Article

Unsupervised Clustering of Multi-Perspective 3D Point Cloud Data in Marshes: A Case Study

Chuyen Nguyen [1]**, Michael J. Starek** [1,2,]*****, Philippe Tissot** [1] **and James Gibeaut** [3]

[1]   Conrad Blucher Institute for Surveying and Science, Texas A&M University-Corpus Christi, Corpus Christi, TX 78412, USA; cnguyen13@islander.tamucc.edu (C.N.); philippe.tissot@tamucc.edu (P.T.)
[2]   School of Engineering and Computing Sciences, Texas A&M University-Corpus Christi, Corpus Christi, TX 78412, USA
[3]   Harte Research Institute, Texas A&M University-Corpus Christi, Corpus Christi, TX 78412, USA; james.gibeaut@tamucc.edu
*   Correspondence: michael.starek@tamucc.edu; Tel.: +1-3618253978

check for updates

Abstract: Dense three-dimensional (3D) point cloud data sets generated by Terrestrial Laser Scanning (TLS) and Unmanned Aircraft System based Structure-from-Motion (UAS-SfM) photogrammetry have different characteristics and provide different representations of the underlying land cover. While there are differences, a common challenge associated with these technologies is how to best take advantage of these large data sets, often several hundred million points, to efficiently extract relevant information. Given their size and complexity, the data sets cannot be efficiently and consistently separated into homogeneous features without the use of automated segmentation algorithms. This research aims to evaluate the performance and generalizability of an unsupervised clustering method, originally developed for segmentation of TLS point cloud data in marshes, by extending it to UAS-SfM point clouds. The combination of two sets of features are extracted from both datasets: "core" features that can be extracted from any 3D point cloud and "sensor specific" features unique to the imaging modality. Comparisons of segmented results based on producer's and user's accuracies allow for identifying the advantages and limitations of each dataset and determining the generalization of the clustering method. The producer's accuracies suggest that UAS-SfM (94.7%) better represents tidal flats, while TLS (99.5%) is slightly more suitable for vegetated areas. The users' accuracies suggest that UAS-SfM outperforms TLS in vegetated areas with 98.6% of those points identified as vegetation actually falling in vegetated areas whereas TLS outperforms UAS-SfM in tidal flat areas with 99.2% user accuracy. Results demonstrate that the clustering method initially developed for TLS point cloud data transfers well to UAS-SfM point cloud data to enable consistent and accurate segmentation of marsh land cover via an unsupervised method.

Keywords: TLS; UAS; structure-from-motion; voxelization; clustering; marsh; K-means

1. Introduction

Given the level of precision required to measure minute changes in marsh elevation over time, survey methods have to be adapted to improve the measurements' accuracy while minimizing the impact on the marsh itself. Spatial characterization of marsh surface elevation change is typically based on intensive field surveys of relatively small areas. Examples of such techniques include a surface elevation table (SET), which can provide a very precise and accurate measurement of sediment elevation (cm to mm scale) in wetlands [1,2]. The limitations are that a SET is a single point source measurement that is tedious to obtain. Alternatively, remote sensing can provide high spatial and temporal resolution for larger areas while greatly reducing the impact on the marsh. Geodetic remote sensing techniques

based on lidar, such as Airborne Lidar Scanning (ALS) or Terrestrial Laser Scanning (TLS), employ active laser ranging to provide a dense sampling of the terrain and land cover. Three-dimensional (3D) point cloud data produced by these various scanning methods will provide a different representation of the underlying terrain and land cover. The resultant point cloud data can be used to geometrically characterize marsh topography and land cover features based on this representation. Numerous studies have demonstrated this potential with airborne lidar data [3–6].

From a ground perspective, TLS can be applied to provide a more precise, accurate, and dense measurement of marsh surface elevation and land cover relative to airborne lidar but at the expense of limited spatial coverage. Two ranging modalities are currently applied for TLS: continuous-wave ranging and time-of-flight ranging. For terrain mapping with TLS, time-of-flight ranging is preferred as it enables measurement at longer distances. This is done by precisely measuring the elapsed time between the emission of a laser pulse and the arrival of the reflection of that pulse to the sensor's receiver [7]. Furthermore, certain TLS systems that employ time-of-flight ranging can perform multiple return echo detection, which may be beneficial for resolving below canopy features. TLS is capable of recording dense topographic datasets and information, with point densities of hundreds to thousands of points per square meter over local spatial extents (e.g., 50–100 hectares or less dependent on the effective range of the scanner and occluding structures). However, the complexity of interactions of the laser pulse with vegetative structures, underlying topography including moist and dry ground, water surface, and other structures can lead to ambiguous information. Furthermore, occlusion of the laser pulse by vegetation and other features is often a nuisance for terrain mapping due to the oblique perspective of TLS [8,9]. Because of the measurement precision and sampling density, TLS has been applied in a variety of studies to monitor marsh elevation and land cover [10–14].

Equipped with miniaturized cameras, Unmanned Aircraft System (UAS) technology provides a new paradigm for aerial surveying in support of studying landform evolution and distinctive features within marsh environments [15–17]. Compared to traditional airborne or satellite remote sensing, UAS provides certain advantages such as rapid deploy capabilities, temporal flexibility, cost reduction at localized geographic scales, and high image resolutions (cm to sub-cm). These imagery data can be processed using Structure-from-Motion (SfM) photogrammetry to generate highly detailed 3D point clouds (e.g., cm to sub-cm average point spacing). SfM is a revolutionary, low-cost, user-friendly photogrammetric technique for three-dimensional surface reconstruction [18]. The SfM method applies a highly redundant bundle adjustment based on matching of features in multiple overlapping images to solve the camera position and scene geometry simultaneously and automatically. The generated point clouds can then be applied to assess marsh environments at a level of spatial detail previously unattainable or not practical using traditional techniques. UAS provides an efficient and convenient platform to implement SfM for 3D mapping of marsh terrain and land cover. In contrast to TLS, SfM implemented with an UAS (UAS-SFM) provides a nadir aerial perspective, which is more beneficial for regularized sampling of land cover and the underlying ground surface. However, SfM is a photogrammetric method that can be susceptible to false parallax stemming from dynamic surfaces, such as water movement or vegetation blowing in the wind, and poor feature matching due to low surface texture resulting in noisy or sparse point clouds over certain terrains [19–21]. Furthermore, UAS-SfM is limited in its ability to measure below canopy compared to active ranging techniques like airborne lidar or TLS, particularly when compared to lidar systems that employ multiple return echo detection. For UAS-SfM to measure the underlying structure within canopy, sufficient sized gaps in the vegetation must be present to enable multi-perspective pixel overlap for 3D reconstruction. This relationship depends on the density and type of vegetation cover, overlap parameters, and ground sample distance of the sensor/camera, which is a function of flying height and sensor characteristics [22].

While there are inherent differences between TLS and UAS-SfM, the challenges shared between these two technologies include high resolution point clouds often resulting in very large and complex point clouds for localized study areas. Extraction of relevant information from such point cloud data

can be quite challenging in natural environments such as wetlands. Given their size and complexity, the point clouds cannot be individually segmented into homogeneous features without the use of intelligent algorithms. To do so, automated methods must be developed and implemented to identify land cover structures in the point clouds prior to the creation of end-products such as digital elevation models (DEMs), fractional vegetation cover, and above ground biomass. In this study, we apply an unsupervised clustering method introduced by the authors in [23] that was developed for TLS point cloud data acquired in wetlands. We adapt it to cluster point clouds acquired by UAS-SfM photogrammetry then evaluate its performance by comparing clustering results for TLS and UAS-SfM surveys of the same marsh scene. The objective is to explore the adaptability and generalizability of the unsupervised clustering method to point cloud data acquired from multiple perspectives (air and ground) and different modalities (UAS-SfM and TLS).

The paper is organized as follows. Section 2 details the study area and datasets. Additionally, Section 2.2.2 provides details on the SfM photogrammetric processing workflow. Section 3 reviews the clustering method introduced in [23] and explains how it is adapted to UAS-SfM point cloud data. Section 4 presents a comparison of clustering performance between TLS and UAS-SfM with emphasis on assessing performance over key land cover features such as tidal flats. Section 5 presents concluding remarks.

2. Study Area and Dataset

2.1. Study Site

The study site is a marsh located on a barrier island along the southern portion of the Texas Gulf Coast, USA, bounded by Corpus Christi Bay, the Laguna Madre, and the Gulf of Mexico called the Mustang Island Wetland Observatory (Figure 1). The study area is located on the bay side of Mustang Island and has a nominal 0.10 m tidal range. It is characterized as a microtidal dominated coast subject to diurnal tides [24]. The width of the barrier island in this region is approximately 2.2 km. Progressing from the Gulf shoreline inland to the Corpus Christi Bay shoreline, upland environment in the region ranges in elevation from approximately 0.52 to 5.49 m (NAVD88); high marsh environment ranges in elevation from 0.2 to 0.8 m; low marsh environment ranges in elevation from −0.1 to 0.3 m; tidal flat environment ranges in elevation from −0.05 to 0.5 m [24]. These elevation ranges overlap because different geo-environments can occur at the same elevation for different locations [23].

The survey area consists of different marsh areas including upland, high marsh, low marsh and tidal flats. The average elevation is highest for upland areas and lowest for low tidal flat areas. The dominant environment of this study area is upland [23]. The dominant vegetation species are *Zchizachyrium littorale* (coastal bluestem) and *Spartina patens* (gulf cordgrass) commonly found growing in mats. The second most prevalent environment of this study area is the tidal flat, which are banks of exposed sediment regularly inundated. Tidal flat surfaces slope gently in elevation from mean high tide water level down toward mean spring tide low water level. High tidal flats are typically only partly submerged at high tide, whereas low tidal flats are located at lower elevation and regularly inundated at high tide. Low, regularly flooded tidal/algal flats are significantly less abundant than high flats in this area. These local tidal flats are designated as wind-tidal flats because flooding occurs mainly due to wind-driven tides [25]. Blue-green algae flourish in low tidal flats after long periods of inundation. In some of the study's tidal flat area, salt marsh vegetation such as *M. litoralis* (shore medick), *Batis maritima* (pickle weed), and *Salicornia spp.* (glasswort) can be found sparingly. Low marsh areas are very high in biologic productivities. More frequently inundated areas near tidal creeks are dominated by *Avicennia germinans* (black mangrove) with *Batis maritima* (pickle weed) following. High marsh environments are the least abundant in the study area. With elevations well above mean high tide, they are rarely inundated. Within the high marsh range, *Monanthochloe litoralis* (shoregrass) and *Salicornia* spp. (glasswort) are the dominant species [24].
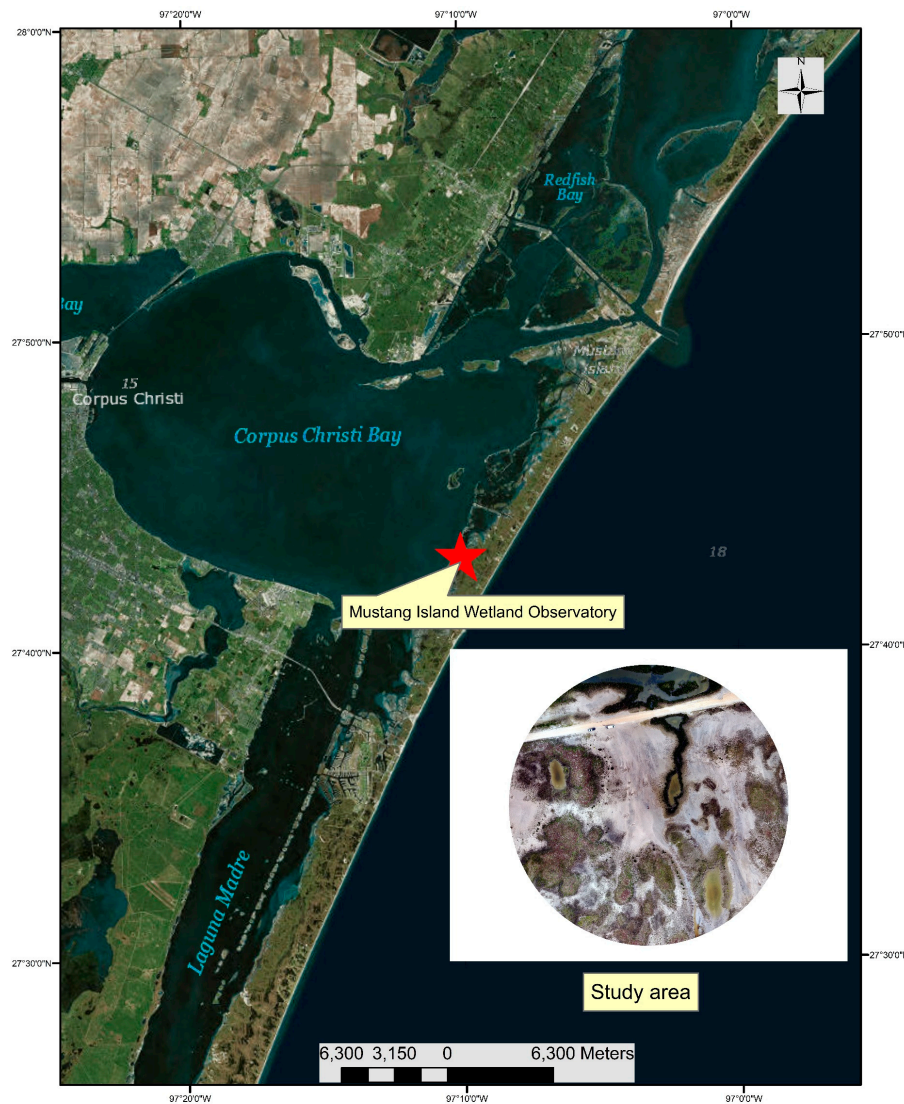
Figure 1. Study area at Mustang Island Wetland Observatory. The circular inlay is the study location imaged from an Unmanned Aircraft System (UAS).

2.2. Dataset

The Conrad Blucher Institute at Texas A&M University-Corpus Christi conducted a TLS survey and UAS survey of the Mustang Island Wetland Observatory on June 9, 2017. Surveys on this date were conducted during a spring low tide to minimize the area of water cover on the marsh surface. To ensure comparability, all data were georefereced in the same coordinate system and co-registered using ground control targets (explained below). Horizontal coordinates were referenced to NAD83 State Plane Texas South meter, while vertical values were referenced to NAVD88 by converting from ellipsoid heights using the GEOID12 model produced by the U.S. National Geodetic Survey [26].

2.2.1. TLS

The TLS survey was conducted with a Riegl VZ-400 scanner. The 1550 nm near-infrared laser pulse used by this system is heavily absorbed by water or typically gets scattered away from the sensor (due to specular reflectance) [27] and hence information cannot be reliably captured from inundated areas. Furthermore, the scanner utilizes online waveform processing to enable multi-return echo detection with up to 10+ echoes per an emitted laser pulse, although such a high number of returns is not expected in a marsh environment. Specifications of the TLS can be found in Table 1.

Table 1. Riegl VZ-400 factory specifications.

Pulse repetition rate	Up to 300,000 kHz
Laser wavelength	1550 nm
Beam divergence	0.3 mrad
Spot size	3 cm at 100 m distance
Range	1.5 m (min), 600 m (max) *
Field of view	100° vertical x 360° horizontal
Repeatability	3 mm (1 sigma @ 100 m range)
Minimum stepping angle	0.0024°

* for natural targets with reflectivity >90%.

The TLS was mounted on a leveling tripod two meters above ground level, and three scans were acquired at three different locations, targeting an approximately 16-hectare area at the study site. Each scan was acquired at a full 360° horizontal field of view using the long-distance ranging mode (600 m average range dependent on surface albedo). The horizontal and vertical stepping angle was set to 20 millidegrees, which resulted in an average point separation of 3.4 cm at 100 m radial distance from the scanner. Six reflector targets (10 cm cylinders) were spread throughout the scan scene and were subsequently used to co-register all three scan locations together and georeference the merged point cloud. All targets were geodetically surveyed with real-time kinematic (RTK) global navigation satellite system (GNSS) positioning using an Altus APS-3 GNSS receiver (Altus Positioning Systems, Torrance, CA, USA). Differential corrections were provided by the Western Data Systems Trimble virtual reference station (VRS) network that offers centimeter-accuracy coordinates. The resultant merged point cloud consisted of points from all three scan locations with a mean density of 2346 points/m^2.

For the study presented in this paper, the merged point cloud data was first clipped at a radial distance of 110 m away from the second scan position (located at the center of the scene) to focus in on areas of higher point density and less scan occlusion from vegetation (Figure 2). Due to the radially decreasing point density, an octree filter was applied to better regularize the point spacing near the scanner in high density areas and reduce point density to save computational time [28]. The point clouds were mapped into voxels, and then the centroid of all points within each voxel were used to extract a point per voxel. The voxel size was set to 2 cm × 2 cm × 1 cm to maintain high spatial detail in the horizontal and vertical components without over redundancy of information [29]. After this process, the TLS point cloud contained 27,109,599 points with a mean density of 776 points/m^2. This voxelization was only applied to filter the point density prior to clustering. It is not related to the multi-scale voxelization procedure for feature extraction described in Section 3.2 below.
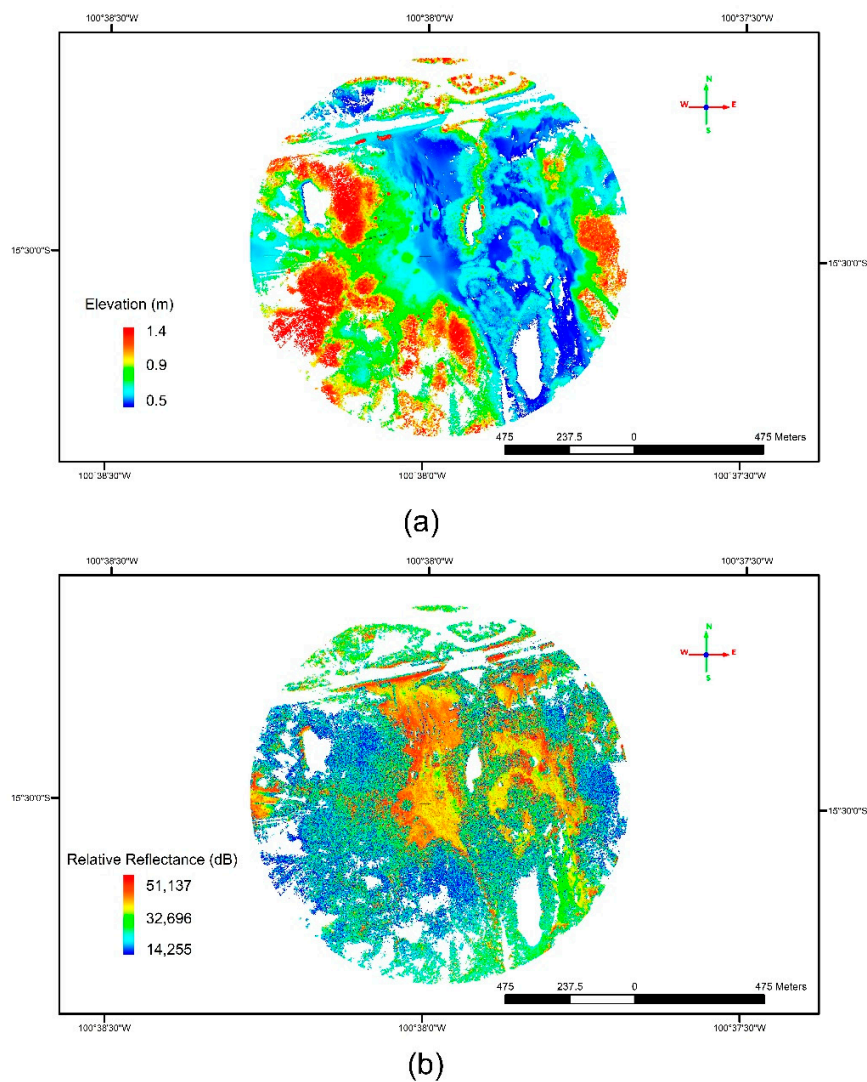
Figure 2. Terrestrial Laser Scanning (TLS) merged point cloud colored by (**a**) elevation and (**b**) relative reflectance.

2.2.2. UAS-SfM

A UAS survey was also conducted at the same time as the TLS survey using a DJI Phantom 4 Pro rotary platform equipped with a 20 MP and 1-inch sensor frame Red-Green-Blue (RGB) digital camera. Imagery was collected at 80% endlap and 80% sidelap flown in a grid pattern with parallel flight lines and a 90-degree (nadir) camera orientation. This high amount of overlap aids the SfM processing explained below. The flight was conducted at an altitude of 50 meters above ground level resulting in an average ground sample distance (GSD) of approximately 1.4 cm. Accurate georeferencing of the acquired imagery is critical for comparison to the TLS point cloud data. Without proper ground control, the raw positional accuracy of the acquired imagery stemming solely from geotagging by the UAS platform's onboard single-channel, non-differential GPS is roughly 1 to 5 m horizontally and can be worse in the vertical. Five aerial ground control targets (1.2 m × 0.8 m planar target with black and white bulls eye pattern) were placed inside the four corners and midpoint of the survey area. The target positions (x, y, z) were surveyed using the same RTK GNSS receiver and differential correction method applied to survey the TLS targets as described above. This enabled horizontal positioning accuracy < 4 cm, which ensured consistent registration with the TLS point cloud data. The Pix4DMapper1.1 software (Pix4D SA, 1015 Lausanne, Switzerland) was used to process the UAS

images with SfM photogrammetry and generate a dense point cloud and orthomosaic of the study site. More details on SfM processing are provided below.

Most small UAS photogrammetric surveys are processed using a technique called Structure-from-Motion (SfM)/Multi-View Stereo (MVS) photogrammetry (or SfM for short). Traditional photogrammetry requires metric cameras precisely calibrated. However, metric cameras are expensive and not conducive for widespread use of UAS for mapping applications. SfM exploits information from multiple overlapping images to extract 3D object information and camera internals negating the need for pre-calibrated metric cameras. SfM derives three-dimensional structure from two-dimensional image sequences through movement of the camera thereby providing different perspective views of the scene. The SfM image processing workflow is summarized as follows [19,20,30]:

- Image sequences are input into the software and a keypoint detection algorithm, such as the scale invariant feature transform (SIFT), is used to automatically extract features and find keypoint correspondences between overlapping images using a keypoint descriptor. SIFT is a well-known computer vision algorithm that allows for feature detection regardless of scale, camera rotations, camera perspectives, and changes in illumination [31].
- A least squares bundle block adjustment is performed to minimize the errors in the correspondences by simultaneously solving for camera interior and exterior orientation. Based on this reconstruction, the matching points are verified and their 3D coordinates calculated to generate a sparse point cloud. Without any additional information, the coordinate system is arbitrary in translation and rotation and has inaccurate scale.
- To further constrain the problem and develop a georectified point cloud, ground control points (GCPs) and/or initial camera positions (e.g., from onboard GNSS) are introduced to constrain the solution. The input GCPs can be used to transform the point coordinates to a real-world coordinate system and to optimize rectification.
- Finally, the interior and exterior orientation for each image are used as input into a MultiView Stereo (MVS) algorithm, which attempts to densify the point cloud by projecting every image pixel, or at a reduced scale. This so called "dense matching" phase can be highly impacted by variations in surface texture as well as the MVS algorithm utilized.

The output from SfM processing is a colorized 3D point cloud. UAS-SfM point clouds are generally considered very high resolution (easily exceeding 1000 pts/m^2) due to the high camera resolution and typical low altitudes at which data are collected. These 3D point clouds can be used to generate a digital surface model (DSM), which can subsequently be used to generate an orthorectified image mosaic (common with SfM software). In contrast to multi-echo detection airborne lidar, UAS-SfM point clouds can be considered single-return or first-surface point clouds as they stem from pixel-to-pixel correspondence. Furthermore, as mentioned in the introduction, SfM methods are image-based and therefore susceptible to false parallax induced from surface movement between overlapping images and prone to poor feature correspondence in scenes of highly uniform texture [19]. As such, land cover reconstruction from UAS-SfM can vary significantly dependent on several factors, most notably the land cover texture itself [32].

The acquired UAS-SfM point cloud data was clipped to the same area as the TLS point cloud (a radial distance of 110 m away from the scanner) (Figure 3). This study's UAS-SfM point cloud contained 38,690,541 points with a mean density of 970 points/m^2. Furthermore, a high resolution orthomosaic (1.4 cm ground sample distance) was generated from the overlapping image sequence. The resultant orthomosaic was used to aid in ground truthing and validate the clustering results.
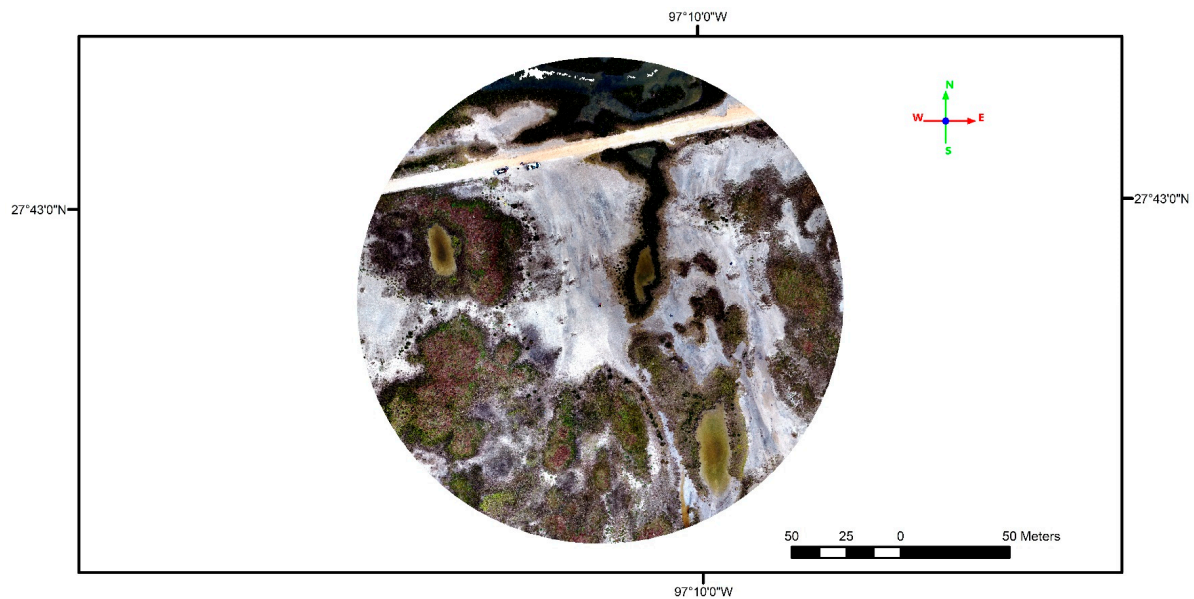
Figure 3. Unmanned Aircraft System based Structure-from-Motion (UAS-SfM) point cloud colorized by the camera's Red-Green-Blue (RGB) brightness values.

3. Method

3.1. Overview

For both point clouds, the TLS and UAS-SfM generated data sets, features are extracted from a combination of point and neighborhood characteristics, the latter generated by statistical measures computed over a voxel. The main idea of the recently developed clustering method for TLS [23] is to apply a multi-scale voxel-based categorization of the 3D point cloud to characterize the complexity and geometry of the terrain. The choice of features is one of the most important factors influencing the performance of a clustering algorithm [33]. Therefore, the first step in adapting the method to UAS-SfM point clouds is to select features representing the complexity and geometry of the terrain well. The set of features include "core" features that can be extracted from any 3D point cloud to describe scene geometry and "sensor specific" features based on per-point recorded attributes unique to the imaging modality. TLS per-point features used include elevation (z), calibrated relative reflectance, and pulse waveform deviation values. In contrast, UAS-SfM per-point features used include elevation (z) and the red, green, and blue pixel brightness values assigned to each point stemming from the UAS's onboard RGB digital camera. These per-point features along with statistical features computed from voxel neighborhoods are then used to generate the feature sets for TLS and UAS. The selection of the features, or feature engineering, is described in Section 3.2.

Another challenge with data clustering is determining the optimal number of clusters (k) that will best segment one's data. The optimization approach implemented in [23] was also followed here for the selection of a good segmentation of the point cloud while using the K-means clustering algorithm. The implementation of the Davies Bouldin algorithm (DB) is further described in Section 3.3. DB values were first computed while segmenting the point cloud in an increasing number of clusters. A discontinuity or change in the graph of DB values vs number of clusters may indicate a natural clustering of the TLS or UAS-SfM point clouds. The optimal number of clusters was selected by combining the DB information computed for each data set with a subjective assessment of the segregated scenes corresponding to the different number of clusters. To ensure a fair evaluation and comparison of the two methods, the same cluster number was used for both datasets.

The algorithms and other computational tasks of this research were developed and implemented using the Matlab 2017b programming software. In particular, the research used the following existing algorithms: octree voxelization, Principal Component Analysis (PCA), DB indexes calculation,

normalization, and K-means. The algorithms designed to compute statistical features from each voxel and to assign these attributes to each point falling inside that voxel were also implemented in Matlab. Both TLS and UAS data were processed on the High Performance Computing (HPC) cluster of Texas A&M University-Corpus Christi. It takes between 24 to 28 hours to process data sets containing about 30 million points. The visual interpretation and quantitative assessments were demonstrated in ArcGIS.

3.2. Feature Engineering

As described above, features are extracted from a combination of point characteristics and multi-scale voxel-based statistics. Firstly, per point features are extracted from both TLS and UAS-SfM point clouds.

- TLS point features are height "z" (elevation), relative reflectance, and waveform deviation. Information about these features are described in [23].
- UAS-SfM point features are height "z" (elevation), Red (R), Green (G) and Blue (B) reflectance values. The R, G, and B pixel brightness values are based on the measurement of reflected radiation from different features on the ground. These values will depend on how the SfM software texturizes the point cloud data using the overlapping pixel values from the onboard digital camera. Although these values are not calibrated reflectance values, spectral response patterns of the land cover across the visual spectrum is captured in the data and this information is expected to be useful for aiding SfM point cloud segmentation of the land cover.

Following the procedure outlined in [23], an octree voxelization algorithm and feature extraction were implemented in the Matlab 2017b programming environment. In this experiment, the number of points in each voxel varied for both the TLS and UAS-SfM data sets. An important difference is that for the TLS data the point density decreases with distance from the instrument location as the voxels' sizes do not change based on the number of points resulting in lower density voxels. Voxels with a number of points below threshold are removed from the data set.

To ensure that TLS and UAS-SfM data are compatible, they are voxelized on the same scales. Their features are generated based on statistical measures computed for two voxel sizes. Voxels could be generated with a variety of aspect ratios. However, in this research we chose to use octree voxelization and preserve the initial point cloud aspect ratio because (1) we want to generate features similar to the method described in [23] to facilitate the comparison of the findings in [23] and this research; (2) octree voxelization with its initial point cloud aspect ratio results in a more computationally efficient process. Similar to [23], both fine and coarse voxel dimensions are also driven by the initial point cloud aspect ratio.

- The large-scale voxel statistical measures help identify the general environment of the voxel. Coarser scale voxels of size 697 × 697 × 7.6 cm were selected to capture broader spatial scale differences between environment types, such as the general location of a voxel within a tidal flat or a generally vegetated area. Tidal flats typically span several meters.
- The finer scale voxels provide information as to finer scale differences that would be averaged out by larger voxels such as differences between types of foliage. The finer scale voxel for this data set, 170 × 170 × 1.9 cm, was selected to match the variability of such parts of the scene and provide the information to the algorithm to potentially differentiate these voxels. For example, salt marsh plants such as *Batis maritima* at the study site are dioecious, perennial sub shrubs with heights in the range of 0.1–1.5 m and a span of 1–2 m for a group of plants [23]. Furthermore, portions of tidal flats will have points concentrated over a thin slice. Additionally, selecting a smaller size for the finer scale voxel would have resulted in less than the required minimum of 10 points imposed for statistical feature extraction for a relatively large number of voxels.

Once voxelization is applied, point cloud statistical features are computed to capture the variability within each voxel characteristic of differing terrain types or other scene features. Firstly, two indexes referred to as curvature 1 and curvature 2 are extracted from both point clouds for each fine and broader scale voxel. These features are computed based on Principal Component Analysis (PCA) describing the general curvature of the point clouds within the voxels. PCA is a common statistical approach used to transform multi-variate data into linearly uncorrelated variables called principal components. Each variance value of the principal components corresponds to its eigenvalue [34]. The PCA technique is applied to the point cloud following the method developed by [34]. The variabilities in the morphometric structures (geometry) are analyzed by applying PCA to the 3D data within each voxel with the resulting eigenvalues derived from the covariance matrix [35]. For each voxel, three eigenvalues, λ_1, λ_2, and λ_3 ($\lambda_1 \geq \lambda_2 \geq \lambda_3$), are derived from the 3D point set of N points, p_i (x_i, y_i, z_i) (i = 1, 2, ..., N). λ_1 quantifies the variance explained by the first dimension, while λ_2 and λ_3 do so for the second and third dimensions respectively. Curvature indexes are derived based on the eigenvalues with $c_i = \lambda_i/(\lambda_1 + \lambda_2 + \lambda_3)$ [36]. The largest curvature value, c_1, is referred to as curvature 1. A voxel with a point cloud associated with a dominant value for c_1 ($c_1 \approx 1$ and $c_{2,3} \approx 0$), has a 1D-like distribution. If a voxel has $\lambda_1 = 0.5$, $\lambda_2 = 0.4$, and $\lambda_3 = 0.1$, the geometry of this voxel can be characterized as a 2D-like surface, while a voxel with $\lambda_1 = 0.4$, $\lambda_2 = 0.3$, and $\lambda_3 = 0.3$ is characteristic of a 3D-like geometry. The study site is part of a salt marsh, and thus a goal of the clustering algorithm is to separate vegetation areas that have 3D-like distributions from tidal flats that have 2D-like distributions. The ratio of λ_3 to λ_2 is used to quantify such differences and is referred to as curvature 2 and expressed by Equation (1) [34].

$$a = \frac{\lambda_3}{\lambda_2} \tag{1}$$

where a = curvature 2.

Finally, standard deviations of the per-point features using all points inside each voxel are calculated because they reflect the complexity of the terrain surface. From the TLS point cloud, the following set of three standard deviation features is extracted for each fine and broader scale voxel: standard deviation of height (usually referred to as surface roughness), standard deviation of reflectance, and standard deviation of waveform deviation. From the UAS-SfM point cloud, the following set of four standard deviation features is extracted for each fine and broader scale voxel: standard deviation of height, standard deviation of reflectance from Red channel, standard deviation of reflectance from Green channel, and standard deviation of reflectance from Blue channel. Standard deviations with the voxel points are calculated because they reflect the complexity and geometry of the terrain surface. When the standard deviation approaches one (assuming the feature is normalized), it indicates a very heterogeneous surface. When a voxel standard deviation approaches zero, it indicates that the surface is very homogeneous. For example, standard deviation of point height "z" for a tidal flat will be smaller and closer to zero compared to a mixed vegetation area for both TLS and UAS-SfM point clouds.

In summary, after the features are extracted, 13 different features characterize each point of the TLS data set: 5 features for each small and broad scale voxelization (10 total) and 3 per-point features. 16 different features characterize each point of the TLS data set: 6 features for each small and broad scale voxelization (12 total) and 4 per-point features. These features have different dynamic ranges and units.

Normalization is an essential preprocessing step in K-means clustering [37]. The methods are based on minimizing the Euclidean distance between the data set points and cluster centroids, which are sensitive to differences in magnitude or scale of the features. Because of the differences in range in features' values, one feature can overpower the other. Normalization prevents the outweighing of some attributes and equalizes the magnitude and variability of all features by standardizing their values from different dynamic ranges into a specified range. There are three common normalization methods; they are Min-max, Z-score, and decimal scaling normalization [37]. In this research, we used Z-score, a Matlab function, to normalize the multiple features. In Z-score normalization, the values for

a feature x are normalized based on the mean and standard deviation of x. For sample data with mean μ and sample standard deviation S, z_{score} of a feature x is computed as:

$$z_{score} = \frac{(x - \mu)}{S} \tag{2}$$

where

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} |x_i - \mu|^2}; \ \mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

3.3. Determination of the Number of Clusters

Many criteria have been proposed for determining an optimal number of clusters; examples are the Dunn, Calinski Harabasz, Silhouette, Gap, and Davies-Bouldin (DB) cluster criterions. These criteria have the common goal of identifying a number of clusters that result in compact, well separated clusters and each has pros and cons in terms of applications. The DB criterion was used in this research because of its reasonable performance in many clustering scenarios relative to more complex metrics [38], and the ability of the method to handle very large data sets with reasonable computing times on a high-performance desktop computer. The latter criteria is also advantageous for eventual real-time operation of an unsupervised clustering method for rapid 3D data segmentation based on the method presented here. The DB indexes were calculated using a Matlab function. The Davies-Bouldin index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation [39]. The goal of the method is to find the number of clusters and data segregation that minimizes this index. The Davies-Bouldin index is defined as Equation (3) below.

$$DB = \frac{1}{k} \sum_{i=1}^{k} max_{j \neq i} \{D_{i,j}\} \tag{3}$$

where $D_{i,j}$ is the within-to-between cluster distance ratio for the ith and jth clusters,

$$D_{i,j} = \frac{(d_i + d_j)}{d_{ij}} \tag{4}$$

and d_i: the average distance between each point in the ith cluster and the centroid of the ith cluster. d_j: the average distance between each point in the jth cluster and the centroid of the jth cluster. d_{ij}: the Euclidean distance between the centroids of the ith and jth clusters. Application of the DB index to select number of clusters is discussed in Section 4.1.

In this research, we used K-means clustering, or Lloyd's clustering algorithm [40], which is an iterative, data-partitioning algorithm that assigns each point of a data set to exactly one of k clusters with the clusters defined by their centroids and the optimal k determined by the Davies-Bouldin criterion. The K-means algorithm utilized here was implemented with Matlab. Refer to our prior work [23] for more details on its implementation. A summary of the entire clustering framework implemented is shown in Figure 4 below.
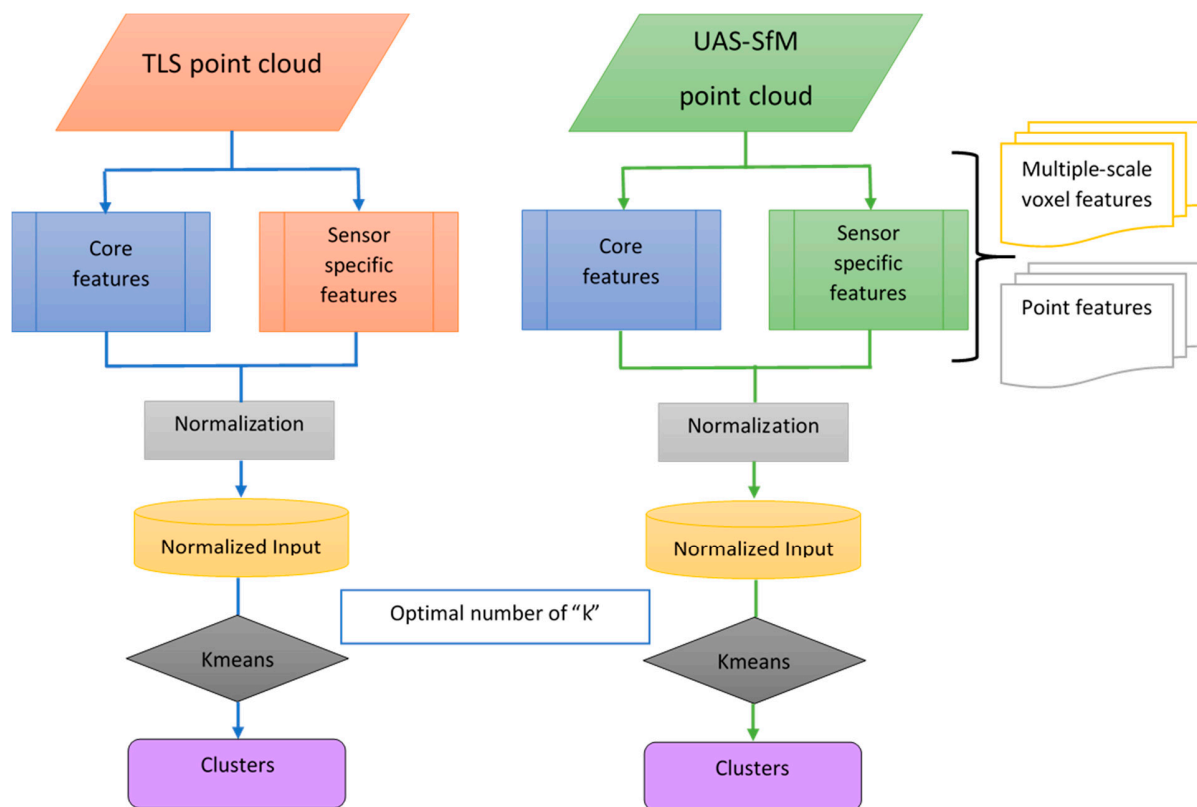
**Figure 4.** Summary of the entire clustering framework.

4. Results and Discussion

4.1. Selection of the Number of Clusters

The K-means clustering algorithm was applied to both TLS and UAS-SfM data while using their respective sets of normalized features described above. Davies-Bouldin values were calculated for 2 to 20 clusters to guide the selection of the number of clusters for the scene for both TLS and UAS-SfM data. For each cluster size, DB indexes were computed for 20 replicates with a maximum of 200 iterations. DB mean with standard error and DB minimum for TLS and UAS are presented in Figures 5 and 6. For both TLS and UAS-SfM, the minimum DB values are obtained for 2 and 3 clusters. This is frequently the case for K-means [41] as large inter-cluster distances typically occur when the number of clusters is low. However, a marsh surface is a complex and heterogeneous environment hence solutions with a larger number of clusters are beneficial as long as the clusters match important features of the scene. Therefore, instead of simply selecting the k = 2 which leads to the minimum DB values, the number of clusters was selected based on the shape of the DB graphs.
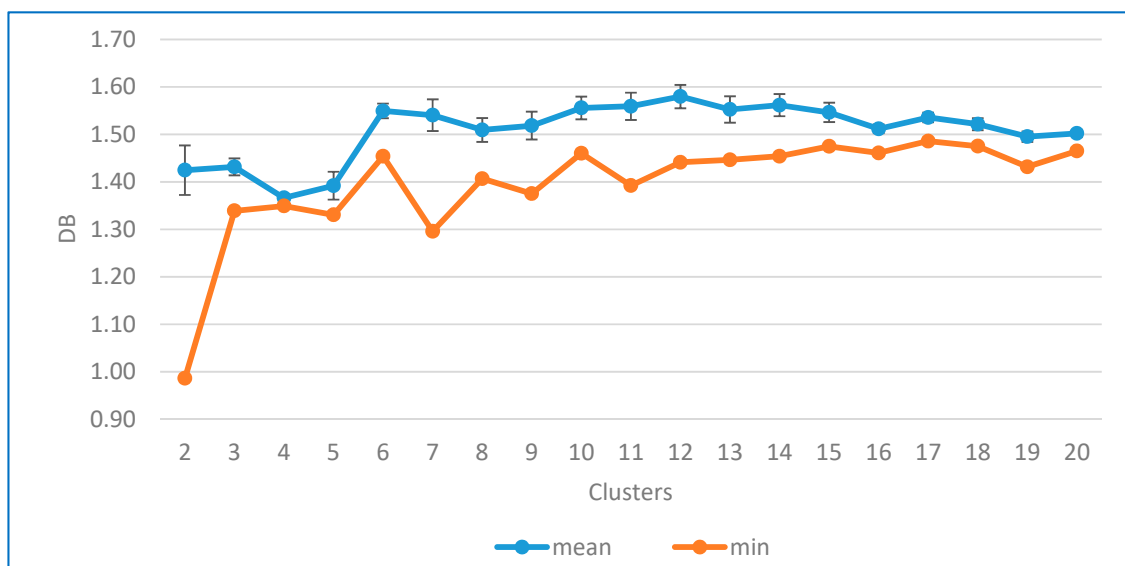
Figure 5. Davies Bouldin (DB) index values for TLS: DB mean with standard error and DB minimum.
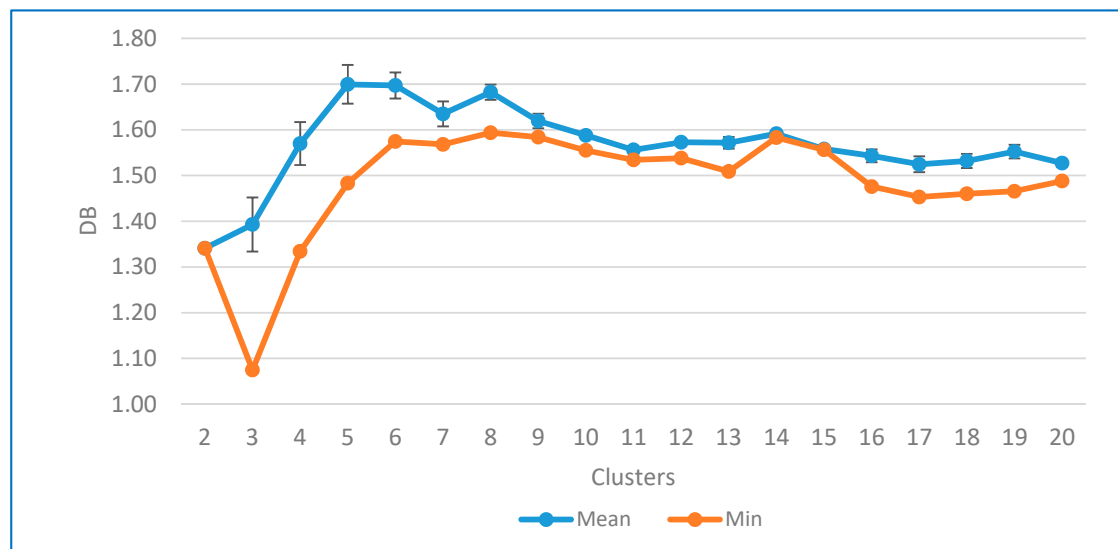
Figure 6. DB index values for UAS-SfM: DB mean with standard error and DB minimum.

The minimum DB values are considered for each cluster size as they correspond to the best clustering solution. After a general rise of their respective minimum DB values, the curves flatten when reaching 6 clusters or larger for the UAS-SfM data (Figure 6). The DB values represent the ratio of average intra-cluster distances over total inter-cluster distances. A flattening of the DB curve indicates that further splitting the data set no longer significantly changes this ratio, i.e. the decreasing average intra-cluster distance is compensated by a similarly decreasing inter-cluster distance. Hence, for both data sets, a clustering solution is sought at the start of these flattening curves.

For the TLS data (Figure 5), minimum DB values are lower for clusters of size 3, 4, 5 and 7 but then increase for clusters of size 6 and larger than 7. The lowest DB values are reached for several solutions (4/20) with a cluster size of seven. For the UAS-SfM data (Figure 6), values of 6, 7, or 8 would be appropriate with similar minimum DB values. To facilitate the comparison, a breakdown of the scene into 7 clusters was selected for both data sets. 20 replications are run for this case k = 7 for both the UAS-SfM and TLS data sets. Out of those 20 replications, the one with the minimum DB value was selected as the K-means cluster solution for further analyses and quantitative assessments.

Comparisons of the clustering accuracy of breakdowns of the scene into different numbers of clusters, including k = 5, 6, 8, are also computed in the discussion for two datasets of the identified marsh areas.

*4.2. Comparative Description of the TLS and UAS-SfM Clusters*

First, we review and compare how clusters are identified in their respective point clouds and then how they are associated to different marsh land cover features. The clustering decomposition of the scene for TLS (7 clusters) and UAS-SfM (7 clusters) into four main land cover groups is presented in Figure 7. One can observe that TLS and UAS-SfM clustered data points represent similar patterns (Figure 7a,b). Both data sets contain contiguous tidal flat areas (orange) surrounded by vegetated areas. Two major differences between the two-point clouds also can be recognized (Figure 7a,b).

- With the advantage of a nadir view, UAS-SfM provides a point cloud of better coverage for this study area, while TLS is limited by the slant scan angle and suffers from vegetation occlusion which results in a less coverage point cloud.
- With RGB camera and photogrammetry technique, the UAS-SfM point cloud can capture the submerged flats in shallow water where the water is clear enough to reconstruct features on the bottom. With TLS using a NIR laser pulse, the pulse is likely bouncing away or sometimes absorbed at the area covered by water. As a result, the TLS provides few points at the submerged flats.
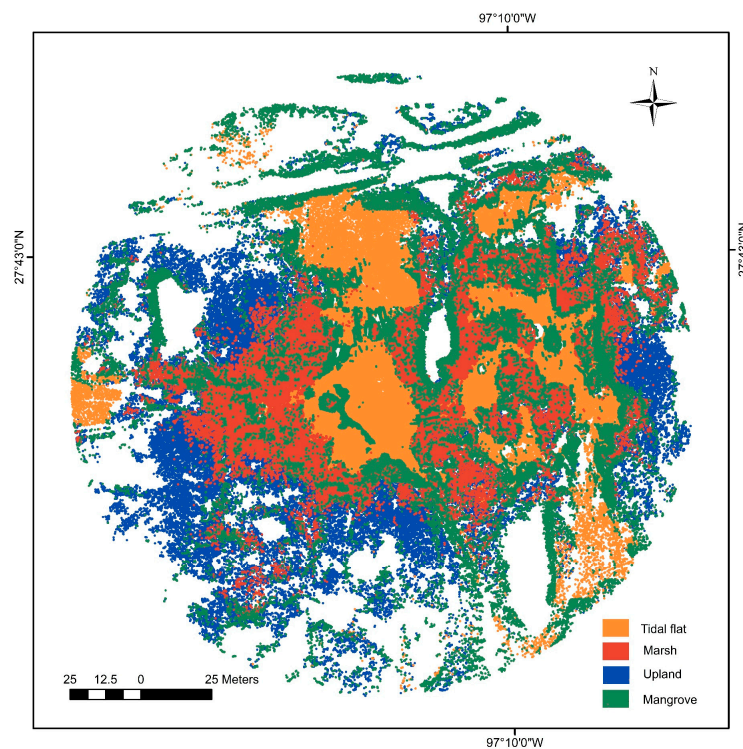
(a)

Figure 7. *Cont.*
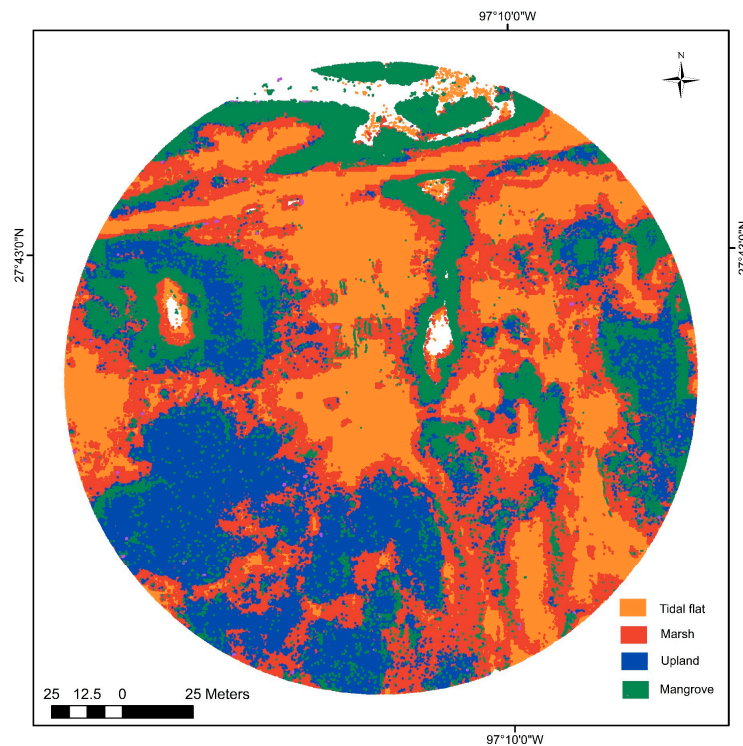
(b)

Figure 7. K-means clustering results of (**a**) TLS and (**b**) UAS-SfM point clouds. Colors represent the 4 different groups: tidal flat areas (orange), mangrove areas (green), upland areas (blue), low and high marshes (red). White areas show data gaps (areas with no points).

For TLS and UAS-SfM, the number of points, point densities, and point coverage are different for each cluster. As a result, the decompositions associated to clusters were computed based on the percentage of that cluster to the total number of points for each point cloud.

For the TLS data set, cluster 3 and cluster 4 represent respectively low flat and high flat areas. For the UAS-SfM data set, cluster 3 represents low flats and submerged flats, while cluster 4 represents high flats. For both point clouds, cluster 3 and 4 are regrouped into one cluster representing all tidal flats. The combination of these two clusters into one simplifies the analysis for submerged, low and high flat areas, which are convoluted and shaped by ever changing coastal conditions. In the TLS point cloud, 31.5% of the data points are identified as tidal flat while 27% of the data points are identified as tidal flats for the UAS-SfM data. The higher percentage of tidal flat points in the TLS data set is due to that all three TLS scan locations are within tidal flats and areas near the scanner benefit from higher density point clouds. The UAS-SfM point cloud density is relatively uniform as the point cloud is generated from the SfM process densified at the ground sample distance of the aerial imagery. This uniform density coverage compared to the occluded TLS representation of the marsh scene can be observed in Figure 7.

Referring to Figure 7 colorization, vegetation surrounding the tidal flats is represented by three main groups: *Avicennia germinans* (black mangrove), upland vegetation, and low and high marsh vegetation.

- *Avicennia germinans* (black mangrove) (green) are dominated and represented as cluster 6 in TLS and cluster 1 in UAS-SfM. The two-point clouds lead to very similar results, 12.1% of the scene for TLS and 11% for UAS. The high vegetation areas are easily identifiable for both methods as they stand out from a top down view (UAS-SfM) or a slant angle view (TLS).

- Upland vegetation (blue) are mainly populated by *Zchizachyrium littorale* (coastal bluestem) and *Spartina patens* (gulf cordgrass). They are represented as cluster 5 in TLS and cluster 2 in UAS. 28.8% of the UAS-SfM point cloud versus only 21.4% of the TLS point cloud are identified as the upland vegetation. These vegetated areas are located away from all three TLS scan positions and occlusions occur more frequently away from the scanner. Without the occlusion effect, UAS-SfM captures a larger fraction of these vegetation areas than TLS.

- *Batis maritima* (pickle weed), *Monanthochloe litoralis* (shoregrass), and *Salicornia* spp. (glasswort) are plants commonly found in the low and high marsh environments at the study site. For the UAS-SfM data set, the marsh vegetation areas (red) are the result of the combination of clusters 5 and 7. In the TLS point cloud, these areas are represented in a single cluster, cluster 1. A higher portion, 35%, of the TLS point cloud was identified as marsh vegetation points as compared to 33.2% for the UAS-SfM data. The higher percentage is likely due to the fact that these vegetated areas are closer to all three TLS scan positions and have an advantage of higher density point clouds.

Both datasets contain few points identified as noise, 0.03% for UAS and 0.05% for TLS. For both UAS-SfM and TLS, noise points are scattered throughout the scene and correspond to vehicles, people, instrument tripod, and reflectors. It is a useful feature of this method that noise is identified as a separate part of the segmentation of the scene.

From visual inspection of Figure 7, the results show a great potential for both methods to segment different portions of the point cloud based on the selected computational features and capture the natural pattern of the study areas. These findings allow us to suggest that the clustering method, which was initially developed for TLS [23], transfers well to UAS-SfM.

Tables 2 and 3 below further summarize the cluster decomposition of the scene for TLS and UAS respectively, including the percentages of the entire point clouds included in each cluster.

Table 2. Cluster decomposition of the scene for TLS.

|	Tidal Flat (orange)	Mangrove (green)	Upland Vegetation (blue)	Marsh Vegetation (red)	Noise (gray)
Clusters	3, 4	6	5	1	2, 7
Number of points	8,452,659	3,242,034	5,738,119	9,370,571	13,270
Percentage of points	31.52%	12.09%	21.40%	34.94%	0.05%

Table 3. Cluster decomposition of the scene for UAS.

	Tidal Flat (orange)	Mangrove (green)	Upland Vegetation (blue)	Marsh Vegetation (red)	Noise (gray)
Clusters	3, 4	1	2	5,7	6
Number of points	10,459,173	4,231,130	11,128,581	12,858,377	13,280
Percentage of points	27.03%	10.94%	28.76%	33.23%	0.03%

4.3. Clustering Accurancy Assessment

To further quantify the cluster partitioning and compare the methods, polygons were created for a portion of tidal flats and vegetated areas in the study area. They were delineated from the co-aligned UAS orthomosaic image based on interpretation by a marsh expert. There are 9 polygons for tidal flats representing 11% of the total scene area, and 13 for vegetated areas representing 29% of the total scene area (Figure 8). These areas were selected at different distances from the TLS scan positions. Different type of marsh environments are included in the polygons: low and high marsh and upland for the vegetated areas and low and high tidal flats for the tidal flat areas. To avoid an imbalanced comparison, no polygons were created for submerged flats. The polygons and clusters were overlaid in order to compute the number of points that are correctly assigned and estimate

the unsupervised classification errors. The comparative assessment was conducted for TLS versus UAS for two types of terrain: exposed ground (tidal flats) and vegetated areas based on all 7 clusters. In addition to the comparison for 7 clusters, comparison of clustering accuracies of 5, 6 and 8 clusters were also performed.
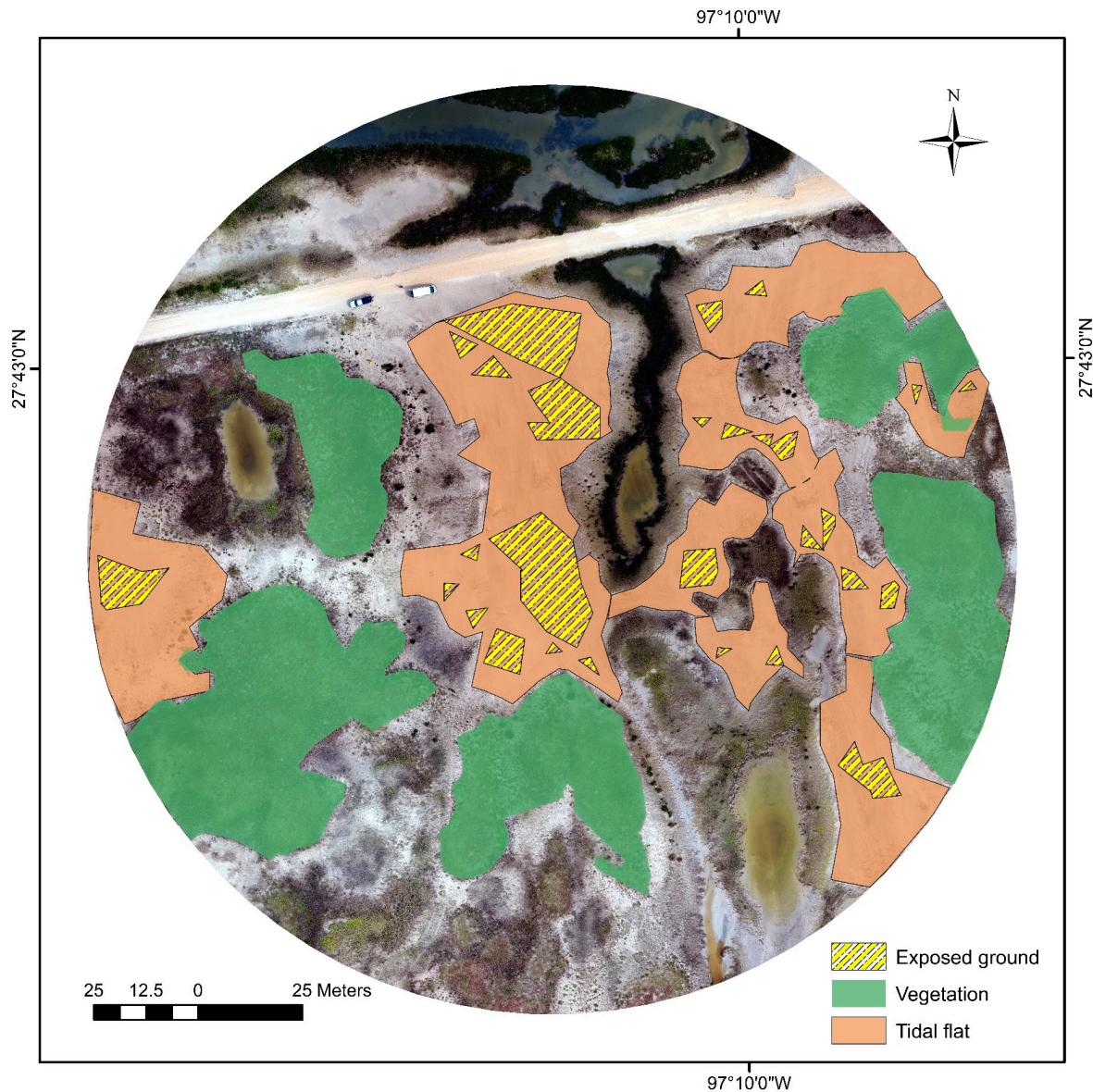
Figure 8. Combined polygon for tidal flats (blue), vegetated areas (green), and exposed ground (yellow) used for accuracy assessments.

The summary of the quantitative assessments, including producer and user accuracies, for TLS and UAS-SfM segmentation results for 7 clusters are described in Tables 4 and 5. The producer's accuracy measures how well a certain area of land cover can be segmented. Its complement is the omission error (producer's accuracy = 100% − omission error). The user's accuracy is indicative of the probability that a point segmented from the whole data set actually represents that category on the ground. Its complement is the commission error (user's accuracy = 100% − commission error).

Table 4. Summary of the quantitative assessments of TLS segmentation results for 7 clusters in identifying tidal flats and areas covered by vegetation.

TLS		Ground Truth (Points)			User's Accuracy
		Tidal Flat	**Vegetated Areas**	**Total**	
Classification by Clusters (points)	Tidal flat	4,359,251	34,543	4,393,794	99.2%
	Vegetated areas	650,614	7,283,037	7,933,651	91.8%
	Total	5,009,865	7,317,580	12,327,445	
Producer's Accuracy		87.0%	99.5%		**Total: 93.3%**

Table 5. Summary of the quantitative assessments for UAS-SfM segmentation results for 7 clusters in identifying tidal flats and areas covered by vegetation.

UAS-SfM		Ground Truth (Points)			User's Accuracy
		Tidal Flat	**Vegetated Areas**	**Total**	
Classification by Clusters (points)	Tidal flat	4,046,110	302,566	4,348,676	93.0%
	Vegetated areas	225,203	15,664,999	15,890,202	98.6%
	Total	4,271,313	15,967,565	20,238,878	
Producer's Accuracy		94.7%	98.1%		**Total: 96.4%**

For the TLS unsupervised clustering results shown in Table 4, about 87.0% percent of all points falling inside the tidal flat polygons are part of tidal flat clusters, while about 13.0% percent are part of vegetation clusters. This corresponds to a 13.0% error of omission. The error matrix shown in Table 4 indicates that the TLS user's accuracy of tidal flat areas is 99.2%, which corresponds to a 0.8% error of commission. The producer of this segmentation can claim that 87.0% of the time a tidal flat was identified as such. A user of this map will find that 99.2% of the time a point that the segmentation identifies as a tidal flat will actually stem from a tidal flat area. For the UAS-SfM unsupervised clustering results shown in Table 5, about 94.7% percent of all points falling inside the tidal flat polygons are part of tidal flat clusters. This corresponds to an approximate 5.3% error of omission for the UAS-SfM tidal flat points. The error matrix shown in Table 5 indicates that the UAS-SfM user's accuracy for tidal flat areas is 93.0%, which corresponds to a 7.0% error of commission.

The producer accuracy of tidal flat areas as shown in Tables 4 and 5 is quite low compared to the user accuracy, especially for the TLS point cloud (87.0%). When considering more closely the tidal flat polygons and the points falling inside these polygons, a significant number of these points belong to small patches of low vegetation, such as small algal mats, scattered in these areas but still identified as tidal flat. To quantify the importance of these cases, the ground truth tidal flat area was further split into two groups of points: (1) points clearly associated with a tidal flat without any visible small vegetation, and (2) areas with small vegetation features. The producer accuracies were then computed for these small exposed ground areas of tidal flats without any visually identifiable vegetation (Figure 8). For TLS unsupervised clustering, 99.98% of all the points falling inside these exposed ground polygons are identified as being part of tidal flat clusters. This indicates that TLS clusters capture the complexity of the true surface and the relatively low producer accuracy of 87.0% was not due to misclassification but the result of the higher clustering resolution of the TLS based method. When analyzing the performance of the UAS-SfM based unsupervised clustering, results indicate that about 94.5% of all points falling inside the broad tidal flat polygons are indeed tidal flats vs 87.0 % for TLS. When analyzing clustering over the smaller exposed ground areas, 99.01% of UAS-SfM points vs. 99.98% of TLS points are part of tidal flat clusters.

When comparing the producer accuracies for tidal flat areas (including vegetated patches) and exposed ground areas (without vegetation present), TLS's clustering provides a higher potential of identifying groups of points belonging to small patches of vegetation as part of clusters that had vegetation as compared to UAS-SfM. Furthermore, TLS provided a higher probability (higher user's

accuracy) that a point segmented from the whole data set actually represents a tidal flat on the ground as compared to UAS-SfM. It is hypothesized, based on the more detailed investigation, that the initial lower producer accuracy of 87.0% is actually an indication of the scattered vegetative component of the tidal flat.

For vegetated areas (Table 4), TLS unsupervised clustering results in about 99% of all points falling inside vegetated polygons being correctly assigned to vegetation clusters while about 1% are assigned to tidal flat clusters. This corresponds to a 1% error of omission for TLS vegetation points. The error matrix shown in Table 4 indicates that the TLS user's accuracy for vegetated areas is 92%, which corresponds to an 8% error of commission. Like TLS, about 98% of all UAS-SfM points falling inside the vegetated polygons are part of vegetation clusters while about 2% are part of tidal flat clusters. This corresponds to a 2% error of omission for UAS-SfM vegetation points. The error matrix shown in Table 5 indicates that the UAS-SfM user's accuracy of vegetated areas is also 98%, which corresponds to an approximate 2% error of commission. The results illustrate that both TLS and UAS-SfM are efficient in identifying vegetated areas with less than 8% in errors of omission and commission for both methods. Producer's accuracies and omission errors show that vegetated areas were slightly better represented from TLS data than UAS-SfM data. User accuracies illustrate that UAS-SfM provided a higher probability that a point segmented from the whole data set represents vegetated area on the ground.

The quantitative results show that the clustering method has different advantages and limitations for each dataset (TLS and UAS-SfM). For example, TLS has better segmentation capability for identification of small patches of vegetation in tidal flat areas while UAS-SfM has more uniform/better coverage. As TLS surveys are typically more expensive than UAS-SfM when accounting for cost of the equipment and the increase in time typically required to complete a TLS survey due to multiple scan positions, it is recommended to add UAS-SfM when possible to a TLS survey for an improved overall accuracy and more uniform, contiguous coverage of the scene from an aerial (nadir) perspective.

The quantitative assessments of the segmentation results for 5, 6, 7 and 8 clusters of both TLS and UAS-SfM were also evaluated to quantify the sensitivity of the methods to the selection of the number of clusters. The summary of their accuracies computed for tidal flat and vegetated areas is presented in Tables 6 and 7.

Table 6. Summary of the quantitative assessment of TLS segmentation results for tidal flats and areas covered by vegetation when increasing the number of clusters.

Number of Clusters	Producer's Accuracy		User's Accuracy	
	Tidal Flat	Vegetated Areas	Tidal Flat	Vegetated Areas
5	93.4%	98.8%	98.1%	95.6%
6	87.0%	99.5%	99.2%	91.8%
7	87.0%	99.5%	99%	91.8%
8	87.6%	99.8%	99.7%	92.2%

Table 7. Summary of the quantitative assessment of UAS-SfM segmentation results for tidal flat and vegetated areas when increasing the number of clusters.

Number of Clusters	Producer's Accuracy		User's Accuracy	
	Tidal Flat	Vegetated Areas	Tidal Flat	Vegetated Areas
5	99.3%	87.7%	68.3%	99.8%
6	87.5%	86.6%	63.6%	96.3%
7	94.7%	98.1%	93.0%	98.6%
8	87.5%	98.0%	92.3%	96.7%

As shown in Table 6, the TLS producer's accuracy for tidal flats is highest for 5 clusters (93%) corresponding to a 7% omission error. For larger clusters, the producer's accuracies decrease to about 87%. While the total number of clusters increases, the number of clusters corresponding to tidal

flats remains at two. The user's accuracies for tidal flat clustering are between 98% and 99% for TLS data. The accuracies correspond to an average of 1.5% error of commission for TLS tidal flat points. Both producer and user's accuracies illustrate that the ability to segment tidal flat does not improve when the number of clusters increases from 6 to 8.

As shown in Table 7, both the producer and user's accuracies resulting from the clustering of UAS-SfM data to identify tidal flats are fluctuating when increasing the number of clusters from 5 to 8. The producer's accuracies are 99% and 95% for 5 and 7 clusters, but they are 88% for both 6 and 8 clusters. This corresponds to 0.7% and 5% errors of omission for 5 and 7 clusters respectively, and about 12% for both 6 and 8 clusters. The user's accuracies are low for 5 (68%) and 6 (64%) clusters, but they improve significantly for 7 (93%) and 8 (93%) clusters.

According to Tables 6 and 7, TLS producer accuracies of tidal flat areas are relatively low compared to vegetated areas for all number of clusters, and UAS-SfM producer and user accuracies for these areas are fluctuating when increasing the number of clusters from 5 through 8. This difference in behavior required us to consider these areas more closely. We observed that a significant number of points belonged to small patches of vegetation, which were initially identified as tidal flat within the ground truth polygons. They were not misclassified by clustering, but were actually part of small vegetative features within tidal flat areas. As explained above for the low producer accuracy of TLS (Table 4), the points resulting from small vegetation scatter can be correctly segmented as vegetation points while they were initially identified as part of the tidal flat ground truth area. These points then skew the producer and user accuracies. To better understand these clustering results, we calculated an additional assessment: (1) exposed ground areas only (within tidal flats) and (2) areas with vegetation (see Figure 8). The summary of the accuracies computed for exposed ground and vegetated areas is presented in Tables 8 and 9.

Table 8. Summary of the quantitative assessment of TLS segmentation results for exposed ground and vegetated areas when increasing the number of clusters.

Number of Clusters	Producer's Accuracy		User's Accuracy	
	Exposed Ground	Vegetated Areas	Exposed Ground	Vegetated Areas
5	99.99%	99.74%	99.12%	100.00%
6	99.98%	99.95%	99.82%	100.00%
7	99.99%	99.95%	99.82%	100.00%
8	99.72%	99.97%	99.91%	99.92%

Table 9. Summary of the quantitative assessment of UAS-SfM segmentation results for exposed ground and vegetated areas when increasing the number of clusters.

| Number of Clusters	Producer's Accuracy		User's Accuracy	
	Exposed Ground	Vegetated Areas	Exposed Ground	Vegetated Areas
5	98.89%	99.82%	98.51%	99.87%
6	99.58%	95.90%	74.51%	99.95%
7	99.01%	99.04%	92.54%	99.80%
8	98.26%	99.05%	92.49%	99.79%

Table 8 shows that TLS producer and user's accuracies for both tidal flat and vegetated areas are very high (99%–100%) and stable when increasing the numbers of clusters from 5 to 8. It corresponds to an average of less than 1% of commission and omission errors. The results show that this method is consistently efficient at identifying the exposed ground and vegetated areas when increasing the number of clusters from 5 through 8. Overall, based on the comparison of the segmentation performances, breakdowns of the scene into 5 to 8 clusters for TLS point clouds are all good solutions to delineate and identify exposed ground areas (within tidal flats) and vegetated areas. The results also

indicate that the selection of the number of clusters for TLS doesn't appear to significantly influence the accuracy of the segmentation of the scene under the class scenarios examined.

As shown in Table 9, UAS-SfM producer's accuracies for exposed ground are also high and steady from 5 to 8 clusters (around 98% and 99%). While user's accuracies for these areas are high (around 98% and 92%) for 5, 7, and 8 clusters. A significant drop (74%) in the user accuracy is observed for 6 clusters. This drop indicates that the selection of the numbers of clusters for UAS-SfM influences the accuracy of the segmentation, especially for the exposed ground. Producer and user's accuracies for vegetated areas vary between 95% and 99% for 5 through 8 clusters. This corresponds to a range of 1% to 5% for the omission errors and < 1% for the commission errors respectively.

To understand the reason behind the drop in user's accuracy for the 6 clusters solution for UAS-SfM shown in Table 9, the variance within each cluster and between clusters were computed for each feature for both TLS (13 features) and UAS-SfM (16 features) over the six clusters. The visualization of statistical comparisons of standardized curvature 2 values (extracted from small voxels) for TLS and UAS-SfM is presented as an example of the difference between feature statistics of TLS and UAS-SfM (Figure 9). Curvature 2 (extracted from small voxels) was selected as the representative example here because it dictates well differences between exposed ground within tidal flat areas versus vegetation patches. It was also the highest ranking, shared feature between TLS and UAS-SfM in terms of separability for the original 7 clusters solution, which is discussed in more detail in Section 4.4 below. As shown in Figure 9, TLS clusters are more compact, while UAS-SfM clusters have a wider range between the boxes' fences and many outliers for the 6 clusters solution. The wider spread of the curvature 2 distributions, including the outliers, could be a reason why UAS-SfM has a drop in user accuracy for the 6 clusters solution. The mean sum of the squared errors (squared difference to the respective cluster's mean) over the six clusters for UAS-SfM is about twice that for TLS, which also indicates that TLS clustering results are more separated and precise as compared to UAS-SfM for 6 clusters solutions. Additional insights can be gained from Figure 9. Clusters with overlapping ranges of feature values (here curvature 2) have distributions with more similar central tendencies compared to those clusters with less overlap. Statistical similarities in feature values for different clusters provides one possible explanation as to how the same land cover type can be split across multiple clusters.

In summary and referring to Tables 6–9, the comparisons also suggest that DB criterion is an insightful guidance for selecting the number of clusters. However, the best segmentation performance should be selected with considerations of DB guidance and also other criteria such as the producer and user's accuracy results and visualizations of the clustered/segmented scene. These methods would have to be applied to more cases to determine if this is a general result.
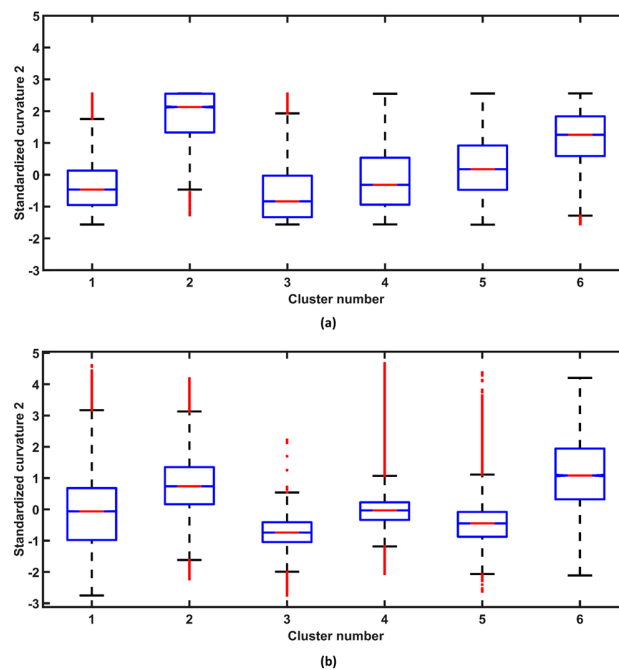
Figure 9. Boxplot showing a comparison of standardized curvature 2 features (extracted from small voxels) for each cluster for (**a**) the TLS point cloud and (**b**) the UAS-SfM point cloud solutions with number of clusters equal to 6. The more compact the box and its fences are, the more similar the points within that cluster. The larger the box and its fences are, including potential outliers in red, the more heterogeneous the cluster. Clusters with overlapping boxes have curvature 2 distributions with similar central tendencies while non-overlapping boxes indicate a good differentiation in curvature 2 values between the clusters.

4.4. Feature Importance

The relative ranking of the features for the 7 cluster segmentation of the scene is quantified using the F statistic. The results calculated for TLS and UAS-SfM are presented in Table 10. The F value is a ratio of variability between clusters over variability within clusters [42]. Therefore, the larger the F value for a given feature relative to the other features, the more separated it is among the scene's clusters. Based on the F statistic, TLS and UAS-SfM clustering results have different rankings for their respective feature distinctness. For both, the combination of features from multi-scale voxels and individual points are important for the clustering of the scene. The *p*-values of all F statistics for both UAS-SfM and TLS data sets are approximately zero indicating significant differences among the clusters but expected given the very large sample sizes.

For the TLS clustering results, the two largest F values are found for curvature 2 computed over the small and large voxels. The high value for this feature is likely due to the partitioning of most of the point cloud into 2D surfaces (tidal flats) and more 3D like surfaces (vegetation). The next four most distinct features based on the F statistics are all computed over large and small voxels: standard deviations of waveform deviation, and standard deviation of reflectance. The relative importance of these features computed over both small and large voxels suggests that capturing the variability at multiple scales through feature engineering is important for a good clustering of the scene.

On the other hand, for the UAS-SfM clustering results, the six largest F values are found for the standard deviation of color values: Blue (B), Green (G), and Red (R) computed over both small and large voxels. The high F values for these features show that the variability of RGB computed for multi-scale voxels is an important influence for the clustering performance. The next six most distinct features are point features (B, G, R and Z) and curvature 2 computed over small and large voxels. The relative importance of these features also suggests that the combination of multi-scale voxels and individual point features are helpful in segmenting the marsh environment. This supports

the hypothesis and findings presented in our prior work [23] on developing the clustering framework with TLS point cloud data, which showed that the combination of TLS voxel and point features helped segmentation of the marsh land cover. Furthermore, it is notable that while the clustering accuracies are relatively similar, the lead features are quite different. The TLS data set is primarily segmented based on the geometrical characteristics (curvature 2) and variability in waveform deviation and reflectance values for neighborhoods (voxels) of each point, while the UAS-SfM data set is primarily segmented based on the variability of the scene's digital color values in the neighborhood of each point.

Table 10. Rankings of feature separability based on computed F statistic values for TLS and UAS-SfM for 7 clusters. σ = standard deviation. For TLS, Z = elevation, R = reflectance, D = waveform deviation. For UAS-SfM, R = Red, G = Green, B = Blue.

TLS		UAS-SfM	
Features	**F Statistic**	**Features**	**F Statistic**
Curvature 2 of small voxel	13,071,500	σ_B of large voxel	7,159,650
Curvature 2 of large voxel	12,032,700	σ_G of large voxel	6,839,710
σ_D of large voxel	9,260,570	σ_B of small voxel	6,086,370
σ_D of small voxel	7,855,800	σ_G of small voxel	6,084,770
σ_R of large voxel	7,274,380	σ_R of large voxel	5,793,100
σ_R of small voxel	6,361,320	σ_R of small voxel	5,715,920
Elevation (Z)	5,485,250	Blue (B)	5,290,480
Waveform Deviation (D)	5,449,800	Green (G)	5,149,100
σ_Z of large voxel	5,110,470	Curvature 2 of small voxel	4,878,490
Reflectance (R)	3,461,560	Red (R)	4,758,860
σ_Z of small voxel	3,062,030	Curvature 2 of large voxel	3,681,740
Curvature 1 of small voxel	1,428,290	Elevation (Z)	2,590,360
Curvature 1 of large voxel	1,422,150	σ_Z of small voxel	2,123,120
		σ_Z of large voxel	1,678,420
		Curvature 1 of large voxel	1,363,320
		Curvature 1 of small voxel	1,152,010

For both TLS and UAS clustering results, the least distinctive features following the F statistics are the curvature 1 indexes computed for both small and large voxels. Looking at Table 10 one can also find that the curvature 1 values are substantially smaller as compared to curvature 2 values. While for this scene such features could be omitted, they should be kept for a general algorithm as curvature 1 has been shown to help identify buildings and other such structures [34].

These statistics' F values are only calculated for a breakdown of the scene into 7 clusters for both the TLS and UAS-SfM data sets. Different numbers of clusters could lead to a different ranking. Similarly, application of the method to different marsh sites leading to potentially different scene complexities captured by the TLS and UAS-SfM point cloud could also lead to different rankings.

5. Conclusions

Three-dimensional point clouds of a marsh environment produced with TLS and UAS-SfM were partitioned using an unsupervised clustering method. Even though the two methods result in very different point cloud representations based on distinctive point cloud densities and characteristics, both are capable of identifying the natural structure of a marsh environment. Without prior training, the method categorizes millions of unlabeled data points into meaningful groups. A segmentation into 7 clusters based on a cluster optimization approach (Section 4.1) identifies similar patterns of tidal flats, mangrove, low marsh to high marsh vegetation, and upland vegetation. The method also identifies non-natural feature noise in the point cloud, which is only 0.03% to 0.05% of the data set, allowing for improvement of the data quality. The upland areas identified through UAS-SfM clustering are larger as compared to TLS clustering results because of occlusions as the distances increase away from the TLS scanner. The comparison of the clustering results presented in Sections 4.2 and 4.3 suggests that TLS and UAS-SfM can produce an accurate identification of tidal flats within different marsh environments. Furthermore, both provide accurate segmentations for vegetated land cover.

Comparisons of the two segmentation results (Section 4.3) demonstrate that clustering based on TLS data is somewhat better suited to distinguish exposed ground and scattered small patches of vegetation within tidal flats as compared to UAS-SfM. The reason could be that the substantial difference between PCA and surface roughness values in exposed ground and vegetated areas provides a prominent feature to distinguish between these two land covers. TLS has a greater potential in differentiating vegetated and exposed ground areas. It could be the case that this scanner's ability to record multiple returns and waveform deviation values plays an important role and provides an advantage for capturing a more complex clustering geometry represented by the marsh's highly varying vegetated land cover. Consequently, the rankings of relative feature importance for TLS and UAS-SfM are substantially different based on computations of the features' F statistics. Curvature 2 of the eigenvalues of the planar surface computed for small and large voxels ranks as the top feature that influences TLS clustering. While the standard deviation of RGB color values computed for small and large voxels significantly affects UAS-SfM clustering. The F statistical values also indicate that features computed over fine and coarser scale voxels are important for the clustering process for TLS and UAS-SfM. The multi-scale voxelization used as part of the feature engineering is efficient to help capture the complexity and geometry of the marsh surface.

The accuracies of clustering results were compared by increasing the number of clusters from 5 to 8 to quantify the effects of different choices for the number of clusters. The comparative results reveal that (1) TLS is more precise in identifying small patches of vegetation within tidal flats, and delineating the exposed ground and areas covered with vegetation as compared to UAS-SfM; (2) UAS-SfM is more sensitive to the selection of number of clusters as compared to TLS. The comparisons also suggest that DB values provide an insightful guidance for selecting the number of clusters. However, the best segmentation performance should be selected with considerations of DB guidance as well as other selections of cluster numbers based on visualizations of the clustered scene and accuracy of the segmentation results for targeting the end user's desired classes. A breakdown of the scene into 7 clusters for TLS and UAS-SfM point clouds is an adequate solution to identify tidal flats and vegetated areas. Overall, the results demonstrate that the multi-scale voxelization approach initially developed for TLS point clouds transfers well to the clustering of UAS-SfM point cloud data. The tresults suggest that this unsupervised clustering method could further apply to other type of 3D point clouds and other types of scenes, hence a general framework for intelligent scene segmentation.

Future work will explore the application of the method at different marsh regimes of varying terrain and land cover, and the stability of the method applied to repeat surveys of the same scene. Because UAS-SfM provides a low-cost and flexible alternative to traditional airborne lidar or TLS for generating dense 3D point clouds, UAS could provide an effective way to explore connections between features and segmentation results for different marsh environments and across repeat surveys.

Author Contributions: C.N. and M.J.S. conceived and designed the clustering framework with input from P.T.; C.N. developed the computational code, performed the experiments, and analyzed the results; M.J.S. and P.T. helped with algorithm development and results interpretation; J.G. provided guidance on experimental design and results interpretation; C.N., M.J.S., and P.T. wrote the paper.

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

References

1. Cahoon, D.R.; Hensel, P.F.; Spencer, T.; Reed, D.J.; McKee, K.L.; Saintilan, N. Coastal wetland vulnerability to relative sea-level rise: Wetland elevation trends and process controls. In *Wetlands and Natural Resource Management*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 271–292.
2. Ford, M.A.; Cahoon, D.R.; Lynch, J.C. Restoring marsh elevation in a rapidly subsiding salt marsh by thin-layer deposition of dredged material. *Ecol. Eng.* **1999**, *12*, 189–205. [CrossRef]
3. Schmid, K.A.; Hadley, B.C.; Wijekoon, N. Vertical accuracy and use of topographic LIDAR data in coastal marshes. *J. Coast. Res.* **2011**, *27*, 116–132. [CrossRef]
4. Wang, C.; Menenti, M.; Stoll, M.P.; Feola, A.; Belluco, E.; Marani, M. Separation of Ground and Low Vegetation Signatures in LiDAR Measurements of Salt-Marsh Environments. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2014–2023. [CrossRef]
5. Rosso, P.H.; Ustin, S.L.; Hastings, A. Use of lidar to study changes associated with Spartina invasion in San Francisco Bay marshes. *Remote Sens. Environ.* **2006**, *100*, 295–306. [CrossRef]
6. Sturdivant, E.J.; Lentz, E.E.; Thieler, E.R.; Farris, A.S.; Weber, K.M.; Remsen, D.P.; Miner, S.; Henderson, R.E. UAS-SfM for coastal research: Geomorphic feature extraction and land cover classification from high-resolution elevation and optical imagery. *Remote Sens.* **2017**, *9*, 1020. [CrossRef]
7. Lemmens, M. Terrestrial Laser Scanning. In *Geo-Information: Technologies, Applications and the Environment*; Springer: Dordrecht, The Netherlands, 2011; pp. 101–121.
8. Starek, M.J.; Mitasova, H.; Hardin, E.; Weaver, K.; Overton, M.; Harmon, R.S. Modeling and analysis of landscape evolution using airborne, terrestrial, and laboratory laser scanning. *Geosphere* **2011**, *7*, 1340–1356. [CrossRef]
9. Starek, M.J.; Mitásová, H.; Wegmann, K.; Lyons, N. Space-time cube representation of stream bank evolution mapped by terrestrial laser scanning. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1369–1373. [CrossRef]
10. Fan, L.; Powrie, W.; Smethurst, J.; Atkinson, P.M.; Einstein, H. The effect of short ground vegetation on terrestrial laser scans at a local scale. *ISPRS J. Photogramm. Remote Sens.* **2014**, *95*, 42–52. [CrossRef]
11. Fan, L.; Atkinson, P.M. Accuracy of digital elevation models derived from terrestrial laser scanning data. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1923–1927. [CrossRef]
12. Friess, D.A.; Moller, I.; Spencer, T.; Smith, G.M.; Thomson, A.G.; Hill, R.A. Coastal saltmarsh managed realignment drives rapid breach inlet and external creek evolution, Freiston Shore (UK). *Geomorphology* **2014**, *208*, 22–33. [CrossRef]
13. Xie, W.M.; He, Q.; Zhang, K.Q.; Guo, L.C.; Wang, X.Y.; Shen, J.; Cui, Z. Application of terrestrial laser scanner on tidal flat morphology at a typhoon event timescale. *Geomorphology* **2017**, *292*, 47–58. [CrossRef]
14. Olagoke, A.; Proisy, C.; Feret, J.B.; Blanchard, E.; Fromard, F.; Mehlig, U.; de Menezes, M.M.; dos Santos, V.F.; Berger, U. Extended biomass allometric equations for large mangrove trees from terrestrial LiDAR data. *Trees Struct. Funct.* **2016**, *30*, 935–947. [CrossRef]
15. Klemas, V.V. Coastal and environmental remote sensing from unmanned aerial vehicles: An overview. *J. Coast. Res.* **2015**, *31*, 1260–1267. [CrossRef]
16. Ramírez-Juidías, E. Study of erosion processes in the Tinto salt-marshes with remote sensing images. *Adv. Image Video Process.* **2014**, *2*, 39–52. [CrossRef]
17. Kalacska, M.; Chmura, G.L.; Lucanus, O.; Berube, D.; Arroyo-Mora, J.P. Structure from motion will revolutionize analyses of tidal wetland landscapes. *Remote Sens. Environ.* **2017**, *199*, 14–24. [CrossRef]
18. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314. [CrossRef]
19. Starek, M.J.; Davis, T.; Prouty, D.; Berryhill, J. Small-scale UAS for geoinformatics applications on an island campus. In Proceedings of the 2014 Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), Corpus Christ, TX, USA, 20–21 November 2014; pp. 120–127.
20. Slocum, R.K.; Parrish, C.E. Simulated imagery rendering workflow for uas-based photogrammetric 3d reconstruction accuracy assessments. *Remote Sens.* **2017**, *9*, 396. [CrossRef]
21. Starek, M.J.; Giessel, J. Fusion of UAS-based structure-from-motion and optical inversion for seamless topo-bathymetric mapping. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 2999–3002.

22. Chu, T.X.; Chen, R.Z.; Landivar, J.A.; Maeda, M.M.; Yang, C.H.; Starek, M.J. Cotton growth modeling and assessment using unmanned aircraft system visual-band imagery. *J. Appl. Remote Sens.* **2016**, *10*, 036018. [CrossRef]

23. Nguyen, C.; Starek, M.J.; Tissot, P.; Gibeaut, J. Unsupervised clustering method for complexity reduction of terrestrial lidar data in marshes. *Remote Sens.* **2018**, *10*, 133. [CrossRef]

24. Paine, J.G.; White, W.A.; Smyth, R.C.; Andrews, J.R.; Gibeaut, J.C. Mapping coastal environments with lidar and EM on Mustang Island, Texas, U.S. *Lead. Edge* **2004**, *23*, 894–898. [CrossRef]

25. Church, J.A.; White, N.J. A 20th century acceleration in global sea-level rise. *Geophys. Res. Lett.* **2006**, *33*, 1944–8007. [CrossRef]

26. Wang, G.; Soler, T. Measuring land subsidence using GPS: Ellipsoid height versus orthometric height. *J. Surv. Eng.* **2014**, *141*, 05014004. [CrossRef]

27. Raber, G.T.; Jensen, J.R.; Hodgson, M.E.; Tullis, J.A.; Davis, B.A.; Berglund, J. Impact of lidar nominal post-spacing on DEM accuracy and flood zone delineation. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 793–804. [CrossRef]

28. Perroy, R.L.; Bookhagen, B.; Asner, G.P.; Chadwick, O.A. Comparison of gully erosion estimates using airborne and ground-based LiDAR on Santa Cruz Island, California. *Geomorphology* **2010**, *118*, 288–300. [CrossRef]

29. Douillard, B.; Underwood, J.; Kuntz, N.; Vlaskine, V.; Quadros, A.; Morton, P.; Frenkel, A. On the segmentation of 3D LIDAR point clouds. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, 9–13 May 2011; pp. 2798–2805.

30. Strecha, C.; Küng, O.; Fua, P. Automatic mapping from ultra-light UAV imagery. In Proceedings of the EuroCOW 2012, Barceloa, Spain, 8–10 February 2012.

31. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

32. Schwind, M.; Starek, M. How to produce high-quality 3D point clouds Structure-from-Motion Photogrammetry. *Gim Int. Worldw. Mag. Geomat.* **2017**, *31*, 36–39.

33. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]

34. Wakita, T.; Susaki, J. Multi-scale based extracion of vegetation from terrestrial lidar data for assessing local landscape. *Int. Arch. Photogramm.* **2015**, *II-3/W4*, 263–270. [CrossRef]

35. Vandapel, N.; Huber, D.F.; Kapuria, A.; Hebert, M. Natural terrain classification using 3-d ladar data. In Proceedings of the 2004 IEEE International Conference on Robotics and Automation, New Orleans, LA, USA, 26 April–1 May 2004; pp. 5117–5122.

36. Carlberg, M.; Gao, P.; Chen, G.; Zakhor, A. Classifying urban landscape in aerial LiDAR using 3D shape analysis. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009; pp. 1701–1704.

37. Visalakshi, N.K.; Thangavel, K. Impact of normalization in distributed k-means clustering. *Int. J. Soft Comput.* **2009**, *4*, 168–172.

38. Bowen, Z.H.; Waltermire, R.G. Evaluation of light detection and ranging (lidar) for measuring river corridor topography 1. *JAWRA J. Am. Water Resour. Assoc.* **2002**, *38*, 33–41. [CrossRef]

39. Davies, D.L.; Bouldin, D.W. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, 224–227. [CrossRef]

40. Lloyd, S.P. Least-Squares Quantization in Pcm. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. [CrossRef]

41. Milan, D.J.; Heritage, G.L.; Hetherington, D. Application of a 3D laser scanner in the assessment of erosion and deposition volumes and channel change in a proglacial river. *Earth Surf. Process. Landf.* **2007**, *32*, 1657–1674. [CrossRef]

42. Montgomery, D.C.; Runger, G.C.; Hubele, N.F. *Engineering Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2009.