Maize Biomass Estimation using UAS-Based Structure from Motion Point Cloud Data and Volumetric Approaches

Jose Manuel Mendoza Reyes

Academic exercise submitted in partial fulfillment of the requirements for the degree of:

Master of Science in Geospatial Technologies

Institute for Geoinformatics

2019

Münster, Germany

Declaration

I wrote this thesis on my own without the help of others, listed all used references and sources, and did not submit this thesis to any other university.

Jose Manuel Mendoza Reyes

Münster, Germany

Name:	Jose Manuel MENDOZA REYES
Matric. Number:	457204
Email:	j_mend04@uni-muenster.de
Hosting University:	University of Münster
Supervisor:	Jan LEHMANN
Co-supervisor 1:	André GROßE-STOLTENBERG
Co-supervisor 2:	Ignacio GUERRERO

"Quiero dedicar este trabajo de investigación a mis amados padres Baltazar Mendoza y Soledad Reyes por su amor y apoyo, sin ustedes jamás habría alcanzado esta meta en mi vida, los amo."

Table of content

Article

Maize Biomass Estimation using UAS-based Structure from Motion Point Cloud Data and Volumetric Approaches

Jose-Manuel Mendoza Reyes [1]

[1] University of Münster; ifmendozar@gmail.com
* Correspondence: ifmendozar@gmail.com

Abstract: Plant phenotypic traits such as biomass work as predictors of important biological outcomes like fitness, disease, and mortality. Plant biomass is an essential parameter for crop management, growth monitoring, and yield estimation. Traditionally, the manual approach known as destructive sampling is the most accurate technique to estimate biomass. However, the large scales of modern agricultural research or production schemes are turning this approach into an impractical technique. Remote Sensing (RS) is a technology that has been applied to facilitate this task. Light Detection and Ranging (LiDAR) is a popular RS tool for evaluating and understanding plant canopy structure due to its accuracy and ability to build 3D point clouds, but its high cost currently limits its application. Structure from Motion (SfM) is a low-cost alternative to LiDAR that involves acquiring images using a digital camera from multiple positions to generate a 3D point cloud similar to that which is produced with LiDAR. In this study, SfM point clouds derived from Unmanned Aerial Systems (UAS) were used to build volumetric models evaluating and comparing two different methodologies for estimating plant biomass in maize (*Zea mays* L.); voxel-counting and convex-hull approach. The voxel-counting approach works by encapsulating the point cloud into volumetric pixels to generate a voxel grid. The volume is estimated by counting the number of voxels with at least one point inside. The convex-hull approach splits the point cloud into two parts and progressively calculate the extreme points to generate a polygon that encompasses the full point cloud. The result of this study showed that volumetric models based on SfM point cloud data are suitable for estimating maize biomass combining with volumetric models. Voxel-counting was more accurate predicting maize biomass (R^2=0.973) than the convex-hull approach. SfM point cloud data coupled with the voxel-counting approach offers a low-cost alternative for providing accurate biomass estimations.

Keywords: remote sensing, structure from motion, UAS, convex-hull, voxel, maize, biomass.

---

1. Introduction

To cope with the growing world population and its increasing demand for food, crop production must be increased in the coming years, thus, improved research techniques need to be updated to have suitable methodologies for evaluating observable physical characteristics in agricultural crops knowing better as crop traits such as height, biomass, and leaf area are needed to enhance crop management. These crop traits work as predictors of important biological outcomes, such as fitness, disease, and mortality [1]. Plant biomass serves as one of the critical inputs in crop growth monitoring and yield estimation modeling, among others in agricultural, ecological, and meteorological applications. Therefore, the accurate estimation of plant biomass remains an active field of research [2]. The most accurate techniques to estimate crop biomass are traditional manual approaches such as destructive sampling, this is a labor-intensive and time-consuming method in which the plants are subjected to constant mechanical stress. In addition, due to the large scales of modern agricultural research and production schemas, collecting biomass using traditional manual approaches has become impractical [4].

To address these issues, remote sensing methods that enhance the speed and accuracy of plant biomass estimation have been developed. Several studies have focused on estimating crop biomass using different sensors. Bending *et al.* [4] applied UAS-based Red-Green-Blue (RGB) imagery to estimate barley biomass, Christiansen *et al.* [5] and Li *et al.* [6] utilized Light Detection and Ranging (LiDAR) for predicting crop biomass. Gao *et al.* [7] implemented Radio Detection and Ranging (RADAR) for evaluating maize biomass. Wang *et al.* [2] Estimated maize biomass with a hyperspectral sensor and Liu *et al.* [8] applied multispectral data to evaluate maize biomass. RS of vegetation is based on the interaction between vegetation and electromagnetic radiation [32]. Vegetation Indices (VIs) are a practical example of these interactions. VIs are numerical values without units that take advantage of this relationship at particular wavelengths of light [32]. The most commonly applied VIs utilize the information about the amount of reflected in the red and near-infrared (NIR) regions of the light spectrum. This information is combined in the form of ratios that is predictive of plant biomass [33]. Although there are many indices that have been developed for crop trait prediction, the Normalized Difference Vegetation Index (NDVI) is one of the most commonly used VI to assess plant biomass and has been applied in several studies. For example, Battude *et al.* [39] and Sharma *et al.* [40] used NDVI to estimate maize biomass, Zhu and Liu [34] and Gonzalez-Alonso *et al.* [37] applied Landsat NDVI time-series and NDVI composites to calculate forest biomass, Ren and Zhou [35], Todd *et al.* [36] and Jin *et al.* [38] implemented NDVI to estimate grasslands biomass.

As an alternative to VIs, LiDAR data has been successfully used in estimating crop biomass due to its accuracy and ability to build 3D point cloud data. LiDAR has been applied for evaluating and understanding plant canopy structure, namely canopy height [9]. LiDAR itself uses multiple laser-based measurement techniques to determinate 3D point cloud coordinates on the surface of an object relative to the instrument [10]. Although LiDAR data has been shown to be suitable for estimating plant biomass, its high cost limits its application. However, there are cheapest alternatives that can build point clouds such as RGB sensor and photogrammetry technique.

Photogrammetry is a well-established method that has been developed and used in cartography since the 1920s [10]. Photogrammetry takes advantage of digital image processing to automatically generate digital elevation models and orthophotos. The Structure from Motion approach (SfM) integrates the feature of computer vision into classical photogrammetry. It involves acquiring images using a digital camera from multiple positions relative to the object of interest to generate a high-resolution point cloud, which can be color-coded using simultaneously acquired RGB values. SfM point cloud data has the potential to overcome two of the challenges associated with LiDAR: cost and occlusion [10]. SfM is a flexible and reliable approach due to its property to yield high-density geometric and radiometric data [11]. This property enables the production of similar products to those derived from LiDAR. Moreover, previous research has tested the SfM approach to validate its 3D properties in the context of measuring vegetation; Morgenroth and Gomez [12] analyzed the tree structure using a 3D image analysis. Miller *et al.* [13] applied a 3D model to evaluate the accuracy of height, diameter and volume measurement in trees. Roberts *et al.* [14] evaluated the stem diameter of trees with 3D imaging. Holman *et al.* [15] used 3D digital surface models of a crop field to estimate plant height in wheat.

The analysis of SfM point cloud data can be divided into two approaches: Raster based by computing a Canopy Height Model (CHM) and Point Cloud-Based (PCB). Most studies have implemented CHM, which converts raw point clouds into regular 2D grids. However, with this approach, the full dimensionality of the 3D point cloud data is not leveraged. Instead, the PCB approach takes advantage of all of the point cloud data. The PCB approach separates the vertical point cloud distribution into different height levels, enabling the construction of 3D canopy and volumetric models [16].

Performance of the SfM technique depends on the point cloud density, which relies on the image quality and the UAS´s platform. Directly related to point cloud density is Nominal Point Density (NPS), which is defined as the representative distance between points in the point cloud. NPS determines the spatial resolution that can be extracted from the point cloud [43]. In the other hand LiDAR point spacing is determined by the scanner frequency (pulses per second) and speed of the UAS [43]. Another component that can determinate the resolution of the SfM point cloud data is the computer workstation [44]. In order for SfM point cloud data to yield accuracy predictions, the images must have a high percentage of overlap. Software applications stitch overlapping images by matching know object to generate 3D points. SfM generates a point derived for each pixel within the photos, and therefore the point density directly reflects the spatial resolution of the collected photos [45].

This research is comparing the ability of voxel-counting and convex-hull approach to estimate biomass. A voxel can be defined as the volume element in a 3-D array [17]. The voxel is the 3D counterpart of the 2D pixel. Each voxel can have a binary value, which represents measurable properties. For example, the value 0 could mean “NODATA” while the value 1 could mean “at least one value.” In contrast to surface representations such as Triangulated Irregular Networks (TINs) or Digital Elevation Models (DEMs), which only give a single Z value for each X and, Y value, voxels are suitable for representing discontinuous surfaces due to their ability to depict multiple Z values for the same X and,Y coordinates, thereby providing better visualization and improving biomass biomass [18]. On the other hand, the convex-hull approach works by dividing the point cloud into two parts by drawing a line between the farthest points of the point cloud. The point with the highest distance to the line is then used to create a triangle that ignores all the points inside. This step is run several times until all points are included and the convex-hull is finished [19].

The voxel-counting and convex-hull approaches were chosen due to their superior performance in using point cloud data to estimate biomass. Several previous studies have been assessed the ability of volumetric models to assess plant biomass. Kim *et al.* [16] calculated voxel-based metrics using LiDAR to estimate plot-level above-ground tropical rain forest biomass. Brüninghoff *et al.* [19] compared the sum of the voxels and convex-hull approaches with LiDAR data to estimate total grassland biomass (fresh and dry). Greaves *et al.* [22] evaluated the voxel counting approach combined with LiDAR to estimate Arctic shrub biomass. Olsoy *et al.* [25] estimated Sagebrush biomass using both the voxel and convex-hull approaches applied to LiDAR data. However, this relationship has not yet been clarified in maize.

The objective of this research is to evaluate the application of volumetric models with SfM point cloud data derived from aerial photography for estimating maize biomass. Voxel-counting and convex-hull approaches are implemented and compared to determine which model is more accurate for predicting maize biomass. This study aimed to address the following research questions:

1. Can SfM point cloud data be used to estimate maize biomass?
2. Which volumetric model is more accurate for predicting maize biomass?

2. Materials and Methods

2.1. Study area

Field measurements were collected at the Bioland-Hofladen Herzkamp field located in the municipality of Havixbeck in the district of Coesfeld in northern North Rhine-Westphalia, Germany. This company has greenhouses and farmlands where they grow different crops to sell in the local store. According to its harvest schedule, the maize field was harvested weeks before the study begins, but as a special request, they left a third part of the whole plantation for research purposes. The variety of cultivated maize "*Zea mays* var. Saccharata" was sown in a spatial arrangement of individual rows. The total area of the experimental plot was 57.6 m^2 (18x3.20m) with the centroid located at latitude 5,758,244.73 and longitude 392,990.41 (UTM Zone 32U) (Figure 1).

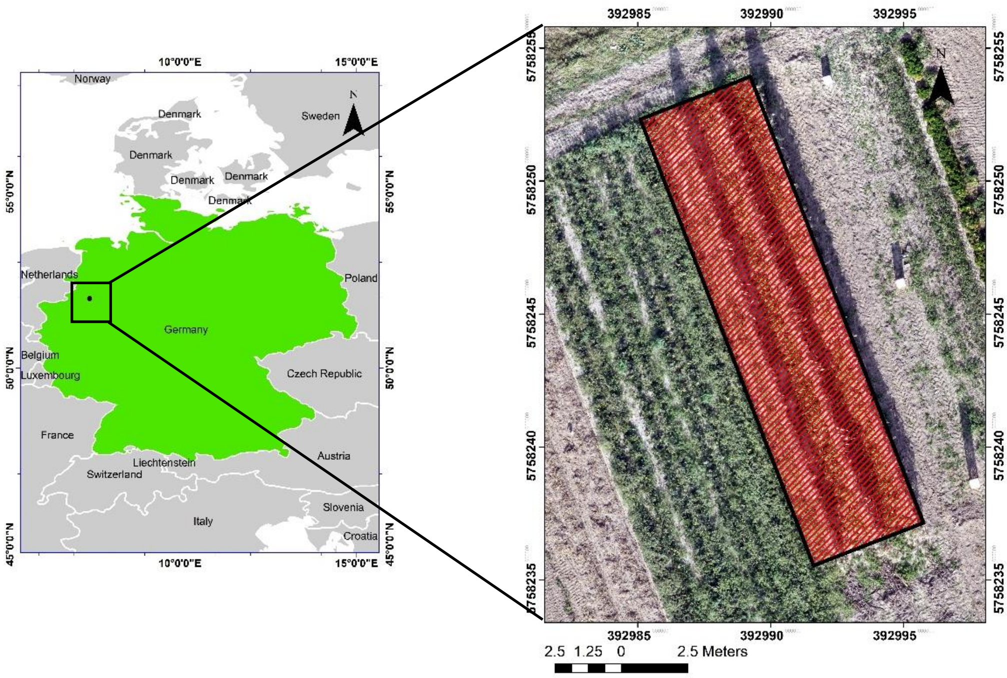


Figure 1. The study area is located in the Northwest part of Germany, in the district of Coesfeld, in collaboration with the agricultural company Bioland-Hofladen Herzkamp. The experimental plot consists of three individual maize rows with an area of 57.6 m^2

2.2. Data acquisition

2.2.1. UAS data

Aerial photographs were acquired with an md4-1000 quadcopter (Microdrones, Siegan, Germany) on October 17th, 2018 at 12h. The flight altitude was 15m. The sensor used was an Alpha a5100 camera (Sony, Minato, Tokyo, Japan) that captures 16-megapixel images in true color (RGB). The focal length and image size were 16mm and, 6000x4000 pixels, respectively. Two flights were carried out in two different orientations, North-South and East-West, with an overlap of 80%, the camera was set to trigger automatically upon reaching flying altitude. To ensure accurate geolocation of the imagery six markers were used as ground control points (GCPs) and positioned in the experimental plot. The position was measured with a high precision Differential Global Position System (DGPS; Trimble, Sunnyvale, CA, USA). Four boxes with heights of 40, 80, 120 and 160 cm were placed in the field for the purpose of ground truthing (Figure 2).

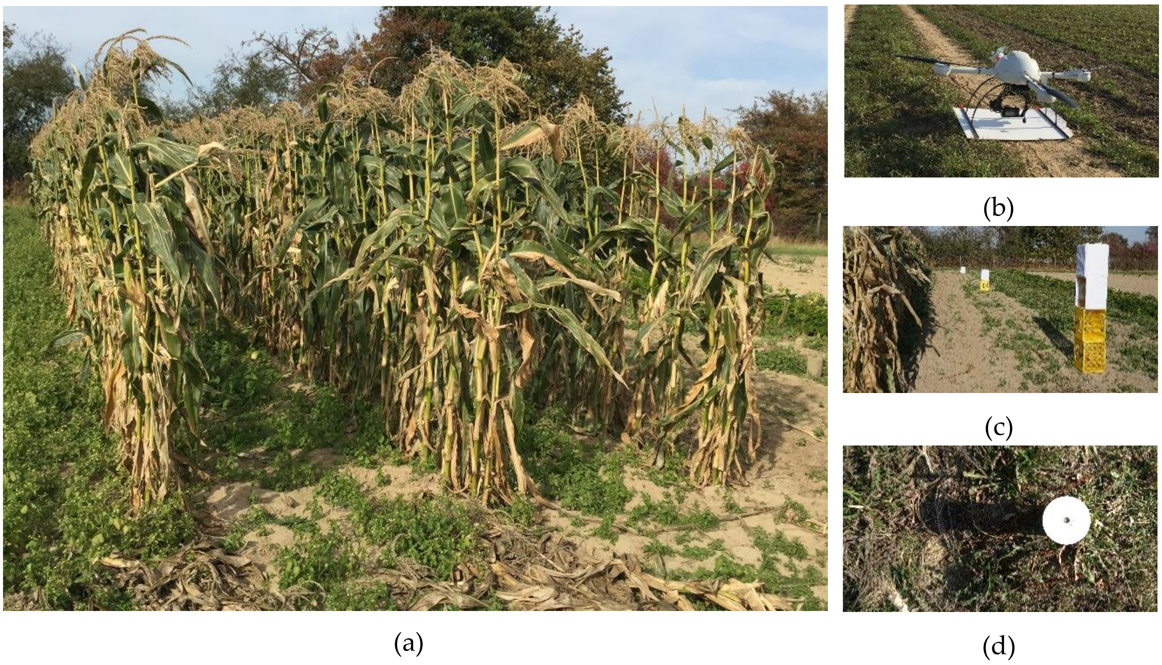


Figure 2. (a) Maize experimental plot; (b) UAS md4-1000 used to collect the aerial photographs; (c) Boxes placed in the field with different heights; (d) White-colored compact discs placed within the plot as GCPs to georeferenced the SfM point cloud data.

2.2.2 Field measurement

The biomass sampling was carried out following the "Yield and Yield Components Guide" [20]. Three harvests were done on October in random positions along the experimental plot while avoiding the border area. The harvests were executed within a rectangle of 3 m in length and one row in width. The fresh biomass was stored in plastic bags and transported to the laboratory immediately following the cut to prevent dehydration.

Once in the laboratory, the sample was weighed to assess total fresh weight, then a subsample of roughly 1kg was collected, weighed and dried in the oven at a temperature of 75°C for 48 h to calculate dry matter content and estimate biomass per unit area (Table 1) (Figure 3).

The field-observed maize biomass can be estimated using the Equation (1).

$$Dry\ matter\ biomass\ yield\ (g/\mathrm{m}^2) = \frac{Fw - M}{A} \quad (1)$$

where Fw is total fresh weight, M is the moisture content and A is the area of harvest. The experimental plot had a harvested biomass of 2874.2 g/m².

Table 1. Field- measured parameters in the field

	Value
Total harvested area (m²)	3.6 m
Moisture (%)	0.65
Moisture content (g)	19719.91
Dry matter biomass yield (g/ m²)	2874.21

(a) (b)

Figure 3. (a) Harvested area of 3 m in length and one row in width; (b) Subsample (1kg) of maize to be weighed and dried to calculate dry matter.

2.3. Data processing

2.3.1. Pre-Processing of UAS data

Pix4Dmapper Pro software (https://www.pix4d.com, Lausanne, Switzerland) (version 3.2.14) was run on a computer Intel (R) Core(TM) i5 Windows 7 64-bit with 16 GB RAM to merge the two flights into one project. The final point cloud data had an average density per m^3 of $2.24e^{06}$ and a total of 262,485,788 3D densified points. The set up parameters are described in the table 2.

Table 2. Technical specifications of the point cloud generated in the software pix4D

Processing option	Features
Image scale	1 (Original image size, slow)
Point density	High
Minimum number of matches	3
3D textured mesh generation	Yes
3D textured mesh resolution	High resolution
Number of calibrated images	489 out of 489

Information is taken from the Pix4D quality report

2.3.2. Processing of point cloud data

The point cloud was first cropped to contain only the area of interest, and then it was registered using the GCPs coordinates (UTM Zone 32N) with the open source software CloudCompare (http://www.cloudcompare.org, Sunnyvale, CA, USA) by applying the Iterative Closest Point (ICP) algorithm. The ICP algorithm takes two point clouds, A and B. It first performs matching by identifying each point in cloud A and its nearest counterpart in cloud B. The distance between all pairs of points is the error metric. The second step is to minimize that error, before ultimately performing the registration [21].

The second step was to classify the points using LAStools (https://rapidlasso.com/lastools/rapidlasso GmbH, Gilching, Germany). The las_ground command was used to classify the points as ground points (class=1) and non-ground points (class=2) for bare-earth extraction, choosing the option "wilderness" for high-resolution point clouds. Further parameters were set to the defaults (units in meters, no threshold, no hills). Then, to normalize the point cloud, the command las_height was applied, which computes the height of the points by triangulating the points into a Triangular Irregular Network (TIN) and calculating the elevation of each point within the TIN. The option "replace_z" was chosen such that the Z value for each point defines the height as the absolute height.

Finally, the resulting normalized point cloud was divided into the boxes using the ground truth information. One group with five areas was used to estimate maize volume, and a second group with five areas was used to estimate maize biomass (Figure 4). The polygons were drawn in ArcMap with areas ranging from 1 to 5 m². These polygons were used to cut the point cloud with the las_clip function. The point cloud subsamples were cleaned and filter in CloudCompare to contain only maize points without ground data or outliers.

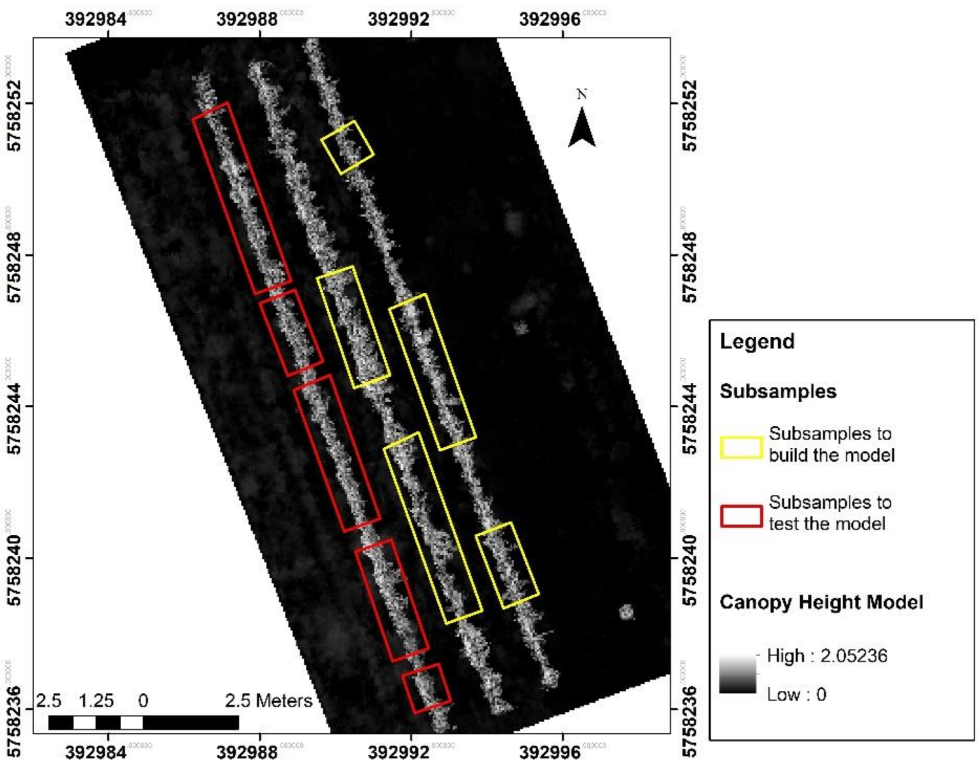


Figure 4. Geolocation of the subsamples used to build the biomass predictor model and to predict the final biomass. Each polygon has an area from 1 to 5m²

2.3.3. Biomass estimation

The processed point cloud was used to calculate maize biomass. Two different volumetric approaches were tested for estimating biomass: voxel-counting and convex-hull. The boxes subsamples were first used to validate the estimated volumes according to the measured volumes. After this step, the volumetric approaches were applied to estimate maize volume using the first group of polygons. The harvested biomass values were scaled to 1 to 5 m^2 and regressed against the volume of the maize to build the biomass predictor model and evaluate its accuracy. Finally, the model predictions were evaluated for the second group of polygons to estimate the biomass of maize and to measure its performance.

2.3.3.1. Voxel-counting approach

The volumetric pixel approach for volume estimation was chosen to take advantage of the 3D property of the point cloud data. A voxel is defined as a volume element in a 3-D array [17]. A script was written in RStudio (https://www.rstudio.com, R Core Team, Vienna, Austria) (R version 3.5.1) based on previous research that has used a counting approach [22]. The point cloud (Figure 5a) is divided into voxels (voxelization) (Figure 5b-d), where all points within the data set are converted into voxel coordinates using the following equations:

$$i = Int\left(\frac{X - X_{min}}{\Delta i}\right) \tag{2}$$

$$j = Int\left(\frac{Y - Y_{min}}{\Delta j}\right) \tag{3}$$

$$k = Int\left(\frac{Z - Z_{min}}{\Delta k}\right) \tag{4}$$

where (i, j, k) are the voxel coordinates into the voxel array, Int is a function to round off at one decimal place to the nearest integer, and (X,Y,Z) represent the point coordinates of the registered SfM point cloud data [17].

The voxel size is determined by estimating the point cloud spacing distance. Las_info was used to estimate this value to be 0.03 cm. The final voxel size was set 0.03x0.03x0.03 cm. After creating voxel coordinates, empty voxels were ignored to create the final voxel model (Figure 5f) and voxels occupied by at least one point were counted to estimate the total volume.

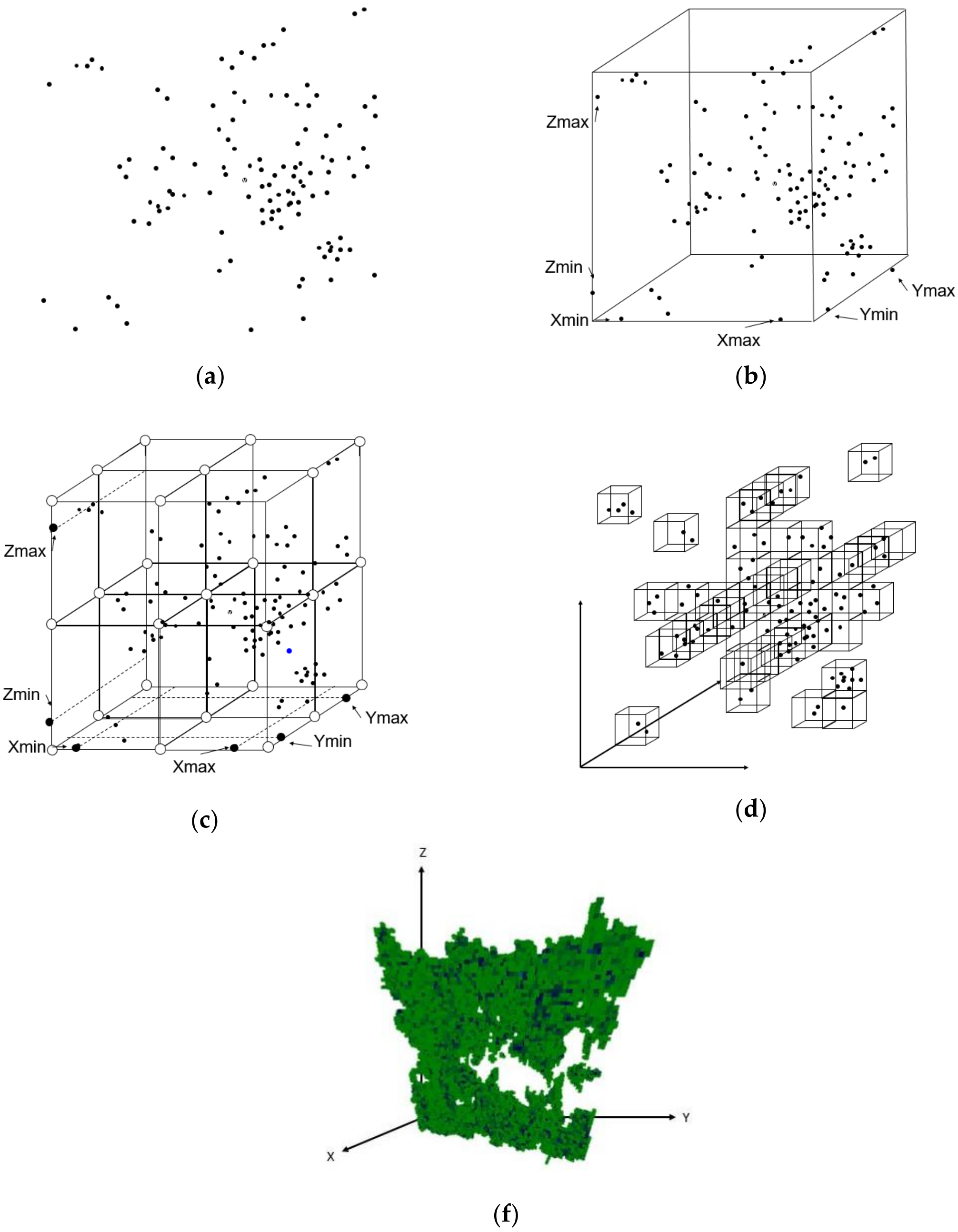


Figure 5. (a) Raw point cloud data; (b) Estimation of the minimum and maximum coordinates values for each axis to draw a bounding box; (c) Grid structure formed by subdividing the minimum bounding box of a point cloud into voxels; (d) Final voxel grid structure with the final voxel size; (f) Real volumetric voxel-model of maize.

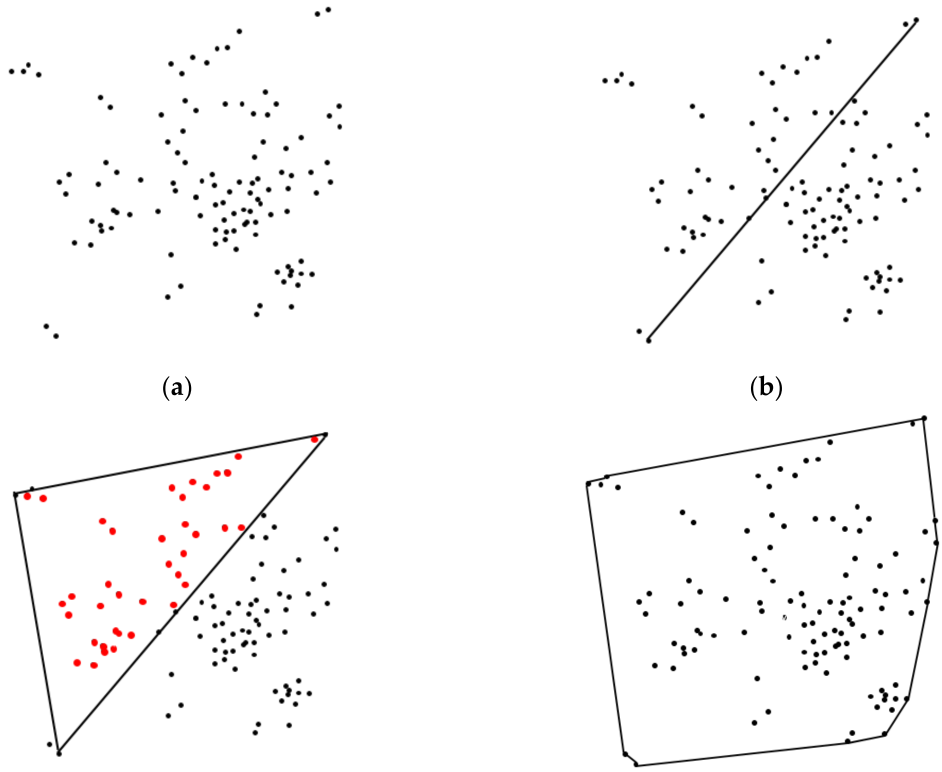
2.3.3.2. Convex-hull approach

The second volumetric approach applied was convex-hull. The convex-hull analysis was performed using the library "rLiDAR" [23] for R. The function chullLiDAR3D computes the 3D convex-hull (and its surface area and volume) with the point cloud data as the input. It uses the quickhull algorithm [24], which is a method that computes the convex-hull of a finite set of points in a plane (Figure 6). It uses a divide and conquer approach, where:

1. The farthest two points are determined and will define the convex-hull polygon.

2. The points are divided into two parts split by a line drawn between the farthest points. Both parts are processed recursively (Figure 6b).
3. The line becomes a triangle by adding the third farthest point to the convex-hull, excluding all points inside this triangle from the convex-hull for further processing steps (Figure 6c).
4. The process is repeated several times for both sides of the dividing line. The process ends when the starting and ending points are reached (Figure 6d).

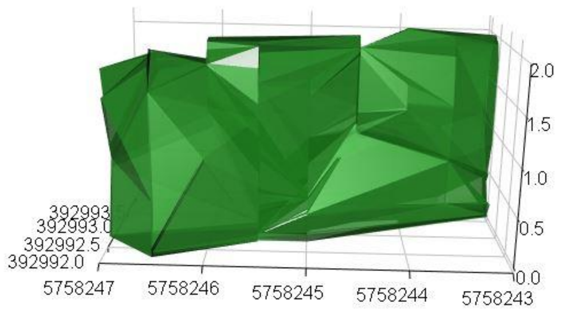
Additionally, K-means was implemented to optimize the convex-hull performance, which narrowed down the empty spaces (gaps) and, improved the convex-hull adjustment to the point cloud. Different numbers of clusters were tested empirically to identify the optimal number of clusters (from 2 to 6 clusters).



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

(**c**) (**d**)

(**e**)



(**f**)

Figure 6. (a) Raw point cloud data; (c) Division of the point clouds into two parts by a line between the farthest points; (d) Determination of the next farthest point to build a triangle in which all remaining points

inside will be ignored for further steps; (e) Final convex-hull after several iterations; (f) Real volumetric convex-hull model of maize.

2.4. Statistical method

The first step consisted of validating the volume of boxes prediction by comparing the error of estimation. This was evaluated using the Sum Squared Errors (SSE) (Equation (5)). SSE estimates the total deviations of the predicted values from the empirical measurements.

$$SSE = \sum_{i=1}^{n} \{w_i (y_i - x_i)^2\} \tag{5}$$

where y_i is the observed data value, x_i is the predicted value from the model. w_i is the weighting applied to each data point, usually $w_i = 1$.

The volume of maize measurements derived from the first step was regressed against the harvested biomass using Ordinary Least Square (OLS). The performance of the linear regression model was evaluated using the adjusted coefficient of determination (R^2).

Finally, the Root Mean Square Error (RMSE) (Equation (6)) was used to assess the final maize biomass predictions for the tested models.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - x_i)^2}{n}} \tag{6}$$

where y_i is the observed data value, x_i is the predicted value

All analyses were carried out using Rstudio.

3. Results

3.1. Volumetric models

The estimation of the boxes volume was quite similar with both approaches. Voxel-based volume estimation had a smaller error ($SSE = 1.7288^{-05}$) that convex-hull. Volume estimations with voxels were closer to the observed volume. Convex-hull-based volume overestimated the volume and its estimation error was slightly higher ($SSE = 0.0155$) than voxel-counting approach (Figure 7).

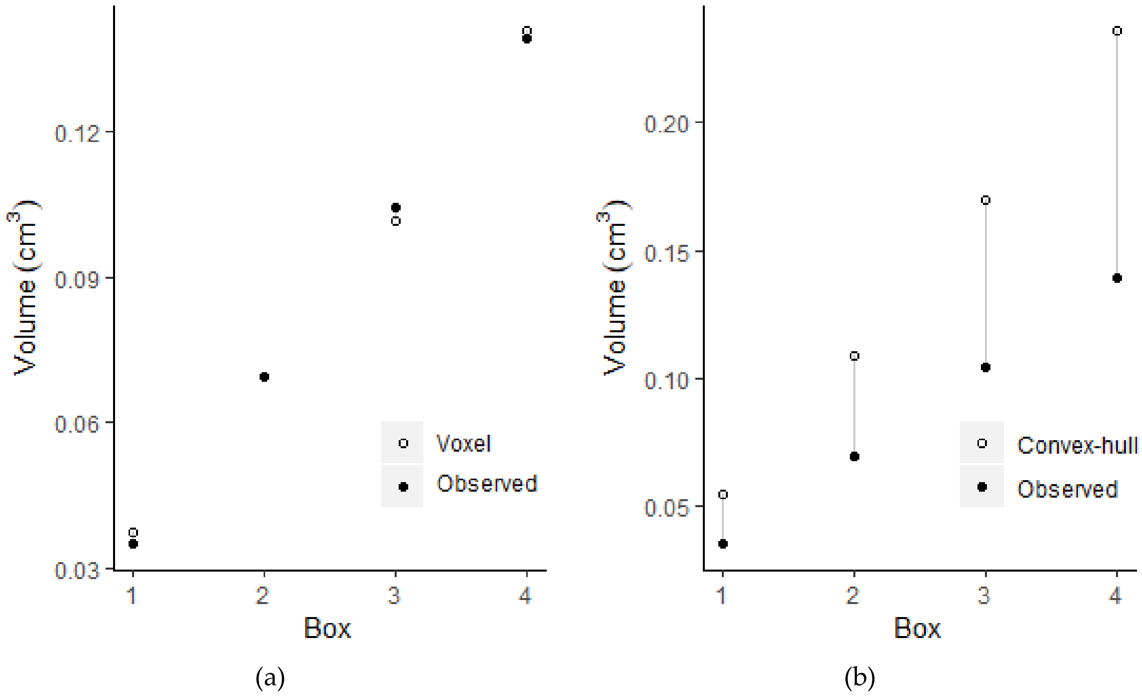


Figure 7. (a) Observed volume against estimated volume with voxel approach, the estimated volumes are closest to the observed volumes; (b) Observed volume against estimated volume with convex-hull approach, the estimated volumes were overestimated.

3.2. Biomass model predictor

The correlation value between the estimated volume and observed biomass was calculated. Estimated volume with convex-hull and observed biomass was slightly lower correlated (0.985) in contrast with voxel volume which showed a higher correlation (0.989) with observed biomass. Regarding the coefficient of determination the voxel biomass predictor model has a value of R^2=0.973 and convex-hull biomass predictor model R^2= 0.961 (Figure 8).

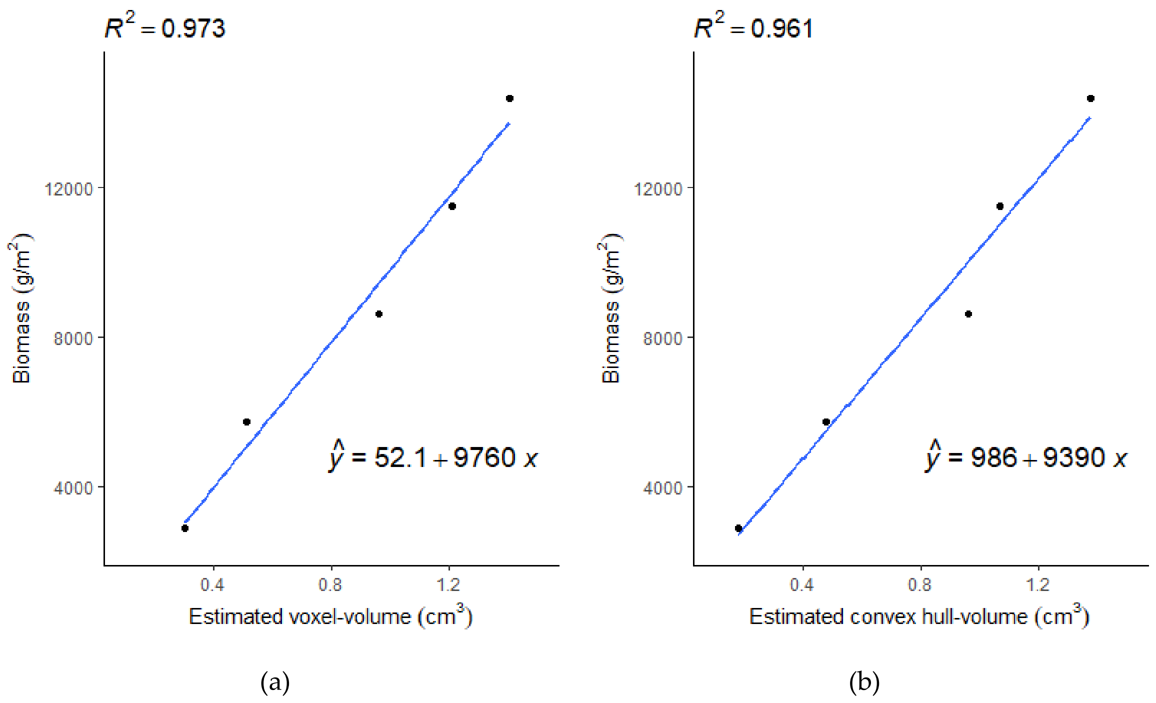


Figure 8. (a) Plot voxel predictor model, the data have 0.989 correlation; (b) Plot convex-hull model predictor model, the data have 0.989 correlation.

3.2.1. Biomass estimation

The voxel-based predictor model yielded a lower RMSE (*RMSE*= 344.5g) than the convex-hull-based predictor model (*RMSE*=1835.6 g). The estimated values using the voxel approached fitted better the linear model. (Figure 9).

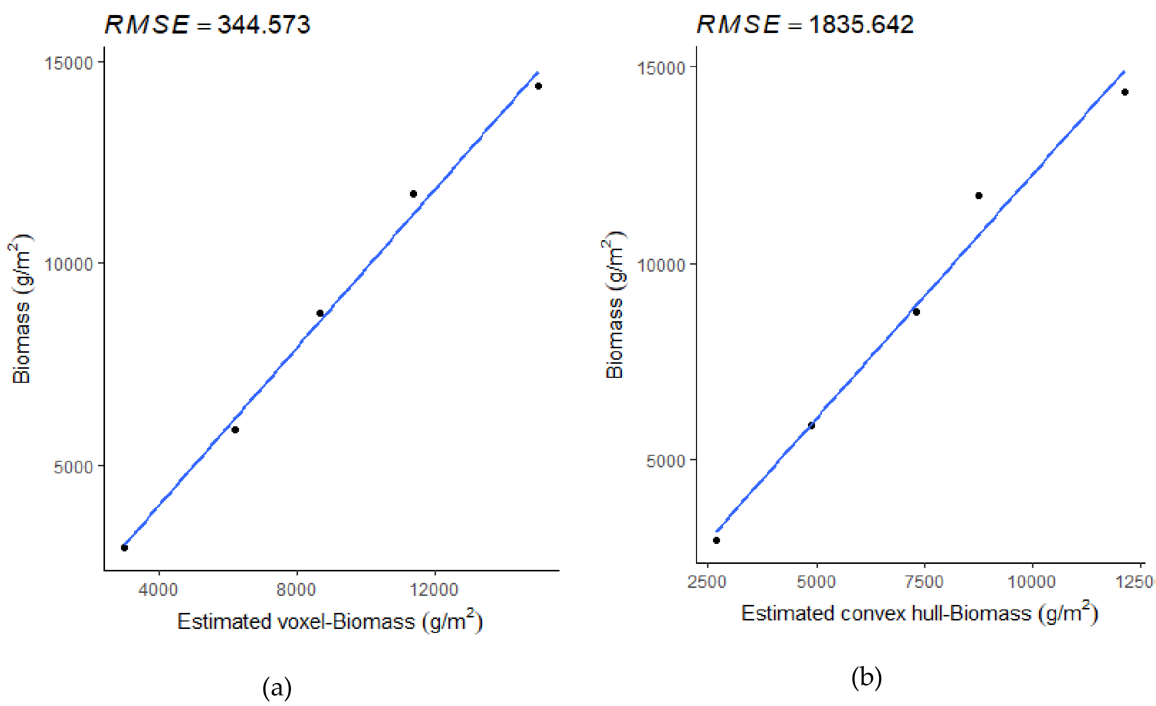


Figure 9. (a) Voxel-based estimated biomass fitted better the observed biomass. (b) Convex-hull-based estimated biomass tends to over and underestimates biomass due to its high RMSE value.

4. Discussion

The result of this study shows that SfM point cloud data obtained with UAS can be processed and is suitable to produce reliable volumetric models to yield accurate maize biomass estimations. Both voxel and convex-hull approaches provided accurate predictions for each subplot. However, the voxel-counting approach was more accurate.

The first aim of this study was to draft a volumetric model predictor based on the SfM point cloud data and two different volume estimation approaches: the convex-hull and voxel-counting. The SfM algorithm generated a high-resolution 3D point cloud similar to that produced using LiDAR. The point density and resolution relied on camera and UAS platform. As SfM point cloud data in this study is based on RGB data, the point cloud can be rendered directly in color, while traditional LiDAR requires the simultaneous acquisition of RGB data. This feature made it easier to interpret and georeferenced because it was faster to locate the GCPs using visual inspection. Whereas with LiDAR this task is more complex due to it only can be colored using the Z value information. Therefore SfM point cloud data is a low-cost alternative to LiDAR data to estimate maize biomass. Some research has proved the effectiveness of SfM point cloud data for estimating biomass. Li *et al.* [29] estimated maize biomass using SfM point cloud from stereo images and Messinger *et al.* [31] applied SfM point cloud to estimate biomass of Amazon forest.

The SfM point cloud data had some gaps where Pix4Dmapper Pro could not match enough features from overlapping images due to occlusion areas where the scan could not reach. The points were classified into ground (6,857,944 points) and non-ground points (45,385,807 points), the presence of other crops and grass between the rows maize increased the number of non-ground points for this reason field preparation task as clearing away herbs is recommended before scanning the area. For this study, this problem was fixed applying significant manual editing of ground versus non-ground points was performed as suggested for Jensen and Mathews [46]. For removing slope effects, the point cloud was normalized by setting Z values as 0 m to all ground points, where the maximum Z value was 2.3 m. The point cloud was split into three subsets. The first subset contained the boxes for model validation accuracy assessment, whereas the second was used to estimate maize volume and correlate them against the observed biomass using scaled values from 1m^2 to 5 m^2. Finally, the third subset was applied to estimate maize biomass.

The boxes that were laid down in the experimental plot as ground truth similar to Ruggles and Clark research [44] comparing the observed volume and estimated volume for each approach to help ensure model quality. Voxel and convex-hull approach both models performed well, but the statistical analysis depicted that voxel volume estimates were closer to the real volume values ($SSE = 1.7288^{-05}$). Possible reason for the deference between models may be rely on voxels based methods assume that no part of the target is occluded [22]. The flatness and no obstacles conditions of the terrain helped to have a good scan of the boxes and maize reducing the occlusion problems and SfM Regarding the convex-hull approach, its volume estimates showed a straight overestimation of the real volumes.

The second aim of this study was to assess the biomass predictor model accuracy. After volume assessment, the maize estimated volume generated with convex-hull and voxel approach was regressed with harvested biomass. The voxel approach had better accuracy than convex-hull approach, this good performance may be linked with the two orientation flights and the low flight altitude which increased the number of points, thus the probability to penetrate fully into the canopy increased as well. On the other hand, convex-hull assumes a consistent biomass-to volume ratio across all the plant [25]. Hence, instead of dividing the point cloud as the voxel approach does convex-hull estimates the overall volume. Therefore, the convex-hull values showed an under and overestimation. This result can be explained by the presence of small gaps spread out along the maize scan. It is important to point out that the optimizer parameter was the number of clusters, the model had good performance using 4 clusters, while using less than 4 and more than 5 clusters the under and overestimation was higher.

There were not found previous research applying the voxel-counting approach to estimate maize biomass. The voxel-counting approach left only 3% of the biomass variation unexplained (R^2=0.97). Similar results are reported from previous research working with voxel volume estimation and vegetation biomass. Graves *et al.* [22] Showed accurate prediction (R^2=0.94) for low-stature Arctic shrubs with Terrestrial Laser Scanning. Olsoy *et al.* [25] reported accurate prediction (R^2=0.86) for dryland shrub biomass with TLS. Brüninghoff *et al.* [19] reported lower but still good accurate (R^2=0.69) working with TLS to estimate *Lupinus polyphyllus* Lindl biomass. Kim *et al.* [16] informed accurate prediction of (R^2=0.87) for a 900 m2 plot of tropical rain forest.

The good performance of volumetric techniques lies in the quality of the point cloud. Therefore, if the scan fails to penetrate deeply into the canopy or the occlusion of a target is high, the voxel-counting approach will treat the lack of points as empty voxels, thus the volume could be underestimated [25]. The convex-hull approache could overestimate the correct volume if the plant has an irregular shape with many gaps, convex-hull will draw a polygon taking into account the gaps as vegetation. Moreover, the presence of outliers by noise in the point cloud, shadowing and wind effect could affect the final estimation of volume. Weak biomass estimations can be overcome addressing different approaches; effect of occlusion can be decreased by scanning a target form multiple directions, ray-tracing, transmission ratio calculations, or point cloud modeling to compensate, but such mitigating strategies can be theoretically and computationally difficult [22]. In order to enhance the volume estimation, noise reduction filter and cluster analysis were implemented optimizing the number of clusters reduced both under-and overestimation of biomass, and is thus recommended as a standard processing step to estimate biomass from SfM point cloud data.

The voxel-volume was confirmed as a powerful predictor for biomass estimation, as demonstrated by [19,22,25]. Nevertheless, there were not found previous research applying volumetric approaches and SfM point cloud data for maize biomass prediction. The existing research is using manly TLS to estimate biomass of different vegetable species [19,22,26] and some others combining TLS with SfM point cloud data [27,29].

Finally, the biomass model prediction was tested using the testing subsamples of 1 to 5 m^2. The voxel model performed better (*RMSE*= 344.573 g/m2) than the convex-hull. Where all the predicted biomass values using the volume estimations with the voxel-counting approach are pretty close to the observed values indicating that voxel volumetric models are suitable to estimate biomass due to its strong relationship with maize harvested biomass. This suggests that SfM point cloud data in combination with volumetric models can be used for biomass assessments, therefore is a reliable source to replace harvest traditional approaches when are difficult to perform.

4.1. Outlook

The work field was carried out in the middle of autumn, most of the crops were already harvested limiting the options were to collect the data and limiting the number of plots to use in the experiment, further studies could increase the number of plots and the number of harvested biomass samples to scale this methodology from small to large scale.

The low-cost LiDAR was considered to use parallel to SfM point cloud data to compare results, but there were many technical drawbacks related its use in the field such as battery life and quality of its pieces that hampered its application. Therefore, future studies could focus on the technical development of low cost LiDAR sensor, and be tested against SfM data especially when vegetation in dense or light conditions are unfavorable.

Besides, it will be interesting to apply this methodology to evaluate different kinds of crops and growth stages to test the presented approach here at different vegetation densities and for a variety of pant shapes

5. Conclusions

This research work has proven that maize biomass can be accurately estimated by using both SfM point cloud data and voxel-counting approach, whether point clouds were gathered at a low flight altitude (15m) and two flight orientations (north-south and east-west) were carried out. Additionally, the evaluation suggests that boxes can be used as a ground reference to validate volume estimations.

Comparing estimated and observed biomass for model validation, the value of RMSE was 344.573 g/m2 providing a high confidence maize estimation. Although the final result has high accuracy, it should be read it carefully due to it just applies to the specific conditions in the experimental plot and the number of subsamples taken.

An additional test is needed applying the same methodology from this small to large scale. Increasing the number of plots, the number of harvests and test different crops to evaluate their effectiveness more broadly.

Acknowledgments: I thank the owner of Bioland-Hofladen Herzkamp Günter Herzkamp for allowed us to gather data on his property and we also thank the ifgicopter group (https://www.uni-muenster.de/Geoinformatics/) for the equipment provided to perform this research.

I thank my supervisor Jan Lehmann for his valuable supervision and time invested in this research, my co-supervisor André Große-Stoltenberg for his valuable advice and Thomas Bartoschek for his support in the field.

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

Appendix

```
Volume_model <- function(listlas,vox.size,cluster.number){
  out.file<- ""
  for (k in listlas){
  las=readLAS(k)#Read the LAS file
  N_points<- nrow(las@data)
  las.df = as.data.frame(las@data) # Extract and convert the LAS properties table into df
  scan= las.df[,c(1,2,3)] # Subset the coordinates from the LAS table
  colnames(scan) = c("x","y","z") # Rename the columne names
  for (i in seq_along(scan)) set(scan, j = i, value = as.numeric(scan[[i]])) # Convert into numeric the
data
  precision = nchar(unlist(strsplit(as.character(vox.size),split = ".", fixed = T))[2]) # Make sure that
precision of the voxel boundaries is appropriate to the voxel size
  xi = round(min(scan$x) - vox.size/2, digits = precision) # Buffer the minimum point value by half the
voxel size to find the lower bound for the x,y, and z voxels
  yi = round(min(scan$y) - vox.size/2, digits = precision)
  zi = round(min(scan$z) - vox.size/2, digits = precision)
  scan$x_vox = ceiling((scan$x-xi)/vox.size) # Assign x, y, and z "voxel coordinates" to each point as a
point attribute
  scan$y_vox = ceiling((scan$y-yi)/vox.size)
  scan$z_vox = ceiling((scan$z-zi)/vox.size)
  scan$xyz_vox = paste(scan$x_vox, scan$y_vox, scan$z_vox, sep = "_") # Combine the x, y, and z voxel
coordinates to identify specific voxels
  scan$xyz_vox = factor(scan$xyz_vox) # Make the unique xyz voxel coordinates into factors
  vertices = chull(scan$x, scan$y) # Find the vertices to calculate the area
  vert.coords = as.matrix(data.frame(scan[vertices, "x"], scan[vertices, "y"])) ## create a matrix of the
vertex coordinates
  Area_m2 <- areapl(vert.coords) # Get the area of the polygon
```

```
las.df2 = as.data.frame(las@data) # Create a new LAS df to process the convex hull
  scan2= las.df2[,c(1,2,3)]
  scan2= data.matrix(scan2)
  clusLAS <- kmeans(scan2,cluster.number) # Run a cluster analysis with a specific number of cluster
  id<-as.factor(clusLAS$cluster) # Add a id column to the cluster datatable
  scan2id<- cbind(scan2,id)
  convex_hull_box <- rLiDAR::chullLiDAR3D(scan2id, plotit = TRUE, col="forestgreen",alpha=0.8) # Run and
plot the convex hull analysis
  convex_hull_vol <- mean(convex_hull_box$crownvolume) # Estimate the mean volume of the convex hull

  Filled_voxes <- nlevels(scan$xyz_vox) # Count the number of voxels with at leats one point inside
  Filled_voxes_m2 <- nlevels(scan$xyz_vox)/Area_m2 # Count the number of voxels with at leats one point
inside per m2
  Voxel_vol <- Filled_voxes*(vox.size*vox.size*vox.size) # Estimate the volume with counted voxels
  out.file<- rbind(out.file,N_points,convex_hull_vol,Area_m2,Filled_voxes,Filled_voxes_m2,Voxel_vol) #
Print the table with the values

  table= out.file[-1] #Format the final table
  table<- as.numeric(table)
  table<- t(table)
  table<- t(table)
  data_table<- as.data.frame(matrix(unlist(table, use.names=FALSE),ncol=6, byrow=TRUE))
  colnames(data_table) <- c("N_points", "Convex_hull_vol","Area_m2", "Filled_voxes", "Filled_voxes_m2",
"Voxel_vol")
  }
 return(data_table)
}
```

References

1. Houle, D.; Govindaraju, D.R.; Omholt, S. Phenomics: the next challenge. Nature Review Genetics 2010, 11, 855–866. [CrossRef]
2. Wang, C.; Nie, S.; Xi, X.; Luo, S.; Sun, X. Estimating the Biomass of Maize with Hyperspectral and LiDAR Data. Remote sens 2017, 9, 11. [CrossRef]
3. Jimenez-Berni, J.A.; Deery, D.M.; Rozas-Larraondo P.; Condon, A.G.; Rebetzke, G.J.; James, R.A.; Bovill, W. D.; Furbank, R. T.; Sirault, X. R. High throughput determination of plant height, ground cover, and Above-Ground Biomass in Wheat with LiDAR. Front Plant Sci 2018, 9, 18. [CrossRef]
4. Bending, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating biomass of barley using crop surface models (CSMs) Derived from UAV-Based RGB imaging. Remote Sens, 2014, 6, 10395-10412. [CrossRef]
5. Christiansen, M.P.; Laursen, M.S.; Jorgensen, R.N.; Skovsen, S. Gislum, R. Ground vehicle mapping of fields using LiDAR to enable prediction of crop biomass, 14th International Conference on Precision Agriculture, Montreal, Quebec, June 24-27th. [CrossRef]
6. Li, W.; Niu, Z.; Huang, N.; Wang, C.; Gao, S.; Wu, C. Airborne lidar technique for estimating biomass components of maize: A case study in Zhangye City, Northwest China. Ecol Indic, 2015, 57, 486-496. [CrossRef]
7. Gao, S.; Niu, Z.; Huang, N.; Hou, X. Estimating the leaf area index, height, and biomass of maize using HJ-1 and RADARSAT-2. Int J Appl Earth Obs Geoinf, 2013, 24, 8. [CrossRef]
8. Liu, J.; Pattey, E.; Miller, J.R.; McNairn, H.; Smith, A.; Hu, B. Estimating crop stresses, aboveground dry biomass, and yield of corn using multi-temporal optical data combined with a radiation use efficiency model. Remote Sens. Environ. 2010, 114, 1167–1177 [CrossRef]
9. Pérez-Harguindeguy N, D.S.; Garnier, E.; Lavorel, S.; Porter, H.; Jaureguiberry, P.; Bret-Harte, M.S.; Cornwell, W. K.; Craine, J.M.; Gurvich, D.E.; Urcelay, C.; Veneklaas, E.J.; Reich, P.B.; Poorter, L.; Wright, I.J.; Ray, P.; Enrico, L.; Pausas, J.G.; de Vos, A. C.; Buchmann, N.; Funes, G.; Quétier, F.; Hodgson, J.G.; Thompson, K.; Morgan H.D. ; ter Steege, H.; Sack, L.; Blonder, B.; Poschlod, P.; Vaieretti, M.V.; Conti, G.; Staver, A.C.; Aquino, S.; Cornelissen, J.H. Corrigendum to: new handbook for standardized measurement of plant functional traits worldwide. Australian Journal of Botany, 64, 715 – 716. [CrossRef]
10. Chandler, J.H.; Buckley, S. Structure from motion (SFM) photogrammetry vs terrestrial laser scanning. In *Carpenter*, M.B. and Keane, C.M. (eds.) Geoscience Handbook 2016: AGI Data Sheets, 5th ed. Alexandria, VA: American Geosciences Institute, Section 20.1 [CrossRef]
11. Martinez-Espejo, I.; Caroti, G.; Piemonte, A.; Riedel, B.; Tengen, D.; Niemeier, W. Structure from motion (SfM) processing of UAV images and combination with terrestrial laser scanning, applied for a 3D-documentation in a hazardous situation. Geomatics, Natural hazards and risk, 2017, 8, 1492-1504. [CrossRef]
12. Morgenroth, J.; Gomez, C. Assessment of the tree structure using a 3D image analysis technique—A proof of concept. Urban For. Urban Green. 2014, 13, 198–203. [CrossRef]
13. Miller, J.; Morgenroth, J.; Gomez, C. 3D modeling of individual trees using a handheld camera: Accuracy of height, diameter and volume estimates. Urban For. Urban Green. 2015, 14, 932–940. [CrossRef]
14. Roberts, J.; Koester, A.; Abd-Elrahman, A.; Hasen, G.; Landry, S.; Wilkinson,B. Terrestrial photogrammetric stem mensuration for street trees. Urban For. Urban Green. 2018, 35, 66-71. [CrossRef]
15. Holman, F.H.; Riche, A.B.; Michalski, A.; Castle, M.; Wooster, M.J.; Hawkesford, M.J. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sens*. 2016, *8*, 1031. [CrossRef]
16. Kim, E.; Kyun, L.; Yoon, M.; Lee, J.; Son, Y.; Salim, K. Estimation of voxel-based above-ground biomass using airborne LiDAR data in an intact tropical rain forest, Brunei. Forest, 2016, 7, 16. [CrossRef]
17. Hosoi, F.; Omasa, K. Voxel-Based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning LiDAR. IEEE Transactions on Geoscience and remote sensing. 2006, 44, 3610-3618. [CrossRef]
18. Stoker, J. Volumetric visualization of multiple return LIDAR data using voxels. Photogrammetric engineering and remote sensing. 2009, 75, 109-112. [CrossRef]
19. Brüninghoff, D.; Astor, T.; Hensgen, F.; Wachendorf, M. Lidar-based estimation of extensive grassland biomass invaded by large-leaved lupine (Lupinus polyphyllus Lindl.). Conference: 27th General Meeting

of the European Grassland Federation, Sustainable meat and milk production from grasslands Cork, Ireland, June 17-21th. 2018, 23, 907-909. [CrossRef]
20. International Maize and Wheat Improvement Center (CIMMYT). Yield and Yield Components, A practical guide for comparing crop management practices, CIMMYT, Mexico 2013, p 22. [CrossRef]
21. Magtalas, M.S.; Aves, J.C.; & Blanco, A.C. Georeferencing UAS derivatives through point cloud registration with archived LIDAR datasets. Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2016, 4,195-199. [CrossRef]
22. Greaves, H.E.; Vierling, L.A.; Eitel, J.U.; Boelman, N.T.; Magney, T.S.; Prager, C.M.; Griffin, K.L. Estimating aboveground biomass and leaf area of low-stature Arctic shrubs with terrestrial LiDAR. Remote Sensing of Environment. 2015, 164, 26-35. [CrossRef]
23. Silva, C.A.; Crookston, N.L.; Hudak, A.T.; Verling, L.A.;Klauberg, C.; Cardil, A. rLiDAR: LiDAR Data Processing and Visualization. 2017. Version 0.1.1. [CrossRef]
24. Berber, C, B.; Dobkin, D, P.; Huhdanpaa, H.The Quickhull Algorithm for Convex Hull. ACM Transaction on Mathematical Software. 1996, 22, 469-483 [CrossRef]
25. Olsoy, P.J.; Glenn, N.F.; Clark, P.E.; Derryberry, D.R. Aboveground total and green biomass of dryland shrub derived from terrestrial laser scanning. Journal of Photogrammetry and Remote Sensing. 2014, 8, 166–173. [CrossRef]
26. Hosoi, F.; Nakai, Y.; Omasa. Voxel tree modeling for estimating leaf area density and wood material volume using 3-D lidar data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2013,5, 115-120. [CrossRef]
27. Cooper, S.D.; Roy, D.P.; Schaaf, C.B.; Paynter, I. Examination of the Potential of Terrestrial Laser Scanning and Structure-from-Motion Photogrammetry for Rapid Nondestructive Field Measurement of Grass Biomass. Remote sensing. 2017, 9, 1-13. [CrossRef]
28. Wijesingha, J.; Moeckel, T.; Hensgen, F.; Wachendorf, M. Evaluation of 3D point cloud-based models for the prediction of grassland biomass. International Journal of Applied Earth Observation and Geoinformation. In press. [CrossRef]
29. Li, W.; Niu, Z.; C, H.; Li, D.; Wu, M.; Zhao, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. Ecological Indicators, 2016, 67, 637-648. [CrossRef]
30. Mlambo, R.; W, I.H.; Gerard, F.; Anderson, K. Structure from Motion (SfM) Photogrammetry with Drone Data: A Low-Cost Method for Monitoring Greenhouse Gas Emissions from Forests in Developing Countries. Forest, 2017, 8, 68. [CrossRef]
31. Messinger, M.; Asner, G.P.; Silman, M. Rapid Assessments of Amazon Forest Structure and Biomass Using Small Unmanned Aerial Systems. Remote sensing, 2016, 8, 615. [CrossRef]
32. Kalaitzidis, C.; Heinzel, V.; Zianis, D. A Review of Multispectral Vegetation Indices for Biomass Estimation. Conference: 29th Symposium of the European Association of Remote Sensing laboratories, Chania, Greece, June 2009, 201- 207. [CrossRef]
33. Fang, H.; Liang, S. Leaf Area Index Models. In Encyclopedia of Ecology; 1nd ed.; Elsevier: Amsterdam, The Netherlands, 2008, pp. 2139-2148. [CrossRef]
34. Zhu, X.; Liu, D. Improving forest aboveground biomass estimation using seasonal Landsat NDVI time-series. Journal of Photogrammetry and Remote sensing, 2015, 102, 222-231. [CrossRef]
35. Ren, H.; Zhou, G. Estimating green biomass ratio with remote sensing in arid grasslands. Ecological Indicators, 2019, 98, 568-574. [CrossRef]
36. Todd, S.W.; Hoffer, R.M.; Milchunas, D.G. Biomass estimation on grazed and ungrazed rangelands using spectral indices. International Journal of Remote Sensing, 1998, 19, 427-438. [CrossRef]
37. Gonzalez-Alonso, F.; Merino-De-Miguel, S.; Roldan-Zamarron, A.; Garcia-Gigorro, S.; Cuevas, J.M. Forest biomass estimation through NDVI composites. The role of remotely sensed data to assess Spanish forest as carbon sinks. International Journal of Remote Sensing, 2006, 27, 5409-5415. [CrossRef]
38. Jin, Y.; Yang, X.; Q, J.; Li, J.; Gao, T.; Wu, Q.; Zhao, Fen.; Ma, Hailong.; Yu, Haida.; Xu, Bin. Remote Sensing-Based biomass Estimation and Its Spatio-Temporal Variations in Temperate Grassland Northern China. Remote Sens, 2014, 6, 1496-1513. [CrossRef]
39. Battude, M.; Bitar, A.A.; Morin, D.; Cros, J.; Huc, M.; Sicre, C.M.; Le Dantec, V.; Demarez, V. Estimating maize biomass and yield over large area using high spatial and temporal resolution sentinel-2 like remote sensing data. Remote sensing of environment, 2016, 184, 668-681. [CrossRef]

40. Sharma, L.K.; Bu, H.; Denton, A.; Franzen, D.W. Active-Optical Sensors Using Red NDVI Compared to Red Edge NDVI for Prediction of Corn Grain Yield in North Dakota, U.S.A. Sensors, 2015, 15, 27832-27853. [CrossRef]
41. Schaefer, M.T.; Lamb, D.W. A Combination of Plant NDVI and LiDAR Measurements Improve the Estimation of Pasture Biomass in Tall Fescue (*Festuca arundinacea* var. Fletcher). Remote Sens, 2016, 8, 109. [CrossRef]
42. Hauglin, M.; Astrup, R.; Gobakken, T.; Naesset, E. Estimating single-tree branch biomass of Norway spruce with terrestrial laser scanning using voxel-based and crown dimension. Scandinavian Journal of Forest Research, 2013, 28, 456-69. [CrossRef]
43. How Dense are You, Any Way. Available online: URL (accessed on 12 February 2019)
44. Ruggles, S.; Clark, J. Comparison of SfM Computer Vision Point Cloud of a Landslide Derived from Multiple Small UAV Platforms and Sensors to a TLS based Model. Journal of Unmanned Vehicle Systems, 2016, 4, 246-265. [CrossRef]
45. Howell, T.L.; Singh, K.K.; Smart, L. Structure from Motion Techniques for Estimating the Volume of Wood Chips. In High Spatial Resolution Remote Sensing: Data, Analysis, and Applications. 1nd ed.; He, Yuhong.; Weng, Qihao.; CRC Press Taylor & Francis Group: Boca Raton, FL, USA, 2018.
46. Jensen, J.L.; Mathews, A. Assessment of Image-Based Point Cloud Products to Generate a Bare Earth Surface and Estimate Canopy Heights in a Woodland Ecosystem, 2016, 8, 50. [CrossRef]