CMSE495 Final Report

Team Go Grape, Go White

Devin Whitlock, Ryan Krupp, Brendan Howard, Alexander Grunewald

April 6th, 2025

# Grape yield estimation from Structure from Motion (SfM) data through vision-based techniques

Grunewald A., Howard B., Krupp R., Whitlock D.


## Abstract

The goal of this project is to develop a novel, automated methodology for accurately estimating the number of fruits within grapevine rows using LIDAR and vision-based techniques. Traditional methods involve sampling a small portion of the vineyard, which leads to large margins of error. While experts can estimate quantities within 5% of actual yield, there is a need for a scalable, real-time solution that can be applied across large vineyards. The project will explore the use of handheld LIDAR (e.g., GEOSLAM) and structure-from-motion techniques to analyze data. The aim is to develop a heuristic model based on the shape, structure, and physical properties of the grape bunches, which can be trained to estimate fruit yield accurately. The method was originally proposed to apply to a wide variety of grapes, though this sample set was later restricted to only purple grapes to assist RGB-based filtering methods.


## Introduction

Wine has been a staple of human culture for millennia, with grape cultivation and harvesting practices evolving alongside winemaking traditions. Accurate estimation of grape yields is critical not only for wine producers to manage logistics, labor, and production but also for ensuring economic sustainability and quality control. Traditionally, grape yield estimation involves manual sampling techniques that assess only small sections of a vineyard, such as the

Average Cluster Weight and Number Method (CV, CW) according to Washington State University [1]. While expert estimators can achieve reasonably accurate results—often within 5% of actual yield—these methods are time-consuming and labor-intensive, as viticulturalists must go out onto the fields and count or weigh the grape clusters manually.

Advances in sensor technology and computer vision, have paved the way for the use of Light Detection and Ranging (LiDAR) and Structure from Motion (SfM) technology as an alternative to traditional grape yield estimation methods as they are inexpensive, scalable, and potentially capable of real-time sampling that could be applied across large vineyards. However, the potential use of LiDAR technology and SfM has not been extensively explored.

Our team aims to explore the use of handheld LiDAR and SfM data to assess the viability of grape yield approximation with the use Cloud Compare, an open-source tool used to analyze point cloud data, R, and Python. We begin by outlining the methods used, starting with an RGB filtering-based grape cluster detection method applied to the LIDAR data. This is followed by a combined RGB filtering and Clustering approach for detecting grape clusters from SfM data as well as a Neural Network classifier. In the Results section, we summarize our findings and evaluate the performance of these techniques on Cabernet Franc and Concord Grapes. Finally, we will discuss our findings and propose future directions of research.

# Methods

## LiDAR RGB Filter

The LiDAR RGB filtering method used in this project was adapted from the approach presented by Jorge Torres-Sánchez et al. [2]. The paper used an open-source R package named LidR, which is mainly used for manipulating and visualizing airborne laser scanning (ALS) data with an emphasis on research & development for forestry and ecology applications. We chose to replicate their approach, since in our opinion, it was a good baseline approach to isolate potential grape cluster candidates. The algorithm is as follows:

1. Decimation: We decimate the point cloud based on the overall average density. We used a 32.8 $ft^2$ grid, which is different from the size used in the paper (5 $m^2$ / 53 $ft^2$).
2. After the point cloud decimation, we removed the ground using the Cloth Simulation Filtering (CSF) algorithm. CSF simulates artificial cloths that attach themselves to the ground (modeling the slopes and elevation changes within the point cloud). Using the interaction of the cloth with the ground, the cloth nodes are used to generate a model of the terrain surface (fig. 1).

3. After obtaining a ground classification, we rasterized the terrain and normalized the height of our point cloud. The reason we do this is to create filtering conditions that allow us to separate areas where no grape clusters can be found. Based on literature, grapes like Riesling or Cabernet Franc are grown in a Vertical Shoot Positioning (VSP) system [2]. Thus, we can expect grapes to be found between heights of 3 ft and 7 ft.
4. Filtering was the last step. We filtered out ground points, heights of unlikely grape-growing regions, and the color of the specific grape. The results can be seen by writing the LidR object to a file and viewing it in Cloud Compare.

**Figure 1:**

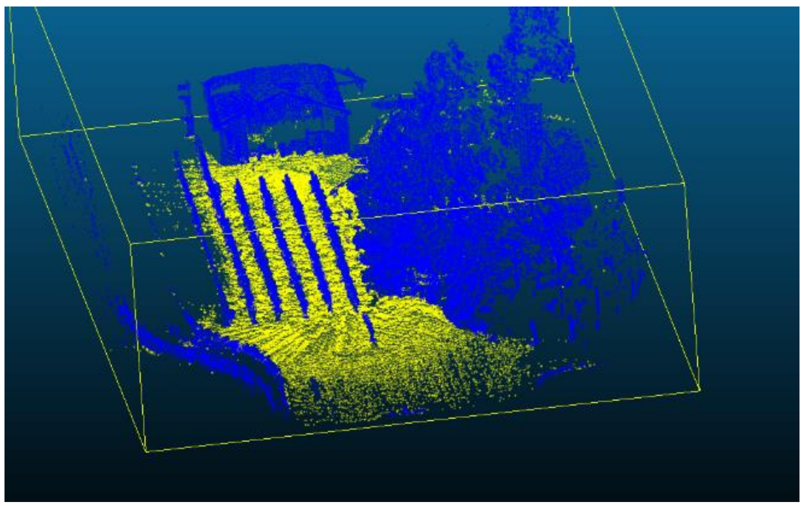

**Figure 1. Ground Classification using CSF algorithm.** Yellow points are ground classified points, and blue points are non-ground points.

## SfM RGB Clustering

Clustering the grapes within the SfM clouds was done using only methods available in Cloud Compare. Once the initial SfM clouds (representing row segments) were processed and uploaded into the Cloud Compare environment, the cluster candidate points were found using RGB filtering. As the cabernet franc grapes within this study retained their strong purple color within the SfM clouds, the RGB filtering was conducted by isolating the green color channel as a scalar field and restricting its range until only the clusters remained. The appropriate green channel range for each SfM cloud was found by incrementally decreasing the isolated channel limit until only the grape clusters remained in the subcloud (a reasonable amount of noise between clusters was deemed acceptable). These subclouds of candidate points were then processed using Cloud Compare's built-in octree-segmentation tool with tuned parameters (octree level = 9; number of points = 2000). The parameter optimizations were completed by roughly maximizing the number of grape clusters contained by the candidate points and minimizing the amount of noise. Resulting clusters were then individually colored for visual distinction and summed across each segment-containing SfM file.

## Vineyard Yield Estimation

The grape yield estimation for the vineyard can begin now that we have the SfM grape clusters isolated from the surrounding foliage in each row segment. Our goal was to determine the volume of points within each of the grape clusters' bounding box. Finding the volume of points would help determine the number of grapes, which in turn would reach the average number of grapes found within each cluster for the given row segment. We committed to this process for the entirety of the SfM row segments until each of the grape clusters detected had a volume and number of grapes assigned to them. Then, for each row segment, the average number of grapes in a cluster was found by taking the average within the entire row segment. Once each row had an average number of grapes in a cluster assigned, then the average number of grapes and average number of clusters within a row segment for the entire SfM cloud set was found. Finally, some simple arithmetic, dimensional analysis, and unit conversions were used to calculate the final grape yield estimation in pounds. These steps included the total number of grapes found in the row segment given these values, an estimation of the number of row segments within the actual row, finding the total number of grapes within the whole row, and converting the weight of one Cabernet Franc grape to the total number to get the final yield result in pounds. Since the SfM cloud data only showcased one side of the row segments, this final number was doubled to compensate for this to get a realistic yield. This number was then compared with the actual grape yield of the row to determine the effectiveness of our methods in producing our own result.

# Results

The octree clustering algorithm used for the SfM data was the only method that reliably produced conceivable grape clusters. These clusters (fig. 2) were plausible (based on canopy height location and general shape) and most resulting clusters correlated with a true grape cluster when layered with the unmodified SfM-point cloud.

**Figure 2:**

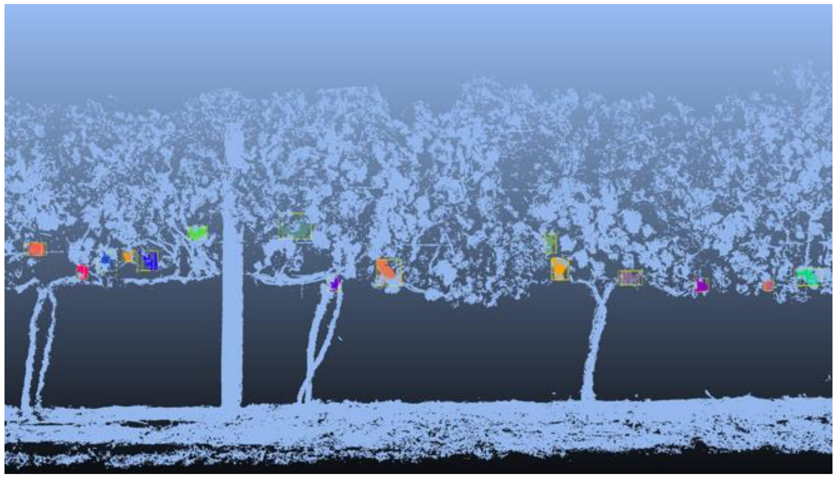

**Figure 2. Clustering Visualization of SfM Point Cloud Section.** The grape clusters found through the octree segmentation methodology were randomly colored and layered on top of the original SfM point cloud (colored differently for contrast).

The grape yield estimation process following the octree clustering algorithm granted a total grape yield of roughly 428 pounds within the Cabernet Franc vineyard row through a process of basic arithmetic, dimensional analysis, and unit conversions. This number derived from the fact that there was an average of 157 grapes within each grape cluster and 19 clusters within each row segment we analyzed. These two numbers multiplied together left us with around 2983 grapes within a row segment on average. Next, the rough dimensional analysis of our row segments and vineyard row gave us an estimation for total number of row segments within the row, which was about 20 segments. Therefore, the number of grapes within the entire row would be approximately 59660. According to Michigan State University, the weight of a single Cabernet Franc grapes ranges from 0.0033-0.0039 lbs [3]. Since 0.0036 lbs. is in the middle of this range, we used this weight to determine that there were roughly 214 lbs. of grapes found within the vineyard row. However, to compensate for the SfM cloud only displaying one side of the row segments, this yield was doubled. The resulting total grape yield ended up being approximately 428 lbs across the entire Cabernet Franc row.

# Discussion

The LiDAR RGB filtering approach did not yield the results we had hoped for. The resulting grape candidates were too sparse and did not resemble true grape clusters (Fig. 3). We suspect this was primarily due to the insufficient point cloud density and the similarity in RGB values between the grapes and surrounding leaves. This made it difficult to reliably distinguish grapes from foliage. We suggest that for a future project, the group should acquire higher quality LiDAR data. This would have to be done at peek growing seasons, which happens generally in October so that grape clusters are ripe enough.

**Figure 3:**

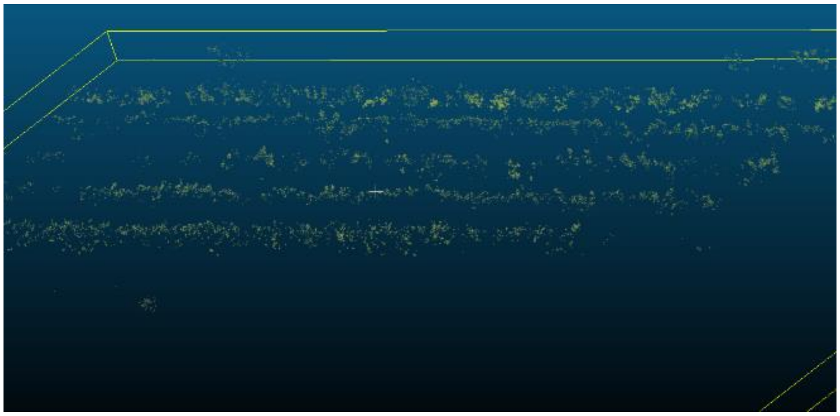

**Figure 3. Final results of RGB color filter approach.** Resulting points are too sparse for DBSCAN and other clustering algorithms to detect grape cluster candidates.

On the other hand, the SfM approach was found to identify grape clusters and predict the total yield for the vineyard with reasonable accuracy. Out of all methods tested within this study, octree segmentation/clustering proved to be the most successful. The actual grape yield of the Cabernet Franc row was given to us beforehand as a reference to see how our estimate performed.  This actual yield was 385 lbs, while our estimate yielded approximately 428 lbs. Now this result could be off for a variety of reasons, such as over classifying the grape clusters in the classification process, too small of a row segment sample size, incorrect estimation of total row segments in vineyard row, incorrect weight of a single Cabernet Franc grape, or overcompensating our SfM cloud data too much by doubling the yield we found originally. However, since we were only about 40 lbs. above the actual yield, this estimation is deemed a successful evaluation of the vineyard's grape yield per row. Overall, this systematic approach allowed us to streamline our grape yield estimation process. Although this approach is not the most precise for estimation, it offers a practical and scalable way to predict grape yields, helping vineyards better manage their resources and improve overall productivity.

Regardless of the small flaws present within the SfM-based approach, it still yielded significantly better results than those obtained from the LiDAR-based methods. In total, it seems that LiDAR as a technology is not as well-suited to grape yield estimation as SfM has been found to be. This difference is likely due to the extremely small scale of grape clusters. With only a limited resolution (at least for handheld applications), LiDAR is unable to capture enough detail to differentiate these miniscule clusters from the rest of the canopy. In contrast to this, the high level of detail that SfM is capable of capturing allows grape clusters to be easily visually differentiated.

The only apparent downside to our approach involves the reliance on RGB filtering—making it nearly impossible to process green grape data. Without an obvious contrast of color from the canopy, our SfM approach is unable to accurately locate any grape clusters. Due to this limitation, we propose that further research in this application should focus on developing a geometry-based filtering technique. By relying on the shape and surface features of the grape clusters, instead of their color, a geometry-based approach could be able to process any kind of grape, regardless of color.

# Alternative Approach: Machine Learning

One of Team Grape's community partners (Phillipe Wernette) has developed a machine learning tool utilizing TensorFlow API that can take point clouds containing vegetation data, and segment said data so that it separates between classes [5]. Multiple programs are used within the model to compute the vegetation indices, their M-Statistics, and Otsu threshold values. M-Statistics is used to denote population mean, and Otsu thresholds minimize the in-class variance of two groups of pixels separated by a thresholding operator. It has been previously used on the SfM grape data before; however, it was not deemed to be quite accurate enough to use in the field. The goal of using this data once again was to see if there was potential for further accuracy, not only compared to the previous attempt but the SfM RGB clustering method. The program was given two data files of SfM cabernet franc grape rows separated between the grape class, and everything outside of said class. The two files were created by taking one file containing all of the data and manually cutting the grape class out of said file. The ML program was then given a new file of a SfM grape row entirely separate from this first, which the ML algorithm attempted to separate the two classes within.

# ML Results

Figure 4

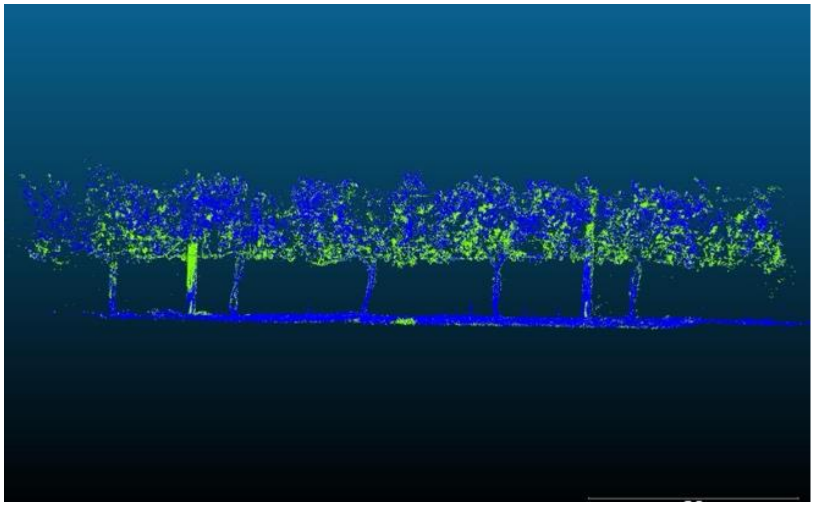

**Figure 4. Neural net predictions for grape classification.** The blue color indicates foliage/canopy, while the green color indicates grape cluster candidates.

As seen in Fig. 4, the neural network was able to classify the points into two distinct classes. Given the nature of the model, one could assume that it was predicting blue to be like the canopy, and green to be its closest approximation of the grapes. Most of the classification for grapes was centered around the underside of the vines, with sparse grouping in other areas such as the second to leftmost post.

# ML Discussion

The data clearly does not pass the eye test for classifying the grapes accurately but does work well as a weak classifier. Comparing the full color file to the resulting classifications, while most if not all cabernet franc grapes do tend to land within the green area, there is a lot of area which is classified as grape that is in actuality canopy. This could be due to the color values still not being differentiable enough for the machine learning approach to find a reasonable difference. Another reason for the lack of accuracy could simply be a lack of training data, which might be solvable by splicing together grape, post, and canopy classes from separate files to achieve higher accuracy or simply having larger files to begin with. The use of this approach can be expanded upon in the future, potentially cascading multiple weak classifiers with the machine learning approach to produce more accurate results. Another possible future exploration could be based on anomaly detection using autoencoders. This approach, which would be similar to the method used in [4], could potentially collapse all the important, and even unrepresented, features into lower dimensionality to more accurately classify the grape clusters.

**References:**


[1] E. Skinkis and M. Moyer, *Vineyard Yield Estimation*, Washington State University, 2014. [Online]. Available: https://www.vineyardteam.org/files/resources/Vineyard%20Yield%20Estimation-%20WSU.pdf

[2] J. Torres-Sánchez, F.J. Mesas-Carrascosa, L. Santesteban, F.M. Jiménez-Brenes, O. Oneka, A. Villa-Llop, M. Loidi, & F. López-Granados, "Grape Cluster Detection Using UAV Photogrammetric Point Clouds as a Low-Cost Tool for Yield Forecasting in Vineyards," *Sensors*, vol. 21, no. 9, pg. 3083, April 2021. [Online]. Available: https://doi.org/10.3390/s21093083

[3] T. Zabadal, "Crop Adjustment in Wine Grapes," SWMREC Special Report no. 24, Michigan State University Extension, revised June 4, 2010. [Online]. Available: https://www.canr.msu.edu/grapes/uploads/files/wine_grape_06-08-10.pdf

[4] Z. Chen, C. K. Yeo, B. S. Lee and C. T. Lau, "Autoencoder-based network anomaly detection," *2018 Wireless Telecommunications Symposium (WTS)*, Phoenix, AZ, USA, 2018, pp. 1-5. [Online]. Available: https://doi.org/10.1109/WTS.2018.8363930.

[5] P. A. Wernette, "Segmenting Vegetation from bare-Earth in High-relief and Dense Point Clouds using Machine Learning". Zenodo, Apr. 12, 2024. [Online]. Available: https://doi.org/10.5281/zenodo.10966854.