Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Automatic flower cluster estimation in apple orchards using aerial and ground based point clouds

Chenglong Zhang [a,1], Christiaan Mouton [b,1], João Valente [b,*], Lammert Kooistra [a], Rachel van Ooteghem [c], Dirk de Hoog [d], Pieter van Dalfsen [e], Peter Frans de Jong [e]

[a] Laboratory of Geo-Information Sciences and Remote Sensing, Wageningen University & Research, Droevendaalsesteeg 3, 6708 PB Wageningen, the Netherlands
[b] Information Technology Group, Wageningen University & Research, Hollandseweg 1, 6706 KN Wageningen, the Netherlands
[c] Farm Technology, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB Wageningen, the Netherlands
[d] Agrosystems Research, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB Wageningen, the Netherlands
[e] Field Crops, Wageningen University & Research, Lingewal 1, 6668 LA Randwijk, the Netherlands

ARTICLE INFO

Article history:
Received 17 November 2020
Received in revised form
28 April 2022
Accepted 9 May 2022

Keywords:
UAV
ground vehicles
structure-from-motion
blossom
machine learning
point cloud

Chemical and mechanical thinning processes have long been used in stone and pome fruit production. During the thinning of apple flowers, growers use chemicals to regulate the tree load. Hand thinning is applied after the June drop to prune trees with excess crop load. The process of thinning can be unpredictable especially in biennial bearing cultivars. Thus, incentives to optimise chemical usage and to reduce expensive manual labour is ever increasing. Ground based machine vision systems have grown in popularity in orchard management due to the level of detail as well as plant coverage they can inspect with. Additionally, unmanned aerial vehicles (UAV) -based remote sensing technology is becoming a popular non-invasive quality inspection solution. This work proposes a framework for combining UAV and ground based RGB image data to detect flowering intensity in a Dutch Elstar apple orchard. The framework, based on point cloud reconstruction, presents automatic point cloud handling techniques as well as automated unsupervised flowering intensity estimation methods. Two linear regression models based on unsupervised machine learning methods were trained and validated from the framework that estimate flowering intensity in the orchard with both models having $R^2 > 0.65$, RRMSE < 20% and p-stat < 0.005 for the correlation between the image derived flower index and the flower cluster number counted in field. The proposed methods provide a novel strategy for guiding flower thinning using simple RGB images and location data only. Moreover, the proposed methods also reveal the flexibility of intra-tree inspection by checking its sub-volumes.

* Corresponding author.
E-mail address: joao.valente@wur.nl (J. Valente).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.biosystemseng.2022.05.004

Nomenclature

UAV	unmanned aerial vehicles
RGB	red, green, blue
SVM	support vector machines
CNN	convolutional neural networks
SD^2_{bio}	biological standard deviation
SD^2_{obs}	observed variance
SE	standard error
WT	west-top observation volume of a tree
WM	west-middle observation volume of a tree
WB	west-bottom observation volume of a tree
ET	east-top observation volume of a tree
EM	east-middle observation volume of a tree
EB	east-bottom observation volume of a tree
RTK	real time kinematic positioning
GPS	global positioning system
PC	point cloud
I_E	images taken from Eastern side using a ground vehicle (.TIFF)
I_W	images taken from Western side using a ground vehicle (.TIFF)
I_U	images taken from a UAV (.jpg)
GV	ground vehicle sensing platform
PC_E	point cloud constructed from I_E only
PC_W	point cloud constructed from I_W only
PC_U	point cloud constructed with I_U only
SOM	spatial orientation model
PCA	principle component analysis
MSAC	M-estimator sample consensus
RANSAC	random sample consensus
FDM	flower detection model
PC_{Hyb}	hybrid point cloud constructed by merging PC_E and PC_W
LOOCV	Leave-one-out cross validation method
UI	user interface
Model-T	flower cluster prediction model for the top volume of a tree
Model-MB	flower cluster prediction model for the sum of the middle and bottom volumes of a tree
R^2	coefficient of determination
RMSE	root mean square error
RRMSE	RMSE relative to mean

1. Introduction

During the process of growing apples, the trees undergo an annual flowering phase in the spring. During this phase, the quantity of flowers can be an indicator to the fruit set (the number of flowers successfully pollinated and becoming fruits) ultimately indicating the resulting yield. At this stage of the cycle the grower would like to perform tree load management to guarantee desired fruit set, which is the management of, in this case, the number of flowers per tree (Zhang et al., 2021). This is where the process of mechanical and chemical thinning, is performed - to control the number of flowers per tree and ultimately the fruit set. Thinning is usually carried out over a rather large period including bloom and post bloom (Greene & Costa, 2013). It is an established method used to control the tree load in order to ensure fruits of a marketable size on a regular basis. A higher than optimal tree load results in higher quantity of a smaller size fruits which is unfavourable for the grower since it leads to less sellable product. Excessive fruit load in apple trees may also result in suboptimal sugar levels, fruit colour and even storage life (Forshey, 1986). Thinning is usually implemented up to the point where fruits are at 18 mm diameter in size. According to Greene and Costa (2013), manual and chemical thinning techniques are used nowadays for stone and pome fruits, but in the case of pome-fruits, specifically apples, the method of choice is mostly chemical thinning to lower labour costs for manual thinning.

Flower and fruit thinning remain unpredictable processes, especially in biennial bearing apple species that can alternate irregularly between high and low bearing seasons (Greene & Costa, 2013). In addition to irregular annual flowering, the fact that individual trees in the orchard vary in flowering intensity, unpredictability is understandable. Since this variation in flowering intensity occurs often, hand thinning, in addition to spraying is still necessary to ensure the correct dosage per tree. It is this unpredictability, the manual estimation of errors and the use of manual labour that holds great potential for improvement. It is therefore beneficial to the grower to know exactly how the intensity of tree flowering or, if possible, how many flower clusters are present per tree. The grower could with this information, for instance, optimise the fruit set per tree, optimise the amount of chemicals used and reduce significantly the manual labour involved in estimating the clusters during flowering and thinning labour. If flower intensity information of individual trees can accurately be accumulated, flowering in terms of spatial variability could also be mapped (allocating for instance each tree with a flower intensity score) and used to make better thinning decisions for individual trees (Farjon et al., 2020). These tree load maps could for instance be used in conjunction with a variable rate sprayer for the thinning process (Krikeb et al., 2017).

Rapid technological advancements have made it more accessible to gather flowering and other plant information in orchards, such as water stress, nutritional status and tree height, in a non-destructive way with the use of machine vision systems on a range of platforms, such as handheld devices, ground vehicles and unmanned aerial vehicles (UAV) (Aggelopoulou et al., 2011; Bargoti & Underwood, 2017; Lopez-Granados et al., 2019; Zhang et al., 2021). With regard to the sensors equipped to the platform within the research focussing on flowers in orchard, only red, green, blue (RGB) and multispectral sensors have been employed, due to the distinct colour feature of flowers against the background (Liakos et al., 2017; Tubau Comas et al., 2019). Theses platforms have their own pros and cons depending on specific application scenario. Though the monitoring conducted with handheld devices is inferior to others when it comes to the data collection efficiency, the data quality, especially the image resolution, enable more features available (Wu et al., 2020). Ground vehicle platforms are also often used in orchard management because of their convenience and data collection efficiency (Dias et al., 2018a; 2018b). For a particular research purpose,

ground vehicle platforms can also provide a unique circumstance that benefits the data collection. For instance, with the help of the artificial lights equipped on the ground vehicle, image collection at night is available which reduces the effects of sunlight significantly (Wang et al., 2021). Current achievements in sensor miniaturization enable successful use of UAVs for orchard management, however, only a few researchers have conducted systematic research into flowering intensity estimation (Zhang et al., 2021). UAVs can not only provide efficient and reliable monitoring for orchards but also a detailed spatial and temporal solution. Research has demonstrated the use of UAVs in flowering intensity estimation and the ground vehicle in flower detection (Horton et al., 2017; Sun et al., 2021). Yet comparison of these two platforms has been subject to considerable discussion. For example, when the spatial information of fruit trees is retrieved, such as the flower cluster number per tree, with 3D point clouds derived from UAV and ground vehicles, the two platforms often show different spatial scales and occlusions because of the different data acquisition locations. Thus there remains a paucity of evidence on which platform is more suitable for flowering intensity estimation in orchards.

Extensive research has been carried out on flower detection, classification and estimation in orchards with machine vision technology. Flower estimation studies have been mostly restricted to estimation at picture level rather than at tree level (Dias et al., 2018a, 2018b; Wu et al., 2020). In recent studies, the detection of flowers has been investigated in two ways, pixel-based and object-based detection methods (Vanbrabant et al., 2020). For pixel-based methods, the first step is the extraction of the pixels of interest. Once these pixels are extracted, flower detection is conducted by the calculation of flower pixel amount or a different fraction. Thresholding techniques are most commonly used within pixel-based methods, where the targeted objects are segmented by transforming a grayscale images into a binary one (Aggelopoulou et al., 2011). This is effective for images with high levels of contrast, and significant outcomes have been observed from previous studies (Horton et al., 2017). Krikeb et al. (2017) reported a R^2 of 0.97 for the correlation between the threshold derived flower areas to the flowering intensity scored given by an expert. A R^2 of 0.86 was observed between the flower intensity and the apple yield for both RGB and multispectral image based detection (Liakos et al., 2017). One disadvantage of thresholding is that the threshold adjustment is needed for a new environment or dataset. By contract, more advanced machine learning techniques are more robust and have attracted interest in the flower detection community. More specifically, the machine learning derived pixel-based methods consist of support vector machines (SVM) (Dias et al., 2018b; Xiao et al., 2014), K-means (Tubau Comas et al., 2019), convolutional neural networks (CNN) (Dias et al., 2018b). For example, a model based on the combination of CNN and SVM proved to be successful with precision and recall rates of 90% (Dias et al., 2018b). Object-based detection is generally conducted with two steps, image segmentation and object classification. Image segmentation algorithms first aggregate pixels into spectrally homogenous objects, and then each of the objects is classified to accomplish the detection (Dias et al., 2018a; Vanbrabant et al., 2020). For the object-based CNN, an image patch is evaluated directly to detect if there is a flower (Chen et al., 2019; Yuan & Choi, 2021). Since object-based methods do the classification at object level, different classes of targeted objects can be determined and labelled during image annotation process. Based on this approach, not only flowers can be detected, but also the specific flowering stage can also be classified (Koirala et al., 2020; Yuan & Choi, 2021).

Earlier studies demonstrated the capability of using machine vision technology for flower detection at picture level. To assist the growers to decide how many flowers needed to be removed for each individual tree, flower quantification at tree level is required to support flower management within the orchard. To our knowledge, there are only five studies focussing on the correlation between the flower estimation derived from image and the exact flower number counted in situ. Three studies were conducted with ground vehicles (Hočevar et al., 2014; Koirala et al., 2020; Wang et al., 2021) and two with UAVs (Tubau Comas et al., 2019; Vanbrabant et al., 2020). Two reported R^2 values of the correlation between the image based flower index and the exact cluster number counted in situ were 0.56 and 0.59 because of the camera capture view, using only top-view and one-side view of the trees, respectively (Hočevar et al., 2014; Tubau Comas et al., 2019). Simple capture view limits the estimation ability of machine vision, especially for the complex-structured fruit trees with more flower occlusion. Fruit trees are larger and have a more complex architecture compared to common agricultural crops, such as cotton, i.e., plant height and flower numbers for fruit trees and cotton, respectively, are 3–4 m versus <1 m and 350–2100 flowers versus 15–20 flowers (Vanbrabant et al., 2020). By contrast, Koirala et al. (2020) employed a dual-view imaging approach for mango panicle estimation at tree level capturing two images from each side of the trees to represent the total number of flowers per tree. Although the highest R^2 of the correlation between the panicle count on images and in-field human counts per tree yielded from the rotational region CNN was 0.86, a high bias was also observed. In addition, the proposed method could not be applied to the flower estimation of apple or pear trees since apple flower estimation suffers from higher flower occlusion and it is much more difficult to label each apple flower cluster in the picture. Wang et al. (2021) reported a CNN-based apple flower estimation, but the ground truth used was part of the tree where a coloured square section marker was used to delineate and limit one counting area from a tree. RGB dense point cloud datasets derived from multi-view images have potential to retrieve tree spatial information, such as the geometric traits, as shown in several studies (Lopez-Granados et al., 2019; Sun et al., 2019; Torres-Sánchez et al., 2018). Vanbrabant et al. (2020) has demonstrated its performance of addressing the pear flower occlusion problems, but the errors in their study could not be explained because of the limitation in ground truth.

This paper aims to develop and evaluate an approach for automated apple flower cluster estimation at tree level. The presented research contributes to the following:

1. Comparison of the capability of four conventional classification algorithms, manual threshold, Otsu

segmentation, K-means and Hierarchical clustering, for flower cluster segmentation;
2. Comparison of the accuracy of UAV and ground vehicle derived point clouds for flowering intensity estimation;
3. Examine the effects of point cloud density on the flowering intensity estimation accuracy;
4. Evaluate the possibility of combining UAV and ground vehicle derived point clouds for flowering intensity estimation.

2. Materials and methods

2.1. Study area and field data

The study area for this project was an apple orchard in Randwijk, Netherlands (51.938, 5.7068 in WGS84 UTM 31U) (Fig. 1). Apple trees in this orchard were planted in 2007 with four-years trees. In total 14 rows were designed, while for each row 101 trees were planted. The average tree height was around 2.3 m in 2018. Row 5 (100 trees) of the orchard is unique from the rest of the orchard with the aim being to intensely monitor the phenotypical change in the row. Hand thinning is needed each year, but the amount of mechanical thinning is very variable. Detailed information about the orchard can be found in Table 1. Ground truth data, flower cluster and floridity, for Row 5 were established by an expert in the orchard. Flower cluster number was counted and recorded manually for 31 random trees in row 5. More specific, for each individual tree, the exact flower cluster number was counted in the top, middle and bottom parts of the tree and documented accordingly. Floridity was determined by giving each tree a flowering intensity rating between 1 and 9 where 1 being no flowers and 9 being extreme flowering intensity. Floridity refers to the intensity at which the tree is flowering while flower clusters are the clustered arrangement of inflorescence.

High within-sample variability can be observed in an orchard with high natural or biological variability (Anderson et al., 2021). To illustrate the within-sample variability in this study, the biological or intrinsic variability, biological standard deviation (SD^2_{bio}), was calculated by decomposing the estimator variance:

$$Var = SD^2_{obs} = SD^2_{bio} + SE^2 \quad (1)$$

where SD^2_{obs} represents the observed variance within the collected samples, and SE represents the standard error.

The ground truth, cluster number and floridity, was acquired on the 24th of May 2018 when trees were fully blooming. And the trees generally scored floridity between 5 and 8.5. Afterwards, an inventory of the collected ground truth was created (Table 2), based on Eq. (1). As mentioned before, floridity is a manually assessed score and was given by the expert in the orchard. The variability within the samples cannot reflect the natural variability. Thus no biological variability, SD^2_{bio}, was calculated for floridity, but the standard deviation and the standard error were provided.

In row 5 of the Randwijk orchard, 100 trees were equipped with observation tape (Fig. 2), which was used by the expert as a frame of reference to compartmentalize trees and ground truth flower count. The observation tape runs throughout the entire row and serves the purpose of maintaining consistency in the counting of flower clusters and floridity of the tree. The tape divides each tree into the top, middle, and bottom observation windows. The tape was placed through the middle of the tree as to additionally divide the trees into East and West halves. Each tree in the row therefore has six observation volumes namely west-top, west-middle, west-bottom, east-top, east-middle and east-bottom (Fig. 2C). The expert used these observation volumes as a guide for assigning floridity and flower clusters scores to the trees and to be able to compare results of different instances and the changes that occur. Only cluster number was counted in detail for these observation volumes. Floridity was assessed from East and West sides and the average was marked as the flowering intensity for each tree. The inventory of the cluster variability within sub-observation volumes was calculated in Table 3.

2.2. Platforms and image acquisition

Images of row 5 were acquired using two platforms with onboard cameras. A ground vehicle equipped with three RGB cameras as well as a real time kinematic positioning (RTK) system, drove through row 5 taking pictures (1 m from the trees) at three different heights (0.75 m, 1.30 m and 1.85 m from the ground respectively). These three cameras were equally spaced vertically to capture, in combination, the full height variation of the trees. The images were taken at a rate of 11 frames per second. The ground vehicle drove through the orchard on the East and West side of row 5. This platform path allowed pictures to be taken from the Eastern and Western side of the trees. The second camera platform was a UAV equipped with a camera and a global positioning system (GPS), taking images of the orchard at a height of approximately 10 m above the orchard canopy. This camera setup offered multiple vantage points of any given tree in the row. Refer to Table 4 for more details regarding the platforms and Fig. 3 for a summary of the data collected. At the time of data acquisition (24 April 2018), weather conditions were cloudy.

2.3. Methods

The approach proposed followed the methodology presented in Fig. 4. To utilise both data sources and merge them, it was decided to assess the trees through the use of photogrammetric point clouds (PCs). It was decided that a PC of row 5 could offer great insight into the spatial variability of the flowering intensity of the trees by allowing not only inter-tree analysis but also intra-tree analysis. This approach proposes to create three separate PCs from three vantage points, handle and analyse them separately and ultimately combine their results at the statistical analysis stage of the framework. Two software, Agisoft Metashape Professional (Agisoft LLC, St. Petersburg, Russia) version 1.6.2 build 10247 and MATLAB R2019b (MathWorks Inc., Natick, MA, USA), were used for the manual and automatic analysis indicated in the framework. This approach allows all three PCs to undergo the same empirical operations without any spatial or colour-space differences influencing their results. It also allowed for the trees in the three PCs to be subdivided and therefore inspected in a more flexible way.

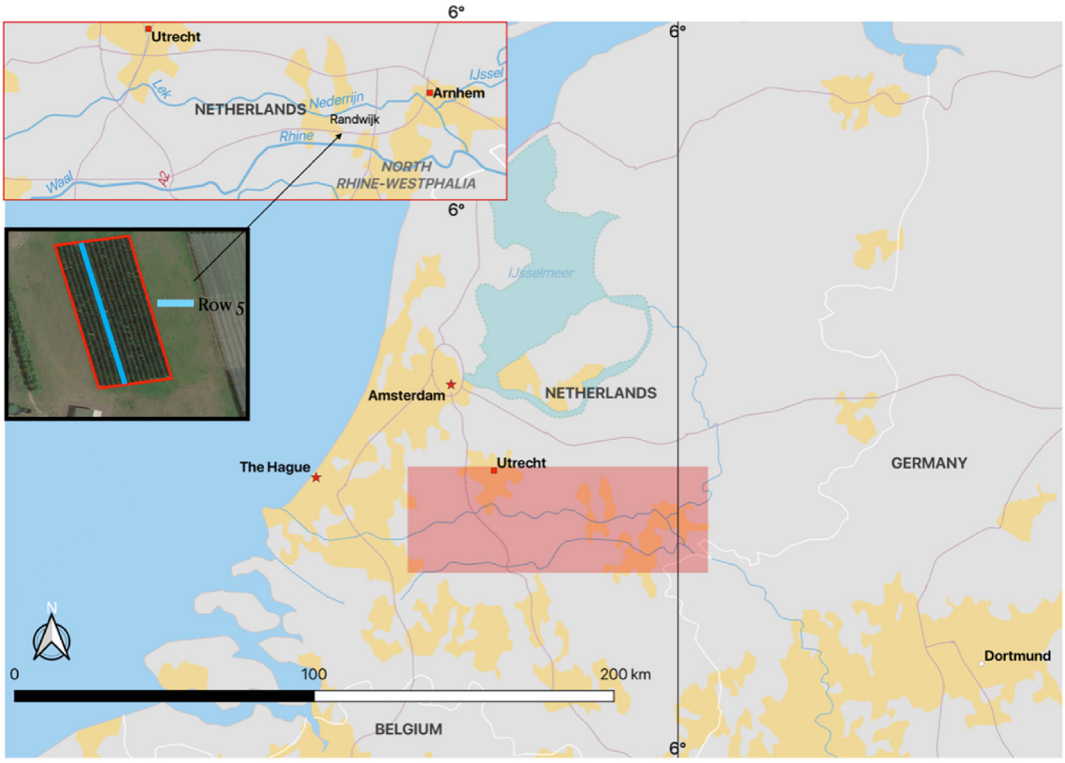

Fig. 1 – The study area and the subject row of this study (row 5). The orchard is delineated with red line and the row 5 is marked with a blue rectangle.

Table 1 – Orchard information and management activities for 2018.

Size	0.47 ha
Variety	Elstar
Rootstock	M9
Row spacing	3.0 m
Tree spacing	1.1 m
Flower thinning chemical	ATS (ammonium thiosulphate)
Fruit hand thinning threshold	110 fruits tree^{-1}
Rows	14

2.3.1. *Image pre-processing*

During the pre-processing phase, raw images and positioning data from the respective camera platforms were manually processed. The altitude of the images was determined by the altitude of the study area and adjusted according to the known height the image was taken. The location data (RTK & GPS) along with the images were loaded into Agisoft Metashape software. Using Agisoft Metashape, location data and image data were used to create point clouds of the row of trees. Two main stages were involved in this process: aligning the images and building the dense cloud. Aligning the photos automatically identified image features, created image pairs and reconstructed camera locations. Moreover the 3D structure of the scene was initialised. Based on the tie points, the images were loaded again and the geometry between the tie points was calculated during the second stage. Detailed processing parameters applied during these two stages are summarised in Table 5. From Fig. 5 it can be seen that the output of the pre-processing phase consists of three versions of point clouds of the same row of trees. Then these point clouds were transferred to .ply format, which is compatible with MATLAB, for the automatic processing in the following two models, spatial orientation model and flower detection model. Density information per point cloud can be found in Table 5. Density difference between the ground vehicle sensing platform (GV) and UAV point clouds can be described as considerable.

2.3.2. *Spatial orientation model*

The input to the spatial orientation model (SOM) (Fig. 6), three point clouds, undergo a process whereby the main goal was to divide each tree in the point clouds into six observation volumes. These volumes and their properties can then

Table 2 – Calculated floridity and cluster variance within trees.

	N (#trees)[a]	Average (# or #cluster/tree)	SD_{obs} (# or # cluster/tree)	SD_{bio} (#cluster/tree)	SE (#cluster/tree)
Floridity	100	6.8	0.7	–[a]	0.1
Total cluster	32	194	64	63	11

[a] "#" represents number; "–" represents no value.

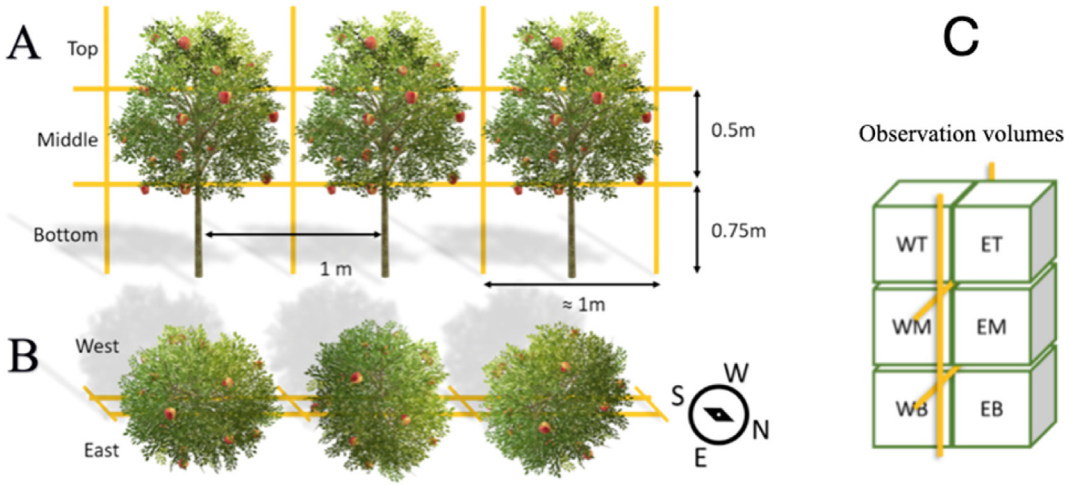

Fig. 2 – Yellow observation tape used for counting purposes. A. Front view; B. Top view; C. Observation volumes. WT: west-top volume of a tree, WM: west-middle volume, WB: west-bottom volume, ET: east-top volume, EM: east-middle volume, EB: east-bottom volume. Note: directional indicator is valid for B only.

Table 3 – Calculated cluster variance within different sub-observation volumes.

	N (volumes)	Average (cluster volume^{-1})	SD_{obs} (cluster volume^{-1})	SD_{bio} (cluster volume^{-1})	SE (cluster volume^{-1})
Top	32	56	30	29	5
Middle	32	75	23	22	4
Bottom	32	62	21	21	4
Total	96	65	26	26	3

be used to detect the floridity and flower cluster counts. As previously mentioned, the aim was to empirically handle the PCs and fuse the results of the flower detection model. The spatial orientation model therefore prepared the observation volumes to be used in the detection phase later in the process.

The first function in the SOM was to down sample the PCs in order to limit computational time and allow for the PCs to be handled more computationally efficient by the succeeding algorithms. The PCs were down sampled randomly from their original count as to not influence the spatial distribution of the input data. To further smooth the point cloud from outliers and reduce the amount of measurement errors and density variations a statistical outlier removal filter was applied (Rodríguez et al., 2018; Sultani & Ghani, 2015). More specific, all points inconsistent with all neighbours were trimmed out by calculating and comparing the mean distances with the interval defined by mean and standard deviation. The global orientation of the PC entering the SOM might be distorted. In a similar fashion to Dong et al. (2020), the global alignment function made use of principle component analysis (PCA) to align the component of highest spatial variance (in this case the length of the PC) to the X-axis. An important function of the SOM was removing the ground beneath the trees (Fig. 7). The ground data points that fall within a specified distance from the ground plane were removed from the PC to leave only the trees remaining. This improves the accuracy of succeeding operations by eliminating the possibility of taking irrelevant data, ground data points, into account.

With ground data points being removed and only trees left the PC now served as input to the PCA alignment function. During this function the PC was aligned in all three spatial dimensions for the second time to offset the result of the first PCA alignment due to ground points. Using PCA, the component with the highest spatial variance (length) was re-aligned to the X-axis. The component with the second highest variance (height) was aligned to the Y-axis. Finally, the component with the least spatial variance (depth) was aligned to the

Table 4 – Camera platform specifications and data size.

	Aerial vehicle	Ground vehicle
Vehicle type	DJI™ Phantom 3 PRO, Shenzhen, China (quad-rotor)	Tractor (generic)
Camera type	FC300X, Shenzhen, China (RGB)	Intel® RealSense™ Depth Camera D435 (RGB-Depth)
Sensor resolution	4000 × 3000	1920 × 1080
Frequency	Variable	11fps
Positioning system	GPS (geotag per image)	RTK (10 Hz)
Acquired images	362	10,355

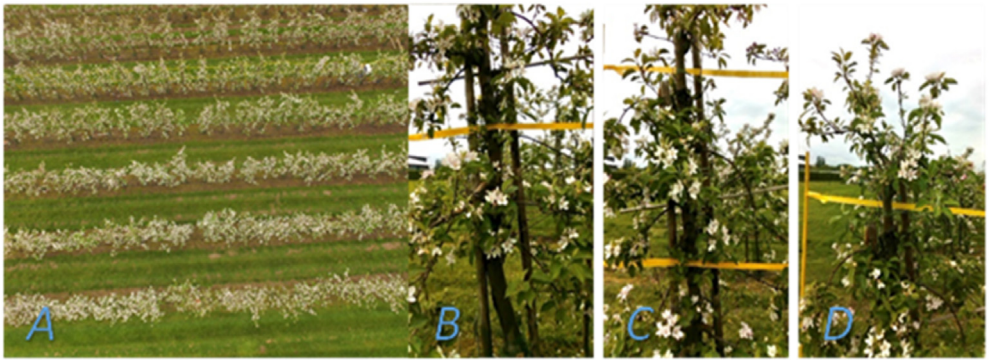

Fig. 3 – Example image data. A: UAV imagery. B: Ground vehicle imagery, camera position: 0.75 m. C: Ground vehicle imagery, camera position: 1.30 m. D: Ground vehicle imagery, camera position: 1.85 m.

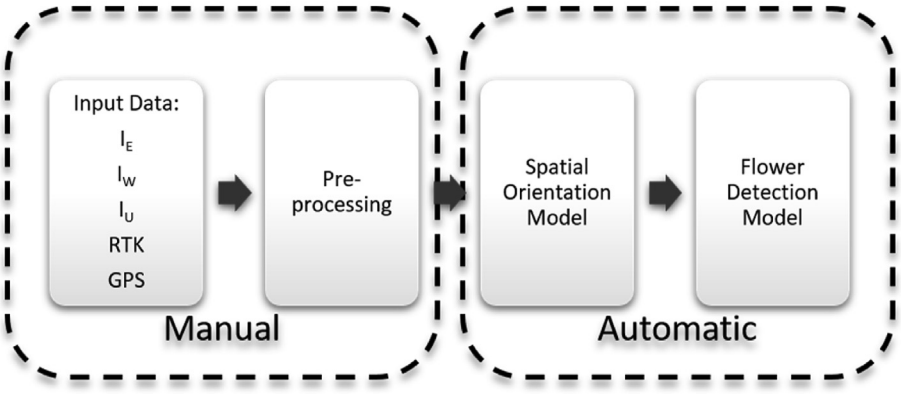

Fig. 4 – Schematic of the proposed framework. I_E and I_W: ground vehicle derived mages taken from Eastern and Western side of a tree, respectively; I_U: images taken from the UAV.

Z-axis (Fig. 8). At this stage of the model, all three input PCs were aligned in the same coordinate system and orientation. Equally important is the distribution of spatial data start from the origin (0) after this function was applied.

The segmentation function automatically isolated all trees based on the vertical observation tape in between trees using a M-estimator sample consensus (MSAC) (Fischler & Bolles, 1981) algorithm, a variant of the random sample consensus (RANSAC). The MSAC was used to fit rectangular planes to the observation tape horizontally and vertically (Fig. 9). Four planes isolated a tree from the rest of the row as well as divide the tree into a bottom, middle and top volume based on the observation tape. The size of these volumes varies from tree to tree depending on how the observation tape was installed in the orchard. The resulting individually segmented trees then passed through to the final function of the SOM. The subject tree enters the final function of the SOM where the tree was divided into six observation volumes (Fig. 10). Every tree in the three input PCs, PC_E, PC_W and PC_U, was processed. The height layers were defined for every tree and the data points that fall within those height layers were divided in an East and West part of the tree. Therefore, at this stage of the framework, for any given tree in row 5, there were three versions with each six subsections.

Table 5 – Parameters of the point cloud generation in Agisoft Metashape and the yielded point cloud density.

	PC_E[a]	PC_W	PC_U
Align photos			
Accuracy	High	High	High
Generic preselection	Disabled	Disabled	Disabled
Key point limit	40,000	40,000	40,000
Tie point limit	4000	4000	4000
Build dense cloud			
Quality	High	High	High
Depth filtering	Mild	Mild	Mild
Calculate point colours	Enabled	Enabled	Enabled
Coordinate system	WGS 84/UTM zone 31N (EPSG: 32631)		
Data points per tree	3.90×10^5	3.30×10^5	0.30×10^5
Tie points per image	158	141	71

[a] PC_E and PC_W: point cloud constructed from the ground vehicle derived Eastern and Western side images of the trees, respectively; PC_U: point cloud constructed from the UAV imagery.

2.3.3. *Flower detection model*

The aim of the flower detection model (FDM) (Fig. 11) was to extract data points, based on their colour properties, that belong to flower clusters. These data points can then be used to find a correlation between the amount of white data points

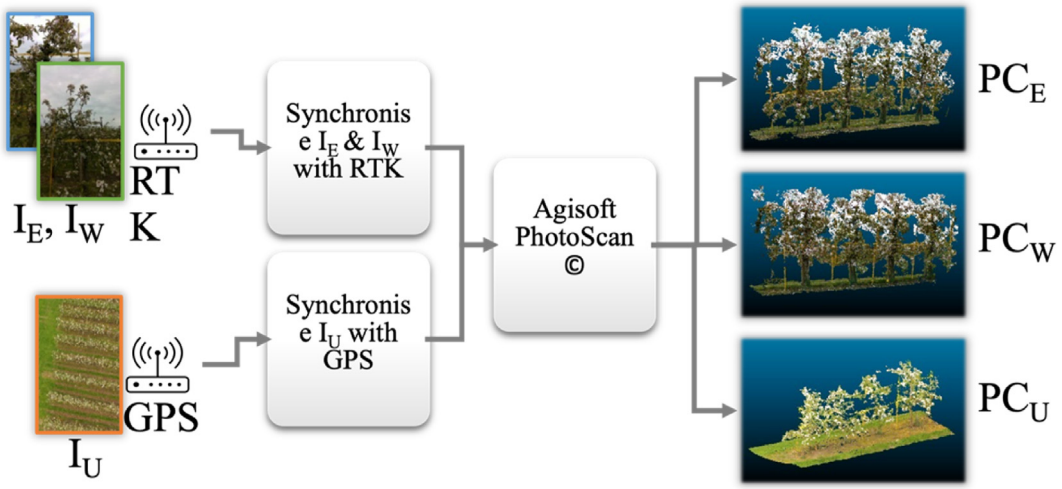

Fig. 5 – Schematic of pre-processing phase with three point clouds as output. I_E and I_W: ground vehicle derived mages taken from Eastern and Western side of a tree, respectively; I_U: images taken from the UAV; PC_E: point cloud constructed from I_E only; PC_W: point cloud constructed from I_W only; PC_U: point cloud constructed from I_U only.

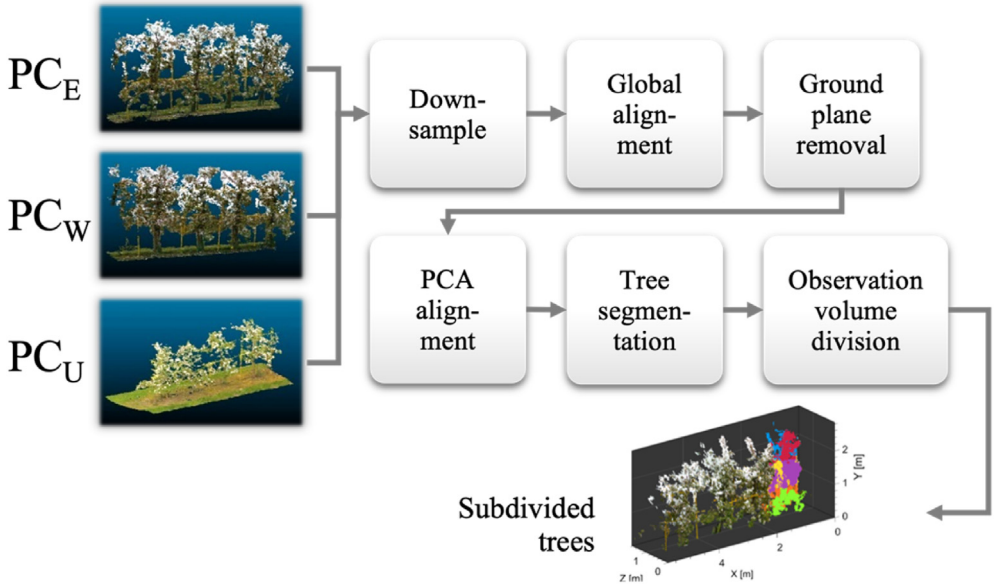

Fig. 6 – Schematic of the spatial orientation model (SOM). PC_E and PC_W: point cloud constructed from the ground vehicle derived Eastern and Western side images of the trees, respectively; PC_U: point cloud constructed from the UAV imagery.

and the ground truth flower clusters in the various sub-sections of the trees.

To segment the data points that belong to white flower clusters in the flower segmentation function, four different methods were used:

1. Manual thresholding
2. K-Means clustering
3. Otsu segmentation
4. Hierarchical clustering

Manual threshold segmentation was achieved by manually inspecting a collection of images (45 from GV and 40 from UAV) from random trees in the orchard. By segmenting all the images manually, average lower and upper thresholds were determined for red, green and blue colour bands that explain the flower cluster pixels. These thresholds were used to segment data points in the point clouds that belong to flower clusters by extracting data points that have the colour properties which match the criteria.

The blue colour band was used for the three automatic methods to prevent the influence of the yellow observation tape on the segmentation result. The flower clusters could not be segmented without including the observation tape when using the red colour band. The blue and green colour band showed some collinearity in segmentation and therefor only

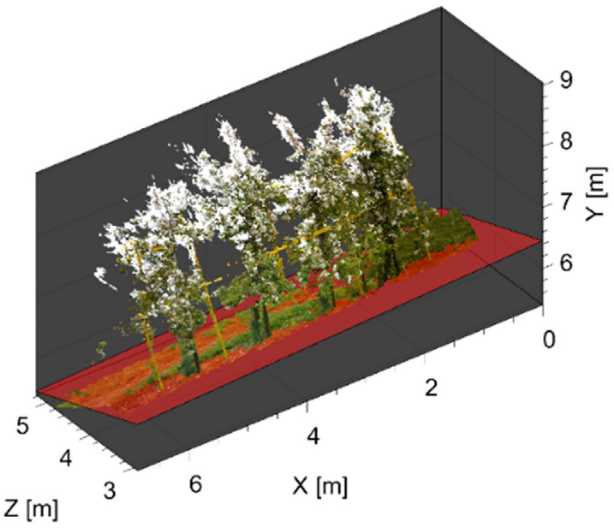

Fig. 7 – Illustration of the ground plane removal function. The red plane represents the ground plane.

blue was used. Using only one colour band also reduces the computation time for more computationally intense algorithms such as hierarchical clustering.

K-Means clustering is an unsupervised machine learning algorithm that finds clusters in data by maximising inter-class variance while minimising intra-class variance. It is also a popular point cloud clustering technique that uses features or attributes from the point clouds (Grilli et al., 2017). The algorithm finds k number of clusters, where k is predetermined. K-means clustering was implemented on the colour properties of the data points from the various subsections to cluster the data points into classes based on the blue colour band. By manual inspecting a collection of images, flower colour feature were found distinguishable in the blue channel. Moreover, flower points in the point cloud showed highest brightness compared with the background, in UAV coloured point clouds. For ground-based point clouds, four components showed significantly different sensitivity to the changes of

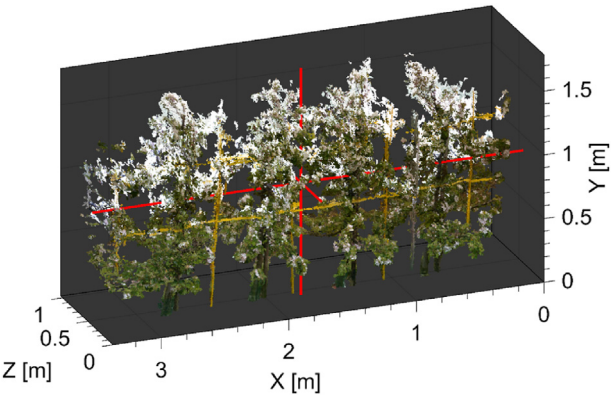

Fig. 8 – Trees fully aligned by PCA alignment function. The red lines indicate the direction of each principle component.

threshold in blue channel, they are the sky, the flowers, the ground with no vegetation and the green objects (leaves and the grass). Therefore, k = 2 was set for PC_U, and k = 4 was determined for ground-based point clouds.

Otsu segmentation is an algorithm that automatically segments data into two parts by automatically defining a threshold between two classes. It has also been proven successful in segmenting point clouds based on RGB information (Jia et al., 2019). Otsu segmentation method divides the data by maximizing the inter-class variance.

Hierarchical clustering is a machine learning algorithm that finds a predefined number of clusters in the data set. It is an unsupervised clustering method that iteratively splits the data set into smaller subsets until every subset contains only one object (Ng & Han, 1994). The hierarchical clustering algorithm iteratively splits the blue colour band data of the subject point cloud to find distinct classes in the data. With the same approach to the determination of k for K-means method, Hierarchical cluster number was predefined as 4 and 2 for ground vehicle and UAV data respectively.

The outcome of flower segmentation consists of only data points belonging to flower blossoms (Fig. 12). The count of data points in the resultant subsections is known as the white index and will be used in the analysis function.

The sub volumes of the trees created in the SOM allow any combination of sub volumes to be used as "puzzle pieces" to build a tree comprising of different point cloud origins. The conclusion can also be made that the Eastern side of PC_E was of higher quality than the Western side since the exposure to the camera was greater and occluding branches and leaves prevent the camera from potentially capturing flower blossoms on the Western side. The same can be said for PC_W. The tree creation function combined the best sides of PC_E and PC_W to create a hybrid combination tree. Therefore, the three eastern segmented subsections from PC_E (east-top, east-middle and east-bottom) and the three western segmented subsections from PC_W (west-top, west-middle and west-bottom) were combined for every tree. Each sub volumes had flower data points only. The output of the tree creation function is explained in a tabulated form (Table 6).

The analysis function used the PC_U and PC_{Hyb} trees as well as ground truth data as inputs to compare the count of the trees to the floridity and the flower cluster ground truth. The aim was to find which segmentation algorithm results in the highest correlation between the white index and ground truth data. Evaluating if the white index of the trees should have a linear relation to the floridity and amount of flower clusters, linear regression was used as the statistical model to find correlation. The combination of data source, flower segmentation method and section of the tree with the highest correlation was used to train and validate a linear regression model for predicting flowering intensity. Due to the limited size of the model, a leave-one-out cross validation (LOOCV) method was used to validate the models. LOOCV is a special case of the resampling procedure K-fold cross validation. It is used to evaluate the performance of machine learning models with small data sets. LOOCV is the case where the fold number equals to the number of observations.

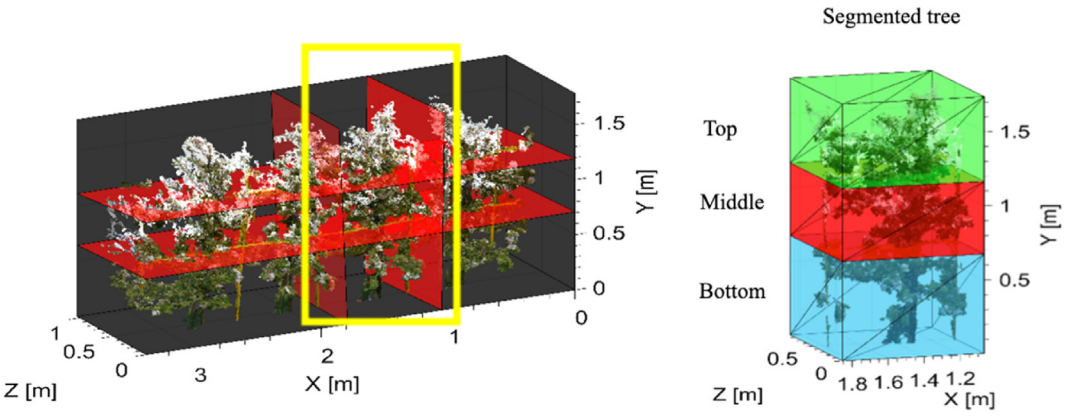

Fig. 9 – Individual tree segmented from row 5. The height layers, defined by observation tape, of the tree in the yellow window are shown on the right.

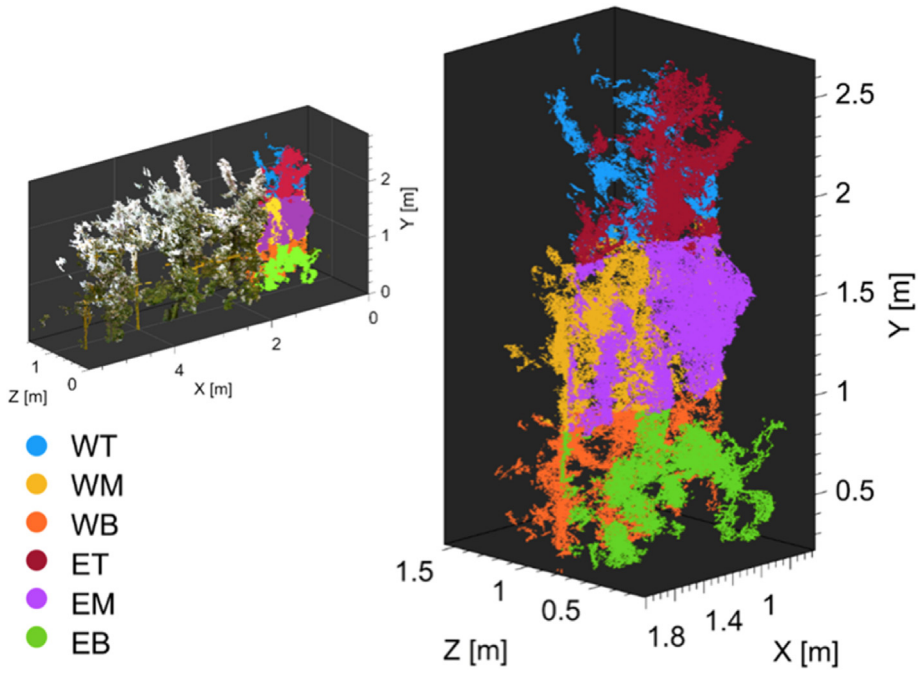

Fig. 10 – Outcome of the SOM model. For each tree, tree points in the six observation volumes are highlighted in different colour. WT: west-top volume of a tree, WM: west-middle volume, WB: west-bottom volume, ET: east-top volume, EM: east-middle volume, EB: east-bottom volume.

3. Results

Two main outcomes of the SOM are the point clouds of individual trees and their subdivided observation volumes. The performance of the individual trees and their subdivided observation volumes segmentation was validated for 107 trees (Fig. 13) and 186 observation volumes (Fig. 14), respectively, by comparing the segmented data points to manually annotated ground truth counts. Outliers are determined for both results. These are trees and observation volumes that could not be spatially segmented properly. Only trees with spatial segmentation accuracy higher than the 25th percentile were used for correlation analysis. A simple user interface (UI) enables the user to select a single tree or its six observation volumes in the row to inspect visually. With this UI, performance of the proposed method on a series of adjacent trees stretching from 17 to 21 was also demonstrated (Fig. 13). In the case of tree 17, inconsistent segmentation results are observed, where the segmentation of East-side is better than that of the West-side. For the result of tree subdivided observation volume segmentation (Fig. 14), it is clear that trees created from GV data can be divided more accurately. This can be attributed to the poor visibility, in the PC_U, of the horizontal yellow observation tape running through the trees. As this is the feature that was used to segment the observation volumes, it makes sense that the performance for PC_U was inferior.

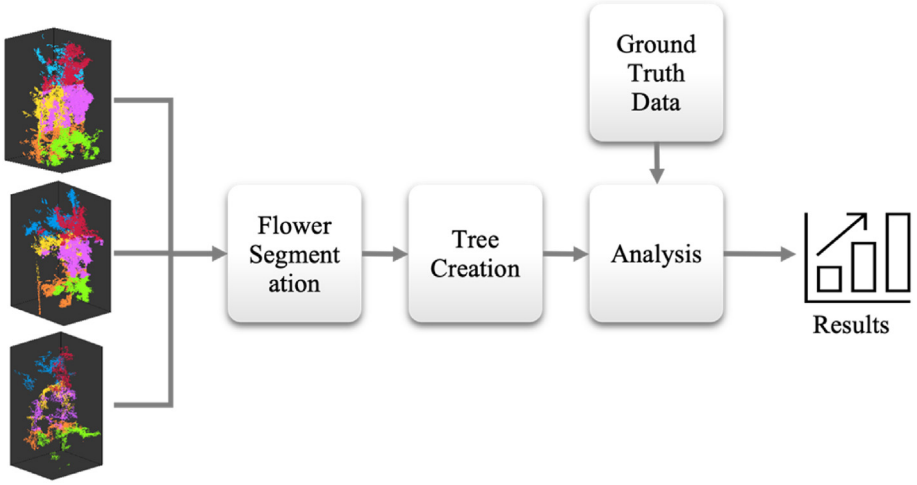

Fig. 11 – Schematic of the flower detection model.

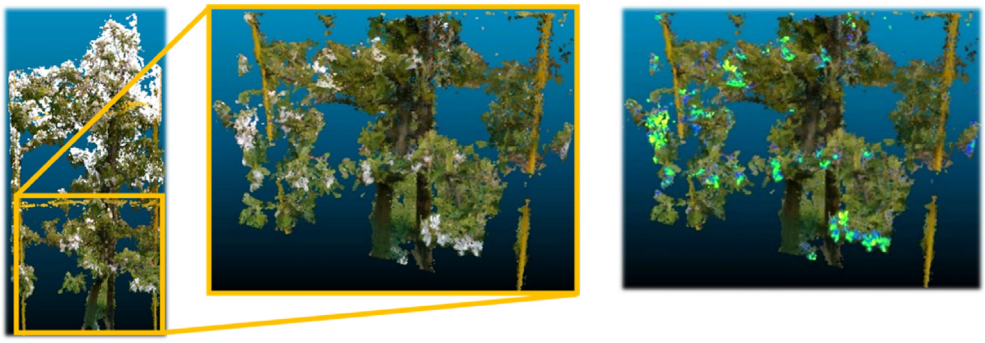

Fig. 12 – Left: coloured point cloud of tree 14. Middle: bottom section of tree 14 with clear flower blossoms present. Right: flower clusters detected from the middle volume using the Otsu segmentation method.

Correlations between white index derived from different segmentation methods and ground truth, flower cluster number and floridity, for PC_{Hyb} and PC_u trees are shown in Table 7. It suggests clearly that for trees in PC_{Hyb} the most effective way to segment flower blossoms from the trees is with Hierarchical Clustering, with the only exception being in the top part of the tree where manual thresholding ($R^2 = 0.67$) supersedes the automatic methods ($R^2 = 0.24$). For PC_U, as hypothesised, optimal method proved to be Otsu segmentation in the blue colour band. However, when the segmentation of the top-middle combination is evaluated, Hierarchical method has a significant performance, with R^2 of 0.78 (Table 7). Floridity is a subjective score given by the expert in the orchards, which is used to describe how intensive the flowering situation is. In general, there is less room for the improvement of the floridity estimation based on PC_U, where no positive correlation is found. And the segmentation performance at tree level is also poor, with a mean R^2 of 0.4. But the PC_U evaluation at upper sub-volumes is significantly good, for example, R^2 for the top and top-middle sub-volume is 0.7 and 0.78, respectively (Table 7).

Results indicate that using Otsu segmentation in the top of the trees with UAV data gives a higher correlation while using Hierarchical clustering with ground vehicle data gives the highest correlation for the middle and bottom of the tree combined. The high correlation can also be checked in details with plots (Fig. 15). For both figures, data points evenly distributed on both sides of the regression line.

To filter out the optimal segmentation methods for flowering intensity estimation, the methods yield the highest correlation with the flowering data, such as the flower estimation at the top volume of the trees, were summarised (Table 8). Additional metrics were calculated for further analysis, such as RMSE and bias. In order to make a comprehensive comparison between PC_U and PC_{Hyb}, the Hierarchical based method was also included. According to the definition

Table 6 – Composition of a tree resulting from the tree creation function.

	PC_U [a]	PC_{Hyb}
Top volume	$ET_{PC_U} + WT_{PC_U}$	$ET_{PC_E} + WT_{PC_W}$
Middle volume	$EM_{PC_U} + WM_{PC_U}$	$EM_{PC_E} + WM_{PC_W}$
Bottom volume	$EB_{PC_U} + WB_{PC_U}$	$EB_{PC_E} + WB_{PC_W}$

[a] WT = west-top volume of a tree, WM = west-middle, WB = west bottom, ET = east-top, EM = east-middle, EB = east-bottom, PC_E = east side of ground vehicle point cloud, PC_W = west side of ground vehicle point cloud, PC_U = UAV point cloud, PC_{Hyb} = the hybrid point cloud constructed from PC_E and PC_W.

of floridity, RMSE 0.4 meet the requirement of practical application when PC_{Hyb}-based Hierarchical method was applied. As the correlation between White index and flower cluster, Otsu and Hierarchical method generally yielded a promising results, though a poor correlation was demonstrated for the cluster estimation at tree level. In addition, small bias was achieved for each units to be estimated. PC_U based estimation provided better performance for the flowers from upper sub-volumes, such as top and top-middle, than that of the PC_{Hyb} based.

The flower cluster prediction model was developed using linear regression as the statistical model. Analysing the results above, two prediction models were developed: one model for the middle + bottom of the tree (Model-MB) and one model for the top of the tree (Model-T). The best performing method for flower cluster detection in the combined middle and bottom part of trees proved to be Hierarchical clustering segmentation in combination with the GV data (PC_{Hyb}). Thus, a linear regression model using Hierarchical clustering was trained and validated (Model-MB). Similarly, a linear regression model using Otsu segmentation was trained and validated for the top of the tree with UAV data (Model-T). The validation results of both models are shown in Table 9.

Although the data sets for both models in Table 9 are limited, the validation suggests that they are statistically significant with P-statistic for both models, lower than 0.05. In terms of relative root mean squared error (RRMSE) both models perform good with values between 10% and 20%. Therefore, an important statistical data fusion was made. Model-MB and Model-T was developed to be used in combination to predict flowering intensity by using each model in a different section of the tree. The validation correlation graph of the two models can be seen in Fig. 16. Within the validation result, three prediction cases yielded a significantly high accuracy, one from the prediction for the middle + bottom volume and two from the top volume. In general, a promising performance was observed for the combination of Model-T and Model-MB.

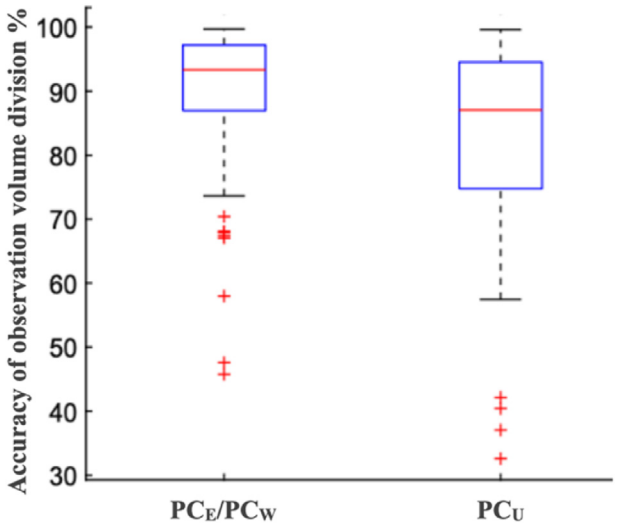

Fig. 14 – Accuracy of the observation volume segmentation. PC_E = east side of ground vehicle derived point cloud; PC_W = west side of ground vehicle derived point cloud; PC_U = UAV derived point cloud.

4. Discussion

In terms of flowering intensity estimation at tree level, this study compared the performance of coloured point clouds derived from two platforms, ground vehicle and UAV. We applied different flower segmentation methods and demonstrated the feasibility of combining UAV and GV imagery to estimate the apple flower clusters. By comparing the flower estimation performance at sub-volume level, the pros and cons of the two point clouds were explored (Table 7). The potential of improving the estimation performance with advanced segmentation methods and larger datasets was also demonstrated.

The major bottleneck for the use of PC_{Hyb} in the preprocessing phase was attributed to the noise in the middle

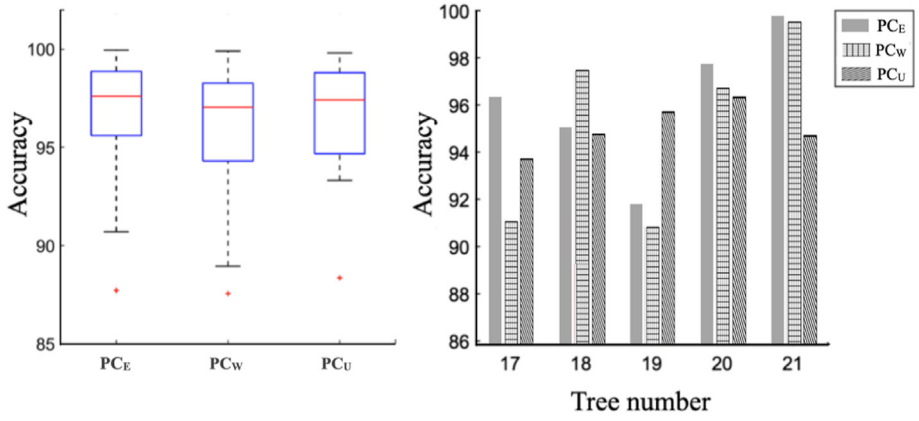

Fig. 13 – Accuracy of the tree segmentation. Left: validation for 107 trees. Right: accuracy of the segmentation to trees ranging from tree 17 to tree 21. PC_E = east side of ground vehicle derived point cloud; PC_W = west side of ground vehicle derived point cloud; PC_U = UAV derived point cloud.

Table 7 – The correlation results for different segmentation algorithms on both PC_U and PC_{Hyb}.

R^2 (white index VS)	Manual threshold		Otsu segmentation (Blue band)		K-means clustering (Blue band)		Hierarchical clustering (Blue band)	
	PC_U	PC_{Hyb}	PC_U	PC_{Hyb}	PC_U	PC_{Hyb}	PC_U	PC_{Hyb}
Floridity	–	0.60	–	0.23	–	0.61	0.21	**0.65**
Total cluster	0.41	0.50	0.43	0.28	0.20	0.52	0.46	**0.61**
Top	0.67	0.41	**0.70**	–	0.33	0.31	0.35	0.24
Middle	–	0.40	–	0.20	0.30	0.15	–	**0.66**
Bottom	–	**0.51**	–	0.36	–	–	–	**0.50**
Top + middle	0.67	0.46	0.67	0.19	0.22	0.44	**0.78**	0.53
Middle + bottom	–	0.51	–	0.27	–	0.47	–	**0.70**

Note: only the first row shows the correlation between white index and floridity, the rest is the correlation with flower cluster number. Values in bold font represent the highest correlation yielded in this study, while cells with no data indicate no positive correlation ($R^2 < 0.1$). PC_U = UAV point cloud. PC_{Hyb} = ground-based hybrid point clouds.

and top part of the trees respectively. The ambiguity that exists in distinguishing similarly coloured flower blossoms from clouds in the sky proved challenging. This is also revealed in Table 7, in which the results concerning the top volume of PC_{Hyb} (R^2: 0.24–0.41) are relatively lower than that of the bottom volume (R^2: 0.36–0.51). Because of the camera view, white data points of white clouds in the sky exist in the PC_{Hyb}, by which the white flower data points could not be extracted precisely. Advanced algorithms need to be tested in follow-up research to deal with this problem (Xu et al., 2018). In addition, a potential solution could be segmenting images firstly rather than a point cloud and using the location of the segmented pixels to determine the location of the data points in the point cloud. These data points can then be labelled as flower points. For the method proposed, however, detailed documentation of camera parameters and setup is required.

In PC_U, on average 71 tie points are present per image compared to 141–158 for the GV images (Table 5). This resulted in a lower quality or less dense point cloud from the UAV data. This explains the inferior performance of the various segmentation models on the UAV data compared to GV data in Table 7. The results of the top of the tree (R^2: 0.33–0.70) compared to the middle (R^2 = 0.30) and bottom (no positive correlation, $R^2 < 0.1$) respectively can be attributed to the same reasons. Tie points in the top of the trees generated with UAV data are far more dense than in the middle and bottom.

This study has shown the potential of flower intensity estimation at tree level with a combination of PC_U and PC_{Hyb} (Fig. 16). However, merging point clouds from multiple views

Table 8 – Correlation between white index and in-situ flowering data, floridity and flower cluster number.

	Methods	PC_U/PC_{Hyb}	R^2	RMSE	bias
Floridity	Hierarchical	PC_{Hyb}	0.65	0.4	0.004
Total cluster	Hierarchical	PC_{Hyb}	0.61	35.3	−1.082
	Hierarchical	PC_U	0.46	29.5	−0.039
Top	Otsu	PC_U	0.70	10.3	−0.026
Middle	Hierarchical	PC_{Hyb}	0.66	9.2	0.769
Bottom	Hierarchical	PC_{Hyb}	0.50	13.9	0.063
Top + middle	Hierarchical	PC_U	0.78	12.0	0.003
Middle + bottom	Hierarchical	PC_{Hyb}	0.70	17.4	1.653

Note: only the first row shows the correlation between white index and floridity, the rest is the correlation with flower cluster number. PC_U = UAV point cloud. PC_{Hyb} = hybrid point clouds constructed from ground vehicle derived point clouds.

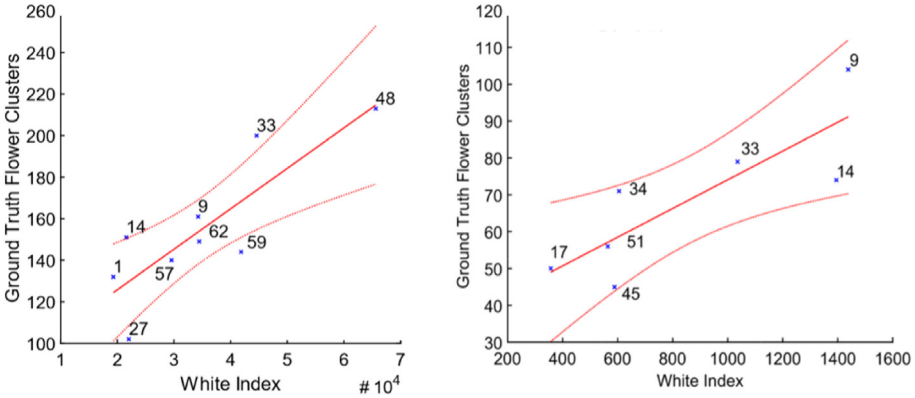

Fig. 15 – Correlation between white index and flower cluster number. Left: white index calculated from bottom + middle volumes of PC_{Hyb} using Hierarchical clustering (R^2 = 0.70, root mean square error (RMSE) = 17.4). Right: white index calculated from top volume of PC_U using Otsu segmentation (R^2 = 0.70, RMSE = 10.3). Blue crosses indicate tree numbers, dotted red lines indicate 95% confidence bounds and solid red line indicates line of best fit. PC_U = UAV point cloud. PC_{Hyb} = hybrid point clouds constructed from ground vehicle derived point clouds.

Table 9 – Validation of linear regression models for flower cluster predictions.

Model – middle + bottom (Model-MB)						
Y = a × X + b	**Estimate**	**Standard error**	**t-statistic**	**p-statistic**	**RMSE**	21.7
b	87.907	19.72	4.458	0.003	**RRMSE**	14.03 (%)
a	0.002	5.27e-04	3.649	0.008	**R^2**	0.66
					R^2 adjusted	0.61
					P-statistic	0.0082
Model – top (Model-T)						
Y = a × X + b	**Estimate**	**Standard error**	**t-statistic**	**p-statistic**	**RMSE**	13.1
b	35.032	11.621	3.015	0.029	**RRMSE**	19.12 (%)
a	0.0392	12.3e-03	3.185	0.024	**R^2**	0.67
					R^2 adjusted	0.60
					P-statistic	0.024

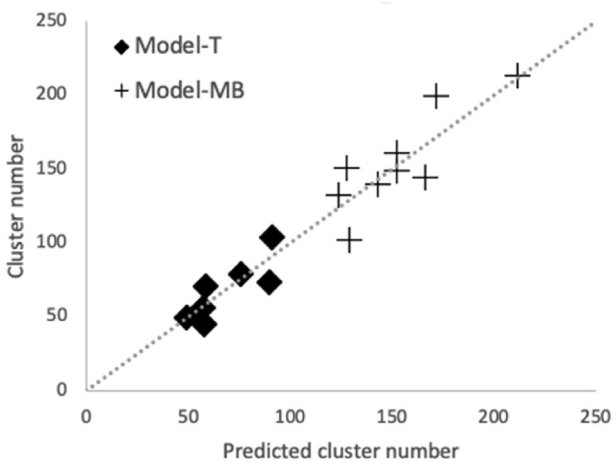

Fig. 16 – Validation of the two flower cluster prediction models. Model-T: flower cluster prediction model for the top volume of a tree; Model-MB: prediction model for the middle + bottom volumes. The grey trendline is a reference line for ideal situation (prediction accuracy = 100%).

required semantic features to be visible and accurately created in 3D space for both point clouds. Using semantic features such as tree trunks is limited in this study since tree trunks are hard to recognize or even not visible in UAV images. Manually inserting reference features in the orchard, visible to both UAV and GV cameras could be useful in spatially merging the two point clouds. An alternative approach to fuse the data between UAV and GV is to use images from both data sources to build a 3D model by feeding image data along with location data from both sources into Agisoft. However, tie points should be found between pairs of images of both sources. And the feature matching algorithms should be able to detect enough features to create tie points between image pairs. In this study, the SURF algorithm could not positively identify matching features between the UAV image and the GV image of the same tree, but has shown good performance for GV images only (Fig. 17). This visually explains the limitation of using both datasets, UAV and ground vehicles datasets, in combination for creating a 3D model in this paper.

Flowering intensity estimation in orchards is still in its infancy. The majority of pre-existing study focuses on the estimation at picture level, which lack the information of the

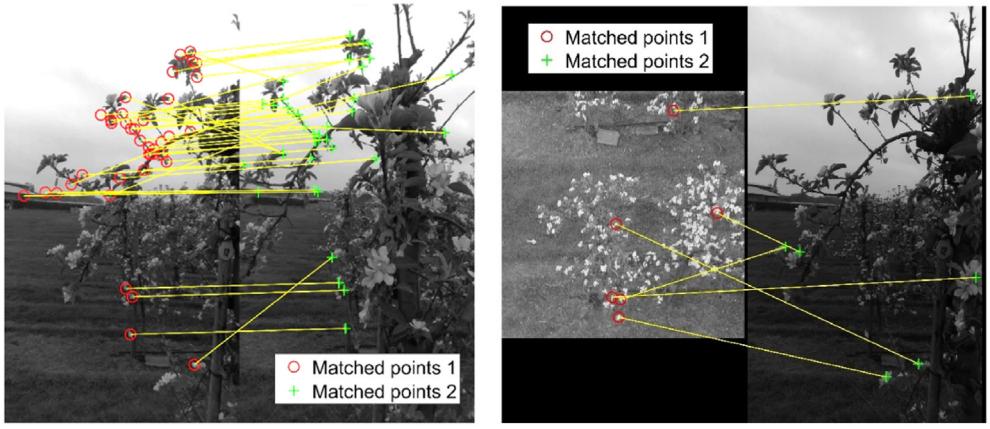

Fig. 17 – SURF algorithm for feature matching in tree 1. Left: feature matching in ground-based images. Right: feature matching between UAV and ground-based images.

real flowers in tree and make it difficult to compare the achievement in this study with. Despite this fact, this study first collected a unique ground truth with not only the total cluster number per tree but also the cluster number in sub-volumes of a tree, overcoming the problems raised by Wang et al. (2021), and further demonstrated the flower estimation at tree level. Compared with the relevant research as mentioned before, the approach proposed is more comparable to Vanbrabant et al. (2020), though the image data used in this study is more complex against them by comparing the average cluster number per tree (198 vs 150). As they mentioned, heavy flowering provides more flower occlusion problems and thus affects the estimation accuracy. The point clouds derived from UAV showed a decrease in the density from the top to the bottom. Whereas the method proposed in this study enable the inspection of the influence of PCs density to the estimation accuracy by comparing the performance from different sub-volumes (Table 7), which further enhance the use of point clouds for fruit tree monitoring. A relative improvement was achieved in this study, while the highest R^2 of the correlation between the image based flower index and the in-field counts of the cluster number achieved by Tubau Comas et al. (2019) and Hočevar et al. (2014) was 0.56 and 0.50 respectively, compared to 0.7 in this study.

The investigated four flower segmentation methods, manual threshold, Otsu, K-means and Hierarchical clustering, have been applied in a variety of flower detection studies (Bhattarai et al., 2020; Dias et al., 2018b; Liakos et al., 2017; Tubau Comas et al., 2019). A comparison of these methods was made and the segmentation results reported from this study were also proved to be reasonably capable of segmenting the flower blossoms (Table 7). In general, the results also returned low bias (Table 8). Otsu segmentation and Hierarchical clustering reported a better performance compared to the other two methods. Because of the camera view, top volume of PC_U is more representative for the flower number in top section of the tree (Table 7). Thus the significant correlation returned from Otsu segmentation, with a R^2 of 0.7, indicates the capability of it for flower point segmentation. Hierarchical clustering shows a stable performance, even for the detection from the bottom volume of PC_{Hyb}, with a R^2 of 0.5. However, unsupervised clustering methods enable the model to be robust against lighting and illumination changes. Different image acquisition locations enable PC_U and PC_{Hyb} show different spatial scales and occlusions. This explain the difference between the floridity and flower cluster estimation performance based on the two type PCs (Table 7). Since floridity was scored by the expert whose view is more close to the camera equipped on the ground vehicle, the floridity estimation from PC_{Hyb} is much better than that of PC_U, with a highest R^2 of 0.65 (Table 7). While a comparable performance of these two PCs was observed for the estimation of flower cluster.

The approach followed in this paper provides a novel solution to estimate the flowering intensity at tree level and yielded a relatively good results (Table 8). It relies on high-quality multi-view imaging, combining UAV and ground based RGB images, and point cloud reconstruction to facilitate the estimation. By constructing point clouds and having GPS/location information available, spatial and even temporal information is readily available. The final products, 3D point cloud representations, provide an understanding of the flowering variability of the whole orchard from not only tree level but also a height based level (Fig. 9 and Table 7). Creating point clouds from two data sources allows flexibility in selecting the best 3D represented regions of trees and analyse them with different segmentation methods based on the height of the tree (Fig. 10). To further validate the approach proposed in this study, higher image resolution will be tested when it comes to the experiment design. Higher resolution RGB cameras could perhaps result in more tie points per image, which contributes to solving the point cloud merging issue discussed above, and therefore point clouds with higher quality will become available. Though aerial images from this study yield a $4{,}000 \times 3000$ resolution, it is far from the advanced RGB camera capable for UAVs. In addition, data size would also be expanded to test the robustness of the method proposed. As shown in Tables 2 and 3, the cluster variance within the orchard is relatively high which limits the robustness of flowering intensity estimation based on small data size. With larger data size, more advanced machine learning technique would also be suitable to test for further improvement of flower estimation, such as the CNN worked on dense point cloud has been reported for cotton bloom detection (Xu et al., 2018).

5. Conclusions

A novel framework based on point clouds derived from UAV and GV images has been designed to automatically estimate flowering intensity in an apple orchard during the full blooming phase of the growth season. The possibility of combing UAV and GV to precisely assess flowering intensity among apple trees, including the spatial variability of flowering intensity in the orchard and even within an individual tree, was demonstrated. Multiple camera angles can in this way be used to complement each other in terms of coverage of certain parts of a tree. Automatic flower cluster estimation at the tree level yielded a R^2 of 0.7, and RMSE lower than 20 for the correlation between the image derived flower index, the white index, and the in-field counts of the cluster number.

The automatic SOM performed well, and high accuracy was achieved in handling the point clouds and spatially segmenting trees and observation volumes. Four flower extraction methods, manual thresholding, Otsu segmentation, K-means and Hierarchical clustering, were explored. Otsu Segmentation and Hierarchical clustering method performed the best for the segmentation in GV and UAV point clouds, respectively. Both models can however be improved with a larger data set. Higher diversity in Floridity scores would also make the linear regression approach for floridity detection more significant.

UAV imagery can be applied for studying less detailed features of the fruit trees. But to analyse detailed features, such as flowering intensity, strict requirements are needed to guarantee the point cloud quality. Depth information can be helpful in improving the accuracy of surface reconstructions in the point cloud, as proven in the study by Dong et al. (2020). For future studies, flying lower and closer to the tree canopy could result in higher quality images for the same resolution

and more tie points per point cloud. We expect that this could also benefit in matching features between the GV images and UAV images and perhaps enable spatial merging of point clouds, if a proper parametrical setup is made for the respective cameras. Grey reference strategies could be implemented to make the model even more robust against illumination differences between UAV and GV data.

Funding

This work was supported by the Netherlands Organization for Scientific Research (NWO) under the framework of the project MARS4Earth [RAAK.PRO03.112].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Aggelopoulou, A. D., Bochtis, D., Fountas, S., Swain, K. C., Gemtos, T. A., & Nanos, G. D. (2011). Yield prediction in apple orchards based on image processing. *Precision Agriculture, 12*(3), 448–456. https://doi.org/10.1007/s11119-010-9187-0

Anderson, N. T., Walsh, K. B., & Wulfsohn, D. (2021). Technologies for forecasting tree fruit load and harvest timing—from ground, sky and time. *Agronomy, 11*(7). https://doi.org/10.3390/agronomy11071409

Bargoti, S., & Underwood, J. P. (2017). Image segmentation for fruit detection and yield estimation in apple orchards. *Journal of Field Robotics, 34*(6, SI), 1039–1060. https://doi.org/10.1002/rob.21699

Bhattarai, U., Bhusal, S., Majeed, Y., & Karkee, M. (2020). Automatic blossom detection in apple trees using deep learning. *IFAC-PapersOnLine, 53*(2), 15810–15815. https://doi.org/10.1016/j.ifacol.2020.12.216

Chen, Y., Lee, W. S., Gan, H., Peres, N., Fraisse, C., Zhang, Y., & He, Y. (2019). Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages. *Remote Sensing, 11*(13). https://doi.org/10.3390/rs11131584

Dias, P. A., Tabb, A., & Medeiros, H. (2018a). Multispecies fruit flower detection using a refined semantic segmentation network. *IEEE Robotics and Automation Letters, 3*(4), 3003–3010. https://doi.org/10.1109/LRA.2018.2849498

Dias, P. A., Tabb, A., & Medeiros, H. (2018b). Apple flower detection using deep convolutional networks. *Computers in Industry, 99*, 17–28. https://doi.org/10.1016/j.compind.2018.03.010

Dong, W., Roy, P., & Isler, V. (2020). Semantic mapping for orchard environments by merging two-sides reconstructions of tree rows. *Journal of Field Robotics, 37*(1), 97–121. https://doi.org/10.1002/rob.21876

Farjon, G., Krikeb, O., Bar Hiller, A., & Alchanatis, V. (2020). Detection and counting of flowers on apple trees for better chemical thinning decisions. *Precision Agriculture, 21*(3), 503–521. https://doi.org/10.1007/s11119-019-09679-1

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381–395. https://doi.org/10.1145/358669.358692

Forshey, C. G. (1986). Chemical fruit thinning of apples. *New York's Food & Life Sciences Bulletin, 116*, 1–7.

Greene, D., & Costa, G. (2013). Fruit thinning in pome- and stone-fruit: State of the art. *Acta Horticulturae, 998*, 93–102. https://doi.org/10.17660/ActaHortic.2013.998.10

Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42W3*, 339–344. https://doi.org/10.5194/isprs-archives-XLII-2-W3-339-2017

Hočevar, M., Širok, B., Godeša, T., & Stopar, M. (2014). Flowering estimation in apple orchards by image analysis. *Precision Agriculture, 15*(4), 466–478. https://doi.org/10.1007/s11119-013-9341-6

Horton, R., Cano, E., Bulanon, D., & Fallahi, E. (2017). Peach flower monitoring using aerial multispectral imaging. *Journal of Imaging, 3*(1). https://doi.org/10.3390/jimaging3010002

Jia, C., Yang, T., Wang, C., Fan, B., & He, F. (2019). A new fast filtering algorithm for a 3D point cloud based on RGB-D information. *PLoS One, 14*, e0220253. https://doi.org/10.1371/journal.pone.0220253

Koirala, A., Walsh, K. B., Wang, Z., & Anderson, N. (2020). Deep learning for mango (*Mangifera indica*) panicle stage classification. *Agronomy, 10*(1). https://doi.org/10.3390/agronomy10010143

Krikeb, O., Alchanatis, V., Crane, O., & Naor, A. (2017). Evaluation of apple flowering intensity using color image processing for tree specific chemical thinning. *Advances in Animal Biosciences, 8*, 466–470. https://doi.org/10.1017/S2040470017001406

Liakos, V., Tagarakis, A., Aggelopoulou, K., Fountas, S., Nanos, G., & Gemtos, T. (2017). In-season prediction of yield variability in an apple orchard. *European Journal of Horticultural Science, 82*, 251–259.

Lopez-Granados, F., Torres-Sanchez, J., Jimenez-Brenes, F. M., Arquero, O., Lovera, M., & de Castro, A. I. (2019). An efficient RGB-UAV-based platform for field almond tree phenotyping: 3-D architecture and flowering traits. *Plant Methods, 15*(1). https://doi.org/10.1186/s13007-019-0547-0

Ng, R. T., & Han, J. (1994). Efficient and effective clustering methods for spatial data mining. In *Proceedings of the 20th International Conference on Very Large Data Bases* (pp. 144–155).

Rodríguez, A. S., Rodríguez, B. R., Rodríguez, M. S., & Sánchez, P. A. (2018). Laser scanning and its applications to damage detection and monitoring in masonry structures. *Long-Term Performance and Durability of Masonry Structures: Degradation Mechanisms, Health Monitoring and Service Life Design*, 265–285. https://doi.org/10.1016/B978-0-08-102110-1.00009-1

Sultani, Z. N., & Ghani, R. F. (2015). Kinect 3D point cloud live video streaming. *Procedia Computer Science, 65*, 125–132. https://doi.org/10.1016/j.procs.2015.09.090

Sun, G., Wang, X., Ding, Y., Lu, W., & Sun, Y. (2019). Remote measurement of apple orchard canopy information using unmanned aerial vehicle photogrammetry. *Agronomy, 9*(11). https://doi.org/10.3390/agronomy9110774

Sun, K., Wang, X., Liu, S., & Liu, C. (2021). Apple, peach, and pear flower detection using semantic segmentation network and shape constraint level set. *Computers and Electronics in Agriculture, 185*. https://doi.org/10.1016/j.compag.2021.106150

Torres-Sánchez, J., López-Granados, F., Borra-Serrano, I., & Peña, J. M. (2018). Assessing UAV-collected image overlap influence on computation time and digital surface model accuracy in olive orchards. *Precision Agriculture, 19*(1), 115–133. https://doi.org/10.1007/s11119-017-9502-0

Tubau Comas, A., Valente, J., & Kooistra, L. (2019). Automatic apple tree blossom estimation from UAV RGB imagery. *International*

Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives, *42*(2/W13), 631–635. https://doi.org/10.5194/isprs-archives-XLII-2-W13-631-2019

Vanbrabant, Y., Delalieux, S., Tits, L., Pauly, K., Vandermaesen, J., & Somers, B. (2020). Pear flower cluster quantification using RGB drone imagery. *Agronomy*, *10*(3). https://doi.org/10.3390/agronomy10030407

Wang, X. (Annie), Tang, J., & Whitty, M. (2021). DeepPhenology: Estimation of apple flower phenology distributions based on deep learning. *Computers and Electronics in Agriculture*, *185*, 106123. https://doi.org/10.1016/j.compag.2021.106123

Wu, D., Lv, S., Jiang, M., & Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. *Computers and Electronics in Agriculture*, *178*. https://doi.org/10.1016/j.compag.2020.105742

Xiao, C., Zheng, L., Sun, H., Zhang, Y., & Li, M. (2014). Estimation of the apple flowers based on aerial multispectral image. In *American society of agricultural and biological engineers annual international meeting 2014* (Vol. 6, pp. 4426–4433). ASABE.

Xu, R., Li, C., Paterson, A. H., Jiang, Y., Sun, S., & Robertson, J. S. (2018). Aerial images and convolutional neural network for cotton bloom detection. *Frontiers of Plant Science*, *8*, 2235. https://doi.org/10.3389/fpls.2017.02235

Yuan, W., & Choi, D. (2021). UAV-based heating requirement determination for frost management in apple orchard. *Remote Sensing*, *13*(2). https://doi.org/10.3390/rs13020273

Zhang, C., Valente, J., Kooistra, L., & Guo, L. (2021). Orchard management with small unmanned aerial vehicles: A survey of sensing and analysis approaches. In *Precision agriculture (issue 0123456789)*. US: Springer. https://doi.org/10.1007/s11119-021-09813-y.