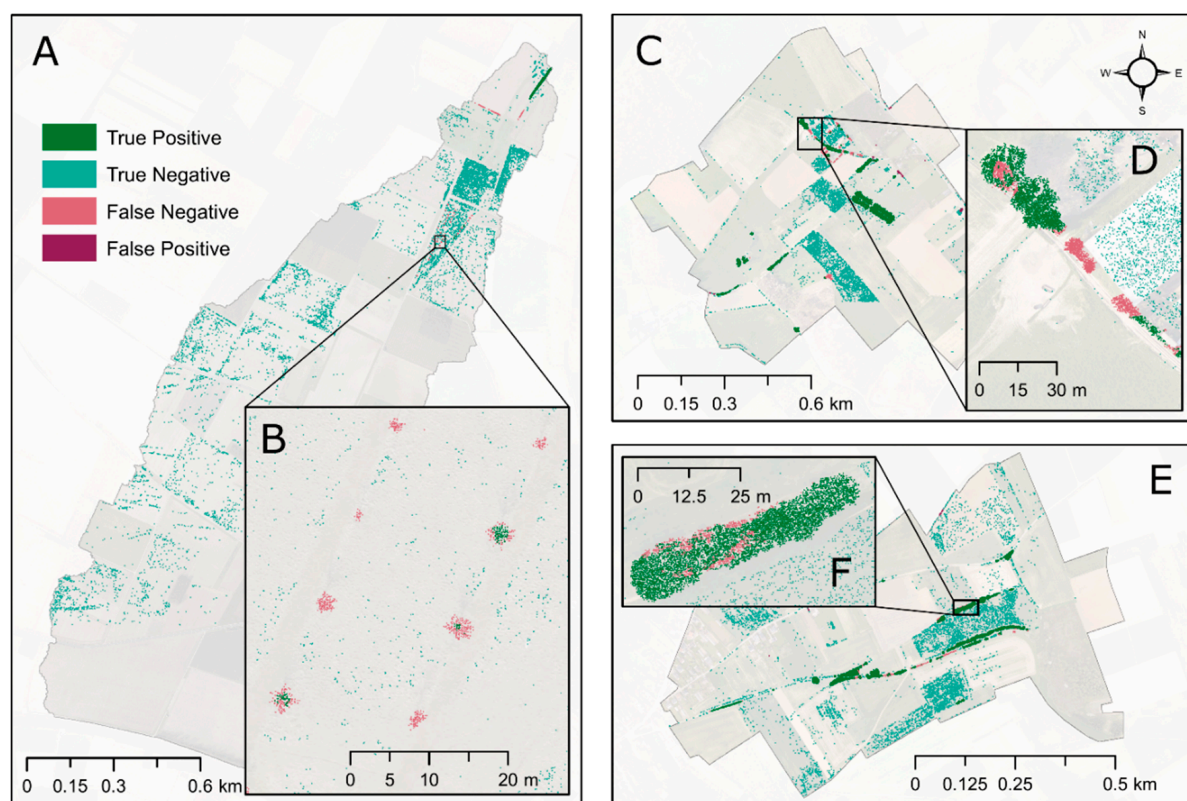


Figure 4. Results of the Random Forest classification of non-ground LiDAR points in SA1 (A,B), SA2 (C,D), and SA3 (E,F). The true positive class (dark green) encompasses the points correctly classified as vLE points. The true negative class (light green) encompasses the points correctly classified as other non-ground points. The false positive class (dark red) encompasses other non-ground points that were falsely classified as vLE points. The false negative class (light red) encompasses other vLE points that were falsely classified as other non-ground points.

4.2. Object Classification

Points that were classified as ‘vLE point’ were clustered by means of the density-based clustering algorithm DBSCAN and segmented into 2D objects by applying the alpha-shape algorithm [48,51]. The resulting modelled 2D objects had a total surface area of 22993.23 m², with 1605.09 m² of vLE objects modelled in SA1, 13899.37 m² modelled in SA2 and 7488.77 m² modelled in SA3. The modelled 2D objects were classified into the classes ‘tree element’ (value 0) and ‘shrub element’ (value 1) by using a Logistic Regression classifier that was trained on a set of reference objects (Table 6). The predictor variables in the Logistic Regression classifier were selected according to the SP₇₀ and according to the RFE feature selection approach.

Table 6. Predictor variables used by the Logistic Regression classifier in SA1, SA2, and SA3. X indicates a predictor variable eliminated due to correlation. Variables are defined in Table 4.

Predictor Variable	Regression Coefficient								
	/	SA1 RFE	SP ₇₀	/	SA2 RFE	SP ₇₀	/	SA3 RFE	SP ₇₀
h_{\min}	−0.12	−0.19	−0.17	−1.07	−1.07	−1.01	−1.20	−1.18	−1.11
h_{\max}	−0.09			0.13			−0.25		
$z_{p\max}$	1.06	1.00	0.95	0.89	0.89	0.76	1.03	0.99	0.93
$z_{p\min}$	−0.32		−0.33	−1.32	−1.45	−1.60	−0.64	−0.64	−0.74

Table 6. Cont.

Predictor Variable	Regression Coefficient								
	/	SA1 RFE	SP ₇₀	/	SA2 RFE	SP ₇₀	/	SA3 RFE	SP ₇₀
Area	0.53	0.50		1.04	0.81	0.77	0.51	0.48	0.51
h _{mean}	−0.68	−0.79	−0.92	−0.87			−0.73	−0.94	−1.05
n _{mean}	0.01			0.44			−0.08		−0.17
P _{Dens}	0.47	0.49		0.51			0.47	0.52	
Z ₅ /z ₁	X	X	X	X	X	X	X	X	X
z ₇ /z ₁	0.78	0.73	0.75	0.16		0.38	0.05		0.15
z ₇ /z ₂	0.59	0.58	0.54	0.42		0.39	−0.02		−0.04
z ₁₀ /z ₁	X	X	X	X	X	X	0.23	0.22	0.14
z ₁₀ /z ₂	0.48	0.69	0.43	1.00	1.33	1.05	0.25	0.23	0.12

The efficiency and accuracy of the Logistic Regression based classification of the vLE objects as either ‘tree element’ or ‘shrub element’ were evaluated by considering the overlapping segments of the reference and modelled 2D objects and by calculating the recall, precision and overall accuracy from the confusion matrix (Table 7). The accuracy assessment showed a precision of between 0.13 and 0.97 for the class ‘Tree element’ and between 0.27 and 1.00 for the class ‘Shrub element’. A recall between 0.10 and 1.00 for the class ‘Tree element’ and between 0.83 and 1.00 for the class ‘Shrub element’ was obtained. The Logistic Regression classifier had an overall accuracy of between 0.32 and 0.95.

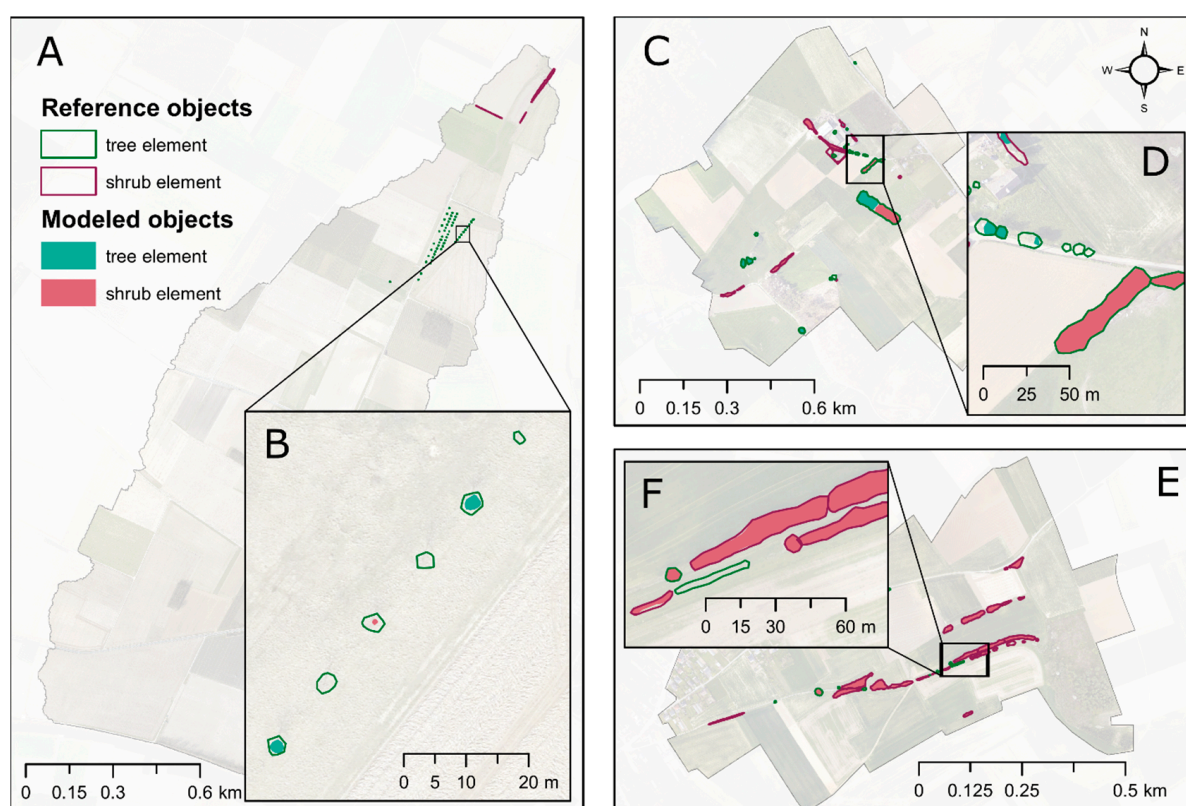
Table 7. Precision, recall, and overall accuracy of the Logistic Regression model.

Study Area	Feature Selection Model	Precision		Recall		Overall Accuracy
		Tree Element	Shrub Element	Tree Element	Shrub Element	
SA1	/	0.13	1.00	0.98	0.83	0.83
	RFE	0.21	1.00	0.98	0.90	0.90
	SP ₇₀	0.14	1.00	1.00	0.83	0.83
SA2	/	0.97	0.38	0.47	0.96	0.60
	RFE	0.89	0.27	0.10	0.96	0.32
	SP ₇₀	0.89	0.27	0.10	0.96	0.32
SA3	/	0.86	0.95	0.28	1.00	0.95
	RFE	0.86	0.95	0.28	1.00	0.95
	SP ₇₀	0.86	0.95	0.28	1.00	0.95

To assess the overall performance of the workflow developed, both the overlapping and non-overlapping parts of the reference and modelled 2D vLE objects were taken into account for calculating the recall, precision, and overall accuracy (Table 8). The accuracy assessment showed a precision between 0.10 and 0.94 for the class ‘Tree element’ and between 0.26 and 0.77 for the class ‘Shrub element’. A recall between 0.09 and 0.43 for the class ‘Tree element’ and between 0.57 and 0.95 for the class ‘Shrub element’ was calculated. Our workflow had an overall accuracy of between 0.26 and 0.73. The reference 2D objects and the modelled 2D objects classified with the Logistic Regression model using all features are shown in Figure 5.

Table 8. Precision, recall, and overall accuracy of the developed workflow.

Study Area	Feature Selection Model	Precision		Recall		Overall Accuracy
		Tree Element	Shrub Element	Tree Element	Shrub Element	
SA1	/	0.10	0.51	0.14	0.57	0.34
	RFE	0.13	0.53	0.14	0.61	0.37
	SP ₇₀	0.14	0.49	1.15	0.57	0.34
SA2	/	0.94	0.37	0.43	0.69	0.49
	RFE	0.78	0.26	0.09	0.69	0.26
	SP ₇₀	0.78	0.26	0.09	0.69	0.26
SA3	/	0.75	0.77	0.23	0.95	0.73
	RFE	0.75	0.77	0.23	0.95	0.73
	SP ₇₀	0.75	0.77	0.23	0.95	0.73

**Figure 5.** Reference 2D objects and modelled 2D objects in SA1 (A,B), SA2 (C,D), and SA3 (E,F).

5. Discussion

5.1. Classification of the LiDAR Points

A Random Forest classifier was used to classify LiDAR ground points either as ‘vLE’ or as ‘other non-ground’. The Random Forest classifier was chosen because it has proven to be successful at classifying LiDAR points as belonging to landscape features [27,40,55]. The Random Forest classifier showed an overall accuracy between 0.92 and 0.97 in the testing phase over the three study areas. This is comparable to the accuracy of 0.97 found by Lucas et al. [27] who used a Random Forest model to distinguish vegetation LiDAR points from other LiDAR points.

A recall of the vLEs between 0.86 and 0.91 was calculated. False negative vLE LiDAR points mainly occurred at the outer ends of a vLE object. A possible explanation for this is the time lag between the recording of the LiDAR data and the ground measurements which could have resulted in reference 2D objects that are larger than they would have been

in case the reference data were recorded simultaneously with the LiDAR data acquiring. Further, the values of the neighborhood features of points at the outer ends of a vLE object are influenced by spatially nearby non-vLE ground points and can therefore be falsely classified as such since the algorithm assumes that the class labels of neighboring points are correlated [41]. Other false negative vLE LiDAR points were observed when few LiDAR data points were located within the vLE reference object. An example of this is a field with 50 young trees in SA1 (Figure 4B). Of these 50 trees, LiDAR points of only ten of these trees were classified as vLE points. The average LiDAR point count of these ten objects is noticeably higher than the average point count of the objects of which the LiDAR points were not correctly classified (212 compared to 77). This suggests that points of objects are classified more correctly when the vegetation object has a more dense point cloud and therefore classification results would improve when the LiDAR data was acquired during the leaf-on period.

A precision of the vLE class between 0.94 and 1.00 was calculated. False positive vLE LiDAR points mainly occurred at field margins. Field margins are typically characterized by a strip of weeds, especially those margins with a fence or terrace slope. Removing LiDAR points in fields with fruit tree orchards increased the precision of the vLE class from 0.10 to 0.94 in SA1 as 15.45% of this area is covered by fruit tree orchards.

5.2. Clustering and Segmentation of the vLE Points

The LiDAR points classified as vLE points by the Random Forest classifier in the testing phase were used to create clusters by means of the density-based clustering algorithm DBSCAN [48]. This algorithm has proven to be successful in clustering LiDAR point data in a number of applications [56–58]. These clusters were subsequently converted into 2D objects (i.e., the modelled objects) by applying the alpha-shape algorithm [51] per cluster of points, with the index α set to 0.10. This resulted in 18, 34, and 12 2D modelled objects for SA1, SA2, and SA3 respectively, while 62, 36, and 40 reference objects were identified in SA1, SA2, and SA3 respectively. In SA1, the difference between the number of modelled and reference objects can be explained by the Random Forest algorithm incorrectly classifying vLE LiDAR points as other non-ground points. When overlaying the modelled and reference objects, 46.43% of the total area of the reference objects of SA1 is missing in the modelled objects. For SA2 and SA3, the surface area of the reference objects missing in the modelled object dataset is only 14.95% and 6.16% respectively. The difference between the number of reference and modelled objects in SA3 is primarily the result of several reference objects being merged into one larger modelled object.

In SA1, the general flow direction goes from southwest to northeast and the modelled 2D objects are mostly positioned along this flow direction. In SA2, the general flow direction goes from northeast to southwest. The modelled 2D objects are position both alongside the flow direction and perpendicular to it. In SA3, the general flow direction goes from west to east. The modelled 2D objects are position both alongside the flow direction and perpendicular to it, creating a natural border between higher positioned agricultural parcels and the stream flow path. The position of the vLE objects along these flow paths indicates that they can be of importance in the catchment's hydrology and that it is meaningful to make a distinction between different types of vLEs that can have a distinct effect on the hydrological cycle. Previously developed applications do not make this distinction between different types of vLEs [27,59].

5.3. Classification of the 2D Objects

A set of features was calculated for each 2D vLE object (Table 4). These features are based on the geometry of the 2D object, the geometric and radiometric characteristics of the LiDAR points within the 2D objects, and the 3D point distribution of the LiDAR points within the 2D object. A Logistic Regression model was trained to calculate the probability of the 2D objects being either a shrub element or not, in which case the vLE object was classified as a tree element. A probability threshold value of 0.5 was applied

to distinguish the two classes. This Logistic Regression model was applied by using all features or a set of features selected by using the SP₇₀ and RFE feature selection models. The selected features varied between the two feature selection models and the three study areas (Table 6). The features Z_{Pmax} and Z_{10}/Z_2 were selected in all three study areas using both feature selection methods while the overall point density (P_{Dens}) was only selected using the RFE feature selection model in SA1 and SA3. This shows the importance of the point distribution in the 3D space when distinguishing between tree and shrub objects. The same is observed when trying to distinguish different tree species [60]. The feature h_{min} was also selected in all three study areas using both feature selection models while h_{max} was omitted in all cases. This indicates there is no clear difference between the height of shrub and tree objects while the height at which foliage starts to grow differs between both object types.

To assess the performance of the classification, the classes of the overlapping segments of the reference and modelled 2D objects were analyzed (Table 7). In SA1, the Logistic Regression classification performed best when only the features selected by the RFE feature selection model were used. However, for all three models, the precision for the 'Tree element' class is low (0.21) meaning a high rate of false positive tree elements was modelled. Because of the large portion of shrub elements in SA1 (90.56% of the total surface area of all reference 2D objects in SA1), this did not translate to a low recall of the 'Shrub element' class or a low overall accuracy. A Logistic Regression model using all features or the features selected using the SP₇₀ feature selection model resulted in a slightly lower overall accuracy (0.83 compared to 0.90). In SA2, the Logistic Regression model performed best when all features were used with an overall accuracy of 0.60. The overall accuracy decreased to 0.32 when features were selected by using the SP₇₀ or RFE feature selection models. In SA2, a low precision for the 'Shrub element' class and a low recall for the 'Tree element' class could be observed when all features were used in the Logistic Regression model (0.38 and 0.47 respectively), meaning a high rate of false positive shrub elements and a high rate of false negative tree elements were modelled. The majority of the reference objects used for training of the tree element class in SA2 (50 out of 64 tree element objects) were young trees that were planted shortly before acquiring the LiDAR data as part of an erosion mitigation project in SA1. As these trees have different characteristics compared to the tree elements located in SA2, this likely influences the classification accuracy of the tree objects in SA2. This finding confirms the importance of a reference dataset that is representative of all objects that are being identified and classified [61]. In SA3, no difference in the performance and accuracy of the Logistic Regression model was found when all or only a selection of the features were used. The overall accuracy of the Logistic Regression models was calculated to be 0.95. The precision for the 'Tree element' class was low (0.28) meaning a high rate of false positive tree elements was modelled in SA3. Because of the large portion of shrub elements in SA3 (92.27% of the total surface area of all reference 2D objects), this did not translate to a low recall of the 'Shrub element' class or a low overall accuracy, as was the case for SA1.

For each Logistic Regression model, the regression coefficients of the predictor variables were calculated (Table 6). The absolute value of a regression coefficient gives an indication of the relative importance of the predictor variable this regression coefficient is linked to. A higher value for predictor variables linked with a negative Logistic Regression coefficient indicates a lower probability of the object being a shrub element, and therefore a higher probability of the object being a tree element. A higher value for a predictor variable linked with a positive Logistic Regression coefficient indicates a higher probability of the object being a shrub element, and therefore a lower probability of the object being a tree element. In general, the sign of the regression coefficients linked with the predictor variables is identical in the three study areas, especially for predictor variables with a high absolute value of the linked regression coefficient. The absolute values of the regression coefficients and therefore the importance of the predictor variables differ between the three study areas, however. In SA1, the predictor variables ' Z_{Pmax} ', ' h_{mean} '

and ' Z_7/Z_1 ' have the highest absolute values. In SA2, the predictor variables ' Z_{Pmin} ', ' h_{min} ', and ' Z_{10}/Z_2 ' have the highest absolute values. In SA3, the predictor variables ' Z_{Pmax} ', ' h_{min} ', and ' h_{mean} ' have the highest absolute values. The differences in importance of the predictor variables between the three study areas can be explained by the differences in the characteristics of the vLE objects in the study areas. In SA1 for example, the vLE objects are less established and are therefore smaller and less dense. The predictor variables ' h_{min} ', ' h_{mean} ', and ' Z_{Pmin} ' are linked with a negative regression coefficient for all Logistic Regression models in the three study areas. The negative correlation coefficients for the predictor variables ' h_{min} ' and ' h_{mean} ' are in line with the expected correlation in Table 4 as shrub elements are expected to have lower minimum and mean normalized height values. The negative correlation coefficient for the predictor variable ' Z_{Pmin} ' is however not in line with the expected correlation in Table 4. It was hypothesized that the relative height at the lowest point density would be lower for tree elements as they are characterized by a leafless trunk and have therefore a low point density near the ground surface. The negative correlation between the predictor variable ' Z_{Pmin} ' and the probability of the object being a shrub element can be explained by the high biomass density of shrub elements. LiDAR pulses are more attenuated at penetrating vLE objects with a high biomass density and therefore the laser beam might not reach the lower portion of the object. This has previously been described in the context of potential errors in LiDAR-derived DEMs [62,63]. This explanation is confirmed by the LiDAR ground point density underneath the reference objects calculated as the average ground point density, which is 13.07 points/m² for shrub elements and 20.78 points/m² for tree elements suggesting the laser pulse is less likely to penetrate the object down to the ground surface for shrub elements.

The predictor variables ' Z_{Pmax} ', ' $area$ ', and ' P_{Dens} ' are linked with a positive regression coefficient for all Logistic Regression models in the three study areas. The positive correlation coefficients for the predictor variables ' $area$ ' and ' P_{Dens} ' are in line with the expected correlation in Table 4 as shrub elements are expected to have a larger area and 3D point density. The positive correlation coefficient for the predictor variable ' Z_{Pmax} ' is however not in line with the expected correlation in Table 4. It was hypothesized that the relative height at the highest point density would be lower for shrub elements since tree elements are characterized by a leafless trunk and have therefore a low point density near the ground surface. The positive correlation coefficient for ' Z_{Pmax} ' could again be explained by the dense biomass of shrub elements that is difficult to penetrate with the laser pulse [62,63]. The sign of the regression coefficient for the predictor variables ' h_{max} ' and ' n_{mean} ' was inconsistent for the three study areas. A negative correlation coefficient for ' h_{max} ' was calculated in SA1 and SA3, which is in line with the expected correlation in Table 4, while a positive correlation coefficient was calculated in SA2. A positive correlation coefficient for ' n_{mean} ' was calculated in SA1 and SA2, which is in line with the expected correlation in Table 4, while a negative correlation coefficient was calculated in SA3.

For the delineation and classification of vLE objects in areas with variations in vLE characteristics, the training dataset should be large enough to be representative of the whole area. Recalibration of the logistic regression model is needed when the model is applied to a new area. This study aids in proving the relevance of making a distinction between different types of vLE objects with unique structural characteristics.

6. Conclusions

We developed a seven step workflow to delineate and characterize vLEs in agricultural landscapes from airborne LiDAR data and applied and evaluated it to three study areas in Flanders, Belgium. Our method performed well when classifying LiDAR points into the classes 'vLE object' and 'other non-ground point' with an overall accuracy between 0.92 and 0.97. The importance of an extensive reference dataset was shown when classifying the modelled 2D objects. Classifying the 2D objects by means of a Logistic Regression model showed an accuracy between 0.32 and 0.95. The overall performance assessment of our workflow showed an accuracy between 0.26 and 0.73. Our workflow is based on

high-resolution LIDAR data which has the advantage that it can be used to characterize the vegetation structure. This information is essential in many applications, including hydrological modelling. Further, we showed that with a limited amount of reference objects, our workflow can be applied in an additional study area. However, classification of the 2D objects should be done in areas with high intra-class homogeneity or the Logistic Regression model should be calibrated per study area in order to classify various vLE types. Our workflow shows that there are differences in the structural composition of the considered vLE types. This proves the importance of considering different vLE types, especially in a hydrological context, as they are often located parallel and perpendicular to flow paths. The developed workflow can be used for supplementing field inventories of vLEs in agricultural landscapes and can aid in keeping existing datasets of landscape elements up-to-date. Additionally, information on the type of vLE can be used in other fields of study like hydrological and ecological modelling. Further, the developed workflow can be adapted to identify other types of vLEs in agricultural landscapes.

Author Contributions: Conceptualization, I.R., J.D., B.S. and J.V.O.; methodology, I.R.; software, I.R.; validation, I.R.; formal analysis, I.R.; writing—original draft preparation, I.R.; writing—review and editing, I.R., J.D., B.S. and J.V.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fonds Wetenschappelijk Onderzoek (FWO), grant number 1SB6819N.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The reference data presented in this study is openly available through https://github.com/IneRosier/vLE_Extraction_from_LiDARPointCloudData accessed on 27 August 2021.

Acknowledgments: The authors would like to thank Ann Cuppens and Rob Hillen for helping with the data collection for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Le Cœur, D.; Baudry, J.; Burel, F.; Thenail, C. Why and how we should study field boundary biodiversity in an agrarian landscape context. *Agric. Ecosyst. Environ.* **2002**, *89*, 23–40. [\[CrossRef\]](#)
2. Favre-Bac, L.; Mony, C.; Ernoult, A.; Burel, F.; Arnaud, J.-F. Ditch network sustains functional connectivity and influences patterns of gene flow in an intensive agricultural landscape. *Heredity* **2016**, *116*, 200–212. [\[CrossRef\]](#)
3. Forman, R.T.; Baudry, J. Hedgerows and hedgerow networks in landscape ecology. *Environ. Manag.* **1984**, *8*, 495–510. [\[CrossRef\]](#)
4. Pelletier-Guittier, C.; Théau, J.; Dupras, J. Use of hedgerows by mammals in an intensive agricultural landscape. *Agric. Ecosyst. Environ.* **2020**, *302*, 107079. [\[CrossRef\]](#)
5. Roelens, J.; Rosier, I.; Dondeyne, S.; Van Orshoven, J.; Diels, J. Extracting drainage networks and their connectivity using LiDAR data. *Hydrol. Processes* **2018**, *32*, 1026–1037. [\[CrossRef\]](#)
6. Horn, R.; Vossbrink, J.; Peth, S.; Becker, S. Impact of modern forest vehicles on soil physical properties. *For. Ecol. Manage.* **2007**, *248*, 56–63. [\[CrossRef\]](#)
7. Klaassen, G.J.; Zwaard, J.J.V.D. Roughness coefficients of vegetated flood plains. *J. Hydraul. Res.* **1974**, *12*, 43–63. [\[CrossRef\]](#)
8. Mérot, P. The influence of hedgerow systems on the hydrology of agricultural catchments in a temperate climate. *Agronomie* **1999**, *19*, 655–669. [\[CrossRef\]](#)
9. Richet, J.-B.; Ouvre, J.-F.; Saunier, M. The role of vegetative barriers such as fascines and dense shrub hedges in catchment management to reduce runoff and erosion effects: Experimental evidence of efficiency, and conditions of use. *Ecol. Eng.* **2017**, *103*, 455–469. [\[CrossRef\]](#)
10. Min, S.-K.; Zhang, X.; Zwiers, F.W.; Hegerl, G.C. Human contribution to more-intense precipitation extremes. *Nature* **2011**, *470*, 378–381. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Rachman, A.; Anderson, S.H.; Gantzer, C.J.; Alberts, E.E. Soil hydraulic properties influenced by stiff-stemmed grass hedge systems. *Soil Sci. Soc. Am. J.* **2004**, *68*, 1386–1393. [\[CrossRef\]](#)
12. Udawatta, R.P.; Anderson, S.H.; Gantzer, C.J.; Garrett, H.E. Agroforestry and grass buffer influence on macropore characteristics. *Soil Sci. Soc. Am. J.* **2006**, *70*, 1763–1773. [\[CrossRef\]](#)
13. Ghazavi, G.; Thomas, Z.; Hamon, Y.; Marie, J.C.; Corson, M.; Merot, P. Hedgerow Impacts on soil-water transfer due to rainfall interception and root-water uptake. *Hydrol. Processes* **2008**, *22*, 4723–4735. [\[CrossRef\]](#)

14. Holden, J.; Grayson, R.P.; Berdeni, D.; Bird, S.; Chapman, P.J.; Edmondson, J.L.; Firbank, L.G.; Helgason, T.; Hodson, M.E.; Hunt, S.F.P.; et al. The role of hedgerows in soil functioning within agricultural landscapes. *Agric. Ecosyst. Environ.* **2019**, *273*, 1–12. [CrossRef]
15. Hill, R.D.; Peart, M.R. Land use, runoff, erosion and their control: A review for Southern China. *Hydrol. Processes* **1998**, *12*, 2029–2042. [CrossRef]
16. Jin, C.-X.; Römkens, M.J.M.; Griffioen, F. Estimating manning's roughness coefficient for shallow overland flow in non-submerged vegetative filter strips. *Trans. ASAE* **2000**, *43*, 1459–1466. [CrossRef]
17. Nunes, A.N.; de Almeida, A.C.; Coelho, C.O.A. Impacts of land use and cover type on runoff and soil erosion in a marginal area of Portugal. *Appl. Geogr.* **2011**, *31*, 687–699. [CrossRef]
18. Barr, C.J.; Gillespie, M.K. Estimating hedgerow length and pattern characteristics in great britain using countryside survey data. *J. Environ. Manage.* **2000**, *60*, 23–32. [CrossRef]
19. Kristensen, S.P. Hedgerow planting activities by danish farmers: A case study from Central Jutland. *Geogr. Tidsskr.* **2001**, *101*, 101–113. [CrossRef]
20. Aksoy, S.; Akcay, H.G.; Wassenaar, T. Automatic mapping of linear woody vegetation features in agricultural landscapes using very high resolution imagery. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 511–522. [CrossRef]
21. Radoux, J.; Chomé, G.; Jacques, D.C.; Waldner, F.; Bellemans, N.; Matton, N.; Lamarche, C.; d'Andrimont, R.; Defourny, P. Sentinel-2's Potential for sub-pixel landscape feature detection. *Remote Sens.* **2016**, *8*, 488. [CrossRef]
22. Sheeren, D.; Bastin, N.; Ouin, A.; Ladet, S.; Balent, G.; Lacombe, J.-P. discriminating small wooded elements in rural landscape from aerial photography: A hybrid pixel/object-based analysis approach. *Int. J. Remote Sens.* **2009**, *30*, 4979–4990. [CrossRef]
23. Tansey, K.; Chambers, I.; Anstee, A.; Denniss, A.; Lamb, A. Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas. *Appl. Geogr.* **2009**, *29*, 145–157. [CrossRef]
24. Lechner, A.M.; Stein, A.; Jones, S.D.; Ferwerda, J.G. Remote sensing of small and linear features: Quantifying the effects of patch size and length, grid position and detectability on land cover mapping. *Remote Sens. Environ.* **2009**, *113*, 2194–2204. [CrossRef]
25. Degerickx, J.; Roberts, D.A.; McFadden, J.P.; Hermy, M.; Somers, B. Urban tree health assessment using airborne hyperspectral and LiDAR imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 26–38. [CrossRef]
26. Zhao, Y.; Hao, Y.; Zhen, Z.; Quan, Y. A region-based hierarchical cross-section analysis for individual tree crown delineation using ALS data. *Remote Sens.* **2017**, *9*, 1084. [CrossRef]
27. Lucas, C.; Bouten, W.; Koma, Z.; Kissling, W.D.; Seijmonsbergen, A.C. Identification of linear vegetation elements in a rural landscape using LiDAR point clouds. *Remote Sens.* **2019**, *11*, 292. [CrossRef]
28. Geerling, G.W.; Labrador-Garcia, M.; Clevers, J.G.P.W.; Ragas, A.M.J.; Smits, A.J.M. Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. *Int. J. Remote Sens.* **2007**, *28*, 4263–4284. [CrossRef]
29. Hou, W.; Walz, U. Extraction of small biotopes and ecotones from multi-temporal rapideye data and a high-resolution normalized digital surface model. *Int. J. Remote Sens.* **2014**, *35*, 7245–7262. [CrossRef]
30. ALV Dataserie Landbouwgebruikspcelen. 2019. Available online: <https://www.geopunt.be/catalogus/datasetfolder/7cc9babc-e021-46bc-abb6-1b74d44b14ea> (accessed on 25 February 2021).
31. Evrard, O.; Persoons, E.; Vandaele, K.; Van Wesemael, B. Effectiveness of erosion mitigation measures to prevent muddy floods: A case study in the Belgian loam belt. *Agric. Ecosyst. Environ.* **2007**, *118*, 149–158. [CrossRef]
32. Septentrio Altus APS3G. Available online: <https://www.septentrio.com/en/products/gnss-receivers/rover-base-receivers/smart-antennas/aps3g> (accessed on 21 February 2020).
33. FLEPOS The Flemish Positioning Service (Flepos). Available online: <https://flepos.vlaanderen.be/> (accessed on 17 August 2021).
34. AIV EODaS Open LiDAR. Available online: <https://remotesensing.vlaanderen.be/apps/openlidar/> (accessed on 17 August 2021).
35. AIV LiDAR Digitaal Hoogtemodel Vlaanderen II—Ruwte Remote Sensing Data. Available online: <https://metadata.vlaanderen.be/srv/dut/catalog.search#/metadata/242ddabc-3cbc-44f0-9623-bee874b29549> (accessed on 21 February 2020).
36. AIV Dataset GRBgis. Available online: <https://www.geopunt.be/catalogus/datasetfolder/7c823055-7bbf-4d62-b55e-f85c30d53162> (accessed on 3 March 2021).
37. AIV Digitaal Hoogtemodel Vlaanderen II, DTM, Raster, 1 m. Available online: <https://download.vlaanderen.be/Producten/Detail/939> (accessed on 30 March 2021).
38. ALV Dataserie Landbouwgebruikspcelen. 2014. Available online: <http://www.geopunt.be/catalogus/datasetfolder/47c5540f-bf7c-45fc-9a74-8e60547cde82> (accessed on 3 March 2021).
39. ALV Dataserie Landbouwgebruikspcelen. 2015. Available online: <https://www.geopunt.be/catalogus/datasetfolder/a5b317f2-1a4c-47df-9b14-bf0cb09de770> (accessed on 3 March 2021).
40. Roelens, J.; Höfle, B.; Dondeyne, S.; Van Orshoven, J.; Diels, J. Drainage ditch extraction from airborne LiDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 409–420. [CrossRef]
41. Weinmann, M.; Jutzi, B.; Hinz, S.; Mallet, C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS J. Photogramm. Remote Sens.* **2015**, *105*, 286–304. [CrossRef]
42. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Contextual classification of Lidar data and building object detection in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 152–165. [CrossRef]
43. Ni, H.; Lin, X.G.; Zhang, J.X. Multiple-primitives hierarchical classification of airborne laser scanning data in urban areas. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *42*, 837–843. [CrossRef]

44. CloudCompare 3D Point Cloud and Mesh Processing Free Software. Available online: <https://www.danielgm.net/cc/> (accessed on 5 March 2021).
45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
46. Chen, C.; Liaw, A.; Breiman, L. *Using Random Forest to Learn. Imbalanced Data*; University California Berkeley: Berkeley, CA, USA, 2004.
47. Scikit Learn Random Forest Classifier. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> (accessed on 17 August 2021).
48. Ester, M.; Kriegel, H.-P.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the KDD-96 Proceedings, international conference on knowledge discovery and data mining, Portland, OR, USA, 2–4 August 1996.
49. Rahmah, N.; Sitanggang, I.S. Determination of optimal epsilon (Eps) value on DBSCAN algorithm to clustering data on peatland hotspots in sumatra. *IOP Conf. Ser.: Earth Environ. Sci.* **2016**, *31*, 012012. [[CrossRef](#)]
50. Schubert, E.; Sander, J.; Ester, M.; Kriegel, H.P.; Xu, X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Trans. Database Syst.* **2017**, *42*, 19. [[CrossRef](#)]
51. Edelsbrunner, H.; Kirkpatrick, D.; Seidel, R. On the shape of a set of points in the plane. *IEEE Trans. Inf. Theory* **1983**, *29*, 551–559. [[CrossRef](#)]
52. Scikit Learn Sklearn Preprocessing StandardScaler. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html> (accessed on 14 June 2021).
53. Scikit Learn SelectPercentile. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectPercentile.html (accessed on 28 August 2021).
54. Scikit Learn Feature Selection RFE. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html (accessed on 28 August 2021).
55. Park, Y.; Guldman, J.-M. Creating 3D city models with building footprints and LIDAR point cloud classification: A machine learning approach. *Comput. Environ. Urban. Syst.* **2019**, *75*, 76–89. [[CrossRef](#)]
56. Ferrara, R.; Virdis, S.G.; Ventura, A.; Ghisu, T.; Duce, P.; Pellizzaro, G. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN. *Agric. For. Meteorol.* **2018**, *262*, 434–444. [[CrossRef](#)]
57. Wang, C.; Ji, M.; Wang, J.; Wen, W.; Li, T.; Sun, Y. An improved DBSCAN method for LiDAR data segmentation with automatic Eps estimation. *Sensors* **2019**, *19*, 172. [[CrossRef](#)] [[PubMed](#)]
58. Wang, X.; Chan, T.O.; Liu, K.; Pan, J.; Luo, M.; Li, W.; Wei, C. A robust segmentation framework for closely packed buildings from airborne LiDAR point clouds. *Int. J. Remote Sens.* **2020**, *41*, 5147–5165. [[CrossRef](#)]
59. Hellesen, T.; Matikainen, L. An object-based approach for mapping shrub and tree cover on grassland habitats by use of LiDAR and CIR orthoimages. *Remote Sens.* **2013**, *5*, 558–583. [[CrossRef](#)]
60. Li, J.; Hu, B.; Noland, T.L. Classification of tree species based on structural features derived from high density LiDAR data. *Agric. For. Meteorol.* **2013**, *171–172*, 104–114. [[CrossRef](#)]
61. Ko, C.; Sohn, G.; Remmel, T.K.; Miller, J.R. Maximizing the diversity of ensemble random forests for tree genera classification using high density LiDAR data. *Remote Sens.* **2016**, *8*, 646. [[CrossRef](#)]
62. Gould, S.B.; Glenn, N.F.; Sankey, T.T.; McNamara, J.P. Influence of a DENSE, LOW-HEIGHT SHRUB SPECIES ON THE ACCURACY of a Lidar-derived DEM. *Photogramm. Eng. Remote Sens.* **2013**, *79*, 421–431. [[CrossRef](#)]
63. Simpson, J.E.; Smith, T.E.L.; Wooster, M.J. Assessment of errors caused by forest vegetation structure in airborne LiDAR-derived DTMs. *Remote Sens.* **2017**, *9*, 1101. [[CrossRef](#)]